Biochimica et Biophysica Acta 1848 (2015) 2065–2074

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Structural and dynamic changes adopted by EmrE, multidrug transporter protein—Studies by molecular dynamics simulation

Monikaben Padariya *, Umesh Kalathiya, Maciej Baginski

Department of Pharmaceutical Technology and Biochemistry, Faculty of Chemistry, Gdansk University of Technology, Narutowicza St 11/12, 80-233Gdansk,Poland

ARTICLE INFO

Article history:
Received 11 February 2015
Received in revised form 25 April 2015
Accepted 18 May 2015
Available online 23 May 2015

Keywords:
Small multidrug resistance (SMR) transporter
EmrE
Multidrug transporter
Heterogeneous lipid bilayer
Molecular dynamics

ABSTRACT

EmrE protein transports positively charged aromatic drugs (xenobiotics) in exchange for two protons and thus provides bacteria resistance to variety of drugs. In order to understand how this protein may recognize ligands, the monomer and asymmetric apo-form of the EmrE dimer embedded in a heterogeneous phospholipid (POPE + POPG) membrane were studied by molecular dynamics simulations. Dimer is regarded as a functional form of the transporter, but to understand molecular aspects of its mode of action, a monomer was also included in our work. We analyzed hydrogen bonds which include inter- and intra-molecular interactions. Analyzing the long-lasting H-bond interactions, we found that water access to the internal transmembrane segments is regulated by residues with aromatic or basic side chains and fluctuating transmembrane helices. Our finding supports that GLU14 in EmrE apo-form is ready to interact or bind with substrate molecule. The analysis of distance center of masses and water entrance area indicate the feasibility of the dimer to undergo induced fit in order to accommodate a ligand. The results indicate that a binding pattern can be formed in the EmrE in such a way that GLU14 binds to the positively charged fragment of a substrate molecule, and other aromatic residues (i.e., TRP63 and TYR40) located in vicinity may accommodate other non-polar parts of substrate molecule. The results of our simulation also allow us to support experimentally testable hypotheses concerning functional inward–outward conformational changes of the protein.

1. Introduction

From origin, living cells had machineries which defend them against toxic compounds, and for this type of survival, the energy-dependent extrusion of toxic compounds has been evolved as a strategy. Due to active extrusion, the concentration of drug molecules in cells decreases, and this develops resistance. Multidrug transporters (MDT) supply one such strategy responsible for drug resistance [1]. EmrE is a multidrug transporter from *Escherichia coli*, and a member of the small multidrug resistance (SMR) family. As one of the smallest known active transporters, EmrE is found in the inner membrane of the bacteria [2]. Due to this size,SMR transporters are useful systems to study the minimum requirements for active transport.

From Cryo-Electron Microscopy (Cryo-EM), X-ray crystallography, and sequence conservation data, a single site alternating access model of an antiparallel dimer with an inward-to-outward-facing conformational exchange during transport has been proposed [3–5]. Structural antiparallel asymmetry is supported by solid-stateNMR measurements [6], solution NMR,FRET experiments [7], as well as mutagenesis experiments [8,9]. Biochemical studies have clearly shown that the membrane embedded, charged, and highly conserved residue Glutamate14 (Glu14 or E14) in transmembrane (TM) helix 1 appeared to coordinate the binding and release of substrate and is also essential for transport [10]. However, the mechanistic details and molecular mode of action of EmrE (MDT) protein, including ligand recognition at the molecular level, remain to be characterized.

As a member of the SMR family and with only 110 amino acids in its structure, EmrE has a small binding region which must be able to accommodate a wide range of substrate molecules within its limited space. Multidrug recognition in this type of small binding pocket has been already conferred in BmrR multidrug-resistant (MDR) transcription factor [11]. In case of BmrR, the same residues of active site interact with different ligand molecules in a highly rigid binding pocket [11]. This is different from the canonical concept of multidrug recognition [12,13], which postulates a key role of flexibility in placing diverse ligands in a specific site. However, the need to carry substrate binding and function is basically different for both transcription factors and transporters. The data obtained from Cryo-EM and NMR studies clearly show that EmrE alters in structure when it binds to substrate molecule [7,14]. Consequently, we assume that flexibility of some regions in EmrE structure is important for multidrug recognition, and this type of nature

* Corresponding author at: Gdansk University of Technology, Department of Pharmaceutical Technology and Biochemistry, Faculty of Chemistry, Narutowicza St 11/12, 80-233 Gdansk, Poland. Phone: +48583471596.
E-mail addresses: monikapadariya1@gmail.com (M. Padariya), umekalat@gmail.com (U. Kalathiya), maciekb@hypnos.chem.pg.gda.pl (M. Baginski).

http://dx.doi.org/10.1016/j.bbamem.2015.05.014
0005-2736/

has been successfully followed by different multidrug transporter proteins which vary in size and functions [12,13].

Although residues which are responsible for interaction with substrate have been revealed primarily by mutagenesis and are significant for structural understanding of these MDR pumps, they may be regarded only as a static snapshots of highly dynamic proteins [15]. Computational studies of MDR pumps for predicting and rationalizing already available data are used more often to analyze their mechanistic details [15]. Clustering analysis of Cryo-EM data shows that there are at least three conformational states adopted by EmrE which suggests that structural flexibility and plasticity is a cornerstone of multidrug recognition and transport for this MDT protein [14,15]. The recalculated X-ray structure of EmrE (PDB ID: 3B5D) [4] concerning topology fits extremely well into the density for the Cryo-EM structure (PDB ID: 2I68) [3]. Additionally, a model derived from the Cryo-EM structure at 7.5 Å resolution and evolutionary constraints has an averageRMSD of 1.4 Å with the C-alpha positions in the revised X-ray structure [4]. Thus, there is excellent agreement between Cryo-EM and X-ray structure of EmrE, showing that EmrE is an antiparallel dimer, and the majority of mutagenic and biochemical data confirm this structure [7–9,16,17].

Therefore, in our study the X-ray structure of EmrE (PDB ID: 3B5D) [4] was used to study the two-model system, EmrE as monomer and as asymmetric dimer in apo-form by molecular dynamics (MD). The main objective of our work is to characterize the structural and dynamic aspect of recognition and binding place of this protein. The molecular details of both EmrE monomer and dimer in apo-form will be of importance as the mechanism of ligand recognition is unknown for this protein. EmrE monomer as such is not regarded as functional unit of this protein, but on the other hand, it has to exist as a step for dimer formation. Therefore, both monomer and apo-form of dimer is worth to study as a possible step in the mechanism of ligand recognition and protein transporter action. The apo-form of dimer is also interesting to study due to proposed ligand-induced fit mechanism of binding. One may expect that our apo model created on the basis of bound structure may undergo kind of reverse induced fit changes toward a real apo-form or may change conformation according to the data observed for apo-form [18]. Thus, the study of apo-form of EmrE dimer enables us to acquire knowledge on how this structure may behave and which residues of the protein are accessible for substrate binding. Moreover, our molecular dynamics study elucidates the structure flexibility and characterizes the functional residues of EmrE protein that are able to make interaction with membrane components and aqueous environment. Results obtained from these MD studies also help to understand what essential elements of EmrE asymmetry structureare important for transport.

2. Materials and methods

2.1. Model building

The C-alpha (CA) coordinates were only available in the structure and were taken from X-ray structure of EmrE protein (PDB ID: 3B5D) [4] bound with TPP^+ compound. The ligand was subsequently removed from this structure. C-alpha coordinates were used to prepare all-atom model of protein within the Discovery Studio Client 3.1 (Accelrys, San Diego, USA) program. The CHARMm forcefield was used to prepare full protein structure. To get proper protonation state of the protein, the following parameters were set up during protein generation using Discovery Studio Client 3.1: pH 7.4, protein dielectric constant equal to 10, and ionic strength equal to 0.145 M. Full protein structure generation was followed by energy minimization protocol using 'smart minimizer' algorithm in Accelrys. Non-bonded list radius was set to 14 and RMS gradient was set to 0.1 in the minimization. The resultant coordinates were used for membrane-protein system preparation.

One may criticize that studied computational model is based on rather poor resolution X-ray structure [4] (resolution R = 3.8 Å) containing only C-alpha chains; however, this X-ray structure and Cryo-EM structure at 7.5 Å resolution have been regarded almost similar [3,4], and anyway, molecular modeling is base on models which only reflect in better or less accurate way the physiological situation. Our model was prepared according to the current state of the art, all missing atoms of residues were added and the whole structure was properly optimized. Such system was embedded in the membrane and fully equilibrated. Analysis of RMSD of transmembrane helices presented in the result section shows reasonable values for this kind of systems and also supports quality of the model used in MD. Moreover, the transmembrane part of the protein is the most interesting concerning recognition of the ligand, and therefore, our computational model and the results are not below quality of the system published in the ref [4] and used by others for discussion of the mode of action of EmrE.

2.2. System setup

For the determination of initial orientation of EmrE protein in membrane, the PPM web server [19] was used, which calculates rotational and translational positions of transmembrane proteins in membranes using their 3D structure as an input. Prepared all-atom structure of EmrE protein was inserted in a membrane mimicking the natural bacterial one, containing zwitterionic phosphatidylethanolamine (POPE) (70% molar concentration) and anionic phosphatidylglycerol (POPG) (30% molar concentration) as membrane components [20]. Both leaflets of the membrane had the same lipid molar content. Each simulated system contained the same amount of lipids, so in case of dimer, the simulated system was bigger. Two different membrane-protein complexes (EmrE monomer-membrane and EmrE dimer-membrane) were prepared using CHARMM-GUI[21] membrane builder, which generates a series of CHARMM inputs necessary to build a protein-membranecomplex for MD simulations. The number of parameters was specified during generation of membrane-protein complex in CHARMM-GUI. Rectangle box type and tetragonal crystal type was specified. Water thickness on the top and bottom of the system was set to 15.0 Å, and a heterogeneous lipid bilayer was generated by the replacement method originally developed by Woolf and Roux [21,22].

The replacement method uses lipid-like pseudo atoms to generate lipid packing around protein and replaces these pseudo atoms with lipid molecules, and short dynamics simulations were performed to maintain the pseudo atoms on each plane. Monte-Carlo method was used as ion-placing method. Penetration of protein surface or lipid ring was not found in the obtained protein-membrane complex from CHARMM-GUI. These final systems were subsequently used for molecular dynamics calculations. In current work for EmrE monomer-membrane system, monomer (protein) was named as Chain P, and for EmrE dimer-membrane system, both monomers were termed with different names (Chain P and Chain U, respectively). The composition of models containing EmrE protein (monomer or dimer) and membrane is described in Table 1 (Fig. S1 in the Supporting Material).

2.3. Molecular dynamics simulation

Molecular dynamics simulations were performed for two different systems (EmrE monomer-membrane and EmrE dimer-membrane) using NAMD [23] with CHARMM27 force field for lipids [24], protein, and TIP3P water molecules. The particle mesh Ewald (PME) [25] was

Table 1
Summary of bilayer structural properties and model systems.

Protein	Lipids (Heterogeneous)	Waters and ions (0.415 M)	Atoms
EmrE monomer	202 POPE + 86 POPG	12110 waters and 128 Na^+ + 42 Cl^-	74226
EmrE dimer (antiparallel)	202 POPE + 86 POPG	12121 waters and 128 Na^+ + 42 Cl^-	75806

used to treat long-range electrostatic interactions with the density of grid points at least 1 Å in all cases and cutoff of 12 Å was applied to treat van der Waals interactions. Periodic boundary condition (PBC) was formed in all directions to avoid finite size effects, and constant pressure and temperature (NPT) ensemble was adopted. Langevin dynamics [26] was applied to maintain constant temperature with damping coefficient (gamma) of 1/ps and the Nose-Hoover Langevin piston method [27] was used to maintain constant pressure with a decay period of 50 fs and an oscillation time of 25 fs. Time step in MD was set to 2 fs, and SHAKE algorithm [28] was used to constrain all the non-polar hydrogen atoms. Equilibration of the whole system was performed for 10 ns at 303.15 K to gradually equilibrate the systems. Final production run was performed for 1000 ns at 303.15 K using an NPT ensemble.

3. Results and discussion

3.1. Overall structural changes

To check stability of the systems and to determine quantitatively the extent of motions, the root-mean-square deviation (RMSD) of each protein was computed along 1000 ns trajectories of the production run for both model systems (Fig. 1). From the analysis of decomposed RMSD values for helices and loops separately in dimer structure of EmrE protein, distortions for helices are much smaller than distortions for loops. All helices except helix1 of both monomers have low values of RMSD(Fig. 1A and B). Higher RMSD values for helix1 may stem from the fact that it is N-terminal helix exposed to lipids, not interacting with other monomer and therefore had more freedom. The overall RMSD values computed for two monomers in dimer were quite similar. The RMSD of the monomer alone in simulation was departed from that of the monomers in dimer and remained higher by about 1 Å for the first 550 ns. For all systems, the overall RMSDs are rather similar for the last 450 ns and remain steady till the end of MD simulations (Fig. 1C). The computed RMSD of overall structure reached up to 5 Å for both monomers in dimer system, but these distortions in RMSD are due to loops flexibility and because of inward and outward conformational changes (Fig. S2). However, loops are less important in our study concerning binding place of EmrE protein as it is located in the region of helices (Fig. S3).

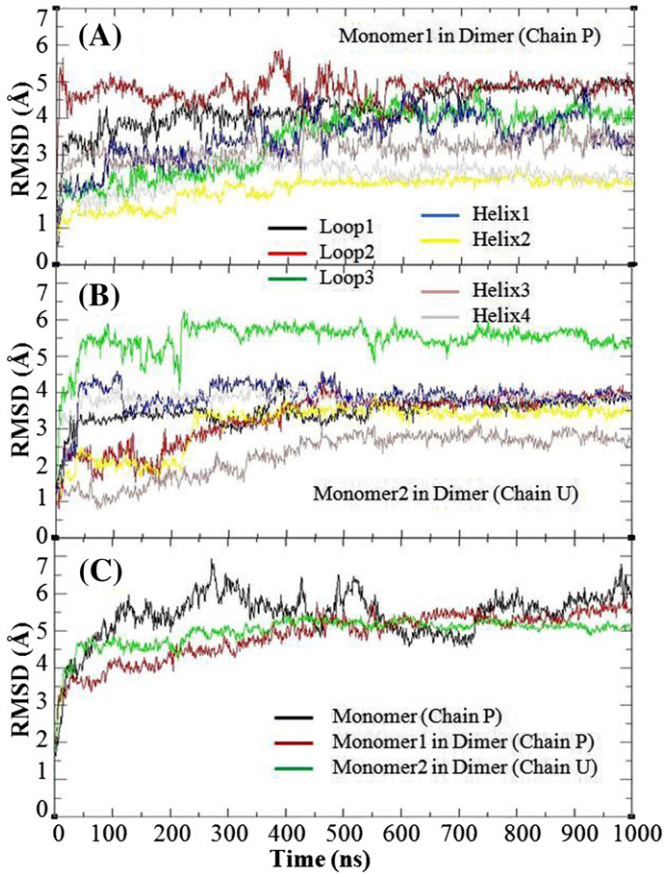

Fig. 1. Decomposed RMSD (root mean squared deviation) for helices **(A)** and loops **(B)** in two monomers (Chain P and Chain U) of EmrE dimer system. Selected residue (res.) range for calculating RMSD for loops and helices are the following: loop1 res. 20–36, loop2 res. 50–61, and loop3 res. 73–88 for both Chain P and Chain U. Helices for Chain P, helice1 res. 6–19, helice2 res. 37–49, helice3 res. 62–72, and helice4 res. 89–105. Helices for Chain U, helice1 res. 4–19, helice2 res. 37–49, helice3 res. 62–72, and helice4 res. 89102. (**C**) Overall RMSD of the EmrE monomer system compared against the two monomers of dimer system. RMSD of the two monomers in dimer does not change much compared to the monomer alone over the course of the simulation.

Flexibility of the protein was also analyzed by checking root-mean-square fluctuation (RMSF) of each residue during the entire production run. It is observed that one monomer (Chain P) which is present in both, monomer and dimer model systems, has more fluctuating region as compared to the second monomer (Chain U) (Fig. 2). More fluctuations for Chain P is observed in three regions, loop between TM2–TM3, TM3–TM4, and a large portion of TM4 showed higher fluctuations. This higher fluctuation of residues in TM4 stems from position of this helix at the protein/lipid edge was what allowed this fluctuation. Moreover, TM4 in Chain P is less fold than that in Chain U. In the second monomer (Chain U) in dimer, all three loops (between TM1 and TM2, TM2 and TM3, and TM3 and TM4) showed fluctuations (Fig. 2). This finding is noteworthy, as in different structural studies of EmrE (MDT) protein, an alternate access mechanism has been proposed in which protein interconvert between conformations during substrate binding and translocation [3,29] as a result of induced fit upon ligand binding. These particular loops and flexible regions of the protein pivot around so as to produce conformational change between the inward- and outward-facing structures. Qualitatively, the RMSFs (Fig. 2) agree well with alteration in conformation as the loops are more flexible relative to TM segments (Fig. 1).

The inward- and outward-facing conformational change is also supported by the distance measurements between residues in two monomers present in dimer. Since the X-ray structure contains less than 105 residues, we used residue LYS22 and ASN102 to compare our results (Fig. S2A and S2B) with the distance changes observed in the studies presented by Morrison *et al.* [7]. Moreover, we selected the residues at top and bottom positions in the same helix, namely, GLY35 and ALA48 from TM2 (Chain P), and ALA87 and VAL98 from TM4 (Chain U), in order to monitor possible inward- and outward-facing conformational change. However in this case, only inward-to-inward (or outward-to-outward) facing conformational change through the state in which both distances are equal was recorded (Fig. S2C and S2D). This lack of full conformational change can be explained by short MD simulation to cover the whole cycle of the conformational change, as well as by the fact that it is apo-form which did not include protons or ligands. Nevertheless, our results agree with recent findings [18] that apo-form of EmrE can also fluctuate along inward- to outward-facing mode.

3.2. Monitoring interactions in EmrE monomer and dimer systems

Many experimental hints suggest that to a lesser extent, a hydrogen-bonding component contributes to the substrate binding of multidrug transporters. On the other hand, the nature of drug binding to some extent may also be influenced by interaction with lipid molecules [30]. We performed a thorough analysis of intra- and intermolecular hydrogen bonds (H-bonds) for protein and contacts between residues of protein and water or lipid molecules. Over the entire 1000 ns time span, a higher number of intramolecular hydrogen bonds were formed between residues of EmrE monomer system than between residues in monomers of EmrE dimer system. However, it is worth to note that

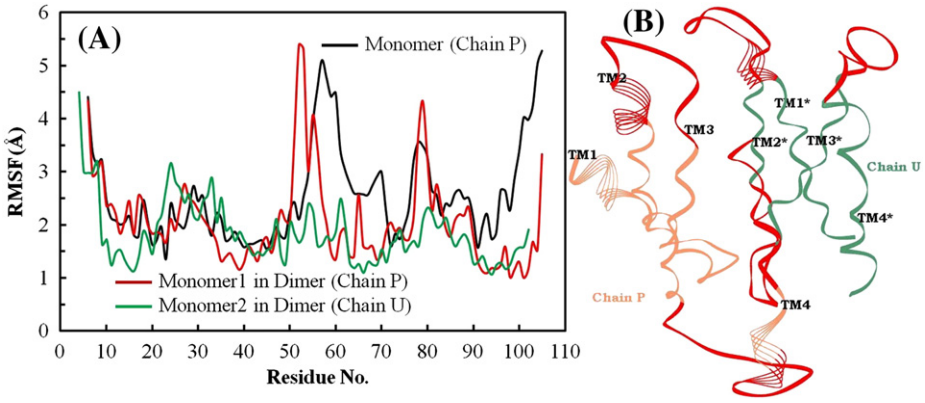

Fig. 2. RMSF of each model system shows transmembrane regions which fluctuate less than the loop region. **(A)** The root mean squared fluctuation of the EmrE monomer and dimer systems during 1000 ns of dynamics. **(B)** Light orange and green regions of the protein are less flexible whereas the regions of loops show the greatest fluctuations (colored in red).

the sum of intermolecular (between two monomers in dimer) and intramolecular interactions of one monomer in dimer gives almost same number of intramolecular H-bonds in monomer-membrane complex. This suggests that some residues forming intramolecular hydrogen bonds in monomer alone at the same time are involved in the intermolecular interactions in dimer when the second monomer is present (Fig. 3A).

The overall number of hydrogen bonds between protein and water in dimer is reasonable, as water molecules penetrated the cavity interface between two monomers in dimer. Whereas, monomer alone in membrane has quite a different pattern of interaction with water molecule (Fig. 3B). The interactions between membrane protein and their lipid bilayer environment are important for the stability and function of such proteins. A comprehensive picture of the interactions between lipid headgroups and protein can be achieved by analyzing H-bonds. We have defined hydrogen bond interactions occurring between protein and lipids, and it is also informative to look at both types of lipid molecules (POPE and POPG) making contacts with monomer system and two monomers present in dimer system (Fig. 3C). After the first ~400 ns, the number of hydrogen bond interactions was rather constant for both dimer and monomer systems, implying successful equilibration of the lipid–protein interactions in the simulation. Monomer system exhibited more hydrogen bond interactions with lipids in comparison of two monomers present in dimer-membrane system. This is understandable as monomer alone has larger surfaces to interact with lipid molecules, whereas monomers in dimer system occupy one side with each other and only the remaining one side can interact with lipid molecules (Fig. 3C).

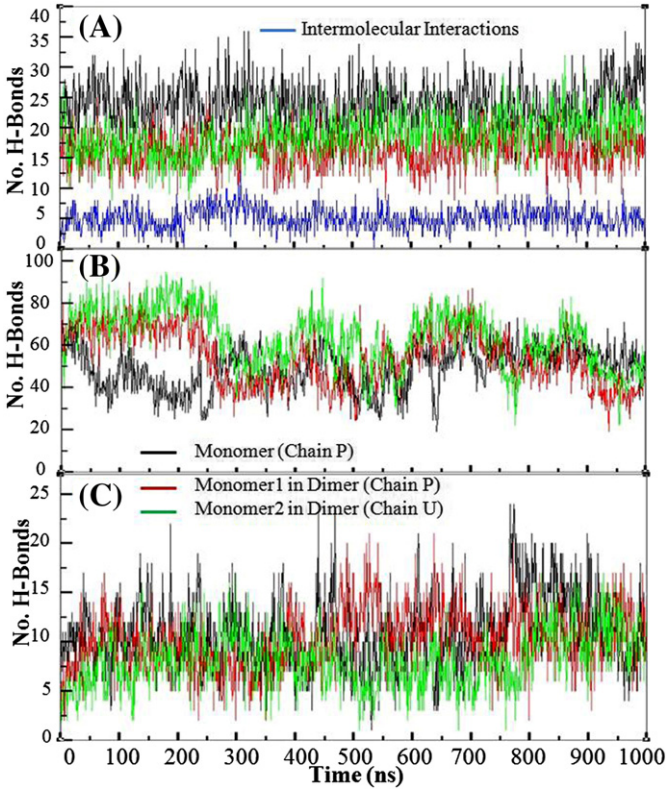

Fig. 3. Hydrogen bond interaction analysis for EmrE monomer (in black color) and dimer systems (red color for monomer1—Chain P) and green color for monomer2—Chain U). Comparison of the number of H-bond contact formed between**(A)** protein–protein (intramolecular and intermolecular) **(B)** protein and water molecules, and **(C)** protein with lipid headgroups. The cutoffs used to define H-bonds are 3.5 Å for the donor-acceptor distance and intermolecular angle values are in range 160°-180° between donor-H-acceptor.

3.3. Interactions between EmrE monomers (Chain P and Chain U) in dimer system

The detailed interactions between individual residues of the two monomers in EmrE dimer were monitored during overall simulation time (1000 ns). Residues of the one monomer (Chain P), which were in contact (<3.5 Å) with the other monomer (Chain U), were listed (Fig. 4). The rapid motion of the residues during simulations resulted in numerous encounters, but most of them were lasting only a short period of time and made a minimal contribution to the protein–protein interaction. To account in a more specific way for it, we considered the occupancy of time (%) for which the specific interaction was formed (Table S1 and Fig. S4). Seven residues in both monomers were involved in interactions between two monomers lasting more than 20% of the total simulation time. Altogether, almost 12 residues of each monomer were found to make contact with each other (≥10% occupancy). It is worth to point out that all these residues were located in the loop region or very near to this region in the structure of EmrE monomers in dimer (Fig. 5A and S5). This is reasonable since these parts of protein structures frequently undergo conformational change when interacting with other proteins in inward- or outward-facing changes [31].

The residues involved in protein–protein interactions show the ability of these specific residues to make interaction between two monomers. Thus, these residues play an important role to achieve the resulting conformational change and flexibility within selected regions of a protein, allowing to adopt new conformations [32]. In this particular case, it may be responsible for inward- to outward-facing conformational change of dimer during transportation of ligand or in apo-form.

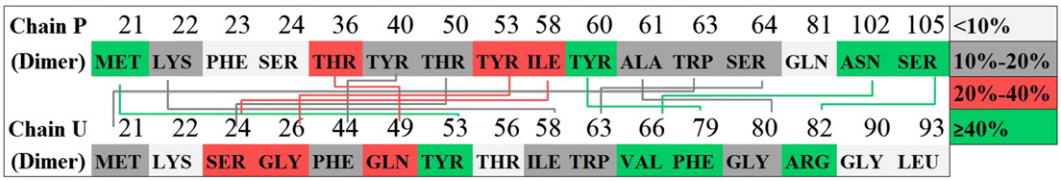

Fig. 4. H-bond contacts between residues of two monomers in EmrE protein dimer system, obtained from the MD simulations. The cutoffs used to define H-bonds are 3.5 Å for the donor-acceptor distance and intermolecular angle values are in range 160°–180° between donor-H-acceptor. Residues forming interactions are colored according to occupancy of H-bond (% of time). Contacts that persist at least ≥ 10% of the examined period of time between residues of two monomers are highlighted by lines.

3.4. Region (residues) of the protein making interaction with lipid headgroups

Protein–lipid interactions are continuously recognized as central for structural and functional study of membrane proteins. There is also evidence for precisely bound lipids that are necessary to achieve biological function of such proteins [33]. However, specific interactions between protein and lipid are particularly difficult to highlight experimentally. Here, we used MD simulations to identify this type of interactions [33]. The specific interactions between individual residues of protein and lipid headgroups were followed over all simulation time. Residues of the protein which were in contact (<3.5 Å) with the lipid molecules were listed (Fig. 6). In order to monitor interaction sites with lipid headgroups, amino acid side chain polarity was taken into account. To this end, it was observed that hydrophilic polar (GLN, THR, TYR, and LYS) (side chain) and non-polar (GLY) residues (main chain) made very stable hydrogen bonds with lipids. And these interactions exhibited lifetime longer than 30% of total simulation time (Table S2 and Fig. S6).

Residues LYS22 and TYR40 made long-lasting interaction with lipid molecule when monomer (Chain P) was alone in membrane. However, both these residues of monomer (Chain P) in dimer system were not interacting at all with lipid molecules since they were involved in protein (Chain P)–protein (Chain U) interaction, and in such situation, they had no interface to come in contact with lipids. Moreover, higher amount of transient H-bonds (occupancy < 20%) occurred between different donors and acceptors (Fig. 6 and Table S2), probably due to the thermal 'breathing' motions of lipid and protein molecules. The residues often found within loops (due to their location within the membrane), exhibiting high flexibility, were found to make long-lasting interaction with lipid molecules in both monomer and dimer system of EmrE protein (Fig. 5B and S5).

It is also worth to compare which residues of monomer, when it is alone, interact with lipids and have the ability to interact with the second monomer when it is present in dimer (Table S3). Surprisingly, only three residues, LYS22, TYR40, and TRP63, exhibit this property. At the same time, these residues are important for ligand binding according to the experimental data (Table 2) [34]. This finding is very unexpected and may even raise the question concerning alternative mechanism of ligand recognition (not only from cytoplasm). One may hypothesize that ligand can be recognized also in the membrane by replacing protein(monomer)–lipid interaction for protein–ligand interaction with simultaneous binding of the second monomer, to facilitate functional structure of the dimer that is able to transport ligand in outward-facing movement. Such mechanism of ligand recognition would be in agreement with 'vacuum cleaner' hypothesis of ligand recognition by membrane protein transporters from the membrane.

The ability of biological membrane components to bend is crucial to understand the interaction between protein and the lipid bilayer. Experimental and computational studies have presented that the lipid molecule in membrane can bend to enable charged or polar residues of proteins to be expose to the lipid headgroups [35]. An efficient computational model obtained from MD simulation of EmrE protein showed that headgroup of POPE and POPG lipid molecule bends slightly to make strong H-bond interaction with TYR40 and ALA96 residues, respectively (Fig. 7A and B).

3.5. Water interacting sites (residues) in EmrE protein

Modeling water–protein interactions is critical to understand the interactions of protein with other molecules or to define the binding site of it with potential ligands. Water molecules form an integral part of

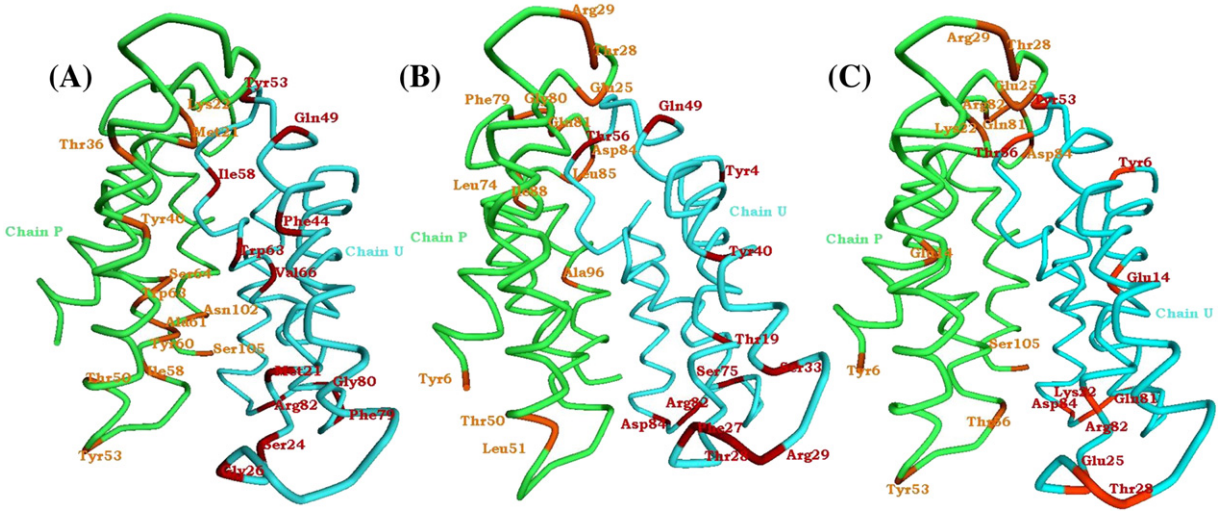

Fig. 5. Residues of protein (EmrE) involved in hydrogen bond interactions. Chain P is represented in green color and Chain U in blue color. **(A)** Residues of one monomer making contact with the second monomer in dimer are colored in red and orange, respectively. Residues having occupancy of H-bond ≥ 10% are represented. **(B)** Residues in EmrE dimer involved in H-bond interaction with lipid molecules and lasting ≥ 10% of total simulation time are selected and highlighted as red and orange color in Chain U and Chain P, respectively. **(C)** Residues of protein-forming long-lasting (≥80% occupancy) H-bond interactions with water molecules at least in one system (monomer or dimer) are colored in red and orange respectively.

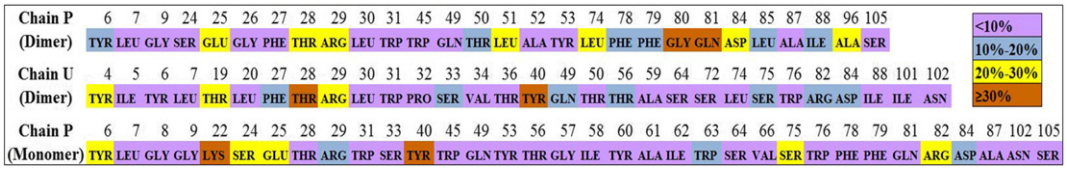

Fig. 6. Hydrogen bond interactions obtained between individual residues of the EmrE protein and the lipid molecules. Residues of the protein, which were, in a contact (<3.5 Å) with the lipid molecules were listed. The cutoffs used to define H-bonds are 3.5 Å for the donor-acceptor distance and intermolecular angle values are in range 160°–180° between donor-H-acceptor. H-bond occupancy time is indicated by different colors for specific residue involved in interaction.

most protein–ligand interactions, and in some cases, water molecules have also been found to be very crucial for ligand recognition [36]. Hydrogen bond interactions between specific residue of protein and water molecules have been followed during 1000 ns simulation time (Table S4, Fig. S7 and S8). Residues of protein that made long-lasting interactions with water molecules (<3.5 Å) were listed (Fig. 8A). Polar or charged residues of protein that were energetically favorable to make contacts with water, and those which exhibited strong interactions (≥80% of total simulation time) were analyzed in both monomers of dimer system and in monomer system alone (Fig. 8B). Mainly, the exposure of hydrophobic residues on the exterior of protein as well as the hydrophilic residues in the hydrophobic core were treated as potential areas ready to interact with external molecules, e.g., ligands. The recognition of potential sites in protein that are ready to interact with other molecules (protein, ligand, and water) may be useful to identify how the protein recognizes ligands and achieves its function [31]. Our thorough analysis of residues that are interacting with water and not involved in protein–protein or protein–lipid interaction provides insight into potential substrate binding site, as in a biological environment where a ligand could replace water molecules present in apo-form.

Twelve protein residues interacting with water nearly all time of simulations were observed (Fig. 8B), among which eight residues (TYR6, GLU14, GLU25, THR28, ARG29, THR56, GLN81, and ASP84) of both monomers were not involved in protein–protein interaction at all (or only partially involved). Only two residues (LYS22 and TYR53) of both monomers in dimer involved in water interactions were also involved in protein–protein interaction for 10–40% of time. ARG82 (Chain U) and SER105 (Chain P) in dimer of individual monomer exhibited similar protein–protein interactions lasting > 40% (Fig. 8B). From analysis of protein–water long-lasting interaction, we have indicated that the residues in loop regions and one residue (GLU14) in binding chamber are critical for the interaction between EmrE protein and water molecules (Fig. 5C and S5).

Table 2
Residues involved in substrate binding identified by biochemical and mutagenesis studies [34] (left column) compared to residues with H-bond interaction in our simulated structure of EmrE protein (apo dimer system) (right columns).

Residues	Chain P	Chain U
LEU7	P–L, P–W	P–L
ALA10	×	×
ILE11	×	×
GLU14	P–W	P–W
GLY17	×	×
THR18	×	P–W
TYR40	P–P	P–L
PHE44	×	P–P
ALA48	×	×
ALA52	P–L	×
TRP63	P–P, P–W	P–P, P–W
LEU93	P–W	P–P

P = Protein, W = Water, L = Lipid.

3.6. Residues of EmrE with aromatic side chains

Aromatic amino acids (i.e., TYR, TRP, and PHE) are believed to anchor the proteins into the membrane through an interaction of their aromatic rings with the lipid headgroups. It has been proposed that TRP and TRY residues "lock" the protein into its correct orientation or level within membrane by forming interactions with lipid headgroups and water molecules in the interfacial region [34]. We first studied the orientations of the aromatic side chains of these amino acids at the interfaces. From the analysis of hydrogen bond interactions of aromatic residues with lipid or water molecules, it has been observed that most of aromatic side chain residues formed interaction either with lipid headgroups or with water molecules. This analysis is consistent with the suggested roles of these aromatic belt residues as membrane protein "anchors" [34]. Interestingly, most of the TYR, TRP, and PHE residues lie in the loop region of EmrE or are oriented so that these residues are nearest to the interfacial region (Fig. 7C). TYR4, TYR6, PHE27, TYR53, TRP76, and PHE78 of EmrE formed H-bond interaction with both water and lipid molecules (Figs. 6 and 8A). TRP31, TYR40, TRP45, and PHE79 were involved in interaction with lipid molecules only (Fig. 6), whereas PHE23, TYR60, and TRP63 were interacting with water molecules only (Fig. 8A).

3.7. Residues with basic side chains in EmrE protein

The residues with other side chains that are thought to play an important role in interaction with lipid headgroups are residues with the basic side chains, which are suggested to "snorkel" to the membrane surface where they can interact with the lipid headgroups. Snorkeling is a way of favorable positioning the charged group without moving the whole residue. Thus, the cost of snorkeling should be taken into account both in experimental and theoretical studies on protein–lipid interactions [37]. As noted, EmrE dimer has six surface-exposed residues with basic side chains (Fig. 7D). LYS22 of both monomers was involved in hydrogen bond interaction with water molecules. For one monomer (Chain U), ARG29 and ARG82 residues were interacting with both water and lipid headgroups, whereas in other monomers (Chain P), only ARG29 interacted only with lipid and ARG82 was not involved in any interaction.

Our studies have also revealed the importance of snorkeling interaction between the residues with basic side chains and lipid headgroups. More detailed examination suggests that residues (with basic side chains) of Chain U were forming more interactions with lipid and water molecules than residues in Chain P. These differences seems to be due to higher flexibility of Chain P relative to Chain U as these specific loop region in Chain P was fluctuating more actively during the whole simulation which disturbed the formation of stable hydrogen bond interactions with lipid or water molecules (Fig. 2B). If one analyzes the orientations of lipid-interacting residues with basic side chains, nearly all these residues are located in the loop region of the protein. Of these two, ARG and one LYS in each monomer lie with their side chains nearly perpendicular to the bilayer during simulation. ARG29 of both monomers lies very close to the lipid headgroup region and in a very mobile loop, which is therefore able to move toward the membrane surface to allow tight interactions with lipid headgroups (Fig. 7D).

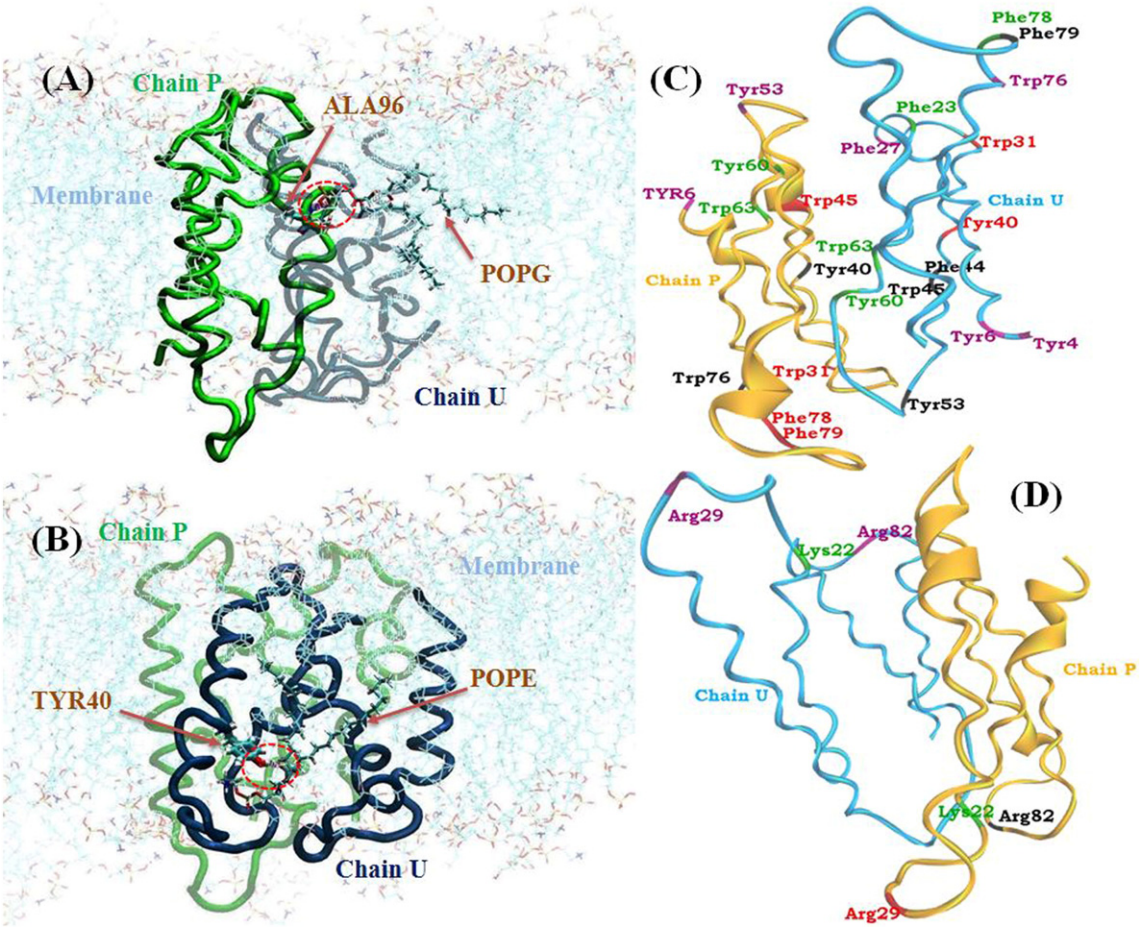

Fig. 7. Illustrative snapshots presenting hydrogen bond interaction analysis of EmrE protein model systems. **(A)** Hydrogen bond interaction pattern of ALA96 residue of Chain P in EmrE dimer with a POPG lipid molecule. **(B)** Hydroxyl group of TYR40 (Chain U in EmrE dimer) H-bonded with the POPE lipid molecule. In case of both ALA96 and TYR40 residues, lipid molecules bend slightly to make strong H-bond interactions. **(C)** Residues of EmrE with aromatic side chains (TRP, TYR, and PHE) at the lipid/water interface. **(D)** Six residues with basic side chains (LYS, ARG) in EmrE dimer. In (**C**) and (**D**), residues involved in interaction with lipid and water molecules are colored in red and green, respectively. Residues that were involved in interaction with both water and lipid molecules are colored in magenta and the residues that were not involved in any interaction are colored in black.

3.8. Comparison between experimental and our current simulated structures

Our MD results can be directly compared with structural data derived from experimental (crystallography, biochemical, and mutagenesis) studies. Twelve residues of each monomer that have been involved in substrate binding and transport were identified previously by biochemical and mutagenesis studies [4,38]. We compared our results of MD study taking into account these residues with the "static snapshot" obtained by experimental studies (Table 2).

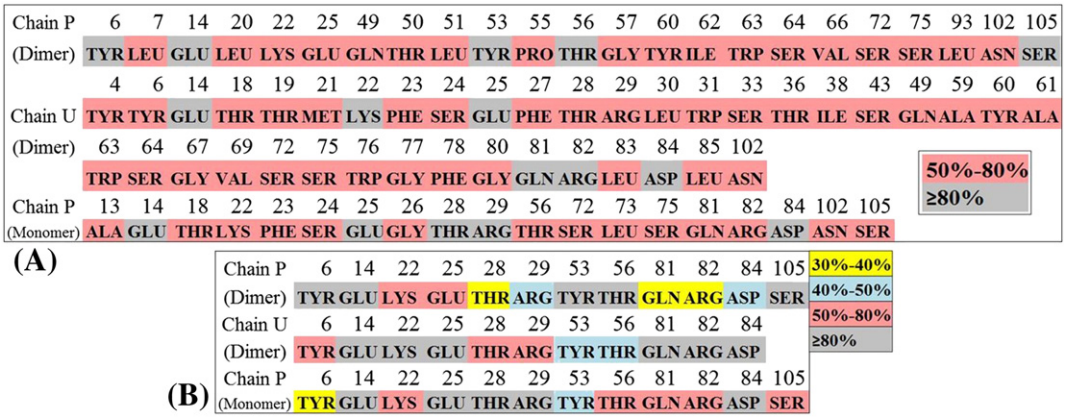

Fig. 8. (A) Interaction of protein (EmrE) with water molecules during the simulations of EmrE monomer and EmrE dimer systems. The cutoffs used to define H-bonds are 3.5 Å for the donor-acceptor distance and intermolecular angle values are in range 160°–180° between donor-H-acceptor. Contacts that persist more than 80% of the simulation time are presented in gray color and hydrogen bond interactions remained between 50% and 80% of the simulation time are presented in pink color. **(B)** Residues involved in H-bond interaction having ≥ 80% occupancy from all systems. Residues exhibiting long-lasting interaction with water in any monomer (protein) and present in all systems but exhibiting less occupancy time (30–40%) are also described (yellow color).

Residues that were common in more than one type of interactions were also listed (Table S5). During 1000 ns simulation time, GLU14 of both monomers were forming hydrogen bond interaction with water molecules for more than 80% occupancy of time. Other four residues (LEU7, THR18, TRP63, and LEU93) were involved in H-bond interaction with water molecules during 50–80% of total simulation time. The results for these four residues (Table 2) are consistent with previous studies as the residues that are interacting with water molecules are also ready to interact with potential substrate. The residues TYR40, PHE44, and ALA52 formed H-bond interaction with lipid headgroups or with other residue of protein. From previous experimental studies, GLY17 suggests a role in conformational changes rather than in substrate binding [38], and during MD studies also, GLY17 was not making interaction at all. Studies described previously suggest that aromatic residues from TM2 and TM3 have also been identified as a part of the binding site and here we showed that aromatic side chains at lipid/water interface are highly involved in H-bond interactions (Fig. 7C).

In TM3, the evolutionary conserved TRP63 has been identified as an essential residue that may be interacting directly with the substrate [38], and during our MD studies, TRP63 was involved either in protein–protein or in protein–water interaction, which also confirms that this residue is crucial in substrate binding. It was speculated that in absence of substrates, GLU14 in the binding cavity is stabilized by interaction with protons or with aromatic residues contributed by TM2 and TM3 (Fig. 2B) [38]. We confirmed in our MD study that GLU14 of both monomers is not involved in protein–protein or protein–lipid interaction at all in the absence of substrate.

3.9. Center of mass distance between two monomers and distance between GLU14 residues of both monomers (dimer system)

It has been confirmed that EmrE binds the substrate molecule near the center of the dimer. Therefore, the distances between binding key residues (GLU14) of each monomer in EmrE dimer, as a measure of binding cleft, are plotted in Fig. 9A. The situation is quite similar when it comes to the distance (center of mass) between overall structures of two monomers in dimer. In both cases, during 1000 ns production run, distances remained steady during the first 300 ns and the last 400 ns, and this distance measures about 2 Å larger for two specific GLU14 residues than that for two monomers in dimer complex (Fig. 9A). For GLU14-CA, the distance increased rapidly to 19 Å within several nanoseconds (from 300 ns to 600 ns) of simulation and was finally stabilized at around 15 Å. This result indicates that during this specific time, GLU14 was favorable to make strong interaction with water molecules and when substrate is present, water could be replaced by substrate. GLU14 is also recommended as an essential residue for substrate binding and a residue forming long-lasting interactions with water. Our simulation results are in good agreement with the experimental studies and verified that fluctuations in the distances between two GLU14 indicate that it is competent to accommodate the substrate molecules that vary in size [4,38].

3.10. Number of water molecules along the translocation pathway suggested in an alternate access model for EmrE protein [2]

To determine the internal water distribution and the dynamic behavior of water, the number of water molecules accounted around residue GLU14 is plotted; using the structures retrieved from entire production runs from EmrE dimer simulation. As shown in Fig. 9B, the number of water entrapped within binding chamber increased to 60, especially at the end of 500 ns simulation time, and this number of water molecules was reduced partially at the end of the simulation period. In view of the above analysis, the EmrE dimer complex is more adept to convert to the inward- or outward-facing conformation (Fig. S2). Secondly, the analysis of distance between two GLU14 and the water

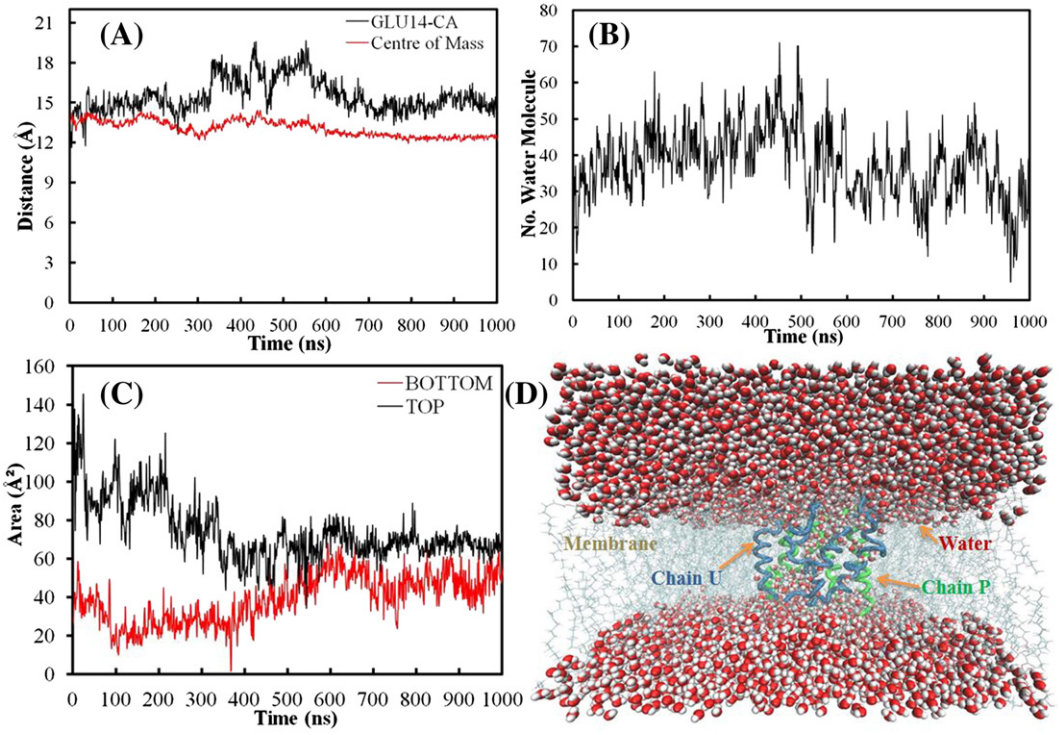

Fig. 9. **(A)** Distances of mass center between two monomers in EmrE dimer and between two GLU14 of each monomer plotted as a function of time, red line for monomers and black for GLU14. Here, GLU14 is selected as it is positioned in the center of substrate binding chamber. **(B)** Number of water molecules present in the translocation pathway as function of time. To calculate this, water molecules surrounding GLU14 of Chain U within 7 Å distance were selected, as the distance between center of mass of two monomers is about 14 Å. **(C)** Plot of water entrance area (Å^2) against time. Area is calculated by selecting three CA atoms of GLY57, GLY26, and GLY77 at the top (black line) in Chain U and same CA atoms of Chain P at the bottom (red). **(D)** Conformation extracted from EmrE antiparallel dimer trajectories, where the solvation of the internal chamber in EmrE protein is observed (*color id*: Chain P in green, Chain U in blue, lipid and water molecules are colored according to their chemical composition: carbon in cyan, nitrogen in blue, oxygen in red, hydrogen in white, and phosphorus in brown).

molecules along GLU14 within 7 Å are parallel to each other (Fig. 9A and B).

Clefts or pockets on the surface of protein interior are crucial for molecular recognition and protein function. Distinct from them are the cavities, described as enclosed space in the interior of the protein. It is of importance to know the average area occupied by water molecules in protein interior and interfaces [39]. We calculated the entrance area for water to quantify the geometrical characteristics of translocation pathway in EmrE dimer (Fig. 9C). Three GLY residues at position 26, 57, and 77 were selected as they are located at the interface of translocation pathway in dimer. Area was calculated by selecting the same residues at the top and the bottom positions in antiparallel dimer (Fig. S9A and S9B). This analysis proposes that area occupied by water molecules in translocation pathway is changing over simulation time and the pathway becomes more compact over simulation time (Fig. 9D). Overall analysis of translocation pathway (distance center of masses and water entrance area) suggests that conformational change is feasible (Fig. S2); however, it could be more pronounced in a period of time longer than 1000 ns time of simulation.

4. Conclusion

We performed a thorough computational study of EmrE multidrug transporter protein, for which transport process is known to work as a single-site alternating access model of an antiparallel dimer [3]. By analyzing the interaction pattern of this protein with lipid or water in monomer and apo-form, we were able to propose a model of this protein for the interaction with substrate or drug molecule. The model offers a reliable explanation on the flexibility of some region of protein for definite time and identifies the residues that are crucial for ligand recognition/affinity and also potential for transport properties.

Results obtained from our MD studies strongly highlight that both GLU14 residues, placed almost in the center of transmembrane part of protein, point toward the binding chamber and appear well placed to form ionic or polar contacts with positively charged substrate molecule. This is also consistent with biochemical and genetic studies describing that GLU14 is definitely required for substrate and proton-dependent transport [10]. Our studies suggest that EmrE could bind with the substrate in such a manner that GLU14 binds to the positively charged part of substrate molecule and other four aromatic residues (i.e., TRP63 and TYR40 of each monomer) located at the same level of surface accommodate aromatic fragments of substrate molecule (present for instance in TPP^+) [29].

Dynamic analysis of H-bond interactions of each residue in protein pointed out a key role of aromatic residues in drug binding, given their ability to stack with aromatic compounds and be involved in electrostatic interactions with charged compounds. It also justifies that a key feature of the binding pocket is the presence of Glutamates and a large number of aromatic residues. Our observation that four protein aromatic residues can be involved in ligand binding in dimer is consistent with experimental data indicating that symmetric TPP^+ type of ligand with four aromatic rings is one of the best binding ligands [29]. This high affinity may be due to complementary interaction with four aromatic protein residues. It is tempting to speculate that location of the residues with aromatic or basic side chains in this central area is critical for proper gating of the translocation pathway.

It is known from previous work that conformational change in this type of protein is an important process to study as it serves as a model for drug transport. The results presented here also suggest that conformational change is achievable by the apo-form or might be linked with a different mechanism for drug recognition and binding. During our MD simulation, kind of inward–outward facing conformational change was recorded what agrees with experimental observations found for apo-form of EmrE [18]. The MD study enhances the knowledge about water-interacting sites in EmrE which comprises TYR6, GLU14, LYS22, GLU25, THR28, ARG29, TYR53, TYR56, GLN81, ARG82, ASP84, and SER105 residues of each monomer. Apart from this, residues of EmrE with aromatic side chains, also forming long-lasting interaction with water or lipid molecules (TYR4, PHE27, TRP76, and PHE78), are defined important to stabilize the conformation of this protein by hydrogen bond interactions.

In conclusion, our studies revealed molecular details concerning interaction pattern and possible conformational changes of apo-form within protein monomers (in dimer), which were not available through experimental methods. This information may help to understand the mechanism of ligand recognition and binding as well as to pose (rise) experimental questions concerning symmetry requirements of protein–ligand interaction. Further studies in this direction by theoretical methods will be also necessary especially including different ligands.

Transparency Document

The Transparency document associated with this article can be found, in the online version.

Acknowledgments

Authors thanks Academic Computer Centre CYFRONET (PL-Grid project), Krakow (Poland), for providing their hardware and software resources on which molecular modeling and molecular dynamics calculations were performed.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbamem.2015.05.014.

References

[1] H. Nikaido, Multidrug resistance in bacteria, Annu. Rev. Biochem. 78 (2009) 119–146.
[2] S. Schuldiner, D. Granot, S.S. Mordoch, S. Ninio, D. Rotem, M. Soskin, C.G. Tate, H. Yerushalmi, Small is mighty: EmrE, a multidrug transporter as an experimental paradigm, News Physiol. Sci. 16 (2001) 130–134.
[3] S.J. Fleishman, S.E. Harrington, A. Enosh, D. Halperin, C.G. Tate, N. Ben-Tal, Quasi-symmetry in the cryo-EM structure of EmrE provides the key to modeling its transmembrane domain, J. Mol. Biol. 364 (2006) 54–67.
[4] Y.J. Chen, O. Pornillos, S. Lieu, C. Ma, A.P. Chen, G. Chang, X-ray structure of EmrE supports dual topology model, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 18999–19004.
[5] I. Ubarretxena-Belandia, J.M. Baldwin, S. Schuldiner, C.G. Tate, Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer, EMBO J. 22 (2003) 6175–6181.
[6] A. Gayen, J.R. Banigan, N.J. Traaseth, Ligand-induced conformational changes of the multidrug resistance transporter EmrE probed by oriented solid-stateNMR spectroscopy, Angew. Chem. Int. Ed. Engl. 52 (2013) 10321–10324.
[7] E.A. Morrison, G.T. DeKoster, S. Dutta, R. Vafabakhsh, M.W. Clarkson, A. Bahl, D. Kern, T. Ha, K.A. Henzler-Wildman, Antiparallel EmrE exports drugs by exchanging between asymmetric structures, Nature 481 (2012) 45–50.
[8] P. Lloris-Garcera, J.S. Slusky, S. Seppala, M. Priess, L.V. Schafer, G. von Heijne, In vivo Trp scanning of the small multidrug resistance protein EmrE confirms 3D structure models, J. Mol. Biol. 425 (2013) 4642–4651.
[9] P. Lloris-Garcera, F. Bianchi, J.S. Slusky, S. Seppala, D.O. Daley, G. von Heijne, Antiparallel dimers of the small multidrug resistance protein EmrE are more stable than parallel dimers, J. Biol. Chem. 287 (2012) 26052–26059.
[10] T.R. Muth, S. Schuldiner, A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE, EMBO J. 19 (2000) 234–240.
[11] S. Bachas, C. Eginton, D. Gunio, H. Wade, Structural contributions to multidrug recognition in the multidrug resistance (MDR) gene regulator, BmrR, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 11046–11051.
[12] H. Wade, MD recognition by MDR gene regulators, Curr. Opin. Struct. Biol. 20 (2010) 489–496.
[13] A.A. Neyfakh, Mystery of multidrug transporters: the answer can be simple, Mol. Microbiol. 44 (2002) 1123–1130.
[14] V.M. Korkhov, C.G. Tate, Electron crystallography reveals plasticity within the drug binding site of the small multidrug transporter EmrE, J. Mol. Biol. 377 (2008) 1094–1103.
[15] K. Wong, J. Ma, A. Rothnie, P.C. Biggin, I.D. Kerr, Towards understanding promiscuity in multidrug efflux pumps, Trends Biochem. Sci. 39 (2014) 8–16.
[16] C.G. Tate, Comparison of three structures of the multidrug transporter EmrE, Curr. Opin. Struct. Biol. 16 (2006) 457–464.
[17] V.M. Korkhov, C.G. Tate, An emerging consensus for the structure of EmrE, Acta Crystallogr. D Biol. Crystallogr. 65 (2009) 186–192.

[18] M.K. Cho, A. Gayen, J.M. Baningan, M. Leninger, N.J. Traaseth, Intrinsic conformational plasticity of native EmrE provides a pathway for multidrug resistance, J. Am. Chem. Soc. 136 (2014) 8072–8080.
[19] M.A. Lomize, I.D. Pogozheva, H. Joo, H.I. Mosberg, A.L. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes, Nucleic Acids Res. 40 (2012) D370–D376.
[20] J.E. Cronan, Bacterial membrane lipids: where do we stand? Annu. Rev. Microbiol. 57 (2003) 203–224.
[21] S. Jo, J.B. Lim, J.B. Klauda, W. Im, CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes, Biophys. J. 97 (2009) 50–58.
[22] T.B. Woolf, B. Roux, Structure, energetics, and dynamics of lipid-protein interactions: a molecular dynamics study of the gramicidin A channel in a DMPC bilayer, Proteins 24 (1996) 92–114.
[23] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (2005) 1781–1802.
[24] S.E. Feller, Molecular dynamics simulations of lipid bilayers, Curr. Opin. Colloid Interface Sci. 5 (2000) 217–223.
[25] T. Darden, D. York, L. Pedersen, Particle mesh Ewald—an N · log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[26] R.W. Pastor, B.R. Brooks, A. Szabo, An analysis of the accuracy of langevin and molecular dynamics algorithms, Mol. Phys. 65 (1988) 1409–1419.
[27] S.E. Feller, Y.H. Zhang, R.W. Pastor, B.R. Brooks, Constant pressure molecular dynamics simulation: the Langevin piston method, J. Chem. Phys. 103 (1995) 4613–4621.
[28] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[29] E.A. Morrison, K.A. Henzler-Wildman, Transported substrate determines exchange rate in the multidrug resistance transporter EmrE, J. Biol. Chem. 289 (2014) 6825–6836.
[30] A. Siarheyeva, J.J. Lopez, C. Glaubitz, Localization of multidrug transporter substrates within model membranes, Biochemistry 45 (2006) 6203–6211.
[31] F. Guo, S.C. Li, M. Wenji, L. Wang, Detecting protein conformational changes in interactions via scaling known structures, J. Comput. Biol. 20 (2013) 765–779.
[32] C.S. Goh, D. Milburn, M. Gerstein, Conformational changes associated with protein–protein interactions, Curr. Opin. Struct. Biol. 14 (2004) 1–6.
[33] M.F. Lensink, C.D. Govaerts, J.M. Ruysschaert, Identification of specific lipid-binding sites in integral membrane proteins, J. Biol. Chem. 285 (2010) 10519–10526.
[34] M.B. Ulmschneider, M.S. Sansom, Amino acid distributions in integral membrane protein structures, Biochim. Biophys. Acta 1512 (2001) 1–14.
[35] K.M. Callenberg, N.R. Latorraca, M. Grabe, Membrane bending is critical for the stability of voltage sensor segments in the membrane, J. Gen. Physiol. 140 (2012) 55–68.
[36] N.M. Levinson, S.G. Boxer, A conserved water-mediated hydrogen bond network defines bosutinib's kinase selectivity, Nat. Chem. Biol. 10 (2014) 127–132.
[37] E. Strandberg, J.A. Killian, Snorkeling of lysine side chains in transmembrane helices: how easy can it get? FEBS Lett. 544 (2003) 69–73.
[38] M. Sharoni, S.S. Mordoch, S. Schuldiner, Exploring the binding domain of EmrE, the smallest multidrug transporter, J. Biol. Chem. 280 (2005) 32849–32855.
[39] R.A. Laskowski, N.M. Luscombe, M.B. Swindells, J.M. Thornton, Protein clefts in molecular recognition and function, Protein Sci. 5 (1996) 2438–2452.