

# SCIENTIFIC REPORTS



OPEN

## Molecular Rationale behind the Differential Substrate Specificity of Bacterial RND Multi-Drug Transporters

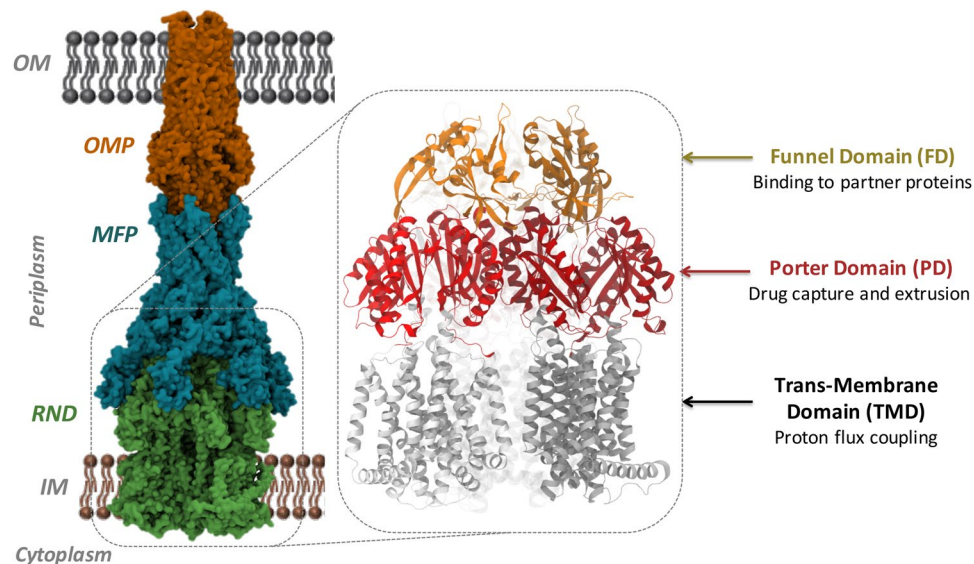
Venkata Krishnan Ramaswamy<sup>1</sup>, Attilio V. Vargiu <sup>1</sup>, Giuliano Malloci<sup>1</sup>, Jürg Dreier<sup>2</sup> & Paolo Ruggerone <sup>1</sup>

Resistance-Nodulation-cell Division (RND) transporters AcrB and AcrD of *Escherichia coli* expel a wide range of substrates out of the cell in conjunction with AcrA and TolC, contributing to the onset of bacterial multidrug resistance. Despite sharing an overall sequence identity of ~66% (similarity ~80%), these RND transporters feature distinct substrate specificity patterns whose underlying basis remains elusive. We performed exhaustive comparative analyses of the putative substrate binding pockets considering crystal structures, homology models and conformations extracted from multi-copy  $\mu$ s-long molecular dynamics simulations of both AcrB and AcrD. The impact of physicochemical and topographical properties (volume, shape, lipophilicity, electrostatic potential, hydration and distribution of multi-functional sites) within the pockets on their substrate specificities was quantitatively assessed. Differences in the lipophilic and electrostatic potentials among the pockets were identified. In particular, the deep pocket of AcrB showed the largest lipophilicity convincingly pointing out its possible role as a lipophilicity-based selectivity filter. Furthermore, we identified dynamic features (not inferable from sequence analysis or static structures) such as different flexibilities of specific protein loops that could potentially influence the substrate recognition and transport profile. Our findings can be valuable for drawing structure (dynamics)-activity relationship to be employed in drug design.

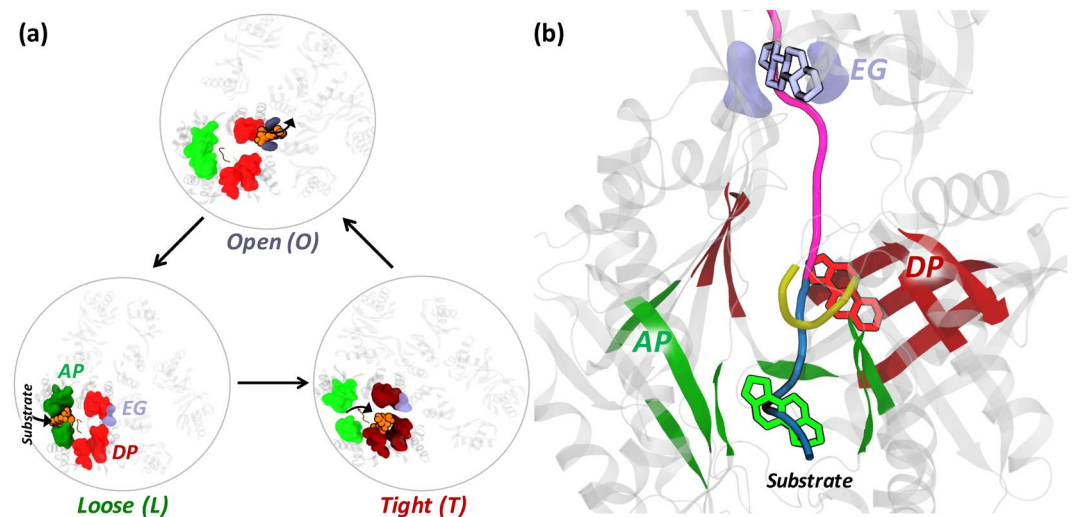
Antimicrobial resistance has re-emerged as one of the major challenges to public health worldwide, especially due to the spread of multidrug-resistant (MDR) or even pan-resistant Gram-negative pathogenic bacteria<sup>1</sup>. The intrinsic drug resistance shown by these bacteria can be largely attributed to the primary barrier imposed by their membranes, endowed with chromosomally encoded molecular filters (porins) and drug efflux pumps<sup>2</sup>. Among these, MDR efflux pumps transport a wide range of structurally dissimilar substrates including antibiotics from various classes, posing a major concern in clinical therapy<sup>3,4</sup>. In particular, the Resistance-Nodulation-cell Division (RND) superfamily members are notoriously known for the extremely wide substrate specificity<sup>3,5-8</sup> and are considered to be involved in both intrinsic and acquired MDR. The RND pump complexes span the entire periplasmic space from the inner membrane (IM) to the outer membrane (OM) by forming tripartite systems<sup>9-13</sup> comprising an RND transporter protein embedded in the IM, an adaptor protein (a.k.a. membrane fusion protein, MFP) located in the periplasmic space, and an outer-membrane protein (OMP) constituting a long alpha-helical and beta-barrel tunnel (Fig. 1).

AcrB is the best characterized RND transporter<sup>8</sup>, and its structure has been solved by several labs both without and with bound substrates and inhibitors<sup>14-18</sup>. Structurally, AcrB is an asymmetric trimer resembling a jellyfish with each protomer comprising a total of 3 domains<sup>8</sup> (Fig. 1): (i) a trans-membrane domain consisting of 12  $\alpha$ -helices embedded in the IM, where energy conversion takes place via proton coupling; (ii) a pore (porter) domain located in the periplasm, where substrate recruitment and transport occur; and (iii) a periplasmic funnel domain, which connects the RND transporter to the OMP via the assembly of MFPs<sup>19</sup> in the constituted pump. It has been proposed that substrate transport in these proteins follows a “functional rotation mechanism”

<sup>1</sup>Department of Physics, University of Cagliari, Cittadella Universitaria, S.P. Monserrato-Sestu km 0.700, I-09042, Monserrato, CA, Italy. <sup>2</sup>Basilea Pharmaceutica International Ltd., Grenzacherstrasse 487, 4058, Basel, Switzerland. Correspondence and requests for materials should be addressed to P.R. (email: [paolo.ruggerone@dsf.unica.it](mailto:paolo.ruggerone@dsf.unica.it))



**Figure 1.** Structure of the RND tripartite pump AcrAB-TolC. AcrAB-TolC pump is shown here with AcrB as the RND component, AcrA as the MFP and TolC as the OMP. In the inset, the general structure of an RND transporter is displayed, where the three main domains (TMD, PD, and FD) are indicated.



**Figure 2.** Proposed functional rotation mechanism of substrate extrusion by RND transporters. (a) Top view illustrating the different conformations assumed by AP, DP and Exit Gate (EG) during the three cycles of the functional rotation mechanism. The substrate (orange) is represented by van der Waals spheres and its pathway along the cycle is indicated by short black arrows. The key regions housing the substrate are coloured dull when binding to the substrate, bright otherwise. (b) Front view of the putative substrate transport pathway from AP to EG going through DP. The pathway is shown as thick tube, coloured in blue and magenta for the stages of the transport cycle associated to the *Loose*→*Tight* and *Tight*→*Open* conformational changes, respectively. The substrate is represented by sticks coloured green, red and iceblue when interacting with the AP, DP and EG (also coloured green, red and iceblue), respectively. The G-loop separating AP and DP is shown as a yellow cartoon.

(Fig. 2) in which concerted but not necessarily synchronous cycling of the protomers occurs through any of the so far identified asymmetric states: *Loose* (*L*) (a.k.a. *Access*) in which substrates bind to a peripheral site named access pocket (AP); *Tight* (*T*) (a.k.a. *Binding*) in which substrates bind to a more deep pocket (DP); and *Open* (*O*) (a.k.a. *Extrusion*) in which the substrate is released into the central funnel leading towards the OMP<sup>14,20,21</sup>. The two pockets (Fig. 2) were previously identified in AcrB as the binding sites responsible for the recognition and selectivity of different molecules<sup>15,22–24</sup>. Namely, the AP and DP have been hypothesized to be responsible for the recognition of high-molecular-mass and low-molecular-mass compounds, respectively<sup>15</sup>. They are separated by a G-rich (a.k.a. switch) loop whose flexibility has been shown to be important for the transport of high-molecular-mass molecules<sup>15,16</sup>.

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

**Table 1.** Antibiotic substrate specificities of the paralog RND transporters AcrB and AcrD from *E. coli*<sup>22, 25, 46, 80–84</sup>. Classes of compounds are indicated, with examples of specific compounds within parentheses (2D chemical structures of these compounds are shown in Supplementary Figure S3).

Antibiotic substrate type (Physicochemical property)	<i>E. coli</i>	
	AcrB	AcrD
Hydrophobic	+	–
Hydrophilic	–	+
Amphiphilic	+	+

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

AcrD is a close homolog of AcrB in *Escherichia coli* with an overall sequence identity (similarity) of nearly 66% (80%) (Supplementary Figs S1–S2). These moderate differences however impact on their substrate specificity patterns, which overlap only partially (Table 1 and Supplementary Fig. S3). While certain substrates like most of the beta-lactam antibiotics are common, macrolides and tetracyclines are transported by AcrB but not by AcrD, which instead exports aminoglycosides, in turn not recognized by AcrB. Categorizing the typical substrates of the two transporters on the basis of their physicochemical properties (Table 2) highlights that they are essentially hydrophobic and hydrophilic for AcrB and AcrD, respectively, while both transporters might shuttle out amphiphilic compounds.

However, such a simplistic classification is not of help either to improve our basic knowledge on RND transporters or in drug design efforts. Achieving a deeper level of information would be highly desirable, and a first step towards this goal consists in mapping the differences in substrate specificities between these two proteins in terms of defined structural, chemical and dynamic features of their putative substrate-binding pockets, whose link has not been traced yet. From a domain-wise perspective, two previous studies attempted to identify substrate recognition site(s) in these RND pumps by using chimeric analysis<sup>22, 25</sup>. The importance of periplasmic loop regions in RND pumps was pointed out by Elkins and Nikaido<sup>25</sup>, Mao *et al.*<sup>26</sup>, Eda *et al.*<sup>27</sup> and Kobayashi *et al.*<sup>22</sup>. In particular, Kobayashi *et al.*<sup>22</sup> identified a few residues in the AP as potential determinants of specificity towards negatively charged beta-lactams (aztreonam, carbenicillin, and sulbenicillin). Namely, by replacing three residues in AcrB with the corresponding ones in AcrD (Q569R, I626R, and E673G), the authors were able to confer to the former transporter the ability to recognize anionic beta-lactams, typical substrates of the latter protein.

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. This void of knowledge traces back mainly to the lack of experimental structures of AcrD and of co-crystal structures of any RND transporter with compounds belonging to beta-lactam, fluoroquinolone or aminoglycoside classes. On the other hand, computational modeling, in particular all-atom MD simulations, have already proven to be insightful in addressing the molecular mechanisms of RND transporters<sup>8, 28–38</sup>. Moreover, given the overall good sequence identity and similarity between AcrB and AcrD of *E. coli*, reliable computational modeling of AcrD and related structure-based studies are possible.

Prompted by this consideration and with the aim to explain in a more deep and informative meaning the substrate specificities of AcrB and AcrD in terms of matching properties with the corresponding (substrate) binders, we performed a systematic comparison of the physicochemical nature of the main putative substrate binding sites (AP and DP) (Fig. 2) between AcrB and AcrD. Importantly, beside crystal structures or homology models, we included for both transporters conformations extracted from extensive multi-copy  $\mu$ s-long molecular dynamics (MD) simulations. This robust computational setup allowed us to extend the realm of structure-function relation to account for subtle interplay between behavior of solvent, charge distribution, structural changes associated with the time-evolution of the system under physiological-like conditions. We characterized and compared the molecular properties (pocket descriptors) of the binding pockets such as their flexibility, accessible binding volume, lipophilic index, electrostatic potential, hydration and multi-functional sites. In particular, we identified dynamic features not inferable from simple sequence analysis, such as the positional flexibility of a loop lining the base of AP and likely playing a key role in regulating access and transport of substrates in these RND transporters. We also pinpointed specific differences in the lipophilic and electrostatic potentials in the binding pockets of these transporters, which complement the physicochemical properties of the known substrates of these pumps and are present also when dynamics of the pocket is accounted for. In particular, in AcrB an electrostatic funnel with negative gradient leading from the periplasm to the centre of the AP shows up in our configurations whilst it is absent in AcrD. Additionally, the DP of AcrB features a more lipophilic character, if compared to AcrD, pointing out the involvement of this pocket as a lipophilicity-based selectivity filter.

The correlation of the different specificity patterns of these two transporters to the dynamic physicochemical and topographical properties of their multi-functional recognition sites could be highly informative for drug design attempts<sup>8</sup>.

## Results

**Sequence Assessment.** Bacteria have an inherent ability to change their genetic makeup and adapt themselves in response to adverse environmental stress. In order to extend the results of our study to Acr proteins extracted from *E. coli* other than the specific one used here, we determined the presence and distribution of conserved regions in AP and DP of all the available AcrB and AcrD sequences in the UniProtKB (October 2016) (<http://www.uniprot.org/blast/>). Shannon entropy<sup>39</sup> was computed for both the Acr proteins and expressed in terms of H factor. This descriptor is commonly used to quantify sequence conservation considering the probability of occurrence of an amino acid at each site in a sequence alignment. Multiple sequence alignment and Shannon entropy analysis together pointed out an overall high sequence conservation of these proteins in all *E. coli* strains. In particular, all the H factors associated with the Shannon entropy were lower than 1, indicating a high degree of conservation.

**Molecular dynamics simulations of AcrB and AcrD.** MD simulations of AcrB were done using the high-resolution asymmetric crystal structure (PDB ID: 4DX5<sup>16</sup>), each protomer being in *Loose*, *Tight*, and *Open* conformations, respectively. Since the structure of AcrD has not been experimentally resolved yet, we generated a homology model of AcrD based on the same AcrB crystal structure (4DX5) as its template using Modeller 9.13<sup>40</sup> (see Methods section). According to RMSD analyses of the complete trimeric protein backbone and of each protomer in relation to the initial structure (Supplementary Fig. S4), we determined the equilibration time of ~0.5 μs to be most suitable for both AcrB and AcrD. The cluster representatives (see respectively Supplementary Figures S5 and S6 for the sampling of AP and DP clusters in AcrB and AcrD along the MD simulations) extracted from the equilibrated trajectories of AcrB and AcrD were used to characterize the distribution of accessible binding volume, molecular lipophilicity, electrostatic potential and multi-functional sites (MFS). Hydration analyses were performed on equilibrium trajectories. Although the level of confidence in homology models cannot be as high as that in experimental structures, we have thoroughly validated the AcrD structures by using state-of-the-art bioinformatic tools (details are reported in Methods and in Supplementary Information, see in particular Supplementary Table S1). The stability of the AcrD model as well as its suitability for subsequent analyses were validated in two independent μs-long MD simulations. This multiple validation of the AcrD model offers a fairly good confidence in the representativity of the clusters extracted from MD trajectories as configurations explored by the system. In the following, we present the results from these analyses on the AP and DP.

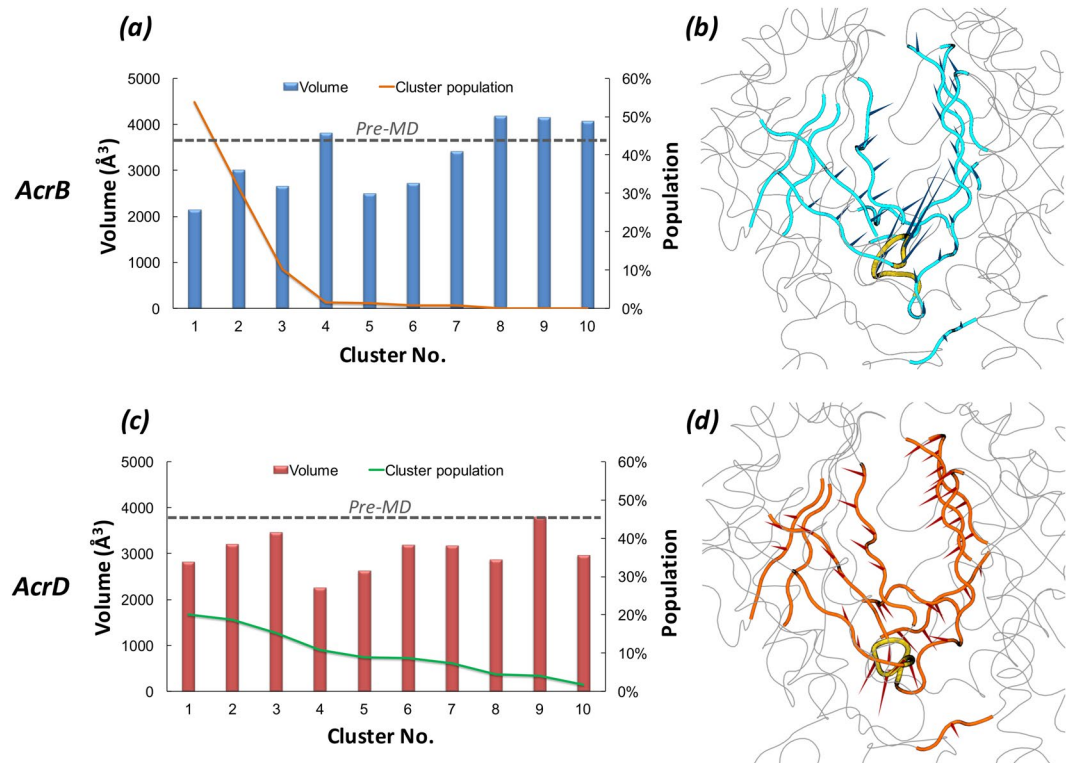
**Access Pocket of the *Loose* protomer.** The AP is a proximally located pocket close to the periplasm in the putative substrate transport pathway of RND pumps (Fig. 2) and also likely the site of recognition for high-molecular-mass compounds<sup>15</sup>. In order to identify if any of the physicochemical properties could be used to differentiate between the APs and in addition to determine which substrate could likely be recognized by these APs, we calculated the following descriptors on the AP of *Loose* protomer in both AcrB and AcrD.

**Pocket Volume and Shape.** 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 putative binding pocket is governed by multiple factors, the primary being the accessible binding volume. Especially with promiscuous proteins like the RND transporters, one would expect that a large binding site with a reasonable degree of plasticity will facilitate binding of molecules with a wide range of sizes. The pocket volumes (Fig. 3a and c and Supplementary Table S2) of pre-MD structures (PDB code 4DX5<sup>16</sup> for AcrB and the final optimized model for AcrD used as starting configuration for MD simulations) and the clusters extracted from MD did not, per se, show any relevant differences. Moreover, both the volumes and the minimal projection areas of AP in AcrB and AcrD are much larger than those of the largest substrates transported by these pumps<sup>41</sup> (Supplementary Table S2). However, principal component analysis performed on the equilibrium MD trajectories revealed a slightly different flexibility of the AP in the two proteins, namely the pocket in AcrB showed larger rearrangements in the loop residues 675 to 678 lining the base of AP (Fig. 3b) whereas in the case of AcrD (Fig. 3d) this region displayed lower flexibility and so was for the entire AP.

**Molecular Lipophilicity Potential (MLP) and Lipophilic Index (LI).** The calculation of the LIs of the AP showed this pocket to be of higher lipophilic nature in AcrB than in AcrD (Table 3 and Supplementary Fig. S7). This is compatible with the (at least partial) hydrophobic character required for AcrB substrates. However, the specific chemical environment of the AP is neither entirely hydrophobic nor entirely polar in both the proteins (Fig. 4a,b). Interestingly, different conformations of the AP displayed similar values of the LIs, so that the relatively higher lipophilicity of AcrB with respect to AcrD turned out to be a robust feature compared to the flexibility of the AP (Table 3). According to the MLP calculated for the representatives of the most populated structural clusters, regions of relatively high lipophilicity for AcrB were located close to the hydrophobic trap (HP-trap) lined by residues F136, F178, F610, F615 and F628<sup>17,29</sup>, and in a region at the border with the putative entrance known as Vestibule<sup>42</sup>. In contrast, no predominant spots were recognizable for AcrD (Fig. 4a,b).

**Electrostatic Potential.** The long range potential due to electrostatic interactions makes them a vital component of molecular recognition between molecules. The electrostatic potentials calculated on the molecular surfaces of the AP of AcrB and AcrD are shown in Fig. 4c,d. The left and right panels collect the results for the pre-MD structures and for the most populated cluster of each system, respectively. Concerning the pre-MD structures (Fig. 4c), positively charged patches were predominant within the AP of AcrD, while the same region of AcrB featured a





**Figure 3.** Volume dynamics of AP in the *Loose* protomer of AcrB and AcrD. (Left Panels) Distribution of the volume of AP within the *Loose* protomer of AcrB (a) and AcrD (c), calculated for the 10 top cluster representatives extracted from equilibrium MD trajectories. Histograms refer to the values of the volume, lines to the relative population of the corresponding clusters. The volumes calculated for the pre-MD structures of AcrB and AcrD are shown as dashed line. (Right Panels) Porcupine plots representing collective motions along the first principal-component eigenvector for AP in AcrB (b) and AcrD (d) simulations shown as arrows ( $>2$  Å) attached to  $C\alpha$  atoms indicating the magnitude of the corresponding eigenvalues. The loop lining the base of AP ('bottom-loop') showing large rearrangement in AcrB but not in AcrD are coloured yellow in both.

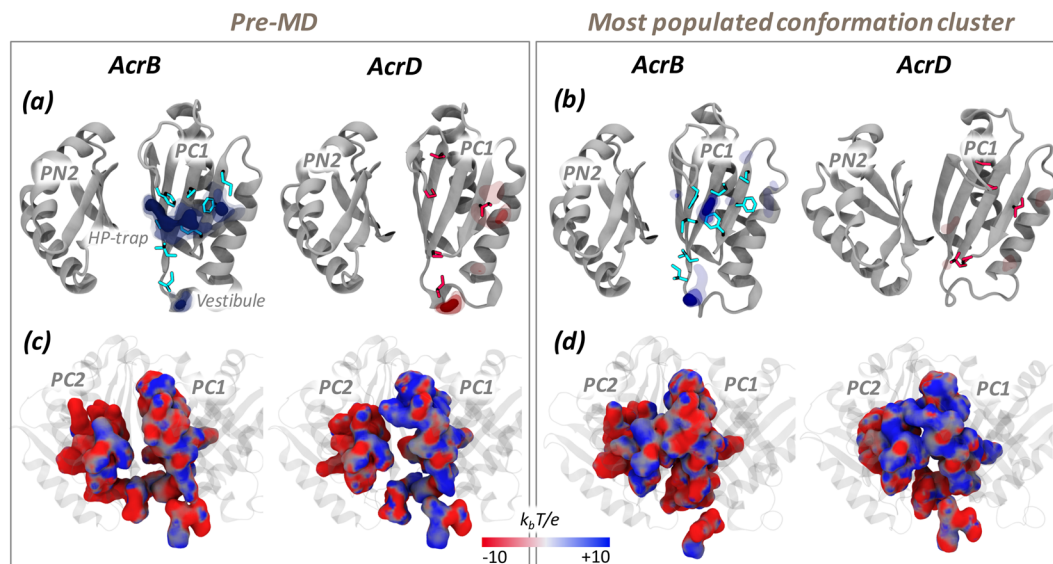
System	LI	
	Pre-MD	MD clusters
AcrB	7.2	$7.0 \pm 1.0$
AcrD	1.2	$1.6 \pm 0.6$

**Table 3.** Lipophilic indexes of AP in the *Loose* protomer of AcrB and AcrD. For AcrB, the pre-MD structure corresponds to the crystal structure identified by PDB code 4DX5<sup>16</sup> while for AcrD it is the final optimized homology model used as starting configuration for MD simulations.

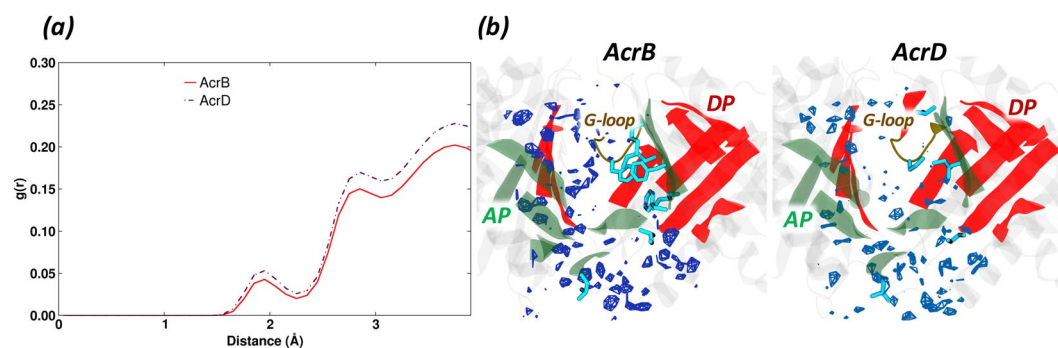
more even distribution of positive vs. negative charges. Importantly, the partial closure of the pockets seen in the MD simulations of the apo proteins did not influence these main findings (Supplementary Fig. S8).

**Hydration Analysis.** Characterizing the hydration profiles around the binding pockets of these proteins helps to effectively understand the molecular mechanism of interaction of water molecules penetrating the pocket in a dynamic manner. The radial distribution function (RDF) profile around the AP residues of AcrB and AcrD were rather similar with only a minor difference in the intensity of hydration (Fig. 5a). The first solvation shell was observed around 1.9 Å in both the proteins with a slightly reduced probability in AcrB. The spatial distribution function (SDF) calculated on the trajectory of the most populated cluster extracted from MD simulations, however, featured no water density spots near the hydrophobic residues in AP of AcrB but showed a higher number of dense regions in AcrD at identical density isovalues (Fig. 5b).

**Deep Pocket of the *Tight* protomer.** The DP is a more deeply located cavity within the putative substrate transport pathway of RND pumps (Fig. 2), and is likely the recognition site for low-molecular-mass compounds<sup>15</sup>. According to the crystal structures, this pocket exists in a collapsed state in the *Loose* and *Open* protomers but is wide open in the *Tight* protomer; therefore, all the analyses concerning this site were performed on the *Tight* protomers of AcrB and AcrD. Based on primary sequence analysis, most of the hydrophobic residues in the DP



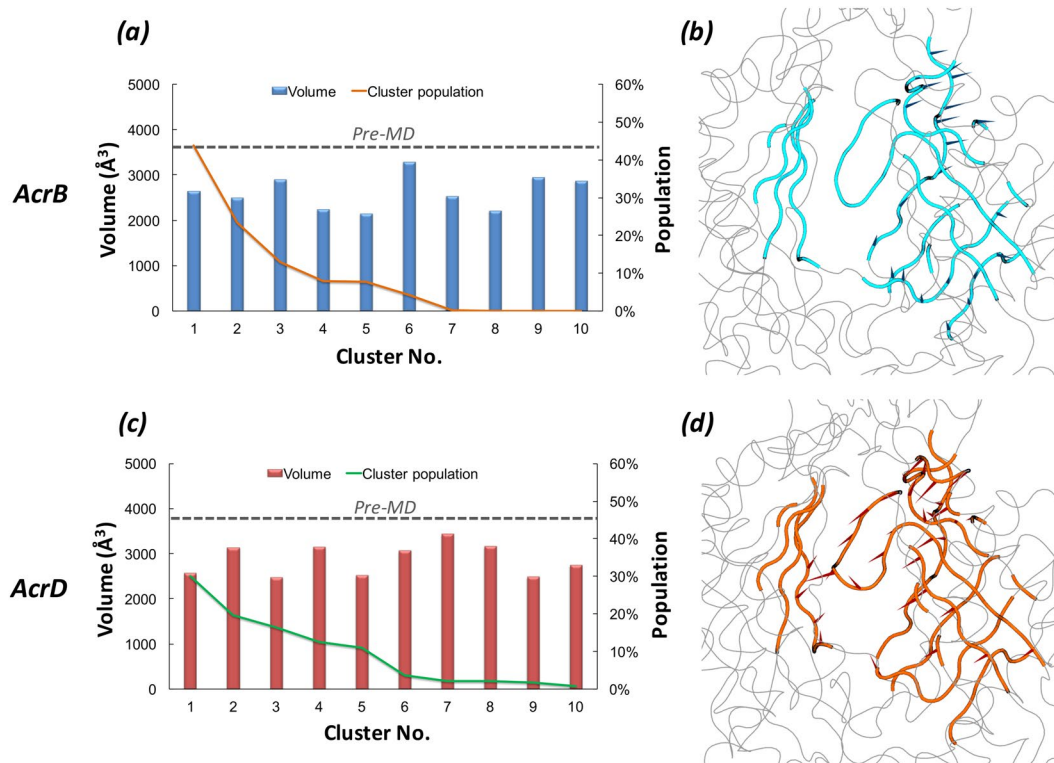
**Figure 4.** MLP and electrostatic potential of AP in the *Loose* protomer of AcrB and AcrD. MLP isosurfaces observed within 4 Å of AP in the *Loose* protomer of AcrB (blue) and AcrD (red) in pre-MD (a) and the representatives of the most populated cluster (b) as seen from the centre of the protomer. The hydrophobic/aromatic residues in the AP are shown as sticks in the structures. Isosurfaces at 0.75 (solid), 0.5 (dark transparent) and 0.25 (light transparent) are shown in blue (AcrB) or red (AcrD). The HP-trap and Vestibule sites are also labeled in the pre-MD structure of AcrB. The electrostatic potential plotted on the molecular surface representation of AP in the Acr proteins in the pre-MD (c) and the most populated cluster representative (d) as seen from the periplasmic front of the protomer. The colour code is red to blue from negative ( $-10 k_B T/e$ ) to positive ( $+10 k_B T/e$ ) potential, where  $k_B$  is the Boltzmann constant, T is the absolute temperature and e is the electron charge.



**Figure 5.** Hydration of AP in the *Loose* protomer of AcrB and AcrD. (a) Comparison of RDF profiles of water oxygen atoms around AP (all atoms) in the *Loose* protomer of AcrB (red solid line) and AcrD (brown dash-dotted line) extracted from the equilibrium MD trajectories. (b) Comparison of SDF of waters within the AP of *Loose* protomer. The SDF was calculated over the configurations forming the most populated cluster of AcrB (left) and AcrD (right). The isosurfaces are shown at density isovalue of 6, meaning that the represented surfaces correspond to 6 times higher average number density of solvent molecules than bulk (see Subsection Hydration in Methods). The AP and DP are marked in green and red while the G-loop in yellow cartoon representations. The hydrophobic/aromatic residues of the pocket are shown as cyan sticks in the respective structures.

of AcrB are replaced by polar/charged amino acids in AcrD (Supplementary Figs S1–S2). The ensuing effects on the physicochemical environment of the DP were thus characterized by the aforementioned pocket descriptors.

**Pocket Volume and Shape.** As for the AP, also the DP showed a partial closure during dynamics yet displaying volumes and minimal projection areas (Fig. 6a and c and Supplementary Table S3) large enough to accommodate its ligands. The cluster distribution was similarly slightly more extended for AcrD than for AcrB. The principal component analysis data showed the DP in AcrD with essential dynamics spread throughout the pocket unlike the less dynamic and more localized motions of DP in AcrB (Fig. 6b and d).



**Figure 6.** Volume dynamics of DP in the *Tight* protomer of AcrB and AcrD. (Left Panels) Volume distribution of DP in the *Tight* protomer of AcrB (a) and AcrD (c) over the simulation timescale. (Right Panels) Porcupine plots of the first principal-component eigenvector for DP in AcrB (b) and AcrD (d) simulations shown as arrows ( $>2 \text{\AA}$ ) attached to  $\text{C}\alpha$  atoms indicating the magnitude of the corresponding eigenvalues.

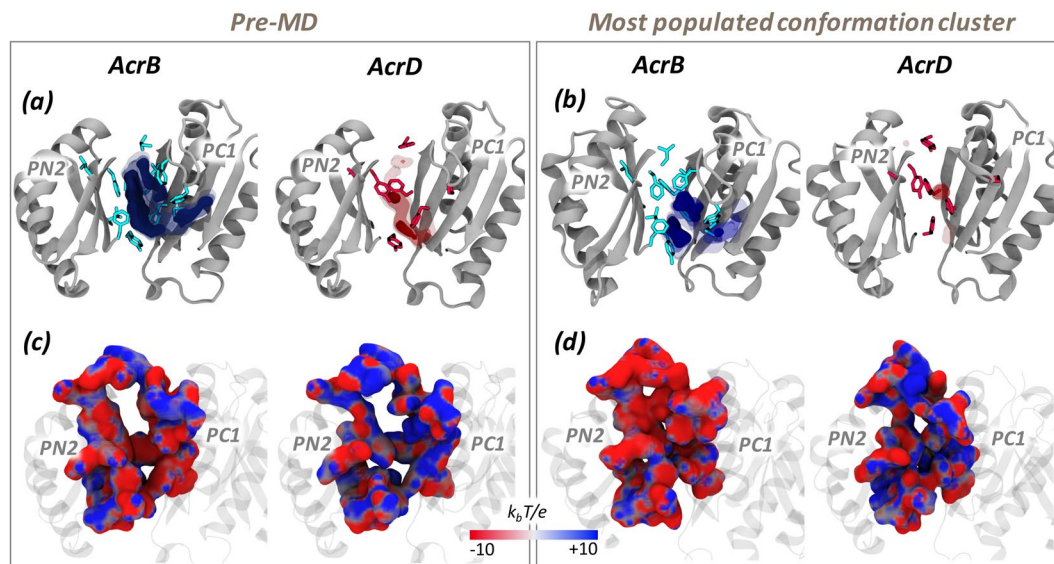
System	LI	
	Pre-MD	MD clusters
AcrB	16.3	$13.1 \pm 3.3$
AcrD	4.9	$1.9 \pm 1.0$

**Table 4.** Lipophilic indexes of DP in the *Tight* protomer of AcrB and AcrD. See Table 3 for further details.

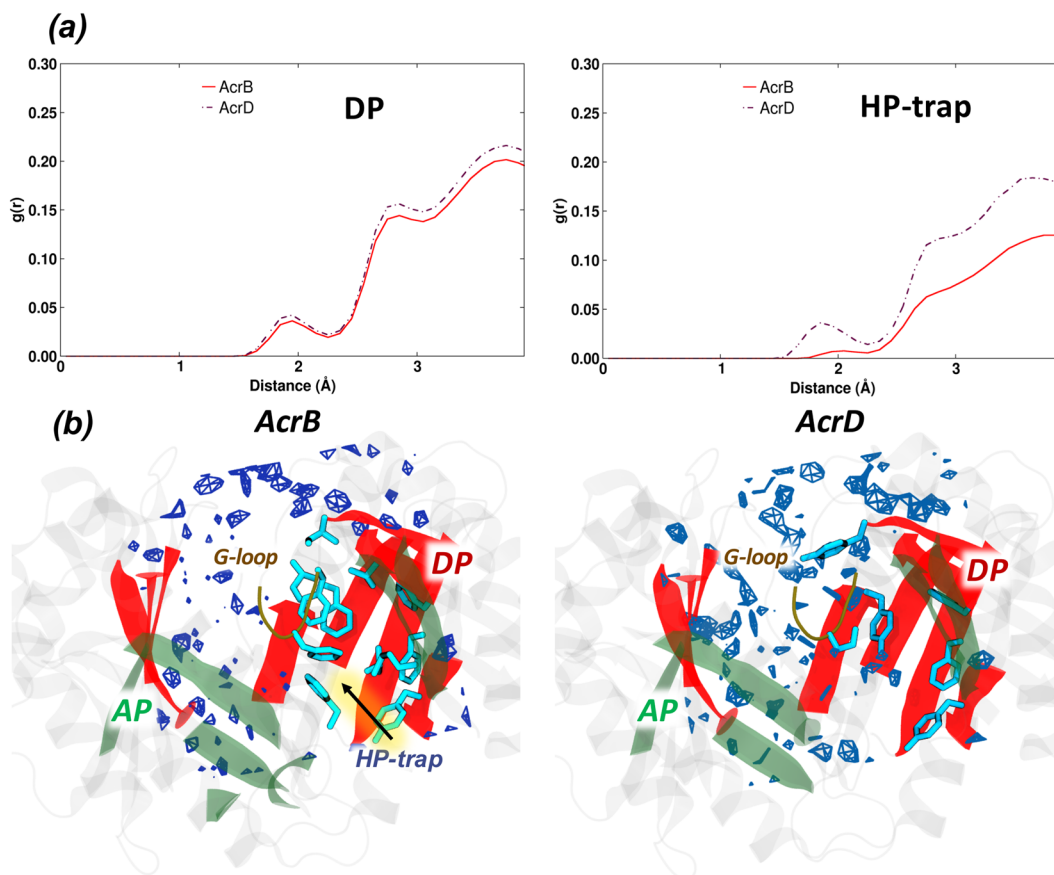
**MLP and LI.** The DP featured larger differences than the AP in the values of the LIs calculated for AcrB and AcrD, despite a reduction in the absolute values when considering the weighted average value extracted from the MD clusters compared to pre-MD structures (Table 4). The LI values indicated a much more prominent hydrophobic character of this pocket in AcrB than AcrD. Most clusters of AcrB were associated with LI values larger than 10, the highest values occurring for clusters 1, 3, 5, and 8 (LI of 14.4, 17.1, 17.0, and 17.2 respectively; see Supplementary Fig. S9). Together, the first three clusters embraced 80% of the conformations sampled by AcrB. In AcrD, the most populated clusters (1 to 5) had LIs ranging from 0.5 to 2.8 (Supplementary Fig. S9). Therefore, as already seen for the AP, the MLP proved to be a robust feature of the pockets, conserved across different conformations assumed in the dynamics. The lipophilic potential surfaces of the pre-MD and the most populated clusters are reported in Fig. 7a,b highlighting the presence of pronounced lipophilic regions in AcrB in comparison to three less extended spots in AcrD.

**Electrostatic Potential.** The electrostatic potential projected on the surfaces of the DP indicated a relatively denser positive environment in AcrD than in AcrB (blue areas in Fig. 7c,d). Noticeable is that the difference was better 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 in AcrD the distribution of areas of negative and positive potentials did not change much and the latter still presented a greater positive component with dispersed negative components. As in the case of AP, the partial closure of the pockets seen in the MD simulations of the apo proteins did not influence these main findings (Supplementary Fig. S10).

**Hydration Analysis.** The RDF profile of water around the DP showed a minor difference in the intensity of the peak between AcrB and AcrD, essentially related to the different hydration of the HP-trap region<sup>17</sup> (Fig. 8a). This is also consistent with the replacement of three out of the five phenylalanine residues in AcrD that are present in



**Figure 7.** MLP and electrostatic potential of DP in the *Tight* protomer of AcrB and AcrD. See Fig. 4 for further details.



**Figure 8.** Hydration of DP and HP-trap in the *Tight* protomer of AcrB and AcrD. (a) Comparison of RDF profiles of water oxygen atoms around the DP (all atoms) and HP-trap (all atoms) in the *Tight* protomer of AcrB (red solid line) and the corresponding regions of AcrD (brown dash-dotted line). (b) Comparison of SDF for waters in the DP calculated over the configurations forming the most populated cluster of AcrB (left) and AcrD (right) illustrating the variation in the immediate environment of the hydrophobic residues. The position of the HP-trap in DP of AcrB is indicated by an arrow. See Fig. 5 for further details.



Structure		Number of MFSs			Total number of MFSs
		AP	DP	Interface/G-loop	
AcrB	Pre-MD	3	2	2	7
	Cluster 1	3	1	1	5
	Cluster 2	2	1	1	4
	Cluster 3	1	2	1	4
	Cluster 4	2	1	1	4
	Cluster 5	2	1	1	4
	<b>Average</b>	<b>2.2</b>	<b>1.3</b>	<b>1.2</b>	<b>5</b>
AcrD	Pre-MD	1	-	1	2
	Cluster 1	1	1	1	3
	Cluster 2	3	-	-	3
	Cluster 3	3	-	1	4
	Cluster 4	3	-	-	3
	Cluster 5	3	1	-	4
	<b>Average</b>	<b>2.3</b>	<b>0.3</b>	<b>0.5</b>	<b>3</b>

**Table 5.** The number of MFSs identified in the binding pockets of AcrB and AcrD before and during MD. See Table 3 for further details.

the HP-trap of AcrB (only F610 and F628 are conserved). Indeed, the SDF clearly displayed a very low probability of hydration near this region in AcrB (Fig. 8b).

**Fragment-Based Binding Site Characterization.** In addition to global physicochemical pocket descriptors discussed above, the ligand-binding properties of protein are governed by the number, strength and spatial distribution of binding energy hot spots<sup>43</sup>. A fragment-based binding site analysis was thus performed employing probes (Supplementary Fig. S11) of different physicochemical features to identify hotspots responsible for specificity by mapping the chemical functionalities on the internal surface of the two proteins.

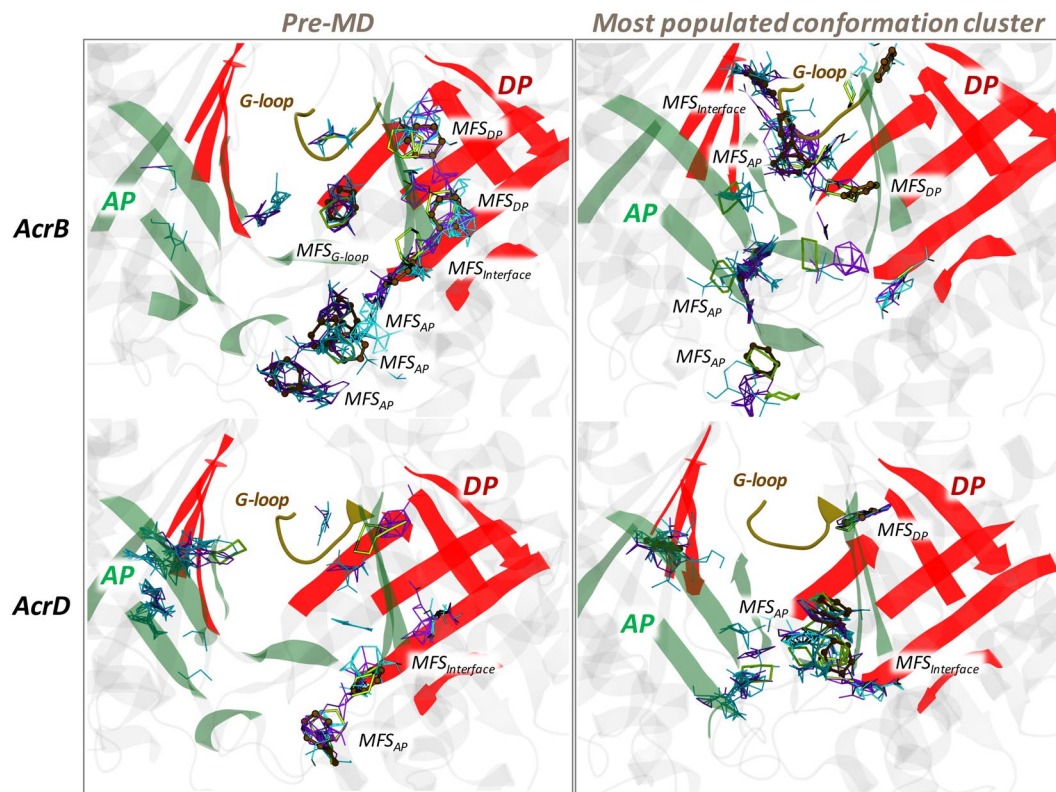
Several multi-functional sites (MFSs) were identified within the binding pockets of AcrB and AcrD. While the AP of both proteins showed consistently higher number of MFSs in comparison to the DP (Table 5), the level of promiscuity became distinct on comparing the MFSs in the latter pocket. The DP in AcrB showed an extended MFS (Fig. 9) in the pre-MD structure where substrates like minocycline<sup>14–16</sup>, doxorubicin<sup>14,16</sup> and inhibitors like P9D<sup>17</sup> and MBX2931<sup>44</sup> were crystallographically resolved. Although closure of the DP during the simulations resulted in the loss of the large extended MFS found in pre-MD, it created other MFS thereby preserving the promiscuity of DP in AcrB as seen in the representatives of the most populated clusters (Fig. 9 and Supplementary Fig. S12). In AcrD, the DP and interface/G-loop showed only a few consensus sites (CSs) and lacked a true MFS in both the pre-MD as well as the clusters sampled during MD. An interesting feature was that the interface between the pockets including the G-loop almost always favored an MFS in AcrB.

## Discussion

AcrB and AcrD are the major RND transporters of *E. coli*. They feature an overall good level of sequence (and likely fold) identity and similarity, and indeed show partly overlapping substrate specificities. However, they have distinct abilities to expel some classes of compounds; for instance, only AcrD recognizes aminoglycosides. The peculiarities of each transporter are likely related to the specific physicochemical features of the main recognition pockets, i.e. the AP and DP. Indeed, these two sites feature a lower degree of sequence conservation compared to the entire protein (see Supplementary Figures S1–S2). In particular, the AP is better conserved than the DP with nearly 60% vs. 40% identical residues, respectively. An inspection of the mismatched residues between AcrB and AcrD showed that the binding pockets of the latter protein are populated with more polar/charged residues than those of the former, likely facilitating the recognition and transport of more hydrophilic molecules by AcrD. However, this hypothesis should undergo a validation through the rationalization of the different substrate susceptibilities in terms of molecular descriptors of the binding pockets. Moreover, the impact of the dynamic nature of these transporters on the pocket environment should also be considered for gaining a more realistic understanding of their differential recognition and transport events as seen *in vitro* or *in vivo*.

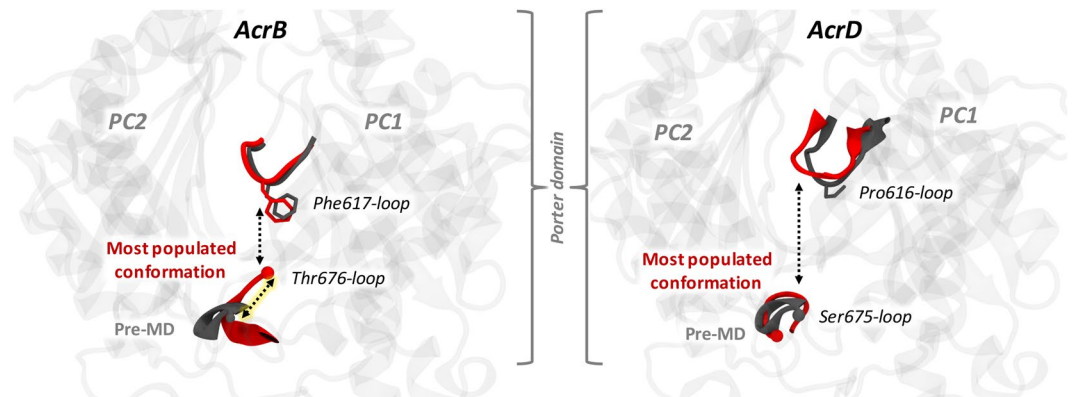
For this reason, in this work we compared several molecular descriptors calculated on the two main putative binding sites (AP in the *Loose* protomer and DP in the *Tight* protomer) within the periplasmic domains of AcrB and AcrD. In addition to experimental structures and homology models, which represent static snapshots of these biologically dynamic systems, we performed our analyses on a set of structures extracted from extensive MD simulations of the apo-proteins for assessing the influence of pocket dynamics. We recall that the MD simulations of AcrD were started from a homology model built using the AcrB structure as template. Clearly, the structures of cluster representatives for the former protein could feature a lower level of confidence than those of the latter. However, the AcrD model was found to be as good as experimental structures using state-of-the-art bioinformatic validation protocols (see Methods). In addition, the MD simulations of AcrD were as stable as those of AcrB, further pointing to the reliability of our findings.

The first descriptor we considered was the pocket volume and shape. However, a pure steric filter for substrates is quite unlikely because of the large volumes of all the pockets considered in the present analysis, which are at



**Figure 9.** MFSs identified in the AP and DP of AcrB and AcrD. MFSs in the AP and DP of pre-MD (left panels) and the most populated cluster representative (right panel) structures of AcrB and AcrD. The binding modes of the different probes are shown as lines for hydrogen-bond donor (cyan), hydrogen-bond acceptor (violet) and aliphatic (yellow), and as CPK for aromatic (ochre) ligands. The AP and DP are marked in green and red, respectively, while the G-loop in yellow cartoon representations. (Note: The categorizing of MFSs here is arbitrary due to indistinct boundaries between the pockets. The sites not labelled as MFS here are all CSs).

least twice as voluminous as the largest compounds transported by AcrB and AcrD. Moreover, the average values of the volumes and minimal projection areas for the AP and the DP of both proteins are, within errors, very similar; therefore, 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 two pockets partly collapsed during the MD simulations with respect to the conformation seen in the X-ray crystal structures and in the homology models of AcrB and AcrD, respectively. The reduction amounted to 30% in both AP and DP of AcrB while it was 15% in the AP and 28% in the DP of AcrD with respect to the pre-MD structures. This behavior is consistent with the findings of Fischer and Kandt<sup>33</sup>, who noticed a closure of the DP in the *Tight* protomer of AcrB in the absence of substrate during shorter MD simulations than those reported here. In coherence with this hypothesis, we also found the DP volume to be at least 1000 Å<sup>3</sup> larger in substrate bound complexes (Supplementary Table S4). 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 80% of the trajectories whilst the distribution is wider for AcrD, especially when AP is considered. Straightforwardly attributing this diversity to dissimilar flexibility might not be completely correct and could hide interesting features associated with the dynamics of the considered regions. For instance, as visualized from the porcupine plots of the first principal component (Fig. 3b and d), the entire AP of 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 to 678 lining the base of AP show larger rearrangements. The dynamicity of this loop (Thr676-loop or hereafter referred to as ‘bottom-loop’) represents a peculiar feature in the AP of AcrB, which is unshared with AcrD. The structures of the cluster representatives of AcrB can be partitioned in two groups featuring “up” and “down” conformation of the bottom-loop (Fig. 10). The most populated cluster is characterized by an “up” conformation, while the crystal structures exhibited only “down” configuration, as for the second most populated cluster representative (Supplementary Fig. S13). 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. With such major conformational shifts of the bottom-loop in AcrB, it is very likely that this loop contributes towards induced fit and minimizes the steric hindrance for the large substrates of AcrB, a hypothesis that is compatible with the larger size of some AcrB substrates that are not transported by AcrD. The importance of this bottom-loop in regulating access to porter domain and its possible active role in substrate transport by pushing compounds towards the hydrophobic binding pocket was already suggested by Fischer and Kandt, who however sampled only “down” conformations in their MD simulations<sup>33</sup>. Moreover, according to Kobayashi and co-workers the mutation E673G (located close to



**Figure 10.** Main conformational states of the Thr676-loop (bottom-loop) in the *Loose* protomer of AcrB (left panel) and of the corresponding Ser675-loop in AcrD (right panel). The conformation of the most populated clusters and the pre-MD structures are shown in red and grey cartoons, respectively. The conformations of the G-loop are also indicated with the same colour code.

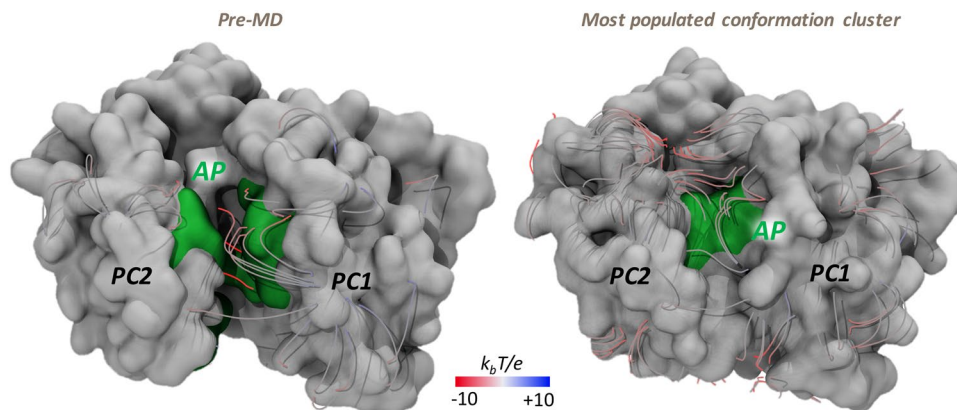
the bottom-loop) in AcrB, in combination with Q569R and I626R, conferred this protein the ability to recognize anionic beta-lactams, which are typical substrates of AcrD. Thus, the analysis of the volumes, although not much enlightening per se, allowed identifying specific structural features more directly involved in the entrance into and transport to a pocket, which might be of relevance in determining substrate specificity.

Next to the volume analysis, we calculated the LIs for the AP and the DP of both transporters in order to quantify how different distribution of hydrophobic residues could affect substrate recognition and how this property is tuned by the dynamics of the protein. Both pockets of AcrB are consistently characterized by higher LI values than those of AcrD, independently of the somewhat important structural changes occurring in these pockets during the MD simulations of the apo-proteins (Tables 3 and 4). The higher lipophilic nature of AP and DP in AcrB when compared to that of the same pockets in AcrD is required for AcrB to provide a favorable environment for its hydrophobic substrates to bind. However, the specific chemical environment of the AP is neither entirely hydrophobic nor entirely polar in both the proteins. Such a dispersed nature binds ligands with different physico-chemical properties by weak polar and hydrophobic interactions<sup>45</sup>, while facilitating easy transport by preventing strong interactions of the substrates with residues of the pocket. Note that the values of the LIs for AP and DP of AcrD are essentially identical whilst in AcrB there is a marked difference between the two sites, the DP being the more lipophilic. This could be an indication of DP being the site where substrates might be differentiated between AcrB and AcrD in terms of their lipophilicity. In other words, the DP could function as a lipophilicity-based selectivity filter for low-molecular-mass compounds. This proposal agrees with previous suggestions based on experimental results of Yamaguchi *et al.*<sup>46</sup>. The difference between the DP of AcrB and AcrD became even more prominent by comparing their molecular lipophilic surfaces (Fig. 7a,b). The MLP isosurfaces are significantly wider in AcrB than in AcrD, which correlates well with the nature of the reported substrates transported by the former protein. Interestingly, the presence of phenylalanines in the G-loop of only 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 results 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.

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<sup>47</sup>. The plot of the SDF around the AP and the DP of both transporters highlight how their different hydrophobic potentials influence hydration profiles. In particular, a lower degree of hydration was seen near the hydrophobic part of the AP in AcrB than in AcrD (Fig. 5b). Even for the DP, our analysis provides a clear evidence of the contribution of the HP-trap in determining the lower hydration of the domain compared to AcrD. While several spots are homogeneously distributed in the SDF calculated for AcrD, the DP featured several zones without hydration in AcrB, especially around the HP-trap (Fig. 8b).

Electrostatic complementarity between the pocket and substrate molecules is essential for initial substrate recruitment and augmentation of their association rate<sup>48</sup>. Therefore, this analysis is of particular interest for the AP, which is more peripheral than the DP. For AcrB, the distribution of positively and negatively charged patches on the molecular surface of this site is fairly homogeneous (Fig. 4c,d). Interestingly, in AcrB an electrostatic funnel with negative gradient leading from the periplasm to the centre of the AP is recognizable in Fig. 11, which could help the long-range recognition of positively charged compounds, and is compatible with the monocationic character of several substrates of AcrB. The electrostatic surface of this site in AcrD reveals instead a marked positive patch on the upper part, also due to the presence of residues like Arg568 and Arg625, which have been recently reported as key residues for specificity of AcrD towards negatively charged molecules like the anionic beta-lactams<sup>22</sup>.





**Figure 11.** Electrostatic funnel converging into AP in the *Loose* protomer of AcrB. Only the strongest field lines are shown and coloured red to blue from negative ( $-10 k_b T/e$ ) to positive ( $+10 k_b T/e$ ) potential ( $k_b$  is the Boltzmann constant,  $T$  is absolute temperature and  $e$  is charge of an electron). AP is shown in green and the rest of the porter domain in white surface representation.

Concerning the DP, the large electrostatic negative environment presented to the incoming substrate in AcrB would favor the recognition and binding of cationic compounds. However, this becomes unfavorable to hydrophilic polycationic aminoglycosides due to the high lipophilicity of this pocket in AcrB functioning as a probable lipophilicity-based selectivity filter as discussed above. In AcrD, a relatively denser positive environment compared to AcrB is identifiable, originating from the electrostatic contributions of amino acids like Arg and Lys replacing their less polar counterparts in AcrB (Supplementary Figs S1–S2). In conjunction with the low lipophilicity of this pocket in AcrD, the observed mosaic-like electrostatic patches provide a favorable binding site for anionic beta-lactams as well as for polycationic aminoglycosides. On the other side, the poor electrostatic and hydrophilic complementarity provided by the DP in AcrB permits the binding of charged molecules like anionic beta-lactams but with far less affinity than that in AcrD.

Multidrug transporters are known to have large, flexible overlapping substrate binding pockets rich in polar and aromatic residues to bind substrate molecules at different locations with different orientations. In alternate terms, these proteins show polyspecificity with no inherent ligand specificity which otherwise could stem from the binding site geometry<sup>49</sup>. Our results suggest that features such as shape, lipophilicity, electrostatic potential and hydration of AP and DP 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. To further establish this, we performed a fragment based binding site characterization using FTMMap, whose philosophy retraces experimental high-throughput and fragment screening methods. Also, as its algorithm does not rely on alternate measures of ligand-binding propensity such as pocket volume, cavity depth or the ability of binding non-polar spheres, the results obtained here can complement those from physico-chemical pocket descriptors discussed above.

As evident from the overall distribution of MFSs (Table 5), both AcrB and AcrD provide multiple binding possibilities with different functionalities as expected from such promiscuous transporter proteins. This is distinct from the limited number of MFSs one would detect in restrained substrate binding sites of other ordinary substrate-receptor systems restricted to a specific class of substrates<sup>45</sup>. In particular, AcrB with its numerous and wide spread MFSs offers a greater level of promiscuity for diverse substrate types than AcrD, which with its smaller localized MFSs puts forth certain prerequisites on the substrates being recognized. While the AP in both proteins showed comparable numbers of MFSs dispersed within the pocket, a clear distinction is noticeable in the DP (as also observed with the pocket descriptors discussed above) where a true and wide spread MFS is seen only in AcrB. The DP of this transporter is clearly a multi-functional (or multidrug binding) site with higher preference towards hydrophobic and aromatic fragments alongside hydrogen-bond donor and acceptor fragments. The non-selective characteristics of weakly polar (by Q176, G616) and weakly hydrophobic (by F178, Y327, F615, F617, F628) interactions are a predominant player towards the promiscuous binding behavior of AcrB DP<sup>30</sup>. The DP of AcrD does not reach the level of multi-functionality seen in AcrB, and shows greater preference for hydrogen-bond donors/acceptors (by R44, D134, N136, T139, Y178, Y327, S614, G615), together with a very limited preference towards hydrophobic (by Y178, Y277, F627) fragments. The MFSs identified here are in good agreement with the data reported for AcrB by Imai *et al.*<sup>45</sup>, and were also found close to the residues identified as crucial for the recognition of anionic beta-lactams by Kobayashi *et al.*<sup>22</sup>, thereby strengthening the reliability of our findings.

As seen from the distribution of MFSs in the various MD clusters (Supplementary Fig. S12), their position is not constant and this dynamicity (attributed by spatial changes in internal cavities caused by peristaltic motions<sup>50</sup>) is most likely important to avoid the substrate from being trapped in a single site and to facilitate its efflux by multisite-drug-oscillation<sup>46</sup>. Additional studies involving substrate bound complexes can provide information on the interaction profile of these homologous Acr pumps with their corresponding substrates. However, redundancy in the residue type in the binding pocket leading to easy adaptability of the binding orientation of



substrates in the presence of mutations as identified by Bohnert *et al.* in AcrB<sup>51</sup> makes these pumps very challenging for studies with simple molecular docking<sup>30</sup>.

## Conclusions and Perspectives

In this study, we performed a comparative analysis of the physicochemical properties like pocket volume and shape, lipophilicity, electrostatic potential, hydration and multi-functional sites of AcrB and AcrD to rationalize their differential substrate specificities. Importantly, these analyses were performed not only on static structures but also on conformations extracted from extensive MD simulations accounting for the impact of protein dynamics. Our results reveal several features in which both the AP and the DP differ considerably between these two transporters. First, the calculated lipophilic potential turned out to be significantly different between the AP and DP of AcrB and between the corresponding pockets of AcrB and AcrD considering even the dynamics of the pockets. In particular, the DP of AcrB is more lipophilic than all other sites, suggesting the possible role of this pocket as a lipophilicity-based selectivity filter. Second, we observed specific differences in the electrostatic environment within the pockets. In particular, the presence of an electrostatic funnel sourcing from the AP of AcrB could be important for the recognition of monocationic compounds by this transporter. Thus, these two properties likely play a central role in governing substrate recognition by and specificity of AcrB and AcrD. Meanwhile, the cavity volume, which essentially remains large enough to accommodate all potential substrate molecules, possibly has an indirect effect on the lipophilic and electrostatic environment along with the distribution of MFS, which altogether with the ensuing hydration within the pocket govern recognition and transport of substrates by these pumps. In addition, specific features like the flip conformations of the bottom-loop and the lipophilic bridge created by Phe617 of G-loop both in AcrB (and not in AcrD), which could not have been identified from simple sequence analysis, are expected to play a key role in the recognition and transport function of 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.

## Methods

**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<sup>52</sup> (UNIPROT ID: P24177), and subsequently searched for the best available template structures bearing homologous relationship to the query sequence using the NCBI-BLAST tool<sup>53</sup> against the Protein Data Bank (PDB) ([www.rcsb.org](http://www.rcsb.org)). AcrB sequence showed the highest identity (~66%) (similarity of ~80%) with least gaps over a maximum sequence coverage; therefore its high resolution crystal structure, 1.9 Å (PDB ID: 4DX5<sup>16</sup>), was chosen as template for modeling AcrD. The two protein sequences were optimally aligned by ClustalOmega<sup>54</sup> and the results were visually inspected to ensure the absence of gaps in important secondary structure regions. Modeller 9.13<sup>40</sup> 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<sup>6</sup>. The resulting models were ranked using discrete optimized protein energy (DOPE)<sup>55</sup> 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<sup>56</sup>. The most reliable model was then selected based on various geometric and stereochemical quality factors evaluated for backbone angles, side chains flips, rotamers, steric clashes etc. using PROCHECK<sup>57</sup>, ERRAT<sup>58</sup>, ProSA<sup>59</sup>, Verify3D<sup>60</sup> programs available in MolProbity<sup>61</sup> and Structure Analysis and Verification Server (<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 also employed on the crystal structure of AcrB for use as reference. Visual inspections were performed with VMD1.9.1<sup>62</sup> and PyMOL<sup>63</sup>.

**Molecular dynamics simulations of AcrB and AcrD.** MD simulations of the crystal structure of AcrB (PDB ID: 4DX5) and of the most reliable homology model of AcrD (see Supplementary Table S1) were carried out using the AMBER14 molecular modeling software<sup>56</sup>. Protomer specific protonation states<sup>18</sup> were adopted with E346 (E346) and D924 (D922) protonated in both *Loose* and *Tight* protomers while deprotonated in the *Open* protomer of AcrB (AcrD). The residues D407 (D407), D408 (D408), D566 were protonated only in the *Open* protomer of AcrB (AcrD). The charge state of the residue L565 of AcrD, corresponding to D566 in AcrB, was not modified of course. 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 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) bilayer patches, solvated with explicit TIP3P water model, and neutralized with the required number of randomly placed K<sup>+</sup> ions<sup>28,29,64</sup>. The ions count was suitably adjusted to account for an osmolarity of 0.15 M KCl. Embedding of the protein into a pre-equilibrated POPE bilayer patch was performed using the PPM server<sup>65</sup> and subsequently the CharmmGUI tool<sup>66</sup>. The lipid residue nomenclature was converted from the CHARMM to AMBER format using the *charmm lipid2amber.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<sup>67</sup>. Periodic boundary conditions were used and the distance between the protein and the edge of the box was set to be at least 30 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 constraints 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 (collision frequency of  $1 \text{ ps}^{-1}$ ) 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<sup>68</sup>. The particle-mesh Ewald (PME) algorithm 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. Coordinates were saved every 100 ps. The ff14SB<sup>69</sup> version of the all-atom Amber force field was used to represent the protein systems while lipid14<sup>67</sup> parameters were used for the POPE bilayer. After equilibration, multi-copy  $\mu\text{s}$ -long MD simulations were performed for each system, namely two  $\sim 3 \mu\text{s}$ -long production simulations for each transporter (for a total simulation time of  $\sim 12 \mu\text{s}$ ). Trajectory analysis was done using *cpptraj* module of AmberTools14 and VMD1.9.1, and graphs were plotted using the *xmgrace* tool.

**Principal component analysis.** To characterize and highlight possible similarities and differences in the collective motions of the binding pockets, we calculated the covariance matrices from the equilibrium trajectory and performed a principal component analysis<sup>70,71</sup>. As customary in principal component analysis, the covariance matrix was constructed taking the three-dimensional positional fluctuations of C $\alpha$  atoms from their ensemble average position (after least-squares fitting to remove rotational and translational motion). Diagonalization of the covariance matrix yields a set of eigenvectors and corresponding eigenvalues, which represent the direction and amplitude of the motion, respectively. The eigenvectors are then ranked according to the decreasing order of their associated eigenvalues, such that the first eigenvector represents the largest contribution to the total fluctuation of the system. To visualize the motions represented by the eigenvectors, the structures from the trajectories can be projected onto each eigenvector of interest [principal component (PC)] and transformed back into Cartesian coordinates. The two extreme projections along each eigenvector can then be interpolated to create an animation or compared to understand which parts of the protein are moving according to that specific eigenvector and to what extent. Usually, (a combination of) the first few principal components are able to represent most of the collective motions (the “essential dynamics”<sup>70</sup>) occurring in an MD simulation among the different regions of a protein.

**Clustering of MD trajectories.** A cluster analysis of the MD trajectories was performed using the average-linkage hierarchical agglomerative clustering method 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 in two separate instances the trajectory based on root mean square deviation (RMSD) (cutoff set to 3 Å) of the AP in the *Loose* protomer and of the DP in the *Tight* protomer. For each protein, the representative structures from each of the 10 top clusters generated in each of the two cases considered (AP in *Loose*, DP in *Tight*) were used to perform quantitative analyses in order to account for dynamical behavior. Except for hydration analyses, all non-protein molecules were stripped from the trajectory during post-processing to reduce additional memory usage and to speed up file processing.

**Pocket descriptors.** The list of the pocket descriptors identified for the present study includes: i) cavity volume and shape; ii) molecular lipophilicity potential; iii) electrostatic potential; iv) site hydration; v) fragment-based binding site characterization. The various pocket descriptors used to characterize the binding site were calculated using specific programs after validating their applicability to RND systems by assessing results against available crystal structures and experimental data, as well as previous computational reports<sup>29, 30, 33, 35, 45, 64</sup>.

**Cavity volume and shape.** Evolution of size and shape of the AP and DP during the MD simulations was examined using the two-probe sphere method of *rbccavity* program bundled in the rDock suite<sup>72</sup>. This allows obtaining detailed information on the pocket volume and plasticity of the site. In this method, the binding site volume was identified by a fast grid-based cavity detection algorithm<sup>73</sup> within a sphere of radius 14 Å, centred over the pockets, using large and small probe radii of 6.0 Å and 1.5 Å, respectively. These radii were found to be optimal for our case after evaluating different combinations and checking through visual inspection their accuracy in predicting volume of the pocket space by keeping the possible inclusion of regions extending outside the pocket of interest at its least. rDock also gives information about the approximate shape of the pocket; we could thus provide approximate values for the minimal cross-sectional area associated to each cavity.

**Molecular lipophilicity potential.** 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 molecule and the solvent system, at that given point. Thus, MLP can be calculated from the fragmental system of logP and a distance function as shown in the following equation<sup>74</sup>:

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

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 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:

$$LI = \frac{\sum MLP^+}{\sum MLP^+ + |\sum MLP^-|} \cdot 100 \quad (2)$$

The lipophilicity of AP in the *Loose* protomer and DP in the *Tight* protomer were qualitatively and quantitatively estimated in this way using MLP Tools<sup>75</sup> plugin available for PyMOL.

**Electrostatic potential.** The electrostatic potential surface maps were computed by APBS<sup>76</sup>, after preprocessing structures of AcrB and AcrD to assign charges and atomic radii using the PDB2PQR server<sup>77</sup>. All electrostatic potential calculations were performed at 0.15 M physiological 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.

**Hydration analysis.** The radial distribution function (RDF) indicates the probability of finding water molecules at a certain distance from a region or residue of interest and is commonly used to analyse 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 is computed from the histogram of the number of solvent particles found as a function of the distance  $R$  from an (ensemble of) atom(s), normalized by the expected number of solvent particles at that distance in bulk. The normalization is estimated from:

$$Density * \left( \left[ \frac{4\pi}{3} (R + dR)^3 \right] - \left[ \frac{4\pi}{3} dR^3 \right] \right) \quad (3)$$

where  $dR$  is equal to the bin spacing, the default *density* value is 0.033456 molecules Å<sup>-3</sup>, which corresponds to density of water approximately equal to 1.0 g mL<sup>-1</sup>. Bin spacing of 0.1 and a maximum bin value of 4.0 was used in this case to calculate the RDF of all water oxygen atoms to each atom of AP in the *Loose* protomer and of DP in the *Tight* protomer 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) of waters around the whole protein was calculated using the Gromacs utility *g\_spatial*<sup>78</sup> on the trajectory frames grouped into the most populated conformational clusters extracted from MD simulations. SDF allows to determine the three-dimensional density distribution of aqueous solution around the binding pockets of the transporters. Density isovalue gives information regarding the relative number densities with respect to the average number density of solvent molecules in bulk. 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.

**Fragment-Based Binding Site Characterization.** The FTMap server<sup>79</sup> implementing the FTSite algorithm is a tool helpful in the identification of binding sites and of the fragments that could be possible source of structure- and fragment-based drug design attempts. The main aim of such fragment-based binding site analysis is to obtain a measure of the ability of the protein (and in particular the pockets under study) to bind a drug-like molecule.

FTMap identifies the important hot spots based on the consensus clusters of 16 standard probes which include molecules varying in size, shape and polarity (Supplementary Fig. S11). Such a diverse library of probes is useful to capture a range of interaction types that include hydrophilic, hydrophobic, hydrogen-bonding and aromatic interactions. The regions where clusters of different probes of the same or different type overlap are marked as consensus (CS) and multi-functional (MFS) sites, respectively, and are ranked based on the number of their clusters. Clusters in close proximity to a top ranked cluster are merged with it and the protein residues within this region become the top ranked putative ligand binding site.

## References

1. World Health Organization. Antimicrobial resistance: global report on surveillance. [http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748_eng.pdf) (January 2016) (2014).
2. Nikaido, H. Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* **264**, 382–388, doi:10.1126/science.8153625 (1994).
3. Li, X.-Z., Plésiat, P. & Nikaido, H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin. Microbiol. Rev.* **28**, 337–418, doi:10.1128/CMR.00117-14 (2015).
4. Blair, J. M. *et al.* AcrB drug-binding pocket substitution confers clinically relevant resistance and altered substrate specificity. *Proc. Natl. Acad. Sci. USA* **112**, 3511–3516, doi:10.1073/pnas.1419939112 (2015).
5. Poole, K. Efflux-mediated antimicrobial resistance. *J. Antimicrob. Chemother.* **56**, 20–51, doi:10.1093/jac/dki171 (2005).
6. Piddock, L. J. Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clin. Microbiol. Rev.* **19**, 382–402, doi:10.1128/CMR.19.2.382-402.2006 (2006).

7. Blair, J. M., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **13**, 42–51, doi:10.1038/nrmicro3380 (2015).
8. Ruggerone, P., Murakami, S., M Pos, K. & Vargiu, A. V. RND efflux pumps: structural information translated into function and inhibition mechanisms. *Curr. Top. Med. Chem.* **13**, 3079–3100, doi:10.2174/15680266113136660220 (2013).
9. Misra, R. & Bavro, V. N. Assembly and transport mechanism of tripartite drug efflux systems. *Biochim. Biophys. Acta, Proteins Proteomics* **1794**, 817–825, doi:10.1016/j.bbapap.2009.02.017 (2009).
10. Symmons, M. F., Bokma, E., Koronakis, E., Hughes, C. & Koronakis, V. The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proc. Natl. Acad. Sci. USA* **106**, 7173–7178, doi:10.1073/pnas.0900693106 (2009).
11. Du, D. *et al.* Structure of the AcrAB-TolC multidrug efflux pump. *Nature* **509**, 512–515, doi:10.1038/nature13205 (2014).
12. Zguruskaya, H. I., Weeks, J. W., Ntrel, A. T., Nickels, L. M. & Wolloscheck, D. Mechanism of coupling drug transport reactions located in two different membranes. *Front. Microbiol.* **6**, 100, doi:10.3389/fmicb.2015.00100 (2015).
13. Song, S., Kim, J.-S., Lee, K. & Ha, N.-C. Molecular architecture of the bacterial tripartite multidrug efflux pump focusing on the adaptor bridging model. *J. Microbiol.* **53**, 355–364, doi:10.1007/s12275-015-5248-4 (2015).
14. Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T. & Yamaguchi, A. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* **443**, 173–179, doi:10.1038/nature05076 (2006).
15. Nakashima, R., Sakurai, K., Yamasaki, S., Nishino, K. & Yamaguchi, A. Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket. *Nature* **480**, 565–569, doi:10.1038/nature10641 (2011).
16. Eicher, T. *et al.* Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc. Natl. Acad. Sci. USA* **109**, 5687–5692, doi:10.1073/pnas.1114944109 (2012).
17. Nakashima, R. *et al.* Structural basis for the inhibition of bacterial multidrug exporters. *Nature* **500**, 102–106, doi:10.1038/nature12300 (2013).
18. Eicher, T. *et al.* Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB. *Elife* **3**, e03145, doi:10.7554/eLife.03145 (2014).
19. Symmons, M. F., Marshall, R. L. & Bavro, V. N. Architecture and roles of periplasmic adaptor proteins in tripartite efflux assemblies. *Front. Microbiol.* **6**, 513, doi:10.3389/fmicb.2015.00513 (2015).
20. Seeger, M. A. *et al.* Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* **313**, 1295–1298, doi:10.1126/science.1131542 (2006).
21. Pos, K. M. Drug transport mechanism of the AcrB efflux pump. *Biochim. Biophys. Acta, Proteins Proteomics* **1794**, 782–793, doi:10.1016/j.bbapap.2008.12.015 (2009).
22. Kobayashi, N., Tamura, N., van Veen, H. W., Yamaguchi, A. & Murakami, S.  $\beta$ -Lactam selectivity of multidrug transporters AcrB and AcrD resides in the proximal binding pocket. *J. Biol. Chem.* **289**, 10680–10690, doi:10.1074/jbc.M114.547794 (2014).
23. Schuster, S., Vavra, M. & Kern, W. V. Evidence of a substrate discriminating entrance channel in the lower porter domain of the multidrug resistance efflux pump AcrB. *Antimicrob. Agents Chemother.* **60**, 4315–4323, doi:10.1128/AAC.00314-16 (2016).
24. Iyer, R., Ferrari, A., Rijnbrand, R. & Erwin, A. L. A fluorescent microplate assay quantifies bacterial efflux and demonstrates two distinct compound binding sites in AcrB. *Antimicrob. Agents Chemother.* **59**, 2388–2397, doi:10.1128/AAC.05112-14 (2015).
25. Elkins, C. A. & Nikaido, H. 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.* **184**, 6490–6498, doi:10.1128/JB.184.23.6490-6499.2002 (2002).
26. 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.* **46**, 889–901, doi:10.1046/j.1365-2958.2002.03223.x (2002).
27. Eda, S., Maseda, H. & Nakae, T. An elegant means of self-protection in gram-negative bacteria by recognizing and extruding xenobiotics from the periplasmic space. *J. Biol. Chem.* **278**, 2085–2088, doi:10.1074/jbc.C200661200 (2003).
28. Schulz, R., Vargiu, A. V., Collu, F., Kleinekathöfer, U. & Ruggerone, P. Functional rotation of the transporter AcrB: insights into drug extrusion from simulations. *PLoS Comput. Biol.* **6**, e1000806, doi:10.1371/journal.pcbi.1000806 (2010).
29. 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.* **133**, 10704–10707, doi:10.1021/ja202666x (2011).
30. Vargiu, A. V. & Nikaido, H. Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **109**, 20637–20642, doi:10.1073/pnas.1218348109 (2012).
31. Collu, F., Vargiu, A. V., Dreier, J., Cascella, M. & Ruggerone, P. Recognition of imipenem and meropenem by the RND-transporter MexB studied by computer simulations. *J. Am. Chem. Soc.* **134**, 19146–19158, doi:10.1021/ja307803m (2012).
32. Fischer, N. & Kandt, C. Three ways in, one way out: Water dynamics in the trans-membrane domains of the inner membrane translocase AcrB. *Proteins: Struct., Funct., Bioinf.* **79**, 2871–2885, doi:10.1002/prot.23122 (2011).
33. Fischer, N. & Kandt, C. Porter domain opening and closing motions in the multi-drug efflux transporter AcrB. *Biochim. Biophys. Acta Biomembr.* **1828**, 632–641, doi:10.1016/j.bbame.2012.10.016 (2013).
34. Ruggerone, P., Vargiu, A. V., Collu, F., Fischer, N. & Kandt, C. Molecular dynamics computer simulations of multidrug RND efflux pumps. *Comput. Struct. Biotechnol. J.* **5**, 1–11, doi:10.5936/csbi.201302008 (2013).
35. Vargiu, A. V., Ruggerone, P., Opperman, T. J., Nguyen, S. T. & Nikaido, H. Molecular mechanism of MBX2319 inhibition of *Escherichia coli* AcrB multidrug efflux pump and comparison with other inhibitors. *Antimicrob. Agents Chemother.* **58**, 6224–6234, doi:10.1128/AAC.03283-14 (2014).
36. Zuo, Z., Wang, B., Weng, J. & Wang, W. Stepwise substrate translocation mechanism revealed by free energy calculations of doxorubicin in the multidrug transporter AcrB. *Sci. Rep.* **5**, 13905, doi:10.1038/srep13905 (2015).
37. Zuo, Z., Weng, J. & Wang, W. Insights into the inhibitory mechanism of D13-9001 to the multidrug transporter AcrB through molecular dynamics simulations. *J. Phys. Chem. B* **120**, 2145–2154, doi:10.1021/acs.jpcc.5b11942 (2016).
38. Ramaswamy, V. K., Cacciottolo, P., Mallocci, G., Vargiu, A. V. & Ruggerone, P. Computational modelling of efflux pumps and their inhibitors. *Essays Biochem.* **61**, 141–156, doi:10.1042/EBC20160065 (2017).
39. Litwin, S. & Jores, R. In *Theoretical and Experimental Insights into Immunology* Vol. 66 (eds Alan S Perelson & Gérard Weisbuch) 279–287 (Springer, 1992).
40. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815, doi:10.1006/jmbi.1993.1626 (1993).
41. Cha, H.-J., Müller, R. T. & Pos, K. M. Switch-loop flexibility affects transport of large drugs by the promiscuous AcrB multidrug efflux transporter. *Antimicrob. Agents Chemother.* **58**, 4767–4772, doi:10.1128/AAC.02733-13 (2014).
42. Husain, F., Bikhchandani, M. & Nikaido, H. Vestibules are part of the substrate path in the multidrug efflux transporter AcrB of *Escherichia coli*. *J. Bacteriol.* **193**, 5847–5849, doi:10.1128/JB.05759-11 (2011).
43. Ciulli, A., Williams, G., Smith, A. G., Blundell, T. L. & Abell, C. Probing hot spots at protein-ligand binding sites: A fragment-based approach using biophysical methods. *J. Med. Chem.* **49**, 4992–5000, doi:10.1021/jm060490r (2006).
44. Sjuts, H. *et al.* Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proc. Natl. Acad. Sci. USA* **113**, 3509–3514, doi:10.1073/pnas.1602472113 (2016).
45. 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* **115**, 8288–8295, doi:10.1021/jp2015758 (2011).



46. Yamaguchi, A., Nakashima, R. & Sakurai, K. Structural basis of RND-type multidrug exporters. *Front. Microbiol.* **6**, 327, doi:10.3389/fmicb.2015.00327 (2015).
47. Zakrzewska, K., Lavery, R. & Pullman, B. The solvation contribution to the binding energy of DNA with non-intercalating antibiotics. *Nucleic Acids Res.* **12**, 6559–6574, doi:10.1093/nar/12.16.6559 (1984).
48. Levy, Y., Onuchic, J. N. & Wolynes, P. G. Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition. *J. Am. Chem. Soc.* **129**, 738–739, doi:10.1021/ja065531n (2007).
49. Neyfakh, A. A. Mystery of multidrug transporters: the answer can be simple. *Mol. Microbiol.* **44**, 1123–1130, doi:10.1046/j.1365-2958.2002.02965.x (2002).
50. Murakami, S. Multidrug efflux transporter, AcrB—the pumping mechanism. *Curr. Opin. Struct. Biol.* **18**, 459–465, doi:10.1016/j.sbi.2008.06.007 (2008).
51. Bohnert, J. A. *et al.* Site-directed mutagenesis reveals putative substrate binding residues in the Escherichia coli RND efflux pump AcrB. *J. Bacteriol.* **190**, 8225–8229, doi:10.1128/JB.00912-08 (2008).
52. The UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, D204–D212, doi:10.1093/nar/gku989 (2015).
53. Madden, T. In *The NCBI Handbook [Internet]* (eds McEntyre Jo & Ostell Jim) (National Center for Biotechnology Information, 2013).
54. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539, doi:10.1038/msb.2011.75 (2011).
55. Shen, M.-y & Salí, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524, doi:10.1110/ps.062416606 (2006).
56. AMBER 14 (University of California, San Francisco, 2014).
57. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291, doi:10.1107/S0021889892009944 (1993).
58. Colovos, C. & Yeates, T. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511–1519, doi:10.1002/pro.5560020916 (1993).
59. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410, doi:10.1093/nar/gkm290 (2007).
60. Eisenberg, D., Lüthy, R. & Bowie, J. U. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* **277**, 396–404, doi:10.1016/S0076-6879(97)7022-8 (1997).
61. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21, doi:10.1107/S0907444909042073 (2010).
62. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, doi:10.1016/0263-7855(96)00018-5 (1996).
63. The PyMOL Molecular Graphics System. Version 1.5.
64. Schulz, R., Vargiu, A. V., Ruggerone, P. & Kleinekathöfer, U. Role of water during the extrusion of substrates by the efflux transporter AcrB. *J. Phys. Chem. B* **115**, 8278–8287, doi:10.1021/jp200996x (2011).
65. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376, doi:10.1093/nar/gkr703 (2012).
66. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865, doi:10.1002/jcc.20945 (2008).
67. Dickson, C. J. *et al.* Lipid14: the amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879, doi:10.1021/ct4010307 (2014).
68. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874, doi:10.1021/ct5010406 (2015).
69. Maier, J. A. *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713, doi:10.1021/acs.jctc.5b00255 (2015).
70. Daidone, I. & Amadei, A. Essential dynamics: foundation and applications. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 762–770, doi:10.1002/wcms.1099 (2012).
71. García, A. E. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* **68**, 2696–2699, doi:10.1103/PhysRevLett.68.2696 (1992).
72. Ruiz-Carmona, S. *et al.* rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput. Biol.* **10**, e1003571, doi:10.1371/journal.pcbi.1003571 (2014).
73. Morley, S. D. & Afshar, M. Validation of an empirical RNA-ligand scoring function for fast flexible docking using RiboDock<sup>®</sup>. *J. Comput. Aided Mol. Des.* **18**, 189–208, doi:10.1023/B:JCAM.0000035199.48747.1e (2004).
74. Gaillard, P., Carrupt, P.-A., Testa, B. & Boudon, A. Molecular lipophilicity potential, a tool in 3D QSAR: method and applications. *J. Comput. Aided Mol. Des.* **8**, 83–96, doi:10.1007/BF00119860 (1994).
75. Oberhauser, N., Nurisso, A. & Carrupt, P.-A. MLP Tools: a PyMOL plugin for using the molecular lipophilicity potential in computer-aided drug design. *J. Comput. Aided Mol. Des.* **28**, 587–596, doi:10.1007/s10822-014-9744-0 (2014).
76. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037–10041, doi:10.1073/pnas.181342398 (2001).
77. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667, doi:10.1093/nar/gkh381 (2004).
78. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25, doi:10.1016/j.softx.2015.06.001 (2015).
79. Kozakov, D. *et al.* The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* **10**, 733–755, doi:10.1038/nprot.2015.043 (2015).
80. Nishino, K. & Yamaguchi, A. Analysis of a complete library of putative drug transporter genes in Escherichia coli. *J. Bacteriol.* **183**, 5803–5812, doi:10.1128/JB.183.20.5803-5812.2001 (2001).
81. Sulavik, M. C. *et al.* Antibiotic Susceptibility Profiles of Escherichia coli Strains Lacking Multidrug Efflux Pump Genes. *Antimicrob. Agents Chemother.* **45**, 1126–1136, doi:10.1128/AAC.45.4.1126-1136.2001 (2001).
82. Nishino, K., Nikaïdo, E. & Yamaguchi, A. Regulation and physiological function of multidrug efflux pumps in Escherichia coli and Salmonella. *Biochim. Biophys. Acta, Proteins Proteomics* **1794**, 834–843, doi:10.1016/j.bbapap.2009.02.002 (2009).
83. Rosenberg, E. Y., Ma, D. & Nikaïdo, H. AcrD of Escherichia coli is an aminoglycoside efflux pump. *J. Bacteriol.* **182**, 1754–1756, doi:10.1128/JB.182.6.1754-1756.2000 (2000).
84. Aires, J. R. & Nikaïdo, H. Aminoglycosides are captured from both periplasm and cytoplasm by the AcrD multidrug efflux transporter of Escherichia coli. *J. Bacteriol.* **187**, 1923–1929, doi:10.1128/JB.187.6.1923-1929.2005 (2005).

## Acknowledgements

We thank K.M. Pos (Goethe University, Frankfurt am Main, Germany) for the critical reading of the manuscript. The research leading to the results discussed here was partly conducted as part of the Translocation Consortium (<http://www.translocation.eu>) and has received support from the Innovative Medicines Initiative

Joint Undertaking under Grant Agreement no. 115525, resources that are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007–2013) and EFPIA companies in kind contribution. V.K.R. is a Marie Skłodowska-Curie fellow within the “Translocation” Network, project no. 607694.

### Author Contributions

V.K.R., A.V.V., G.M., J.D., P.R. designed experiments. V.K.R. performed experiments. V.K.R., A.V.V., G.M., P.R. analysed the results. V.K.R., A.V.V., G.M., J.D., P.R. discussed the results and wrote the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-08747-8](https://doi.org/10.1038/s41598-017-08747-8)

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017