รายงานวิจัยฉบับสมบูรณ์

โครงการ การศึกษากลไกการขนส่งยา carbapenem ผ่านโปรตีนขนส่งใน *Pseudomonas aeruginosa* สำหรับการพัฒนาและเพิ่มประสิทธิภาพยาต้านเชื้อด้วยระเบียบวิธีทางคอมพิวเตอร์

โดย ผศ.ดร. ประภาศิริ พงษ์ประยูร
ภาควิชาเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

มิถุนายน 2560

สัญญาเลขที่ TRG5880230

รายงานวิจัยฉบับสมบูรณ์

โครงการ การศึกษากลไกการขนส่งยา carbapenem ผ่านโปรตีนขนส่งใน *Pseudomonas aeruginosa* สำหรับการพัฒนาและเพิ่มประสิทธิภาพยาต้านเชื้อด้วยระเบียบวิธีทางคอมพิวเตอร์

ผศ.ดร. ประภาศิริ พงษ์ประยูร
ภาควิชาเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัยและ
มหาวิทยาลัยเกษตรศาสตร์

รูปแบบ Abstract (บทคัดย่อ)

Project Code : TRG5880230
(รหัสโครงการ)

Project Title : Exploring the mechanism of carbapenem permeation through the outer membrane channel from Pseudomonas aeruginosa for future drug development: Molecular dynamics studies
(ชื่อโครงการ)

Investigator : ผศ.ดร. ประภาศิริ พงษ์ประยูร มหาวิทยาลัยเกษตรศาสตร์
(ชื่อนักวิจัย)

E-mail Address : fsciprpo@ku.ac.th

Project Period : 2 years
(ระยะเวลาโครงการ)

Abstract

Pseudomonas aeruginosa is a leading nosocomial human pathogen. The major bottleneck to fight against these bacterial is their impermeable outer membrane (OM). Only specific substrates can penetrate through OM of *P. aeruginosa* via substrate-specific porins, so they become one of the most problematic drug-resistant pathogens. Carbapenems are the most effective drugs against infection with *P. aeruginosa*. One of active carbapenems used for *P. aeruginosa* is imipenem (IMI), employing the Outer membrane carboxylate channels D1 (OccD1) as an entry way. Unlike IMI, ertapenem (ERTA) was found to show weak activity due to its permeability problem. To date, no microscopic evidence can explain why IMI is preferred over ERTA. Therefore, we primarily conducted Molecular Dynamics (MD) Simulations to discover the behaviours of these drugs inside OccD1 comparing to the ligand-free state. We reported here another possible binding site in the constriction region close to the side pore opening. Overall, both drugs employ the core lactam part to tether themselves in the binding site,

whereas the tail guides a permeation direction. L132 and F133 seem to be key interactions for the core attachment. Approximately, at least 4 hydrogen bonds are required for drug binding. The direction of L2 motion also plays a role. The inward flipping traps IMI in the constriction area, while shifting towards a membrane of L2 allows ERTA contacts more water and consequently gets expelled to the protein mouth. The opening of L2 seems to facilitate the rejection of ERTA.

เนื้อหางานวิจัยประกอบด้วย วัตถุประสงค์ วิธีทดลอง ผลการทดลอง สรุปและวิจารณ์ผลการทดลอง และข้อเสนอแนะสำหรับงานวิจัยในอนาคต

Keywords : Molecular dynamics simulations, OccD1, imipenem, ertapenem
(คำหลัก)

Introduction

The gram-negative bacteria employ water-filled porins with various grades of selectivity to transport nutrients and ions across their outer membrane [1-3]. Especially, porins were found to be an entrance for many antibiotics [4-7]. Based on their degrees of substrate selectivity, 2 groups of porins can be divided. The first group is general porins (e.g. OmpC and OmpF) mediating the passage of a range of solutes based on their size, while the second is substrate-specific porins containing unique binding site for substrate [8]. For notorious human pathogens such as *Pseudomonas aeruginosa*, they have impermeable membrane due to the absence of general porins [1]. Consequently, the uptake of nutrients and ions in *P.aerugonosa* can be occurred through substrate-specific porins (e.g. the uptake of phosphate is employed by phosphate-specific OprP channel [9]). The substrate-specific channels have a narrow pore with high affinities for substrate binding and recognition [10]. In *P.aeruginosa*, the uptake of most small and water-soluble metabolites are conducted by the Occ (Outer membrane carboxylate channel) protein family. To date, 19 proteins are in this family and divided into 2 subfamilies (OccD and OccK). Both subfamilies facilitate the transport of carboxylate-containing solutes. The OccD family prefers basic amino acids, whereas the cyclic compounds are specific for the OccK family [10]. Especially, OccD1 (or OprD), the first member of the OccD family, was found to facilitate the uptake of positively charged amino acids, small peptides, and importantly some carbapenems [11,12].

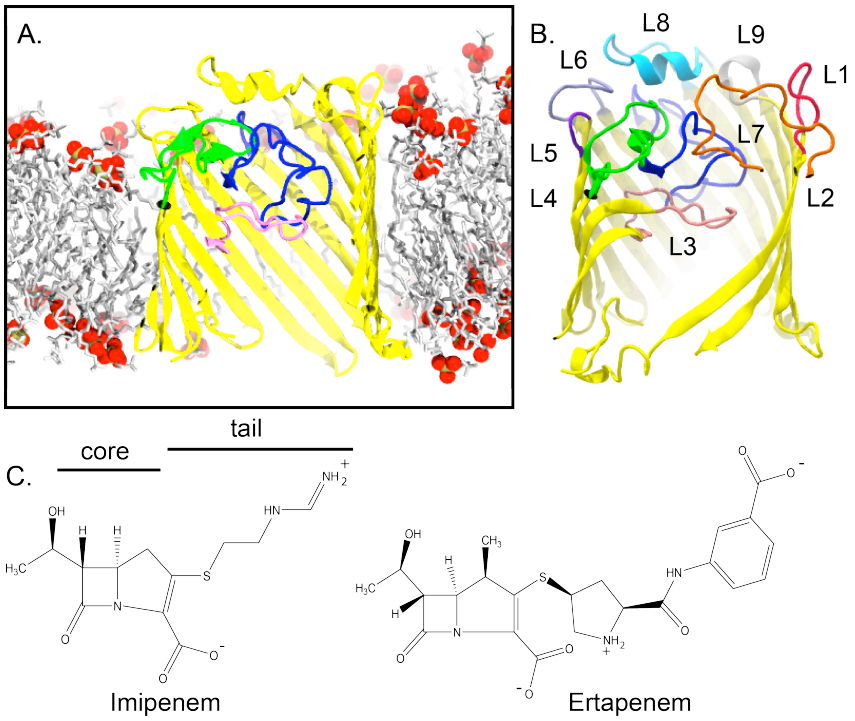


Figure 1 OccD1 embedded in a membrane is shown in (A) where the cross-sectional view of OccD1 with extracellular loops are displayed in (B). The constriction loop L3 is coloured in pink. The chemical structures of imipenem (IMI) and ertapenem (ERTA) are shown in (C).

Recently, the x-ray structures of some Occ proteins have been determined [13,14,10,15]. Both subfamilies share similar common architecture. They contain 18 β-strands connected by extracellular loops and periplasmic turns (Figure 1). The extracellular loop L3 is packed at the mid of the pore to form the so-called constriction site for solute filtering (the pink loop in Figure 1B). Recently, the crystal structures of monomeric OccD1 have been solved and it was found to be an entrance for carbapenem drugs [15,10,13]. In this work, we therefore focus on OccD1 and study its role in drug transport. Apart from L3, OccD1 also has the extracellular loop L7 inserted inside the pore and the pore opening region on the barrel between L2 and L3 (Figure 1B and 3A). Both L3 and L7 create the constriction zone. Moreover, a row of basic pore-lining residues forming the basic ladder and lysine residues are also key to the interaction with substrates [15]. To better understand the molecular selectivity and recognition, the knowledge of how natural substrates or drugs bind to the constriction area is crucial. In a recent work, Parkin *et al.* explored the mechanism of arginine permeation through OccD1 in atomic detail using computational techniques [16]. They

succeeded in predicting the pathway of arginine passage. L2 and L7 were also found to be vital for solute translocation. A year later, Isabella *et al*. conducted both experimental and computational work to probe the uptake of a series of substrates and carbapenems through OccD proteins (OccD1 and OccD3) [11]. The entries of both common carbapenems (imipenem (IMI) and meropenem) were captured there. They reported that imipenem employed the same permeation route as arginine's one, whilst meropenem followed glycine-glutamate's permeation path [11]. Both imipenem and meropenem are commonly used for an infection from *P.aeruginosa*, but fast antibiotic resistances remain [17]. Thus, new potential antibiotics are required to fight against *P.aeruginosa*. Ertapenem (ERTA) is a new carbapenem, but it shows only weak activity against *Pseudomonas* although it shares the same lactam pharmacophore as imipenem and meropenem [11,18] (Figure 1C and D). This reflects the key role of substitution group at the core in successful drug passage. To better understand the mechanisms of drug binding in *P.aeruginosa*, we explore the binding affinities of both good and poor drugs. IMI is used as a reference for a good drug and its behaviour is then compared to ERTA (a poor drug). Primarily, the binding behaviours of both drugs in the constriction area are studied here. To explore why ertapenem shows low activity against *Pseudomonas* microscopically, Molecular Dynamics (MD) simulations were conducted in this work. MD simulations have been commonly performed in many previous studies to explore the structural and dynamic properties of many porins [19-21]. This technique is successfully used to explore many solute (such as antibiotics) passages in comparison with experiments [22-25,2,20]. In this study, we therefore used MD simulations to explore ERTA and IMI activities inside the constriction area. The binding advantages and disadvantages of both drugs will be revealed in atomic level. Key protein-drug interactions will also be pointed out. The knowledge obtained here is useful for future design of potential drugs for the treatment of infections from *P.aeruginosa*.

Objectives

1. To reveal the mechanisms of current carbapenems (imipenem, ertapenem, and doripenem) translocation through the major entryway, OccD1 protein channel at atomic level.

2. To extract and compare structural and dynamic advantages and disadvantages of each carbapenem structure (imipenem, ertapenem, and doripenem) in permeation process. This information will allow us to understand the strength and

weakness of each drug structure in permeation efficiency. This understanding will serve as a key guideline for future drug development

Methodology

The coordinates of OccD1 structure was obtained from the PDB website (PDB ID: 3SY7). The protonation states of charged amino acids were set at physiological pH. The OccD1 protein was embedded in a pre-equilibrated (10 ns) dimyristoyl-phosphatidylcholine (DMPC) membrane. The protein was inserted into a membrane by using water-accessible surface of OccD1 to remove lipids and following by short steered MD simulations of the solvated proteins as explained in detail by Faraldo-Gomez *et al.* [1]. Each system contains 258 lipids and 28,708 SPC water molecules. To neutralise the system, some counter ions were added. The energy minimization of 1000 steps were performed to remove bad contacts using steepest descent algorithm followed by 10 ns protein restraint equilibration runs with a force constant of 1000 $kJmol^{-1}nm^{-2}$. For drugs, the 3D structures of both drugs were obtained from the PDB data bank (PDB code: 3UPN for IMI and 4GCS for ERTA). Both IMI and ERTA topologies were obtained using PRODRG webserver [2].

The protein-membrane snapshot at 10 ns after equilibration was used as an initial protein structure for further drug binding studies. Three of ligand-free (LF), imipenem-bound (IMI), and ertapenem-bound (ERTA) systems were set. Each drug was docked near the constriction site using Autodock4 [3]. To obtain a pose of each drug, numbers of torsions were set after detecting the root. A grid parameter file (gpf) was prepared where a ligand can move freely in the grid box (dimensions of 90×90×36 $\mathring{A}$ with a spacing of 0.375 $\mathring{A}$). The Lamarckian Genetic Algorithm (LGA) was used. The docking was carried out by setting default parameters for random number generator, energy, step size, and output. A maximum of ten different conformations was employed for the ligands. The ligand pose with the lowest binding energy was selected for MD simulations. For all protein-drug complexes, the 1000 steps of energy minimization were conducted using steepest descent algorithm followed by 10 ns protein restraint equilibration runs with a force constant of 1000 $kJmol^{-1}nm^{-2}$. Then, we performed the 100 ns production runs. Two copies with different random seeds were run for each system. All results here are an average between 2 copies.

The simulation protocols were performed following the published literature with modification [4]. We employed the GROMACS 4.5 simulation package (www.gromacs.org) [5] with GROMOS53a6 force fields for all components. The particle mesh Ewald (PME) techniques [6] with a Fourier spacing of 0.12 nm and a short range cut-off of 1 nm were used for electrostatic treatment. The simulations were conducted in the constant number of particles, pressure, and temperature (NPT) ensemble. The Berendsen algorithm at 1 bar with a coupling constant τ_p=1 ps was used for pressure coupling. The temperature of the protein, DMPC, drug, and solution were coupled separately using the v-rescale thermostat [7] at 300 K with a coupling constant τ_t=0.1 ps. The time step of 2 fs was used for integration. The coordinates were recorded every 2 ps.

All outcomes were analysed by GROMACS tools and in-house codes. The hydrogen bonds were computed using g_hbond with default parameters (The hydrogen-donor-acceptor cutoff angle is 30º and the cutoff radius (X-acceptor) is 0.35 nm). VMD was used for visualization and graphic images [8].

Results and Discussions

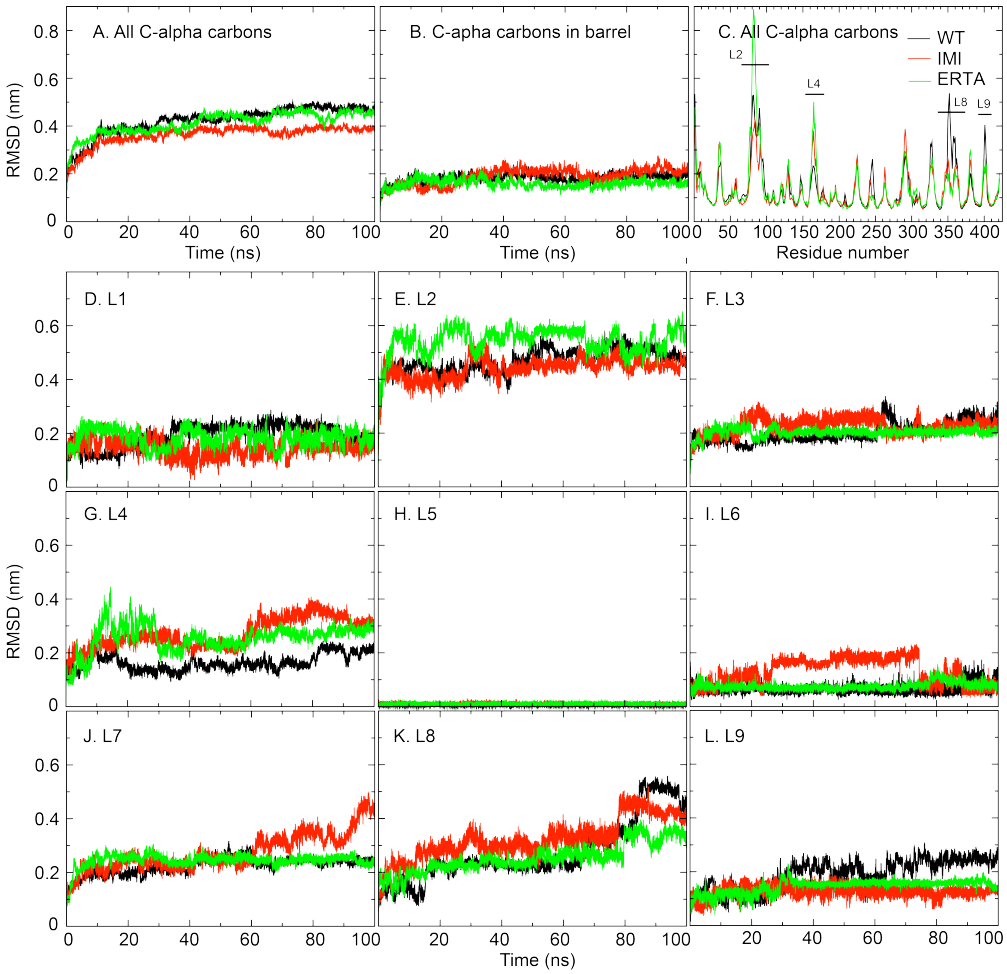


Figure 2 A set of C-alpha RMSDs is shown in (A), (B), and (D)-(L). (C) is the C-alpha RMSFs. Ligand-free (LF), IMI-bound (IMI), and ERTA-bound (ERTA) OccD1 proteins are shown in black, red, and green, respectively. The initial structure at 0 ns was used as a reference for both RMSD and RMSF calculations.

To explore the structural change and flexibility, we compute the root mean-square deviations (RMSD) and fluctuations (RMSF) of the OccD1. A set of RMSDs are calculated by comparing the movement of atoms with initial coordinates at t=0. Generally, all C-alpha RMSDs show that binding to IMI slightly increases overall protein stability (Figure 2A). However, comparing to all-atom RMSDs (~0.4 nm in Figure 2A), the much lower RMSDs of barrel observed in Figure 2B (~0.2 nm) indicate the stable barrel region. The major structural fluctuation seems to be from loop regions. This agrees well with previous work [33]. Moreover, the presence of drugs (both IMI and ERTA) does not affect barrel stability. Generally, the least flexible loop is L5 due to its shortest length (Figure 2H). The constriction loops L3 appear to be kept less mobile in

all cases in order to maintain a function as a selective filter. The large flexibilities are observed in the long extracellular loops L2, L4, and L8 (high RMSFs and RMSDs in Figure 2C, 2E, 2G, and 2K). The displacements of L2, L4, and L8 can be seen as cartoon representatives in Figure 3. Especially, L2 shows the highest loop flexibility (RMSDs ~0.4-0.6 nm and RMSFs ~ 0.9 nm) due to its long length (15 residues). Previous work reported that this long loop L2 could either interact with a membrane or stay closer to a protein [16], but no detail on how it responds to the existence of drugs is provided. In a presence of drugs, we find that L2 is initially aligned nearly parallel to the protein axis before it folds back to the protein in LF and IMI, whilst it stays in contact with a membrane surface in ERTA (Figure 3). ERTA binding makes the motion of L2 differ from LF and IMI. The outward displacement of L2 is observed in ERTA (Figure 3C). This movement generates more space for ERTA to escape to the bulk. On the contrary, shifting toward the pore entrance allows L2 to protect IMI from leaving. Apparently, the L2 relocation plays a role in drug binding affinity. The equivalent L2 motions are observed in the repeat simulation of each system (Figure S1 in supplementary). This observation can be confirmed by the number of contacts in Figure 3C. Clearly, high numbers of lipid and water contacts indicate L2 in ERTA is more lipid- and water-exposed (~45 protein contacts, ~55 lipid contacts, and ~2000 water contacts), whereas L2 of LF and IMI are packed at the OccD1's mouth (approximately ~75 protein contacts, ~35 lipid contacts, and ~1300 water contacts) (Figure 3D). Additionally, our finding supports previous experimental studies where the role of L2 in IMI uptake is highlighted [34,35]. Not only does our study agree well with previous work, but also further find the role of L2 in trapping IMI inside the pore. In case of L7, an increase in loop fluctuation is only observed when binding to IMI (Figure 2J). Binding to IMI permits the L7 reorientation (the dark blue loop in Figure 3B). This displacement may initiate the IMI passage. L2 shifting towards the OccD1's mouth and reorienting L7 conformation inside an eyelet may serve as a signal to begin the IMI uptake process. However, more further experimental studies are needed.

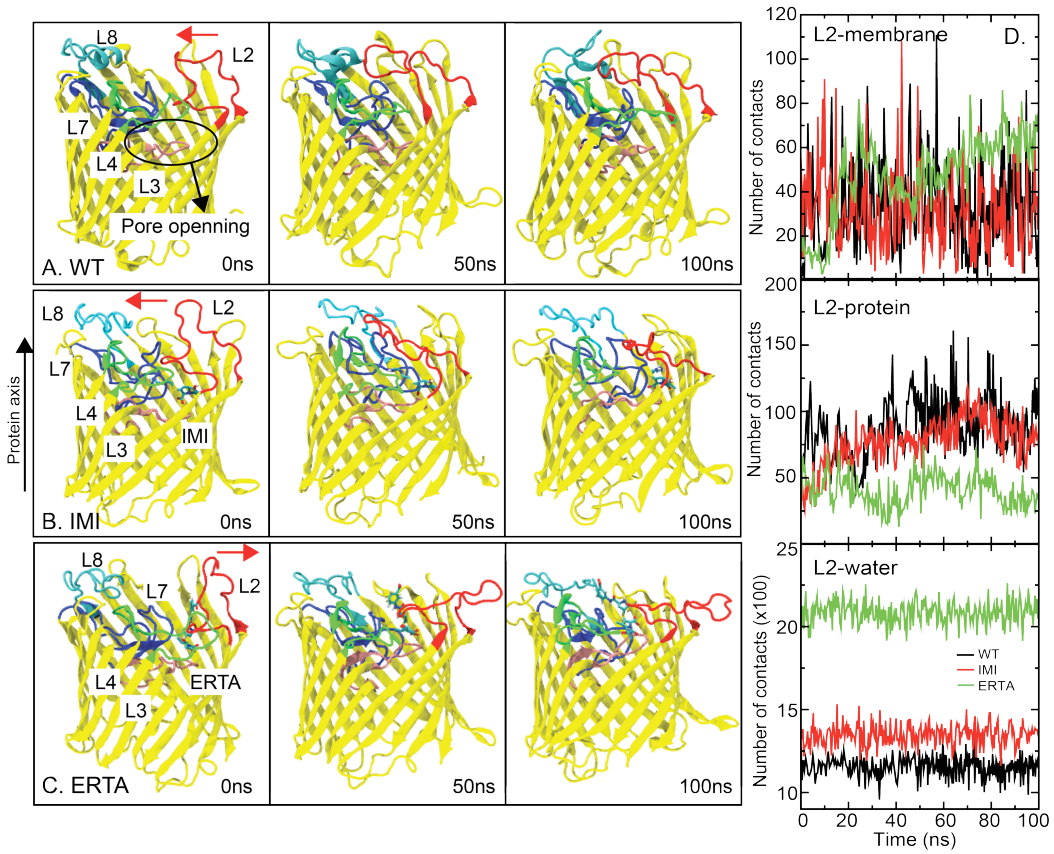


Figure 3 (A)-(C) are the protein shapshots at times 0ns, 50ns, and 100ns. Key loops, L2, L3, L4, L7, and L8 are coloured in red, pink, green, blue, and cyan, respectively. The red arrow displays the direction of L2 movement. (D) shows average numbers of water, protein, ad membrane contacts with loop L2.

Considering IMI and ERTA, both are docked into the mid of the pore. In this work, we focus only on how both IMI and ERTA behave in the constriction region. Since binding to the constriction site is the key part for successful drug permeation, we here aim at pointing out key amino acids and interactions that help OccD1 recognize each drug. Also, advantages and disadvantages of IMI and ERTA binding in the eyelet will be explored in detail here. At the beginning, both drugs are placed at the same position in the eyelet at z ~ 6 nm (Figure 4A and 4B). We discover another possible drug binding site in the constriction area near the side pore opening (a green circle in Figure 4B). No full permeation is observed for both. Within 20 ns, the displacements of both drugs are observed (Figure 4A). IMI drugs in both simulations (IMI1 and IMI2) migrate to the tip of L3 (z~ 6-6.5 nm in Figure 4A and 4B), while ERTA in ERTA1 and ERTA2 are exiled to z ~ 6.5-7 nm, which is at the mouth of OccD1 (Figure 4A and 4B). Both drugs similarly

have at least 2 hydrogen bonds with the protein, but ERTA shows more water-exposed. ERTA gets approximately surrounded by ~6 water molecules, while only ~2 waters hydrogen bond to IMI (Figure 4C). Comparing to IMI, higher water contacts and larger size allow ERTA to easily get expelled to the mouth.

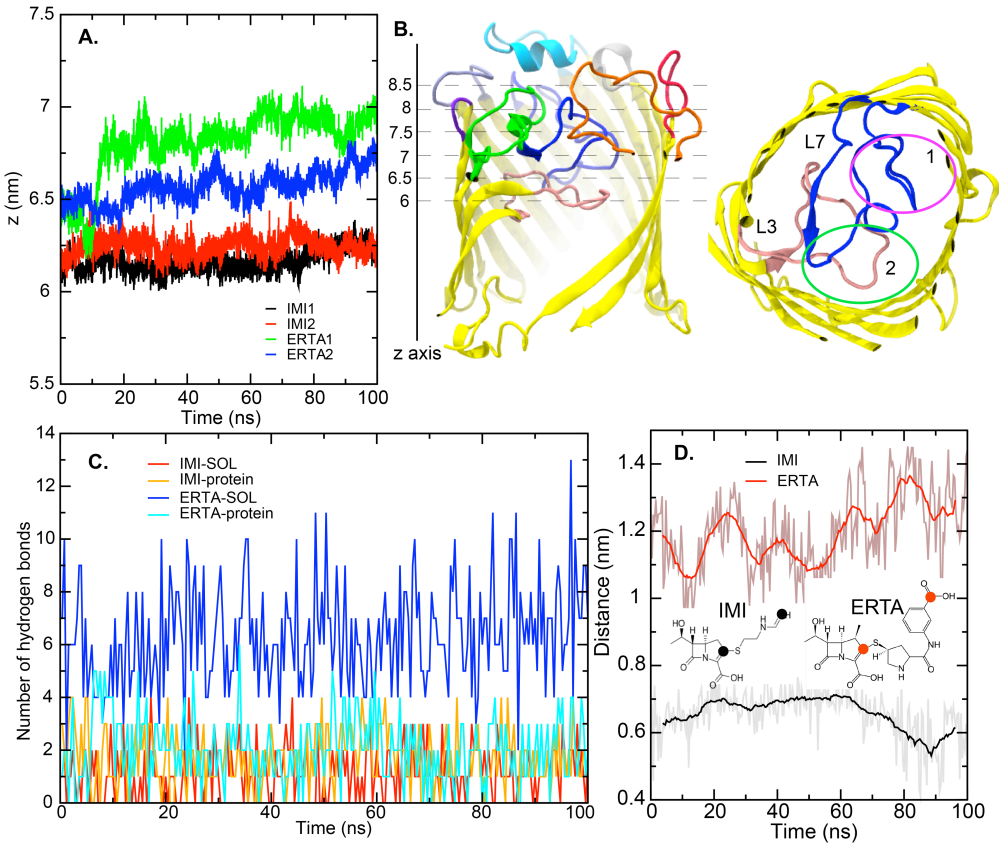


Figure 4 Positions of IMI and ERTA along z axis in all simulations (A). (B) the left is the cross-sectional protein structure while the right is the top view where 1 (pink circle) is the binding site observed by Khalid et al. [4] and 2 (green circle) is the second site that we observed in this work. The numbers of drug-protein and drug-water hydrogen bonds are shown in (C). (D) is the distance between 2 given atoms on each drug. The 2 given atoms are coloured in black for IMI and red for ERTA.

Table1 Percentages of protein-drug hydrogen bonds (%Hb) occurring during simulations. Only residues that form > 10% hydrogen bonds are included here. The %Hb was computed from the frequency of Hb occurring throughout all frames in each simulation.

ERTA1		ERTA2		IMI1		IMI2	
Residue	%Hb	Residue	%Hb	Residue	%Hb	Residue	%Hb
N89	15.2	R94	17.6	D41	67.9	T43	20.0
G91	18.0	D95	14.4	T43	29.4	D95	19.6
L132	52.8	R131	26.8	L132	45.6	L132	27.6
F133	53.2	L132	72.4	F133	47.4	F133	51.6
Q135	23.6	F133	82.8				
		S290	40.0				
		G291	22.2				
		A292	12.8				

To get deeper microscopic view, the drug-protein hydrogen bonds are computed. L132 and F133 can strongly interact with all drugs. Especially, hydrogen bonding to L132 (52.8% in ERTA1 and 53.2% in ERTA2) and F133 (72.4% in ERTA1 and 82.8% for ERTA2) seem to be more preferable for ERTA (Table 1). Additionally, hydrogen bonds with T43 (29.4% for IMI1 and 20% for IMI2), L132 (45.6% for IMI1 and 27.6% for IMI2), and F133 (47.4% for IMI1 and 51.6% for IMI2) are important for IMI binding in both simulations (Table 1). Totally, IMI seems to have 4 hydrogen bonds. Both IMI1 and IMI2 similarly form hydrogen bonds with T43, L132, F133 and one more bond with aspartate. IMI1 has an extra interaction with D41 (67.9%) where IMI2 gains one more interaction from D95 (19.6%). Even though both IMI1 and IMI2 interact with aspartate, these aspartate residues are from different locations. The carboxylic group on the lactam core is stabilised by hydrogen bonding with the backbones of L132 and F133 and interaction with T43. The core part appears to be less mobile and sits nicely inside the pore due to the bonds with L132, F133, T43, and either D41 or D95 in this region. With 4 observed interactions, IMI seems to employ this lactam core to stick in the constriction area. On the other hand, another end of IMI is more mobile. It can flip back and forth therefore permanent hydrogen bonds cannot be captured here. With its

flexibility, the IMI tail may serve as a guide to find the right pathway and trigger IMI permeation. This tail fluctuation observed here agree well with previous MD study where the mobile ARG side chain was captured during transport [16]. Comparing to IMI, the higher number of hydrogen bond donors and acceptors allow ERTA to form more hydrogen bonds with OccD1. Like IMI, the lactam core appears to play a role in attaching to the constriction area. The lactam part mainly interacts with L132 (52.8% in ERTA1 and 72.4% in ERTA2) and F133 (53.2% in ERTA1 and 82.2% in ERTA2) (Table 1). Some partial interactions with N89 (15.2%), G91 (18%), and Q135 (23.6%) in ERTA1 and R94 (17.6%) and D95 (14.4%) in ERTA2 are observed (Table 1). Besides, the pyrrolidine end appears to point towards the pore entrance (Figure 5C and 5D). This end in ERTA2 can form hydrogen bonds with S290 (40%), the backbones of G291 (22.2%) and A292 (12.8%) on L7 and R131 (26.8%) on L3, whilst that of ERTA1 is exposed to water (Figure 5C and D). Both ERTA1 and ERTA2 adopt similar conformation where the lactam core sits inside the constriction area and gets stabilised by residues on L2 and L3, whereas the pyrrolidine end is flexible. Both IMI and ERTA share similar binding pattern. They employ the lactam core to stick inside the eyelet region and then use the other end to direct the passage. Comparing to IMI, the fluctuated distance between the heavy atom at the end and the carbon on the core observed in ERTA indicates the mobile tail (Figure 4D). Nonetheless, ERTA is larger and more water-accessible due to ~4 water-drug hydrogen bonds higher than IMI observed in Figure 4C. This makes ERTA more difficulty to penetrate into the pore. The loss of protein contacts allows ERTA's side chain to float towards the Occd1's mouth. The chain of ERTA aligns parallel to a protein axis, whereas the lactam part attached to the constriction region (Figure 5C and 5D). Unlike ERTA, the IMI's tail seems to lie down on the L3 loop region (Figure 5A and 5B). The major problem of ERTA is its bulky, mobile, and more water-exposed tail. With a lactam base, the aromatic tail seems to be less preferable for OccD1. Additionally, the opening of the loop L2 also facilitates the rejection of ERTA. Considering drug-L2 interactions, ERTA appears to form more hydrogen bonds with L2 due to a closer distance to the pore mouth, while IMI shows less bonds (Figure S2 in supplementary information). L2 in both LF and IMI are packed on top of the pore entrance, whereas that in ERTA moves towards the membrane causing more water exposure to ERTA and consequently the loss of ERTA-protein

contacts (Figure 3A-C). This clearly reflects the significance of L2 dynamics on drug binding.

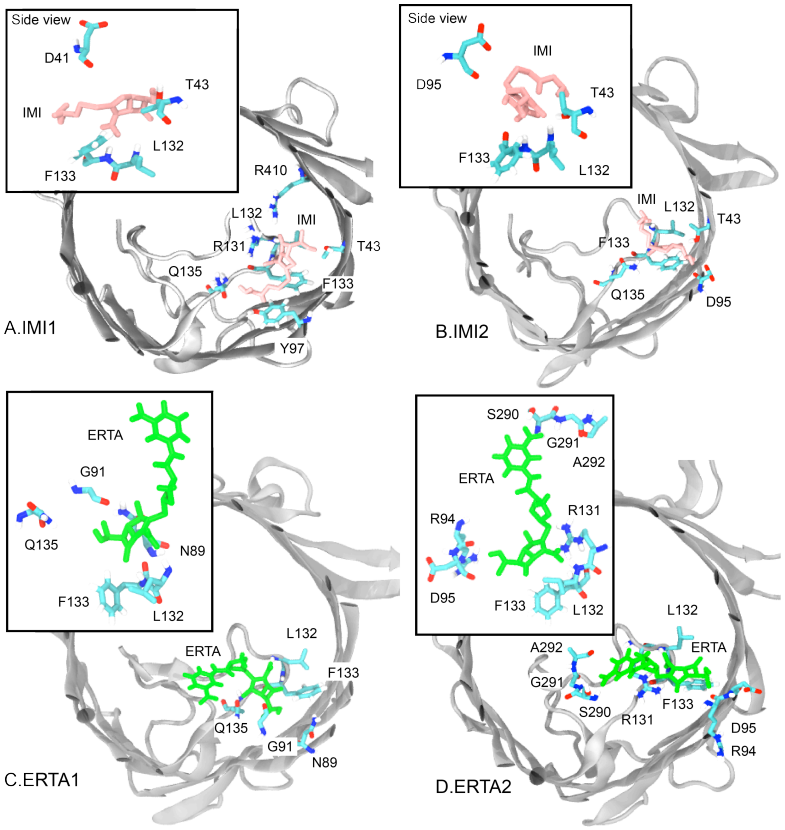


Figure 5 (A)-(D) Cartoon representations of OccD1 with bound IMI (Pink) and ERTA (green) in all simulations. The key amino acids are in licorice format.

Conclusion

The work described here displays the behaviours of two drugs inside the OccD1's constriction region. We found the new possible binding site in the L3 region close to the side pore opening. Our modelling efforts can primarily show differences in binding properties of both IMI and ERTA. Both drugs employ their core lactam part to stick to the constriction site near L3, whilst the ability to exist inside depends on their tail. Combining highly water-exposed and mobile tail with L2 opening appears to aid the exit of ERTA. Seemingly, the tail's properties are one of keys to successful drug translocation. Like ARG, OccD1 seems to prefer the positively-charged tail of IMI over the negatively charged one on ERTA. Initially, entering the cell is one of major obstacles for ERTA. Furthermore, the dynamics of L2 and L7 observed here emphasizes the role of such loops in drug recognition and passage. Our finding can primarily explain why

ERTA gets rejected. Nonetheless, questions on how protein mechanically controls the dynamics and responses of L2 and L7 to accept or reject substrates or drugs at the beginning of permeation process remain largely open. Further studies are required.

A previous work successfully employed Steered Molecular Dynamics (SMD) simulations to track the solute permeation pathway and reveal some key interactions [16]. With the force applied, this method may not allow the solute to explore all possible binding regions. Only major binding site and interactions may be captured. With conventional MD simulations utilised in this work, although no full permeation mechanism is observed, we find another possible drug binding site near the side pore opening area and the partial mechanism of drug acceptance/rejection at the pore entrance. Some key movements and interactions are reported. This finding should primarily help scientists to design and develop potential drugs with higher permeability.

Suggestions for future research work

This work provides an insight on how IMI is preferred to ERTA. There are some key microscopic mechanisms observed such as loop movements and protein-drug interaction network. We found that the tail of drug plays a role in drug recognition and binding therefore this information will be useful for future design and development of more potent carbapenem. Moreover, key residues 132 and 133 are also suggested here to be crucial for drug binding. Further mutagenesis work will be required for confirmation.

Supplementary Information

PROBING BINDING AFFINITIES OF IMIPENEM AND ERTAPENEM WITH OUTER MEMBRANE CARBOXYLATE CHANNEL D1 (OCCD1) FROM *P.AERUGINOSA*: SIMULATION STUDIES

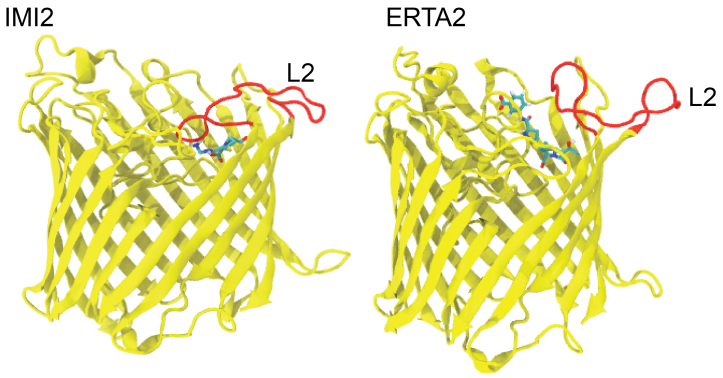


Figure S1 Cartoon views of OccD1 with bound drug at 100 ns. L2 is coloured in red.

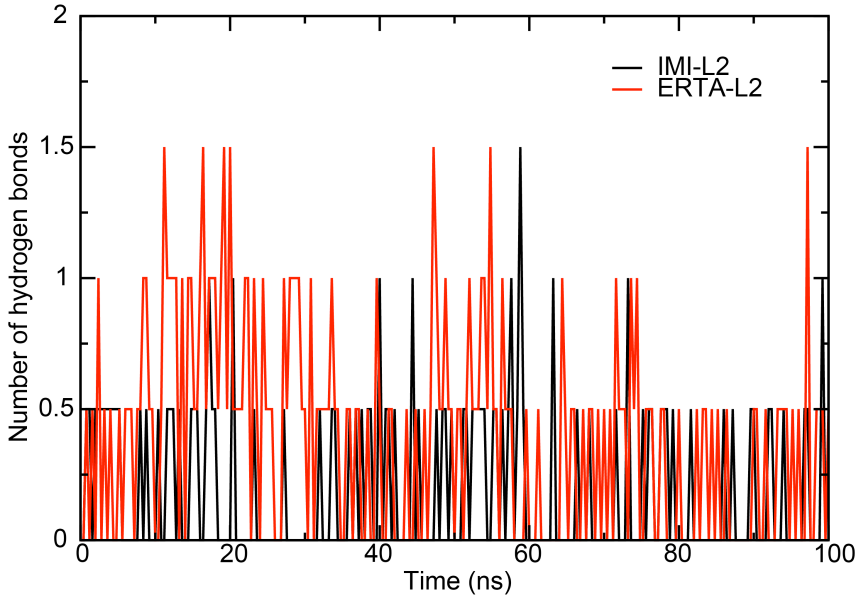


Figure S2 Average hydrogen bonds between drugs and the extracellular loop L2

References

1. Faraldo-Gomez JD, Smith GR, Sansom MS (2002) Setting up and optimization of membrane protein simulations. Eur Biophys J 31 (3):217-227

2. Schuttelkopf AW, van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr 60 (Pt 8):1355-1363. doi:10.1107/S0907444904011679
S0907444904011679 [pii]

3. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. J Mol Recognit 9 (1):1-5. doi:10.1002/(SICI)1099-1352(199601)9:1<1::AID-JMR241>3.0.CO;2-6

4. Niramitranon J, Sansom MS, Pongprayoon P (2016) Why do the outer membrane proteins OmpF from E. coli and OprP from P. aeruginosa prefer trimer? Simulation studies. J Mol Graph Model 65:1-7. doi:10.1016/j.jmgm.2016.02.002

5. Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. Journal of Molecular modeling 7:306-317

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

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

8. Humphrey W, Dalke A, Schulten K (1996) VMD - Visual Molecular Dynamics. J Mol Graph 14:33-38

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ (ระบุชื่อผู้แต่ง ชื่อเรื่อง ชื่อวารสาร ปี เล่มที่ เลขที่ และหน้า) หรือผลงานตามที่คาดไว้ในสัญญาโครงการ

งานวิจัยหลักที่ได้รับทุนวิจัยจาก สกว.

ชื่อเรื่อง Probing binding affinities of imipenem and ertapenem with outer membrane carboxylate channel D1 (OccD1) from *P.aeruginosa*: Simulation studies

ผู้แต่ง Kamolrat Somboon, Jitti Niramitranon, and Prapasiri Pongprayoon*

วารสาร: Journal of molecular modelling

DOI: 10.1007/s00894-017-3400-2

งานวิจัยรองที่มาจากทุนวิจัยจาก สกว.

1. ชื่อเรื่อง Multiscale simulation studies of geometrical effects on solution transport through nanopores

ผู้แต่ง Atthaphon Chaimanatsakun, Deanpen Japrung, and Prapasiri Pongprayoon*

วารสาร: Molecular simulation

DOI: 10.1080/08927022.2017.1334881

2. ชื่อเรื่อง Exploring the interactions of a DNA aptamer with human serum albumins: simulation studies

ผู้แต่ง Wanwisa Panman, Deanpen Japrung, and Prapasiri Pongprayoon*

วารสาร: Journal of biomolecular Structure and Dynamics

DOI: 10.1080/07391102.2016.1224733

3. ชื่อเรื่อง Why do the outer membrane proteins OmpF from E.coli and OprP from *P.aeruginosa* prefer trimer? Simulation Studies

ผู้แต่ง Jitti Niramitranon, Mark SP Sansom, and Prapasiri Pongprayoon*

วารสาร: Journal of molecular graphics and modelling

DOI: 10.1016/j.jmgm.2016.02.002

Journal of Molecular Modeling

PROBING BINDING AFFINITIES OF IMIPENEM AND ERTAPENEM WITH OUTER MEMBRANE CARBOXYLATE CHANNEL D1 (OCCD1) FROM P.AERUGINOSA: SIMULATION STUDIES

--Manuscript Draft--

Manuscript Number:	JMMO-D-17-00114R3
Full Title:	PROBING BINDING AFFINITIES OF IMIPENEM AND ERTAPENEM WITH OUTER MEMBRANE CARBOXYLATE CHANNEL D1 (OCCD1) FROM P.AERUGINOSA: SIMULATION STUDIES
Article Type:	Original paper
Keywords:	Outer membrane proteins; Outer membrane carboxylate channels; MD simulations; Imipenem; Ertapenem; Carbapenem
Corresponding Author:	Prapasiri Pongprayoon, DPhil Kasetsart University Faculty of Science Bangkok, THAILAND
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Kasetsart University Faculty of Science
Corresponding Author's Secondary Institution:	
First Author:	Kamolrat Somboon, MSc
First Author Secondary Information:	
Order of Authors:	Kamolrat Somboon, MSc
	Jitti Niramitranon, DPhil
	Prapasiri Pongprayoon, DPhil
Order of Authors Secondary Information:	
Funding Information:	Thailand Research Fund (TRG5880230) — Dr Prapasiri Pongprayoon
	Kasetsart University Research and Development Institute — Dr Prapasiri Pongprayoon
Abstract:	Pseudomonas aeruginosa is a leading nosocomial human pathogen. The major bottleneck to fight against these bacterial is their impermeable outer membrane (OM). Only specific substrates can penetrate through OM of P. aeruginosa via substrate-specific porins, so they become one of the most problematic drug-resistant pathogens. Carbapenems are the most effective drugs against infection with P. aeruginosa. One of active carbapenems used for P. aeruginosa is imipenem (IMI), employing the Outer membrane carboxylate channels D1 (OccD1) as an entry way. Unlike IMI, ertapenem (ERTA) was found to show weak activity due to its permeability problem. To date, no microscopic evidence can explain why IMI is preferred over ERTA. Therefore, we primarily conducted Molecular Dynamics (MD) Simulations to discover the behaviours of these drugs inside OccD1 comparing to the ligand-free state. We reported here another possible binding site in the constriction region close to the side pore opening. Overall, both drugs employ the core lactam part to tether themselves in the binding site, whereas the tail guides a permeation direction. L132 and F133 seem to be key interactions for the core attachment. Approximately, at least 4 hydrogen bonds are required for drug binding. The direction of L2 motion also plays a role. The inward flipping traps IMI in the constriction area, while shifting towards a membrane of L2 allows ERTA contacts more water and consequently gets expelled to the protein mouth. The opening of L2 seems to facilitate the rejection of ERTA.
Response to Reviewers:	Response to the Reviewer2

	Reviewer #2: There is still a minor issue with one of the answers to my comments. The grid spacing for the docking protocol is stated as "0.375" but with no units. I presume this is in Angstroms? **>>**Thank you. We have added the unit as suggested. This also needs to be rewritten in better English: "The steps of energy minimization and equilibration for drug-bound systems were employed using the same condition as an earlier protein setup, but in the equilibration step, both protein is restraint" **>>** Thanks for your comment. Sorry about this again. We have rewritten them as seen below (p.5); “For all protein-drug complexes, the 1000 steps of energy minimization were conducted using steepest descent algorithm followed by 10 ns protein restraint equilibration runs with a force constant of 1000 kJmol-1nm-2.” Thanks so much for your suggestion.

PROBING BINDING AFFINITIES OF IMIPENEM AND ERTAPENEM WITH OUTER MEMBRANE CARBOXYLATE CHANNEL D1 (OCCD1) FROM *P.AERUGINOSA*: SIMULATION STUDIES

Kamolrat Somboon[1], Jitti Niramitranon[2], and Prapasiri Pongprayoon[1,3,4]*

[1]Department of Chemistry, Faculty of Science, Kasetsart University, Chatuchak, Bangkok, 10900, Thailand.

[2]Department of Computer engineering, Faculty of Engineering, Kasetsart University, Chatuchak, Bangkok, 10900, Thailand

[3]Computational Biomodelling Laboratory for Agricultural Science and Technology (CBLAST), Kasetsart University, Bangkok, 10900 Thailand

[4]Center for Advanced Studies in Nanotechnology for Chemical, Food and Agricultural Industries, KU Institute for Advanced Studies, Kasetsart University, Bangkok 10900, Thailand.

*To whom correspondence should be addressed at

E-mail: fsciprpo@ku.ac.th

Telephone: +66-2562-5555

Fax: +66-2579-3955

For submission to: Journal of molecular modeling

Abbreviations:

MD	Molecular dynamics
OccD1	Outer membrane carboxylate channel D1
OMP	Outer membrane protein
IMI	Imipenem
ERTA	Ertapenem

Keywords: Outer membrane proteins, MD simulations, Imipenem, Ertapenem, Carbapenem

Abstract

Pseudomonas aeruginosa is a leading nosocomial human pathogen. The major bottleneck to fight against these bacterial is their impermeable outer membrane (OM). Only specific substrates can penetrate through OM of *P. aeruginosa* via substrate-specific porins, so they become one of the most problematic drug-resistant pathogens. Carbapenems are the most effective drugs against infection with *P. aeruginosa*. One of active carbapenems used for *P. aeruginosa* is imipenem (IMI), employing the Outer membrane carboxylate channels D1 (OccD1) as an entry way. Unlike IMI, ertapenem (ERTA) was found to show weak activity due to its permeability problem. To date, no microscopic evidence can explain why IMI is preferred over ERTA. Therefore, we primarily conducted Molecular Dynamics (MD) Simulations to discover the behaviours of these drugs inside OccD1 comparing to the ligand-free state. We reported here another possible binding site in the constriction region close to the side pore opening. Overall, both drugs employ the core lactam part to tether themselves in the binding site, whereas the tail guides a permeation direction. L132 and F133 seem to be key interactions for the core attachment. Approximately, at least 4 hydrogen bonds are required for drug binding. The direction of L2 motion also plays a role. The inward flipping traps IMI in the constriction area, while shifting towards a membrane of L2 allows ERTA contacts more water and consequently gets expelled to the protein mouth. The opening of L2 seems to facilitate the rejection of ERTA.

Introduction

The gram-negative bacteria employ water-filled porins with various grades of selectivity to transport nutrients and ions across their outer membrane [1-3]. Especially, porins were found to be an entrance for many antibiotics [4-7]. Based on their degrees of substrate selectivity, 2 groups of porins can be divided. The first group is general porins (e.g. OmpC and OmpF) mediating the passage of a range of solutes based on their size, while the second is substrate-specific porins containing unique binding site for substrate [8]. For notorious human pathogens such as *Pseudomonas aeruginosa*, they have impermeable membrane due to the absence of general porins [1]. Consequently, the uptake of nutrients and ions in *P.aerugonosa* can be occurred through substrate-specific porins (e.g. the uptake of phosphate is employed by phosphate-specific OprP channel [9]). The substrate-specific channels have a narrow pore with high affinities for substrate binding and recognition [10]. In *P.aeruginosa*, the uptake of most small and water-soluble metabolites are conducted by the Occ (Outer membrane carboxylate channel) protein family. To date, 19 proteins are in this family and divided into 2 subfamilies (OccD and OccK). Both subfamilies facilitate the transport of carboxylate-containing solutes. The OccD family prefers basic amino acids, whereas the cyclic compounds are specific for the OccK family [10]. Especially, OccD1 (or OprD), the first member of the OccD family, was found to facilitate the uptake of positively charged amino acids, small peptides, and importantly some carbapenems [11,12].

[Figure 1 here]

Recently, the x-ray structures of some Occ proteins have been determined [13,14,10,15]. Both subfamilies share similar common architecture. They contain 18 β-strands connected by extracellular loops and periplasmic turns (Figure 1). The extracellular loop L3 is packed at the mid of the pore to form the so-called constriction site for solute filtering (the pink loop in Figure 1B). Recently, the crystal structures of monomeric OccD1 have been solved and it was found to be an entrance for carbapenem drugs [15,10,13]. In this work, we therefore focus on OccD1 and study its role in drug transport. Apart from L3, OccD1 also has the extracellular loop L7 inserted inside the pore and the pore opening region on the barrel between L2 and

L3 (Figure 1B and 3A). Both L3 and L7 create the constriction zone. Moreover, a row of basic pore-lining residues forming the basic ladder and lysine residues are also key to the interaction with substrates [15]. To better understand the molecular selectivity and recognition, the knowledge of how natural substrates or drugs bind to the constriction area is crucial. In a recent work, Parkin *et al.* explored the mechanism of arginine permeation through OccD1 in atomic detail using computational techniques [16]. They succeeded in predicting the pathway of arginine passage. L2 and L7 were also found to be vital for solute translocation. A year later, Isabella *et al*. conducted both experimental and computational work to probe the uptake of a series of substrates and carbapenems through OccD proteins (OccD1 and OccD3) [11]. The entries of both common carbapenems (imipenem (IMI) and meropenem) were captured there. They reported that imipenem employed the same permeation route as arginine's one, whilst meropenem followed glycine-glutamate's permeation path [11]. Both imipenem and meropenem are commonly used for an infection from *P.aeruginosa*, but fast antibiotic resistances remain [17]. Thus, new potential antibiotics are required to fight against *P.aeruginosa*. Ertapenem (ERTA) is a new carbapenem, but it shows only weak activity against *Pseudomonas* although it shares the same lactam pharmacophore as imipenem and meropenem [11,18] (Figure 1C and D). This reflects the key role of substitution group at the core in successful drug passage. To better understand the mechanisms of drug binding in *P.aeruginosa*, we explore the binding affinities of both good and poor drugs. IMI is used as a reference for a good drug and its behaviour is then compared to ERTA (a poor drug). Primarily, the binding behaviours of both drugs in the constriction area are studied here. To explore why ertapenem shows low activity against *Pseudomonas* microscopically, Molecular Dynamics (MD) simulations were conducted in this work. MD simulations have been commonly performed in many previous studies to explore the structural and dynamic properties of many porins [19-21]. This technique is successfully used to explore many solute (such as antibiotics) passages in comparison with experiments [22-25,2,20]. In this study, we therefore used MD simulations to explore ERTA and IMI activities inside the constriction area. The binding advantages and disadvantages of both drugs will be revealed in atomic level. Key protein-drug interactions will also be pointed out. The knowledge obtained here is useful for future design of potential drugs for the treatment of infections from *P.aeruginosa*.

Materials and Methods

The coordinates of OccD1 structure was obtained from the PDB website (PDB ID: 3SY7). The protonation states of charged amino acids were set at physiological pH. The OccD1 protein was embedded in a pre-equilibrated (10 ns) dimyristoyl-phosphatidylcholine (DMPC) membrane. The protein was inserted into a membrane by using water-accessible surface of OccD1 to remove lipids and following by short steered MD simulations of the solvated proteins as explained in detail by Faraldo-Gomez *et al.* [26]. Each system contains 258 lipids and 28,708 SPC water molecules. To neutralise the system, some counter ions were added. The energy minimization of 1000 steps were performed to remove bad contacts using steepest descent algorithm followed by 10 ns protein restraint equilibration runs with a force constant of 1000 $kJmol^{-1}nm^{-2}$. For drugs, the 3D structures of both drugs were obtained from the PDB data bank (PDB code: 3UPN for IMI and 4GCS for ERTA). Both IMI and ERTA topologies were obtained using PRODRG webserver [27].

The protein-membrane snapshot at 10 ns after equilibration was used as an initial protein structure for further drug binding studies. Three of ligand-free (LF), imipenem-bound (IMI), and ertapenem-bound (ERTA) systems were set. Each drug was docked near the constriction site using Autodock4 [28]. To obtain a pose of each drug, numbers of torsions were set after detecting the root. A grid parameter file (gpf) was prepared where a ligand can move freely in the grid box (dimensions of 90×90×36 Å with a spacing of 0.375 Å). The Lamarckian Genetic Algorithm (LGA) was used. The docking was carried out by setting default parameters for random number generator, energy, step size, and output. A maximum of ten different conformations was employed for the ligands. The ligand pose with the lowest binding energy was selected for MD simulations. For all protein-drug complexes, the 1000 steps of energy minimization were conducted using steepest descent algorithm followed by 10 ns protein restraint equilibration runs with a force constant of 1000 $kJmol^{-1}nm^{-2}$. Then, we performed the 100 ns production runs. Two copies with different random seeds were run for each system. All results here are an average between 2 copies.

The simulation protocols were performed following the published literature with modification [21]. We employed the GROMACS 4.5 simulation package (www.gromacs.org) [29] with GROMOS53a6 force fields for all components. The particle mesh Ewald (PME) techniques [30] with a Fourier spacing of 0.12 nm and a

short range cut-off of 1 nm were used for electrostatic treatment. The simulations were conducted in the constant number of particles, pressure, and temperature (NPT) ensemble. The Berendsen algorithm at 1 bar with a coupling constant τ_p=1 ps was used for pressure coupling. The temperature of the protein, DMPC, drug, and solution were coupled separately using the v-rescale thermostat [31] at 300 K with a coupling constant τ_t=0.1 ps. The time step of 2 fs was used for integration. The coordinates were recorded every 2 ps.

All outcomes were analysed by GROMACS tools and in-house codes. The hydrogen bonds were computed using g_hbond with default parameters (The hydrogen-donor-acceptor cutoff angle is 30º and the cutoff radius (X-acceptor) is 0.35 nm). VMD was used for visualization and graphic images [32].

Results and Discussion

[Figure 2 here]

To explore the structural change and flexibility, we compute the root mean-square deviations (RMSD) and fluctuations (RMSF) of the OccD1. A set of RMSDs are calculated by comparing the movement of atoms with initial coordinates at t=0. Generally, all C-alpha RMSDs show that binding to IMI slightly increases overall protein stability (Figure 2A). However, comparing to all-atom RMSDs (~0.4 nm in Figure 2A), the much lower RMSDs of barrel observed in Figure 2B (~0.2 nm) indicate the stable barrel region. The major structural fluctuation seems to be from loop regions. This agrees well with previous work [33]. Moreover, the presence of drugs (both IMI and ERTA) does not affect barrel stability. Generally, the least flexible loop is L5 due to its shortest length (Figure 2H). The constriction loops L3 appear to be kept less mobile in all cases in order to maintain a function as a selective filter. The large flexibilities are observed in the long extracellular loops L2, L4, and L8 (high RMSFs and RMSDs in Figure 2C, 2E, 2G, and 2K). The displacements of L2, L4, and L8 can be seen as cartoon representatives in Figure 3. Especially, L2 shows the highest loop flexibility (RMSDs ~0.4-0.6 nm and RMSFs ~ 0.9 nm) due to its long length (15 residues). Previous work reported that this long loop L2 could either interact with a membrane or stay closer to a protein [16], but no detail on how it responds to the existence of drugs is provided. In a presence of drugs, we find that L2

is initially aligned nearly parallel to the protein axis before it folds back to the protein in LF and IMI, whilst it stays in contact with a membrane surface in ERTA (Figure 3). ERTA binding makes the motion of L2 differ from LF and IMI. The outward displacement of L2 is observed in ERTA (Figure 3C). This movement generates more space for ERTA to escape to the bulk. On the contrary, shifting toward the pore entrance allows L2 to protect IMI from leaving. Apparently, the L2 relocation plays a role in drug binding affinity. The equivalent L2 motions are observed in the repeat simulation of each system (Figure S1 in supplementary). This observation can be confirmed by the number of contacts in Figure 3C. Clearly, high numbers of lipid and water contacts indicate L2 in ERTA is more lipid- and water-exposed (~45 protein contacts, ~55 lipid contacts, and ~2000 water contacts), whereas L2 of LF and IMI are packed at the OccD1's mouth (approximately ~75 protein contacts, ~35 lipid contacts, and ~1300 water contacts) (Figure 3D). Additionally, our finding supports previous experimental studies where the role of L2 in IMI uptake is highlighted [34,35]. Not only does our study agree well with previous work, but also further find the role of L2 in trapping IMI inside the pore. In case of L7, an increase in loop fluctuation is only observed when binding to IMI (Figure 2J). Binding to IMI permits the L7 reorientation (the dark blue loop in Figure 3B). This displacement may initiate the IMI passage. L2 shifting towards the OccD1's mouth and reorienting L7 conformation inside an eyelet may serve as a signal to begin the IMI uptake process. However, more further experimental studies are needed.

[Figure 3 here]

Considering IMI and ERTA, both are docked into the mid of the pore. In this work, we focus only on how both IMI and ERTA behave in the constriction region. Since binding to the constriction site is the key part for successful drug permeation, we here aim at pointing out key amino acids and interactions that help OccD1 recognize each drug. Also, advantages and disadvantages of IMI and ERTA binding in the eyelet will be explored in detail here. At the beginning, both drugs are placed at the same position in the eyelet at z ~ 6 nm (Figure 4A and 4B). We discover another possible drug binding site in the constriction area near the side pore opening (a green circle in Figure 4B). No full permeation is observed for both. Within 20 ns, the displacements of both drugs are observed (Figure 4A). IMI drugs in both simulations

(IMI1 and IMI2) migrate to the tip of L3 (z~ 6-6.5 nm in Figure 4A and 4B), while ERTA in ERTA1 and ERTA2 are exiled to z ~ 6.5-7 nm, which is at the mouth of OccD1 (Figure 4A and 4B). Both drugs similarly have at least 2 hydrogen bonds with the protein, but ERTA shows more water-exposed. ERTA gets approximately surrounded by ~6 water molecules, while only ~2 waters hydrogen bond to IMI (Figure 4C). Comparing to IMI, higher water contacts and larger size allow ERTA to easily get expelled to the mouth.

[Figure 4 here]

Table1 Percentages of protein-drug hydrogen bonds (%Hb) occurring during simulations. Only residues that form > 10% hydrogen bonds are included here. The %Hb was computed from the frequency of Hb occurring throughout all frames in each simulation.

ERTA1		ERTA2		IMI1		IMI2	
Residue	%Hb	Residue	%Hb	Residue	%Hb	Residue	%Hb
N89	15.2	R94	17.6	D41	67.9	T43	20.0
G91	18.0	D95	14.4	T43	29.4	D95	19.6
L132	52.8	R131	26.8	L132	45.6	L132	27.6
F133	53.2	L132	72.4	F133	47.4	F133	51.6
Q135	23.6	F133	82.8				
		S290	40.0				
		G291	22.2				
		A292	12.8				

To get deeper microscopic view, the drug-protein hydrogen bonds are computed. L132 and F133 can strongly interact with all drugs. Especially, hydrogen bonding to L132 (52.8% in ERTA1 and 53.2% in ERTA2) and F133 (72.4% in ERTA1 and 82.8% for ERTA2) seem to be more preferable for ERTA (Table 1). Additionally, hydrogen bonds with T43 (29.4% for IMI1 and 20% for IMI2), L132 (45.6% for IMI1 and 27.6% for IMI2), and F133 (47.4% for IMI1 and 51.6% for IMI2) are important for IMI binding in both simulations (Table 1). Totally, IMI seems to have 4 hydrogen bonds. Both IMI1 and IMI2 similarly form hydrogen bonds with T43, L132, F133 and one more bond with aspartate. IMI1 has an extra interaction

with D41 (67.9%) where IMI2 gains one more interaction from D95 (19.6%). Even though both IMI1 and IMI2 interact with aspartate, these aspartate residues are from different locations. The carboxylic group on the lactam core is stabilised by hydrogen bonding with the backbones of L132 and F133 and interaction with T43. The core part appears to be less mobile and sits nicely inside the pore due to the bonds with L132, F133, T43, and either D41 or D95 in this region. With 4 observed interactions, IMI seems to employ this lactam core to stick in the constriction area. On the other hand, another end of IMI is more mobile. It can flip back and forth therefore permanent hydrogen bonds cannot be captured here. With its flexibility, the IMI tail may serve as a guide to find the right pathway and trigger IMI permeation. This tail fluctuation observed here agree well with previous MD study where the mobile ARG side chain was captured during transport [16]. Comparing to IMI, the higher number of hydrogen bond donors and acceptors allow ERTA to form more hydrogen bonds with OccD1. Like IMI, the lactam core appears to play a role in attaching to the constriction area. The lactam part mainly interacts with L132 (52.8% in ERTA1 and 72.4% in ERTA2) and F133 (53.2% in ERTA1 and 82.2% in ERTA2) (Table 1). Some partial interactions with N89 (15.2%), G91 (18%), and Q135 (23.6%) in ERTA1 and R94 (17.6%) and D95 (14.4%) in ERTA2 are observed (Table 1). Besides, the pyrrolidine end appears to point towards the pore entrance (Figure 5C and 5D). This end in ERTA2 can form hydrogen bonds with S290 (40%), the backbones of G291 (22.2%) and A292 (12.8%) on L7 and R131 (26.8%) on L3, whilst that of ERTA1 is exposed to water (Figure 5C and D). Both ERTA1 and ERTA2 adopt similar conformation where the lactam core sits inside the constriction area and gets stabilised by residues on L2 and L3, whereas the pyrrolidine end is flexible. Both IMI and ERTA share similar binding pattern. They employ the lactam core to stick inside the eyelet region and then use the other end to direct the passage. Comparing to IMI, the fluctuated distance between the heavy atom at the end and the carbon on the core observed in ERTA indicates the mobile tail (Figure 4D). Nonetheless, ERTA is larger and more water-accessible due to ~4 water-drug hydrogen bonds higher than IMI observed in Figure 4C. This makes ERTA more difficulty to penetrate into the pore. The loss of protein contacts allows ERTA's side chain to float towards the Occd1's mouth. The chain of ERTA aligns parallel to a protein axis, whereas the lactam part attached to the constriction region (Figure 5C and 5D). Unlike ERTA, the IMI's tail seems to lie down on the L3 loop region

(Figure 5A and 5B). The major problem of ERTA is its bulky, mobile, and more water-exposed tail. With a lactam base, the aromatic tail seems to be less preferable for OccD1. Additionally, the opening of the loop L2 also facilitates the rejection of ERTA. Considering drug-L2 interactions, ERTA appears to form more hydrogen bonds with L2 due to a closer distance to the pore mouth, while IMI shows less bonds (Figure S2 in supplementary information). L2 in both LF and IMI are packed on top of the pore entrance, whereas that in ERTA moves towards the membrane causing more water exposure to ERTA and consequently the loss of ERTA-protein contacts (Figure 3A-C). This clearly reflects the significance of L2 dynamics on drug binding.

[Figure 5 here]

Conclusion

The work described here displays the behaviours of two drugs inside the OccD1's constriction region. We found the new possible binding site in the L3 region close to the side pore opening. Our modelling efforts can primarily show differences in binding properties of both IMI and ERTA. Both drugs employ their core lactam part to stick to the constriction site near L3, whilst the ability to exist inside depends on their tail. Combining highly water-exposed and mobile tail with L2 opening appears to aid the exit of ERTA. Seemingly, the tail's properties are one of keys to successful drug translocation. Like ARG, OccD1 seems to prefer the positively-charged tail of IMI over the negatively charged one on ERTA. Initially, entering the cell is one of major obstacles for ERTA. Furthermore, the dynamics of L2 and L7 observed here emphasizes the role of such loops in drug recognition and passage. Our finding can primarily explain why ERTA gets rejected. Nonetheless, questions on how protein mechanically controls the dynamics and responses of L2 and L7 to accept or reject substrates or drugs at the beginning of permeation process remain largely open. Further studies are required.

A previous work successfully employed Steered Molecular Dynamics (SMD) simulations to track the solute permeation pathway and reveal some key interactions [16]. With the force applied, this method may not allow the solute to explore all possible binding regions. Only major binding site and interactions may be captured. With conventional MD simulations utilised in this work, although no full permeation mechanism is observed, we find another possible drug binding site near the side pore

opening area and the partial mechanism of drug acceptance/rejection at the pore entrance. Some key movements and interactions are reported. This finding should primarily help scientists to design and develop potential drugs with higher permeability.

Acknowledgements

We would like to thank for financial supports from Thailand Research Fund (code: TRG5880230) and Kasetsart University Research and Development Institute (KURDI). We also thank Prof. Supa Hannongbua.

References

1. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev 67:593-656
2. Malek K, Maghari A (2007) Translocation and interactions of L-arabinose in OmpF porin: A molecular dynamics study. Biochem Biophys Res Commun 352 (1):104-110. doi:S0006-291X(06)02419-3 [pii]
10.1016/j.bbrc.2006.10.183
3. Galdiero S, Falanga A, Cantisani M, Tarallo R, Della Pepa ME, D'Oriano V, Galdiero M (2012) Microbe-host interactions: structure and role of Gram-negative bacterial porins. Curr Protein Pept Sci 13 (8):843-854
4. Nestorovich EM, Rostovtseva TK, Bezrukov SM (2003) Residue ionization and ion transport through OmpF channels. Biophys J 85 (6):3718-3729. doi:S0006-3495(03)74788-2 [pii]
10.1016/S0006-3495(03)74788-2
5. Basle A, Qutub R, Mehrazin M, Wibbenmeyer J, Delcour AH (2004) Deletions of single extracellular loops affect pH sensitivity, but not voltage dependence, of the Escherichia coli porin OmpF. Protein Eng Des Sel 17 (9):665-672. doi:gzh078 [pii]
10.1093/protein/gzh078
6. Todt JC, Rocque WJ, Mcgroarty EJ (1992) Effects of Ph on Bacterial Porin Function. Biochemistry 31 (43):10471-10478
7. Kojima S, Nikaido H (2013) Permeation rates of penicillins indicate that Escherichia coli porins function principally as nonspecific channels. Proc Natl Acad Sci U S A 110 (28):E2629-2634. doi:1310333110 [pii]
10.1073/pnas.1310333110
8. Naveed H, Jimenez-Morales D, Tian J, Pasupuleti V, Kenney LJ, Liang J (2012) Engineered oligomerization state of OmpF protein through computational design decouples oligomer dissociation from unfolding. J Mol Biol 419 (1-2):89-101. doi:S0022-2836(12)00224-0 [pii]
10.1016/j.jmb.2012.02.043
9. Moraes TF, Bains M, Hancock RE, Strynadka NC (2007) An arginine ladder in OprP mediates phosphate-specific transfer across the outer membrane. Nat Struct Mol Biol 14 (1):85-87
10. Eren E, Parkin J, Adelanwa A, Cheneke B, Movileanu L, Khalid S, van den Berg B (2013) Toward understanding the outer membrane uptake of small molecules by Pseudomonas aeruginosa. J Biol Chem 288 (17):12042-12053. doi:10.1074/jbc.M113.463570

11. Isabella VM, Campbell AJ, Manchester J, Sylvester M, Nayar AS, Ferguson KE, Tommasi R, Miller AA (2015) Toward the rational design of carbapenem uptake in Pseudomonas aeruginosa. Chem Biol 22 (4):535-547. doi:10.1016/j.chembiol.2015.03.018
12. Shen J, Pan Y, Fang Y (2015) Role of the Outer Membrane Protein OprD2 in Carbapenem-Resistance Mechanisms of Pseudomonas aeruginosa. PLoS One 10 (10):e0139995. doi:10.1371/journal.pone.0139995
13. Biswas S, Mohammad MM, Patel DR, Movileanu L, van den Berg B (2007) Structural insight into OprD substrate specificity. Nat Struct Mol Biol 14 (11):1108-1109. doi:Doi 10.1038/Nsmb1304
14. Biswas S, Mohammad MM, Movileanu L, van den Berg B (2008) Crystal structure of the outer membrane protein OpdK from Pseudomonas aeruginosa. Structure 16 (7):1027-1035. doi:S0969-2126(08)00184-6 [pii]
10.1016/j.str.2008.04.009
15. Eren E, Vijayaraghavan J, Liu JM, Cheneke BR, Touw DS, Lepore BW, Indic M, Movileanu L, van den Berg B (2012) Substrate Specificity within a Family of Outer Membrane Carboxylate Channels. PLoS Biol 10 (1). doi:Artn E1001242
Doi 10.1371/Journal.Pbio.1001242
16. Parkin J, Khalid S (2014) Atomistic molecular-dynamics simulations enable prediction of the arginine permeation pathway through OccD1/OprD from Pseudomonas aeruginosa. Biophys J 107 (8):1853-1861. doi:10.1016/j.bpj.2014.08.035
17. Livermore DM, Mushtaq S, Warner M (2005) Selectivity of ertapenem for Pseudomonas aeruginosa mutants cross-resistant to other carbapenems. J Antimicrob Chemother 55 (3):306-311. doi:10.1093/jac/dki009
18. Quale J, Bratu S, Gupta J, Landman D (2006) Interplay of efflux system, ampC, and oprD expression in carbapenem resistance of Pseudomonas aeruginosa clinical isolates. Antimicrob Agents Chemother 50 (5):1633-1641. doi:Doi 10.1128/Aac.50.5.1633-1641.2006
19. Tieleman DP, Berendsen HJ (1998) A molecular dynamics study of the pores formed by Escherichia coli OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. Biophys J 74 (6):2786-2801. doi:S0006-3495(98)77986-X [pii]
10.1016/S0006-3495(98)77986-X
20. Pongprayoon P (2014) How do the protonation states of E296 and D312 in OmpF and D299 and D315 in homologous OmpC affect protein structure and dynamics? Simulation studies. Comput Biol Chem 53PB:226-234. doi:10.1016/j.compbiolchem.2014.10.006
21. Niramitranon J, Sansom MS, Pongprayoon P (2016) Why do the outer membrane proteins OmpF from E. coli and OprP from P. aeruginosa prefer trimer? Simulation studies. J Mol Graph Model 65:1-7. doi:10.1016/j.jmgm.2016.02.002
22. Kumar A, Hajjar E, Ruggerone P, Ceccarelli M (2010) Structural and dynamical properties of the porins OmpF and OmpC: insights from molecular simulations. Journal of Physics-Condensed Matter 22 (45):-. doi:Artn 454125
Doi 10.1088/0953-8984/22/45/454125
23. Singh PR, Ceccarelli M, Lovelle M, Winterhalter M, Mahendran KR (2013) Antibiotic permeation across the OmpF channel: modulation of the affinity site in the presence of magnesium. J Phys Chem B 116 (15):4433-4438. doi:10.1021/jp2123136
24. Ziervogel BK, Roux B (2013) The Binding of Antibiotics in OmpF Porin. Structure 21 (1):76-87. doi:S0969-2126(12)00410-8 [pii]

10.1016/j.str.2012.10.014
25. Mahendran KR, Hajjar E, Mach T, Lovelle M, Kumar A, Sousa I, Spiga E, Weingart H, Gameiro P, Winterhalter M, Ceccarelli M (2010) Molecular Basis of Enrofloxacin Translocation through OmpF, an Outer Membrane Channel of Escherichia coli - When Binding Does Not Imply Translocation. J Phys Chem B 114 (15):5170-5179. doi:Doi 10.1021/Jp911485k
26. Faraldo-Gomez JD, Smith GR, Sansom MS (2002) Setting up and optimization of membrane protein simulations. Eur Biophys J 31 (3):217-227
27. Schuttelkopf AW, van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr 60 (Pt 8):1355-1363. doi:10.1107/S0907444904011679
S0907444904011679 [pii]
28. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. J Mol Recognit 9 (1):1-5. doi:10.1002/(SICI)1099-1352(199601)9:1<1::AID-JMR241>3.0.CO;2-6
29. Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. Journal of Molecular modeling 7:306-317
30. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys 98 (12):10089-10092
31. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81 (8):3684-3690
32. Humphrey W, Dalke A, Schulten K (1996) VMD - Visual Molecular Dynamics. J Mol Graph 14:33-38
33. Parkin J, Khalid S (2014) Atomistic Molecular-Dynamics Simulations Enable Prediction of the Arginine Permeation Pathway through OccD1/OprD from Pseudomonas aeruginosa. Biophys J 107 (8):1853-1861. doi:10.1016/j.bpj.2014.08.035
34. Ochs MM, Bains M, Hancock REW (2000) Role of putative loops 2 and 3 in imipenem passage through the specific porin OprD of Pseudomonas aeruginosa. Antimicrob Agents Chemother 44 (7):1983-1985
35. Li H, Luo YF, Williams BJ, Blackwell TS, Xie CM (2012) Structure and function of OprD protein in Pseudomonas aeruginosa: from antibiotic resistance to novel therapies. Int J Med Microbiol 302 (2):63-68. doi:10.1016/j.ijmm.2011.10.001

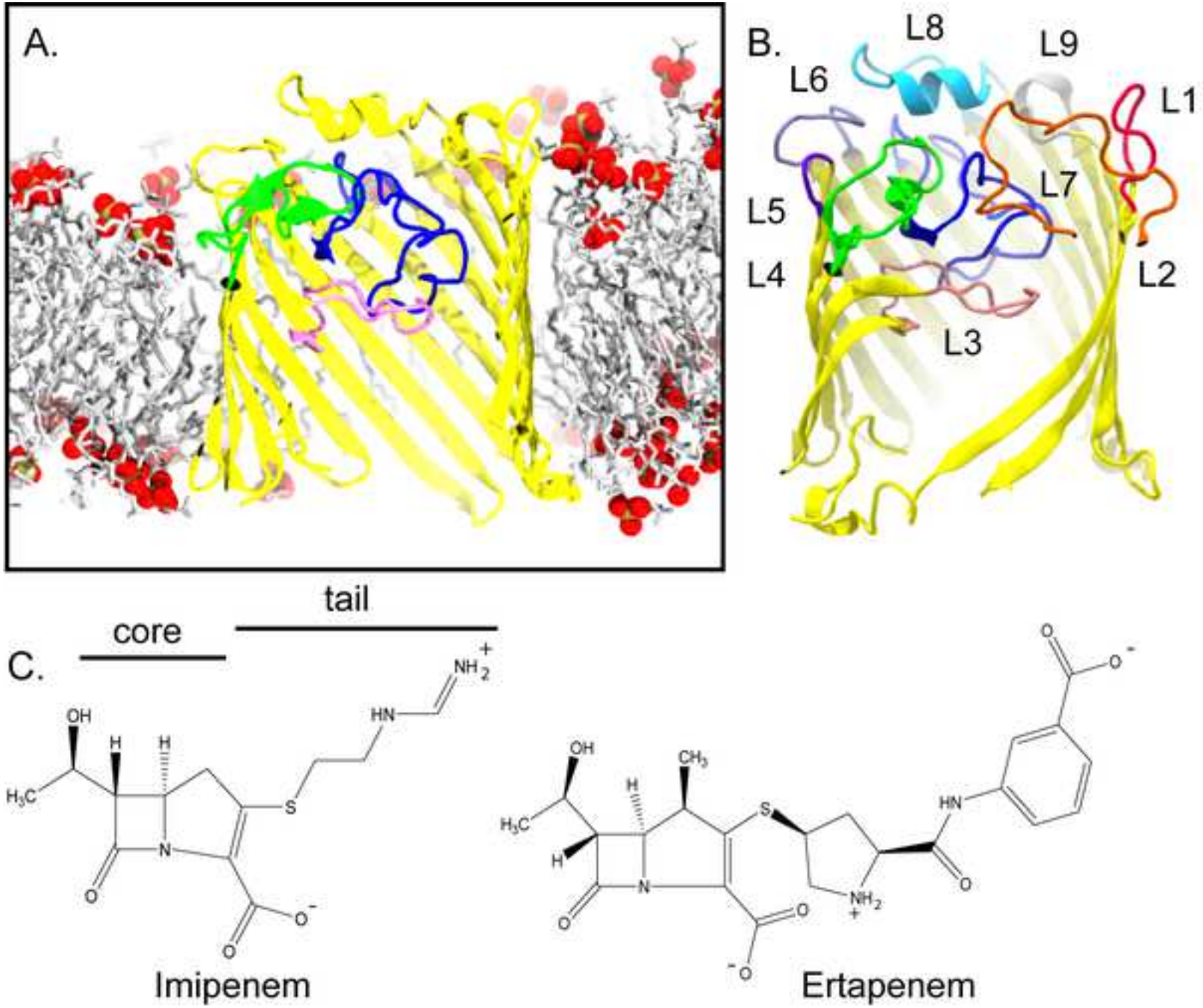


Figure 2

Click here to download colour figure Fig2-mix-rmsd-all-loops.jpg

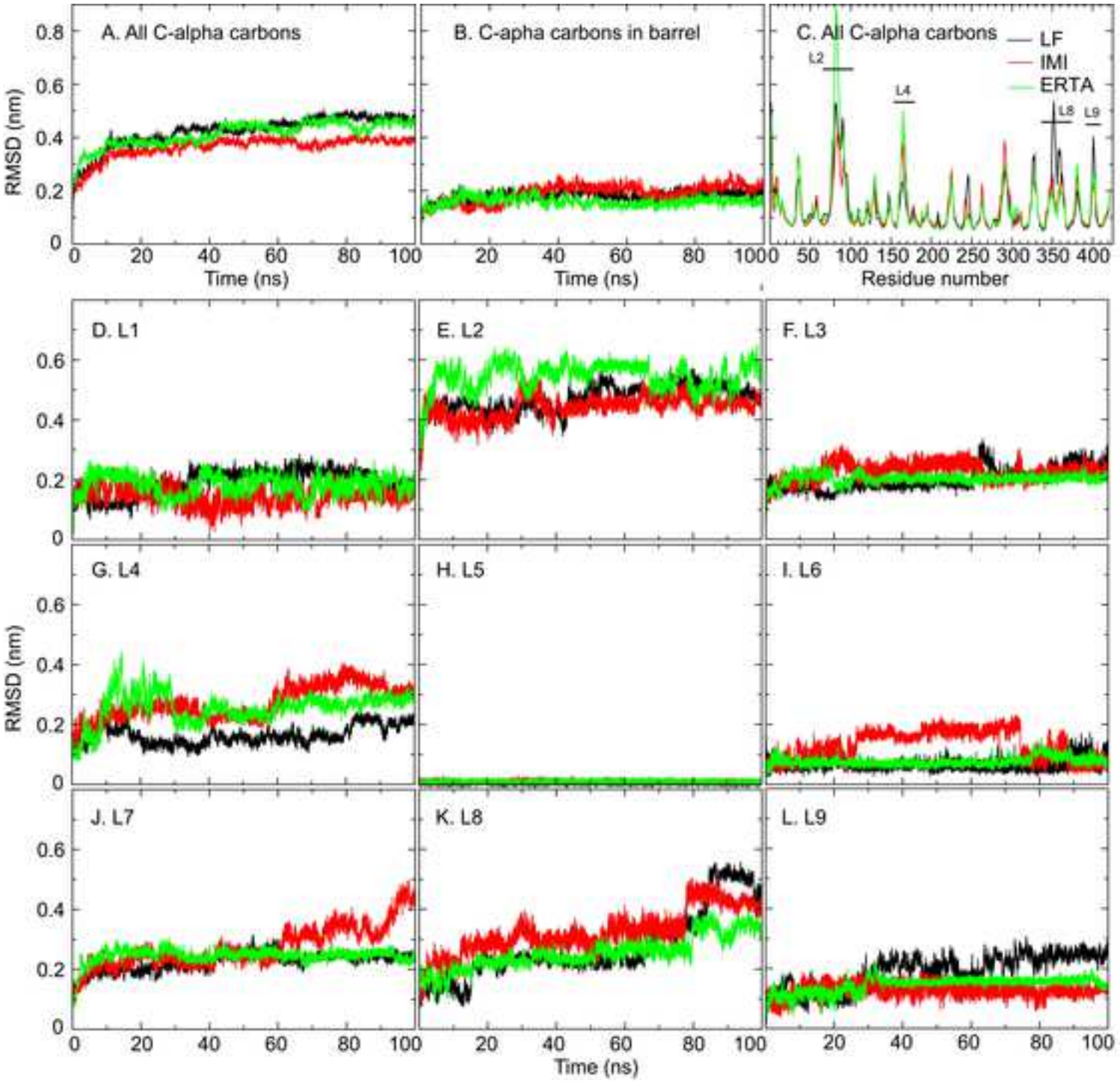


Figure 3

Click here to download colour figure Fig3-loop motion.jpg

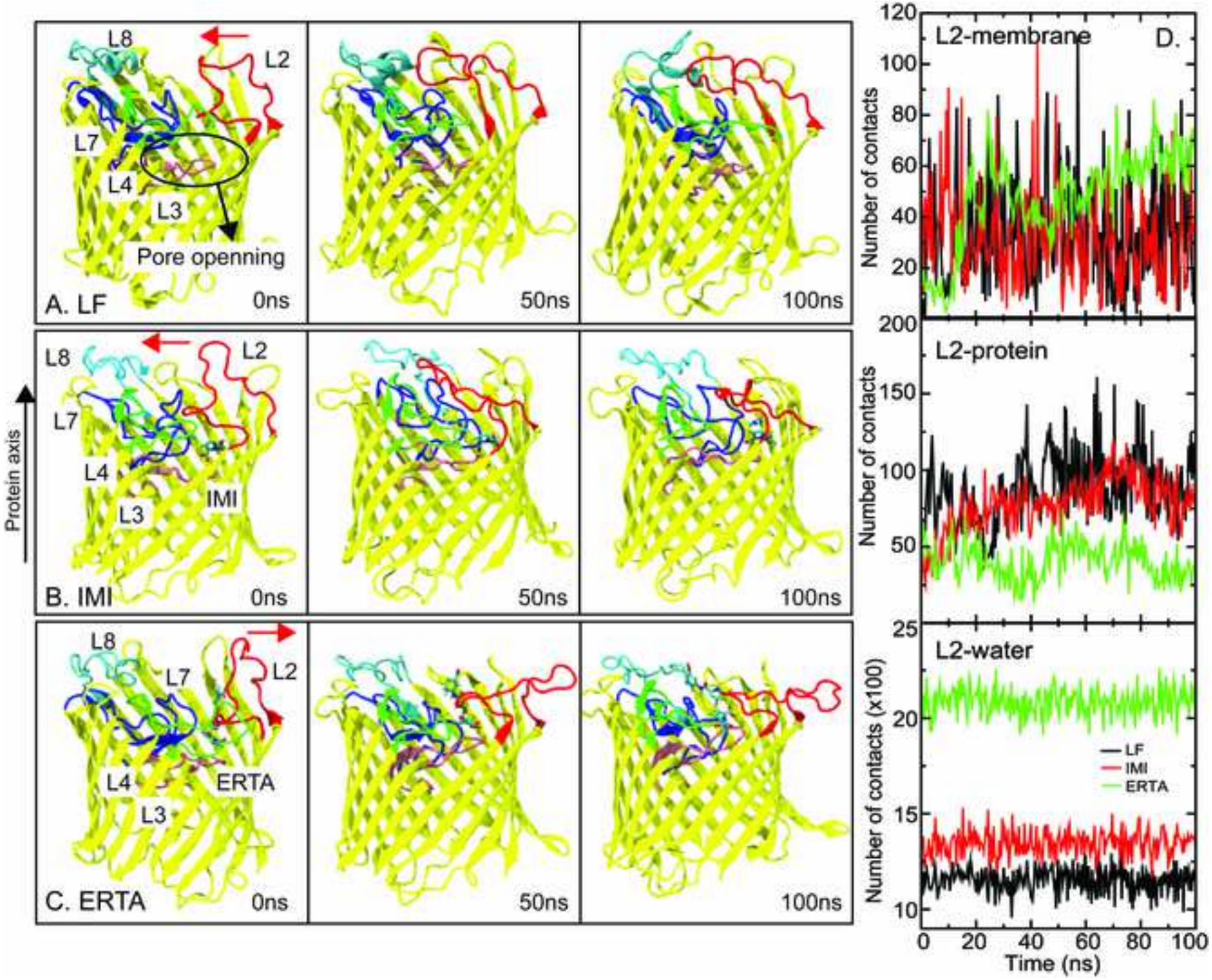


Figure 4

Click here to download colour figure Fig4-all-z-IMI-ERTA.jpg

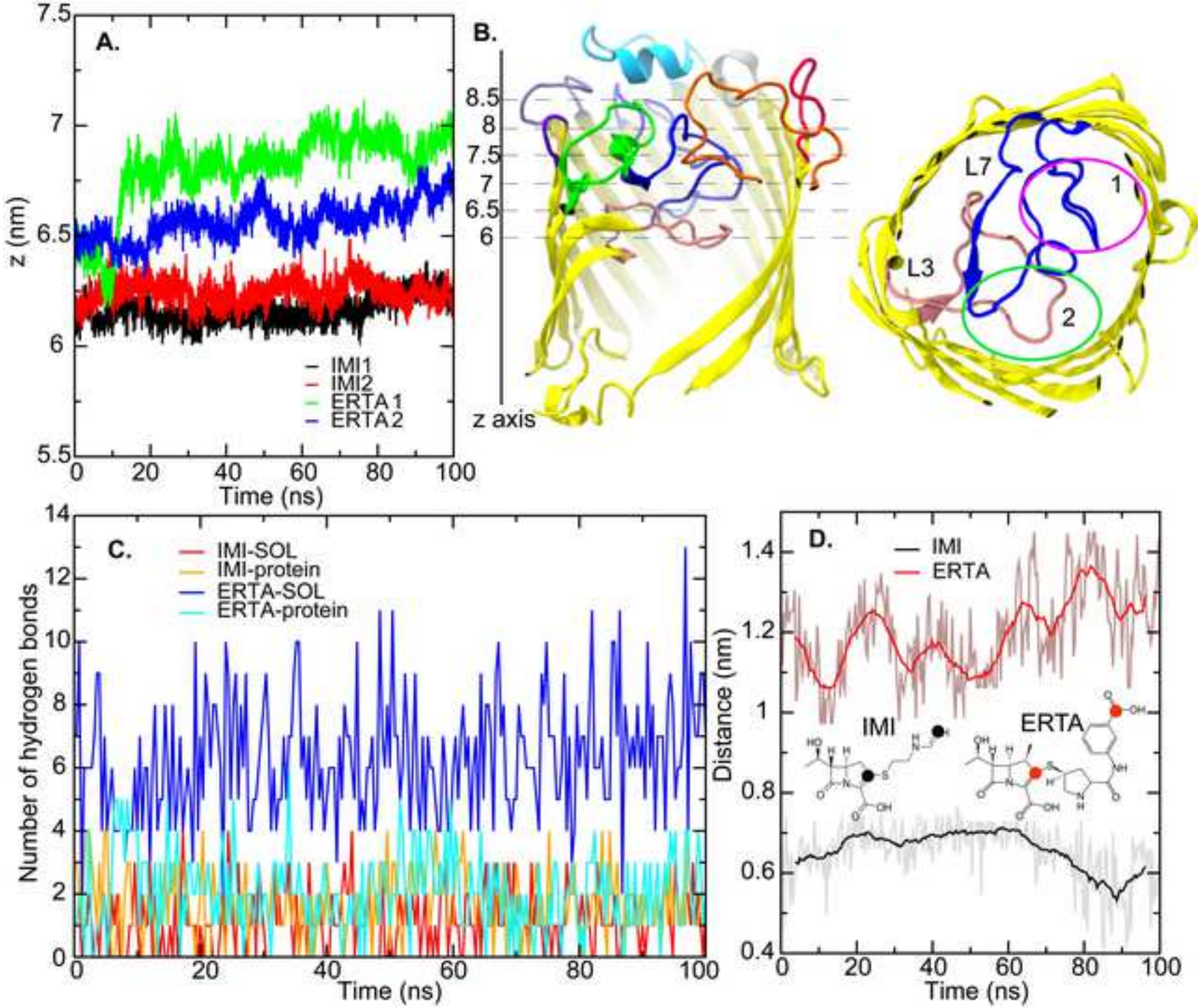
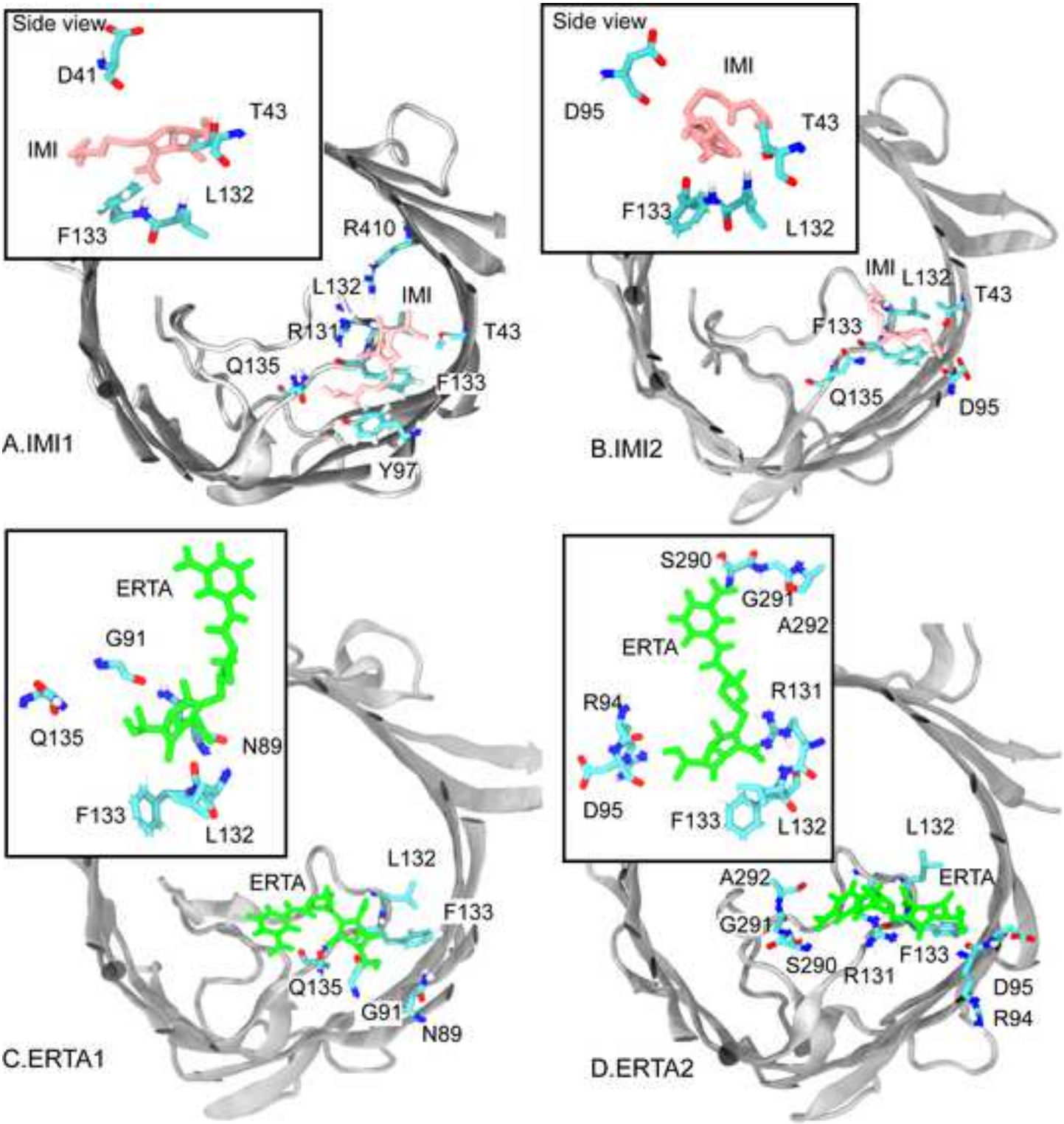


Figure5

Click here to download colour figure Fig5-Residues.tif



PROBING BINDING AFFINITIES OF IMIPENEM AND ERTAPENEM WITH OUTER MEMBRANE CARBOXYLATE CHANNEL D1 (OCCD1) FROM P.AERUGINOSA: SIMULATION STUDIES

Kamolrat Somboon[1], Jitti Niramitranon[2], and Prapasiri Pongprayoon[1,3,4]*

Figure captions

Figure 1 OccD1 embedded in a membrane is shown in (A) where the cross-sectional view of OccD1 with extracellular loops are displayed in (B). The constriction loop L3 is coloured in pink. The chemical structures of imipenem (IMI) and ertapenem (ERTA) are shown in (C).

Figure 2 A set of C-alpha RMSDs is shown in (A), (B), and (D)-(L). (C) is the C-alpha RMSFs. Ligand-free (LF), IMI-bound (IMI), and ERTA-bound (ERTA) OccD1 proteins are shown in black, red, and green, respectively. The initial structure at 0 ns was used as a reference for both RMSD and RMSF calculations.

Figure 3 (A)-(C) are the protein shapshots at times 0ns, 50ns, and 100ns. Key loops, L2, L3, L4, L7, and L8 are coloured in red, pink, green, blue, and cyan, respectively. The red arrow displays the direction of L2 movement. (D) shows average numbers of water, protein, ad membrane contacts with loop L2.

Figure 4 Positions of IMI and ERTA along z axis in all simulations (A). (B) the left is the cross-sectional protein structure while the right is the top view where 1 (pink circle) is the binding site observed by Khalid et al. [4] and 2 (green circle) is the second site that we observed in this work. The numbers of drug-protein and drug-water hydrogen bonds are shown in (C). (D) is the distance between 2 given atoms on each drug. The 2 given atoms are coloured in black for IMI and red for ERTA.

Figure 5 (A)-(D) Cartoon representations of OccD1 with bound IMI (Pink) and ERTA (green) in all simulations. The key amino acids are in licorice format.

Supplementary Information

PROBING BINDING AFFINITIES OF IMIPENEM AND ERTAPENEM WITH OUTER MEMBRANE CARBOXYLATE CHANNEL D1 (OCCD1) FROM *P.AERUGINOSA*: SIMULATION STUDIES

Kamolrat Somboon, Jitti Niramitranon, and Prapasiri Pongprayoon*

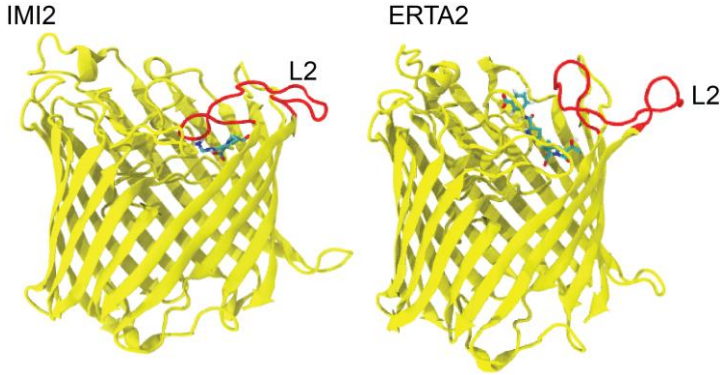


Figure S1 Cartoon views of OccD1 with bound drug at 100 ns. L2 is coloured in red.

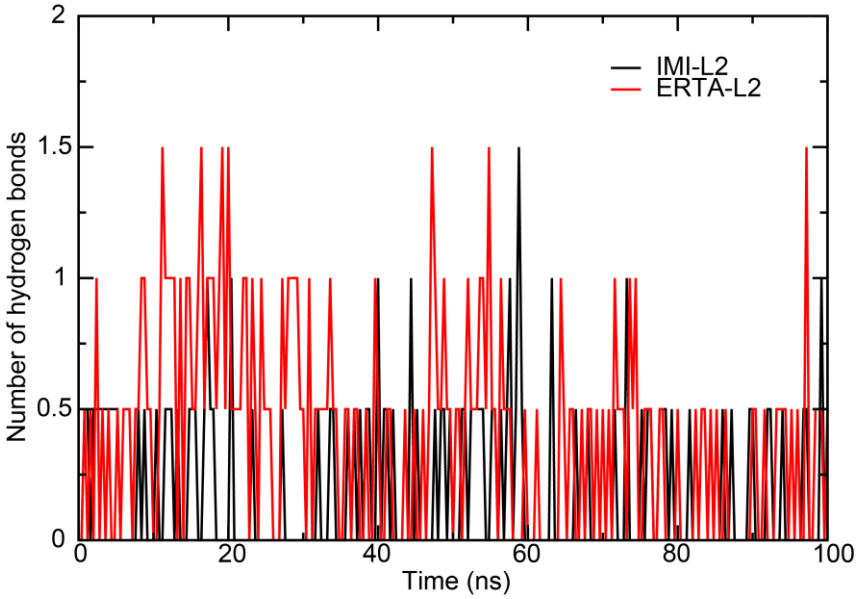
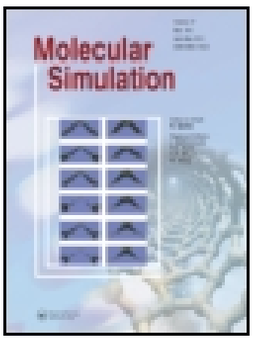


Figure S2 Average hydrogen bonds between drugs and the extracellular loop L2



Molecular Simulation

ISSN: 0892-7022 (Print) 1029-0435 (Online) Journal homepage: http://www.tandfonline.com/loi/gmos20

Multiscale simulation studies of geometrical effects on solution transport through nanopores

Attaphon Chaimanatsakun, Deanpen Japrung & Prapasiri Pongprayoon

To cite this article: Attaphon Chaimanatsakun, Deanpen Japrung & Prapasiri Pongprayoon (2017): Multiscale simulation studies of geometrical effects on solution transport through nanopores, Molecular Simulation, DOI: 10.1080/08927022.2017.1334881

To link to this article: http://dx.doi.org/10.1080/08927022.2017.1334881

MOLECULAR SIMULATION, 2017
https://doi.org/10.1080/08927022.2017.1334881

Taylor & Francis
Taylor & Francis Group

Multiscale simulation studies of geometrical effects on solution transport through nanopores

Attaphon Chaimanatsakun[a], Deanpen Japrung[b] and Prapasiri Pongprayoon[c,d,e]

[a]Faculty of Engineering at Sriracha, Department of Mechanical Engineering, Kasetsart University, Sriracha, Thailand; [b]National Nanotechnology Center, National Science and Technology Development Agency, Pathumthani, Thailand; [c]Faculty of Science, Department of Chemistry, Kasetsart University, Bangkok, Thailand; [d]Center for Advanced Studies in Nanotechnology for Chemical, Food and Agricultural Industries, KU Institute for Advanced Studies, Kasetsart University, Bangkok, Thailand; [e]Computational Biomodelling Laboratory for Agricultural Science and Technology (CBLAST), Kasetsart University, Bangkok, Thailand

ABSTRACT

Due to intrinsic properties, solid-state nanopores are widely used in nanopore technology. Different geometries (cylindrical (CY), hourglass (HG) and conical (CO)) of artificial nanopores have been fabricated and studied. Each was found to promote different transport abilities experimentally. To explore such pore effects, the combination of finite element (FE) and molecular dynamics (MD) simulations with applied electric filed (150 mV) were performed. The dimension of anion-selective protein pore was used as a nanopore template. Different pore geometries with a narrowest diameter ranging from 1.8 to 1.8 µm were studied here. Firstly, we found that the narrowest regions at a pore orifice in CO and constriction site in HG maximise water velocity and consequently control a water flow rate. Secondly, CY triggers the highest water flux, but low ion selectivity, whilst the funnel-like geometries (HG and CO) enhance the ion selectivity significantly. Both HG and CO show similar degrees of permeant flux and selectivity. The orifice and constriction site in CO and HG are the main player for selectivity and permeation control. Thirdly, the transport properties are tuneable by changing the flow direction in asymmetric CO pore. The tip-to-base flow in CO obviously promotes stronger anion selectivity than the base-to-tip one.

ARTICLE HISTORY
Received 9 January 2017
Accepted 15 May 2017

KEYWORDS
Nanopores; GROMACS; finite element; ion flux; molecular dynamics simulations

1. Introduction

Nanopores have long been a topic that attracts many research groups worldwide because of their abundant existence in nature and that they show highly interesting features. The small-sized porous structure leads to an important attribute of nanopores, that is, the very high ratio between their interior surface area and volume. Moreover, the interactions between nanopores and ions/molecules are within the range of this nanoscale confinement, which leads to a unique circumstance when the substrates are transported through the pores.

In biological systems, ions and molecules are transferred between different compartments via pore-forming membrane proteins. Examples include outer membrane protein F (OmpF), potassium channel (KcsA), *Mycobacterium Smegmatis* Porin A (MspA) and α-haemolysin (αHL). Working against certain gradients, the transport through these passive channels and active pumps can be in a controlled manner, ie for a specific ion and/or in a particular direction. The major driving force for nutrient and ion uptake is the electrical membrane potential across a membrane. This generates the electric field and electric current causing an ion gradient in a passive transport and triggering gating process in some ion channels. The ion specificity and selectivity is existed when the ion rectification takes place. Understanding these phenomena, such as ion selectivity and ion current rectification (ICR), of these porins and ion channels are thus important physiologically. To better understand such process in a microscopic view, molecular dynamics (MD) simulations with applied external electric field were successfully used to mimic a transmembrane potential [1–3]. Many previous studies reported that a constant external electric field applied in MD is a reliable tool for mimicking an electromotive force, exerted by a voltage difference [1,4].

In the past decade, many have attempted, with some degree of success, to synthesis artificial structures that mimic the unique properties of biological pores. The produced solid-state nanopores can withstand a wider range of external conditions (eg temperature, pH, ionic strength and applied voltage), and hence are suitable in many more applications. On the other hand, the reproducibility of structure precisely at the atomic level, as in the biological pores, is difficult and a major drawback for these man-made ones. However, the design and fabrication of synthetic nanopores (solid-state nanopore) to achieve the desired functions are currently of huge interest. Recent applications in the field of nanofluidic devices, semiconductors and sensitive sensors certainly drive this area to advance even more rapidly.

CONTACT Prapasiri Pongprayoon fsciprpo@ku.ac.th
Supplemental data for this article can be accessed here https://doi.org/10.1080/08927022.2017.1334881.

The different designs of artificial nanopores have been fabricated and studied for applications in nanopore technology [5–11]. Cylindrical (CY), hourglass (HG) and conical (CO) shapes are common pore geometries used currently. Each of these shapes was found to promote different transport abilities. With the same pore dimension, the hourglass shape was found to be occupied by less water molecules and trigger fast water transport, when compared with a straight pore [12,13]. Moreover, a funnel-shaped pore displays high ion rectification properties due to its asymmetric current–voltage characteristics [14,15]. Also, the degree of rectification can be tuned by modifying surface charge patterns on pore surface [9,16]. So, such rectifying pores become of interest for future semiconductor development. Attempts to relate the ion transport properties to geometrical characteristics have been investigated for cylindrical, conical and hourglass pores [13,17–22]. The mechanism of the ionic transport and rectification in asymmetric channels has become a subject of theoretical model based on experimental and numerical methods [8,13,17,19,23–25]. Recently, a protein transport through cylindrical, cuboidal and pyramidal mesopores has been studied using solvent-free coarse-grained models [26]. They highlighted the role of pore asymmetry on protein selectivity. Moreover, the different types of uniformly charged pattern on CY surface were found to exhibit different behaviours of water and ion conductance [27]. Both studies illustrate an importance of pore geometries and surface charges on transport and selective properties. However, the atomistic details on such properties remain insufficient. In this work, we thus employed multiscale numerical simulations to explore solution permeation through micro/nanopores with various geometries and sizes in microscopic level. Advantages and disadvantages of each pore shape on transport and selective properties are also highlighted here.

To obtain a molecular insight into transport behaviours, we employ MD simulations for sub-nanometre pores and finite element (FE) model for microchannels. We performed MD simulations on a sub-nanometre simplified model pore with a diameter of 1.8 nm in a NaCl solution, since previous studies found that the water flow start to exhibit a unique feature with this pore diameter [6,28,29]. Furthermore, this pore dimension is as similar as that of outer membrane protein channels (OMPs) which control transport of nutrients and ions into the cell. With their high selectivity, bio-inspired nanopores becomes of interest. Pore dimension and surcharge charges with biomimetic patterns play a role in such selectivity so the tailor-made properties are crucial for biomimetic design of nanopores. Although most synthetic nanopores (ie carbon nanotube and silicon nitride) are negatively charged due to a fabrication condition, many previous studies successfully display the feasibility of tuning their surface charges and pore cavity to meet desired transport properties [30–33]. In this work, the hourglass-shaped (HG) pore dimension and charge pattern of anion-selective OMP was initially used as a template to construct a simplified HG pore. Other pore geometries (CO and CY) were built based on a HG dimension (diameters of 1.8 nm at the pore entrance/exit and ~1–1.1 nm at a construction site for CO and HG). Moreover, both hydrophobic and hydrophilic models are investigated here by MD simulations to understand the effects of non-uniform charged on transport properties. To gain macroscopic picture, the 10-, 100- and 1000-fold larger pores are also studied here using computational fluid dynamics (CFD). Only water behaviours in hydrophobic pores were studied in CFD part. The water flow in these CFD models is generated by setting an inlet water velocity. In our MD work, we used an optimal condition of applying an external electric field of 150 mV mimicking a transmembrane potential so as to compare our results with biological pore activities.

2. Materials and methods

2.1. FE models

The 3D simulations of nanopore systems were conducted using the commercial FE software ANSYS FLUENT (Faculty of Engineering, Kasetsart University). The models are shown in Figure 1. All models (HG, CY and CO) are constructed with the dimension as seen in Table 1. Their thickness are 30, 300, 3000 nm. Each pore is connected to two reservoirs of pure water. Since previous experimental studies observed the laminar flow inside nanopores [5,17], we thus used this flow type with a Newtonian fluid here.

The fluid properties (density = 998.2 kg/m^3 and viscosity = 1.003×10^{-3} kg/m*s) were set. An entrance and exit of the analysis domain were defined by inlet and outlet boundary conditions. The low inlet velocity of 0.001 m/s was set and the outlet was defined as outflow in FLUENT. The CFD models were discretised using hexahedral mesh. A fine mesh was applied in the region of the pore because of its small dimension. The region of the domain far away from the pore was applied with coarse mesh.

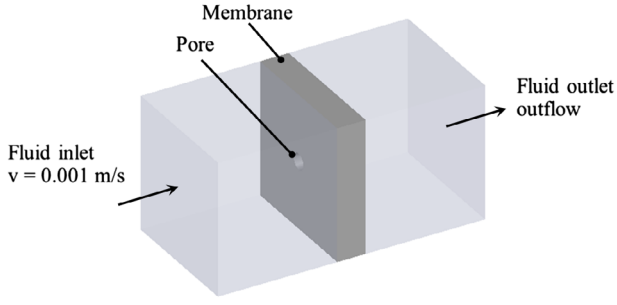


Figure 1. (Colour online) Geometry and boundary condition of the simulation models.

Table 1. Dimensions of pores. Numbers (1–3) represent the positions on each model where the diameters are measured (the location can be seen in Figure 1). The list of numbers (10, 100, 1000) in each bracket is a set of scale factors telling the enlargement of pore diameter from the smallest nanopore (1.8 nm).

	Diameter (nm)								
	Model 1 (10)			Model 2 (100)			Model 3 (1000)		
Model	1	2	3	1	2	3	1	2	3
Cylindrical pore (CY)	18			180			1800		
	Volume: 8100 nm^3			Volume: 8.1 × 10^6 nm^3			Volume: 8.1 × 10^9 nm^3		
	Height: 32 nm			Height: 320 nm			Height: 3200 nm		
Hourglass-shaped pore (HG)	18	11	18	180	110	180	1800	1100	1800
	Volume: 5500 nm^3			Volume: 5.5 × 10^6 nm^3			Volume: 5.5 × 10^9 nm^3		
	Height: 32 nm			Height: 320 nm			Height: 3200 nm		
Conical pore (CO)	18	14	10	180	140	100	1800	1400	1000
	Volume: 4600 nm^3			Volume: 4.6 × 10^6 nm^3			Volume: 4.6 × 10^9 nm^3		
	Height: 29 nm			Height: 290 nm			Height: 2900 nm		

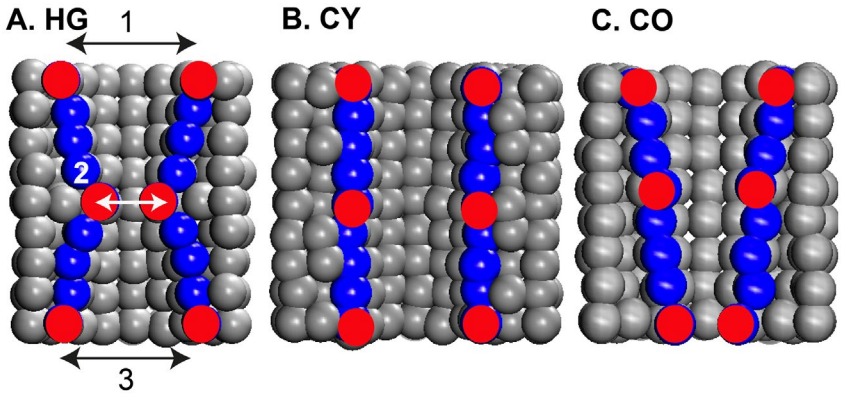


Figure 2. (Colour online) Three pore geometries used in this work. (A) is hourglass-shaped pore where (B) and (C) are cylinder and cone, respectively. The pore shapes are represented by the blue beads, while the red dots (with the numbers 1–3) indicate the locations where the pore dimension is measured in each model.

Table 2. Dimensions of nanopores. Numbers (1–3) represent the positions on each model where the diameters are measured (the location can be seen in Figure 1).

	Diameter (nm) Model 4 (1)		
Model	1	2	3
Cylindrical pore (CY)		1.8	
		Volume: 8.1 nm^3	
		Height: 3.2 nm	
Hourglass-shaped pore (HG)	1.8	1.1	1.8
		Volume: 5.5 nm^3	
		Height: 3.2 nm	
Conical pore (CO)	1.8	1.4	1.0
		Volume: 4.6 nm^3	
		Height: 2.9 nm	

2.2. MD simulations

2.2.1. Model construction

Building upon earlier studies [34–37], all CY, HG and CO nanopores embedded in a toy membrane were constructed from arrays of hydrophobic (methane-like) particles. Each model comprises concentric rings of methane particles with a van der Waals radius of 0.195 nm. All methane-like particles were harmonically restrained to their initial position (with a force constant k_0 = 10,000 kJ·mol^{-1}·nm^{-2}) in order to preserve the pore shape. The GROMOS force fields were used in this model. Using this approach, the models were designed to mimic the dimension and size of anion-selective OMPs and makeable solid-state nanopores. The pore dimensions for each system are shown in Table 2. Two sets of charged and uncharged pores were set. For hydrophilic or charged pores, total surface charges of +6e (+0.5e per particle) were assigned to given particles at the entrance, exit and constriction site of each pore (all positions can be seen as red dots in Figure 2. Each system contains a membrane with embedded pore solvated by electrolyte solution of 1 M NaCl. Counter-ions were added to neutralise the system. Each system contains SPC water molecules of 16,456 in CY, 16,118 in CO and 16,048 in HG.

2.2.2. Simulations protocols

For each system, a 5-ns equilibration of methane-like particles restrained dynamics simulation was conducted with a force constant of 1000 kJmol^{-1} nm^{-2}. To mimic a transmembrane potential, a constant external electric field of 0.05 V/nm (equivalent to a potential of 150 mV across a membrane) perpendicular to the membrane plane (z axis) was applied in all salt solution systems. The 30-ns unrestrained MD production run were conducted with a repeat (each system has a different randomised velocity from a Maxwell distribution at the beginning of a simulation). After a production run, the total energy of each simulation was investigated to ensure a convergence of all simulations. All results are an average from two simulations.

The simulations were performed using the GROMACS 4.5 simulation package (www.gromacs.org) [38] with an extended united atom version of the GROMOS96 force field [39]. To relax steric conflicts generated during set-up, all energy minimisations used up to 1000 steps of steepest descent. Long-range electrostatic interactions were treated using the particle mesh Ewald method [40] with a short range cut-off of 1 nm, a Fourier spacing of 0.12 nm and fourth-order spline interpolation. All simulations were performed in the constant number of particles, pressure and temperature ensemble. The temperature of membrane, solvent and ions were each coupled separately using the Berendsen thermostat [41] at 300 K with a coupling constant τ_t = 0.1 ps. The pressure was coupled using the Berendsen algorithm at 1 bar with a coupling constant τ_p = 1 ps. The time step for integration was 2 fs. Coordinates were saved every 2 ps for subsequent analysis.

The results were analysed by GROMACS and locally written code. Molecular graphic images were prepared using VMD [42].

3. Results and discussions

In biological systems, the protein channels (eg (OMPs)) are highly selective. With an hourglass pore dimension and charged pore-lining amino acids, they show high substrate-selective properties. Since we are here interested in the effects of pore geometry on transport properties, we initially constructed a HG nanopore with the same dimension as substrate-selective OMP's (a diameter of ~1.1 nm at a constriction). Comparing to solid-state experiments, conical and cylindrical pores were able to be fabricated. Then, CO and CY pores were constructed in this work based on the HG dimension. Not only nanopores, but microscale pores are of our interest. The same set of pore geometries and dimensions (with applied scaling factors of 10,

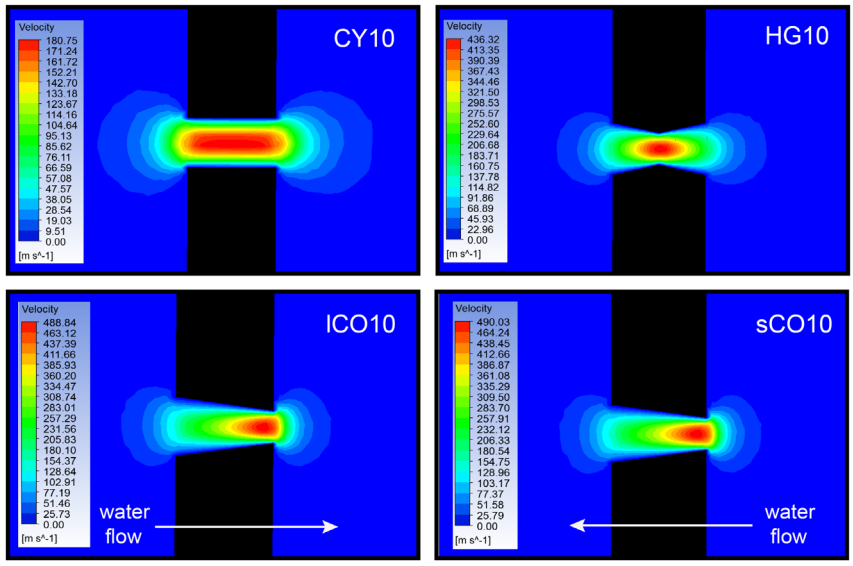


Figure 3. (Colour online) Simulated velocity profiles for water during translocation through 18-nm-radius nanopores with different pore geometries (the whole membrane is shown as a black band).

Table 3. Mean velocity and maximum velocities (max velocity) computed from FE method for each pore shape. The mean velocity was computed along the pore axis with standard deviations. Both sCO and lCO are conical pores, but they provide different directions of permeant flow. sCO represents the tip-to-base flow, whereas it is opposite for lCO. The list of numbers (10, 100, 1000) at the end of each abbreviation is a set of scale factors telling the enlargement of pore size from the smallest nanopore (1.8 nm).

Model	Mean velocity (m/s)	Max velocity (m/s)	Model	Mean velocity (m/s)	Max velocity (m/s)
CY10	100.90 ± 52.90	180.75	lCO10	220.99 ± 119.57	488.9
CY100	1.02 ± 0.51	1.90	lCO100	1.74 ± 1.16	4.45
CY1000	0.013 ± 0.004	0.019	lCO1000	0.021 ± 0.02	0.051
HG10	218.30 ± 103.4	436.33	sCO10	221.49 ± 119.57	490.06
HG100	1.79 ± 1.06	4.34	sCO100	1.74 ± 1.17	4.42
HG1000	0.022 ± 0.01	0.044	sCO1000	0.021 ± 0.012	0.051

100 and 1000) were then continued using. The FE method was selected to explore dynamics of water in microchannels.

At the beginning, we employed FE calculation. In order to reflect flow behaviours mainly affected by pore geometry, all FE models are set uncharged. Only water flow in micropores (18–1800 nm in diameter) is studied here. It was reported in previous studies that the continuum model is appropriate to investigate transport properties through a pore with diameter >1.8 nm [43,44] due to the fluctuation in measured flow velocity. With the same pore geometries, sub-nanometre behaviours of water and electrolyte ions inside each pore are instead investigated in parallel by MD simulations.

To observe water behaviour inside various pores, the contour maps showing velocity distribution are plotted. With different pore shapes, the water velocity is position-dependent. The maximum water velocity is reached at the narrowest region inside each tube. The fastest water flow is observed along the cross-sectional centre of the pore in CY10 (180.75 m/s), the orifices in both sCO10 (490 m/s) and lCO10 (488.9 m/s) and the mid of HG10 (436.3 m/s) (Figure 3 and Table 3). The cone (CO) appears to promote the fastest water flow followed by an hour glass-shaped (HG) pore, while a cylinder (CY) provides the slowest (Table 3). When expanding pore size, the larger pores (CY100, CY1000, lCO100, lCO1000, sCO100, sCO1000, HG100 and HG1000) still display the same behaviours. The cylindrical shape still generates the slowest water flow, while the cone causes the fastest water speed due to the smallest constriction (Tables 2 and 3). The order of maximum water speed is CY < HG < CO. Nonetheless, with constriction-containing geometries, both HG and CO give similar mean velocities along a pore (~220 m/s for CO10 and HG10, ~1.7 m/s for CO100 and HG100 and ~0.02 m/s for CO1000 and HG1000 in Table 3). This finding highlights the role of a constriction on controlling flow pattern inside a pore. Apparently, HG and CO shapes enhance water velocities leading to high water transport rate, while CY gives the slowest velocity. Not only do our FE results agree well with previous studies that HG and CO generate the faster water flow than CY [12,18,19,45–47], but also suggest that CO maximises the water flow velocity (Table 3). As a rule, the water velocity is proportional to pore size. The smaller the pore is, the greater the water velocity is obtained. When scaling up pore sizes, the magnitude of flow is consequently increased by scale factors of 10, 100 and 1000, respectively (Table 3). Considering the velocity distribution inside a pore, the water velocity gradually decreases and vanishes near the pore surface in all cases due to the laminar flow set-up (Figure 3). For symmetric HG and CY pores, the velocity distributions of water at their ends are similar (Figure 3). In case of CO, it is interesting that the flow direction does not affect water velocities. Either base-to-tip (lCO) or tip-to-base (sCO) flow provides similar maximum water velocity of ~490 m/s at the tip (orifice) (Table 3). This finding is also observed in larger lCO and sCO pores (Figure S1 in supplemental information). So, the orifice in CO appears to be important for flow rate control.

In case of sub-nanometre pores, we employed MD simulations to explore their transport behaviours in atomic detail. Both hydrophobic and anion-selective model pores mimicking a dimension of anion-selective outer membrane protein are investigated under the condition of 1 M NaCl solution and an external electric field of 150 mV. Constant total energies of all

systems were observed to confirm the convergence of all simulations (Figure S2 in supplementary information). In Figure 4, the density profiles of all species (Na^+, Cl^- and water particles) across the simulation box are computed. For hydrophobic pores, even though none of ion and water density is detected in Figure 4(A), some water fluxes are present (Figure S3 in Supplementary information). Too low water occupancy can be the bottleneck to calculate correct density. Apparently, all hydrophobic pores can conduct water with different degrees of permeability. The hydrophobic CY shows highest water flux, while the rest show similar degrees of water fluxes (Figure S3 in Supplementary information). These results agree well with previous well-known work by Hummer et al. [48]. Unlike hydrophilic nanopores, none of ion can translocate through all hydrophobic nanopores. This reflects the crucial role of pore-lining charges on ion transport in nanopores. Nonetheless, the water and ion flow in small and hydrophobic pores can be observed when high-pressure condition is applied [12,13,18]. Compared to our work, the common cellular potential (150 mV) used here is not sufficient to trigger ion transport through a tiny hydrophobic pore with radius of 0.9 nm. Recently, the ion and water flow through a sub-10-nm nanopore has been generated with no external electric field applied by a presence of charged protein channels [49]. This experiment displays an importance of charged surface on solution permeation. Furthermore, the charge density has also been reported to be crucial for stable ion conductance [50]. In this study, the effects of the charge surface on transport properties are also considered where the non-uniform charge pattern of anion-selective OMP

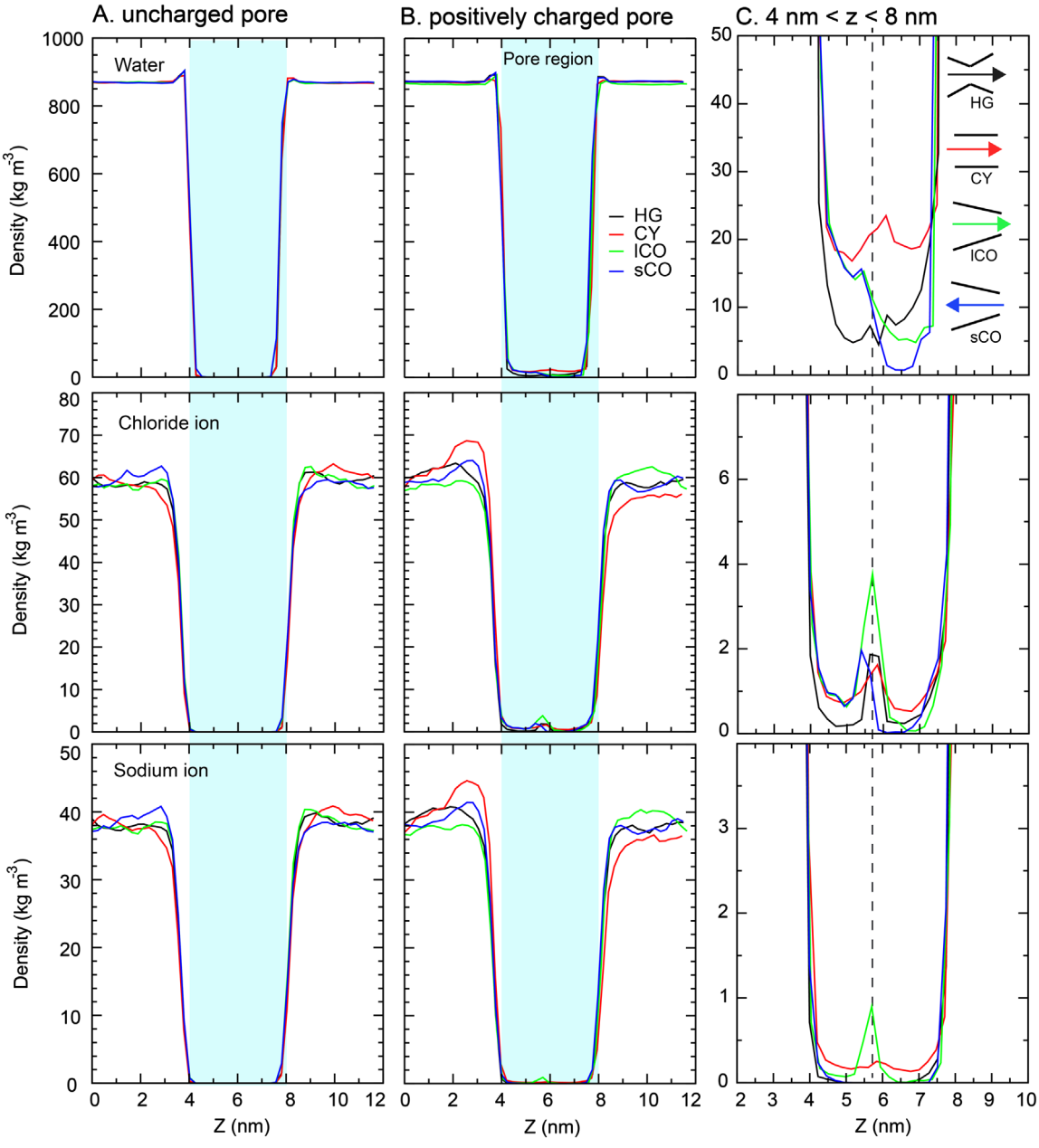


Figure 4. (Colour online) Water and ion density profiles across the simulation box are shown in (A) for uncharged and (B) for positively charged nanopores. The density profiles of each species inside each charged pore are magnified in (C) (4 nm < *z* < 8 nm). The pore region is labelled as a blue band. The pore region is labelled as a blue band. The dashed line in (C) represents the centre of each pore.

was used as a template. When non-uniform charges are set, all species appear to accumulate near the inlet mouth due to the drag from applied electric field across the simulation box (Figure 4(A)). The pattern of Na^+ density at the entrance and exit of a channel is similar to that of anions (Cl^-) due to electroneutrality requirements. This pattern is also observed in the density of whole ions (Figure S5 in Supplementary Information). Owning to the anion-selective pore chloride is the major ions travelling through the pore, so the pattern of whole ion density is dominated by that of Cl^-. All ion appears to accumulate mostly at the mid of all pores (Figure S5 in Supplementary Information). In case of a water passage in charged pores, different pore geometries conduct different degrees of water occupancy (Figure 4(C)). The larger the pore, the more water inside. CY shows the highest water density inside its pore due to the largest pore cavity. This finding can be confirmed by the highest water flux in Figure 6. The lowest water density is observed at the orifices of CO pores (Figure 4(C)). Interestingly, even though the geometries of lCO and sCO are identical, the water profiles inside them are different. At the orifice, lCO displays higher water accumulation, unlike sCO (Figure 4(C)). This can primarily imply the important role of the flow direction on transport properties. Considering ion density, due to anion-selective pore, the accumulation of Cl^- ions is observed in all models (Figure 4(C)). Mostly, chloride ions are found at the mid of all pores. Especially, the highest chloride density observed in lCO (Figure 4(C)) indicates the best environment for chloride accumulation. Nonetheless, CY and lCO still allow the Na^+ uptake. Na^+ ions seem to equally disperse in CY, but instead they get packed at the mid of lCO pore (Figure 4(C)). Such Na^+ accumulation is present despite the fact that both CY and lCO are positively charged. This demonstrates that surface charges inside both pores are insufficient to reject Na^+. For sCO and HG, although the Na^+ density cannot be seen inside a pore (Figure 4(A) and (C)), a permeation of small Na^+ portion is observed. This finding will be discussed in the next section. Interestingly, both lCO and sCO share the same geometry, but their transport behaviours are different. The tip-to-base flow (sCO) makes a conical pore more chloride-selective. This finding is roughly supported by the presence and absence of Na^+ density in lCO and sCO (Figure 4(A) and (C)) and smaller number of Na^+ in sCO when compared to those of base-to-tip flow (lCO) (Figure 5(C) and (D)). Even though a shortage of Na^+ ions is observed after 25 ns in lCO (Figure 5(C)), its number is retrieved after 30 ns (see Figure S6 in Supplementary Information). Our results highlight and capture the importance of flow direction on transport properties in a conical pore. Considering ion selectivity, slightly larger number of Cl^- than Na^+ residing in CY indicates weak anion-selective property (Figure 5(B)). Unlike CY, very small number of interior Na^+ ions observed demonstrates strong anion selectivity in hourglass and conical pores (Figure 5). Approximately, ~3 Cl^- ions for HG and CO and ~4 Cl^- ions for a straight CY pore are found inside, whilst less than 2 Na^+ are detected in each pore geometry except CY (Figure 5). The ratio of Na^+/Cl^- inside can approximately reflect the degree of Cl^- selectivity. The highest number of Na^+ found in Figure 5 illustrates the poorest anion selectivity of CY. Comparing sCO and lCO, it is interesting that lCO shows more Cl^- accumulation, but lower Cl^- selectivity. Entering the pore at the orifice side seems to clearly enhance chloride selectivity. This finding implies a key role of orifice on selectivity monitoring. Combining with FE results, we emphasise that the orifice are crucial for controlling and monitoring a flow rate and importantly pore selectivity.

To better understand transport properties, the ion cumulative flux is measured in Figure 6. With a positively charged pore, Cl^- ions are a major charge carrier. When they move, they also drag water molecules towards positive potential. Therefore, the water flows to the same direction as Cl^- (from −E to +E in Figure 6). Consequently, the direction of Na^+ flux is opposite.

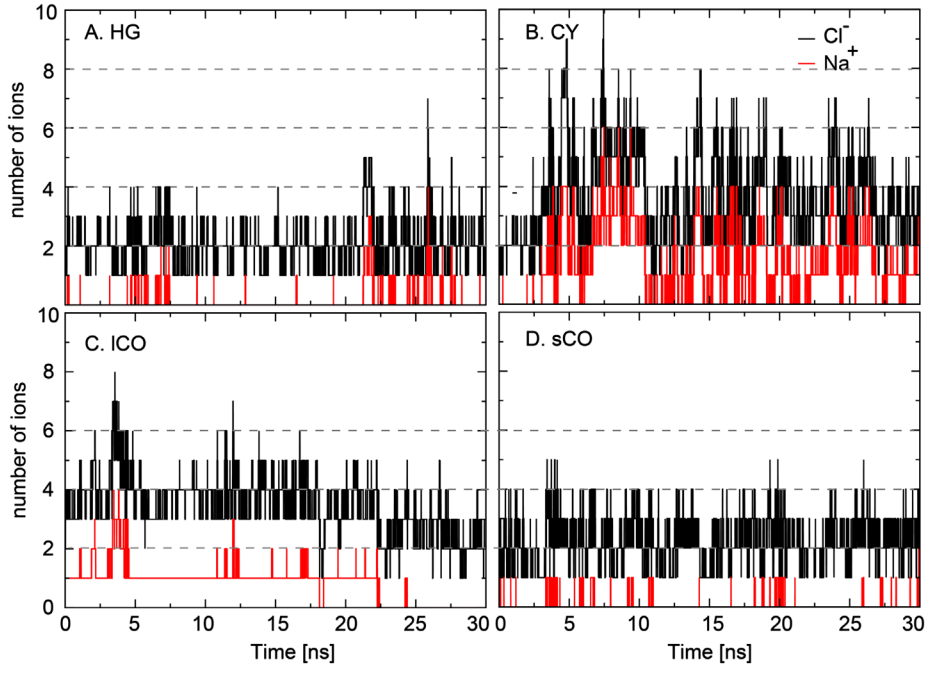


Figure 5. (Colour online) Number of sodium and chloride ions residing in each pore.

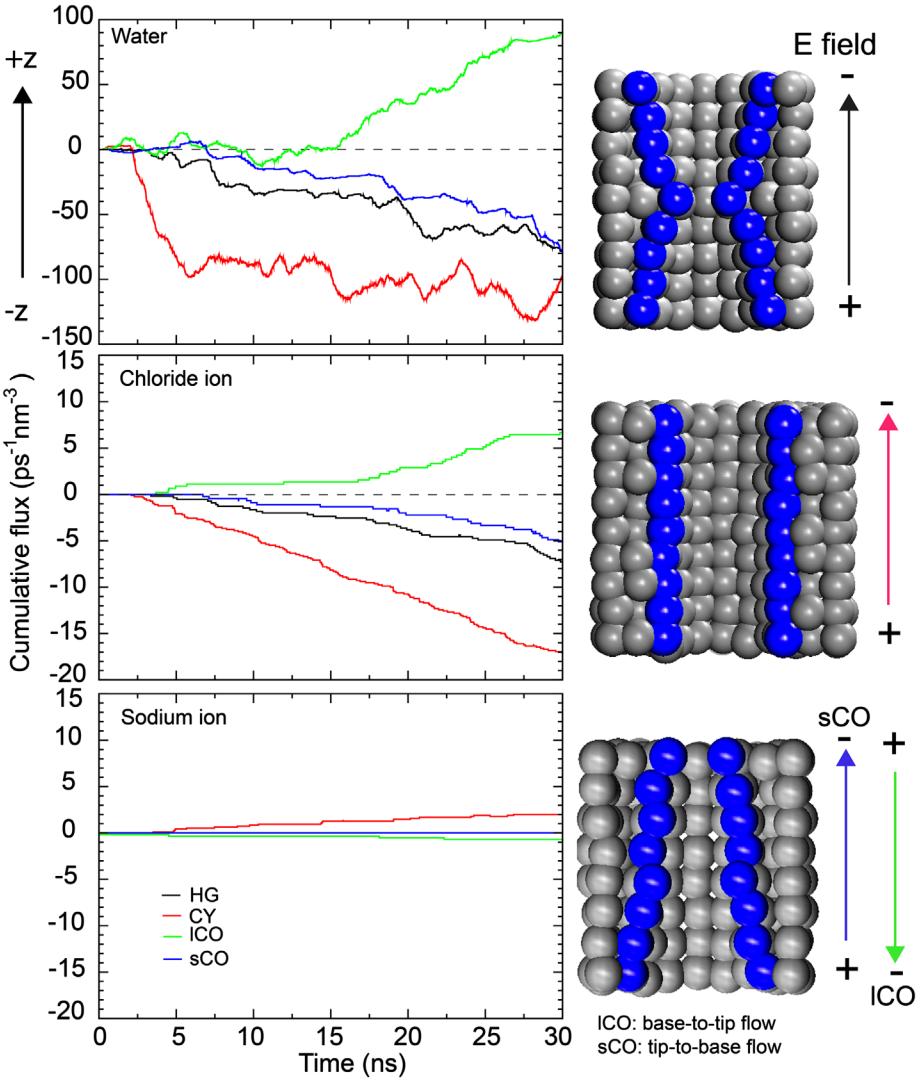


Figure 6. (Colour online) Cumulative fluxes of water and ions through all types of nanopores are shown on the left and the right.

The negative and positive signs of cumulative fluxes in Figure 6 have no physical meaning. They only represent a direction of permeant flow along *z*-axis. Clearly, all nanopore geometries activate the asymmetric ion flux owning to a non-uniform positively charged surface. Based on the cumulative flux in Figure 6, the magnitude of fluxes for each species can be ranked here: (i) CY > lCO > HG~sCO for water flux; (ii) CY > HG~lCO > sCO for Cl^- ions; (iii) CY > lCO > HG~sCO for sodium. Figure 6 shows that the permeant translocation requires at least 3 ns to occur in CY, whereas hourglass and conical pores take longer time (~5 ns). It appears that the solution flux can be quickly triggered in a straight pore. In general, the water flux through nanopores is increased in most of pore types except CY. CY promotes high and fast water flux at the beginning (~100 ps^{-1} nm^{-3} in Figure 6) before a near plateau in water flow develops. Large number of water molecules gradually rush into CY pore and then become nearly constant, while others are quickly filled by water (Figure S4 in supplementary information). The flow behaviour observed in CY is in a good agreement with a previous study where a constant water flow is observed [11]. The flat water flux in CY demonstrates similar rate of forward and backward flows of water. lCO can also promote high water flux at the end of simulation although a minute flow of water is observed at the beginning. Unlike CY, both water fluxes in HG and sCO are linearly increased (Figure 6).

For ions, all pores allow the accumulation of cations with various portions (Figure 5). The radial distribution functions of ion–water contacts are in good agreement with previous nanopore studies (Figure S7 in supplementary information) [51]. The level of cation flow implies how anion-selective the pore is. To better understand the flow behaviours, the cumulative fluxes as a function of time were computed in Figure 6. Each flux was measured per cubic nanometres. In Figure 6, the highest water and ion conductance is observed in CY. CY allows similar ratio of Cl^- and Na^+ translocations resulting in the least anion-selective among all (Figure 6). Despite the fact that the accumulation of Na^+ is observed in all pores, only HG and sCO can show the rejection of cation transport (Figure 5(B) and (C), 6). Seemingly, HG and sCO pores are the most chloride-selective. Comparing HG and sCO behaviours, more water flux is observed in HG due to higher Cl^- flux (7 ps^{-1} nm^{-3} for HG and 5 ps^{-1} nm^{-3} for sCO) before getting the same level as sCO at the end of simulations

(~80 ps^{-1} nm^{-3} in Figure 6). Higher Cl^- flux in HG implies the rapid Cl^- flow. These high selectivity and fast transport may be one of the reasons why most of outer membrane proteins have hourglass-like pore structure. Apart from pore shapes, the detailed geometrical properties like pore radius, length and angle are also found to be important for transport rate [12,13,20,25,43]. Especially, some previous studies show that different pore opening angles activate different transport properties [12,18]. So, this emphasises the necessity for finely tuning the geometry of nanopores for future nanopore technology.

Considering conical pores, it is interesting that entering the pore from different ends can significantly alter transport properties of a pore. Figure 6 clearly shows that the base-to-tip flow (lCO) yields higher ion flow (eight ions in total) than the tip-to-base one (sCO) (totally, five ions). This entrance effect is also observed in a recent study [17]. The higher ion flow in lCO includes both cation and anion, but the larger portion of anion permeant makes lCO anion-selective. On the contrary, sCO clearly inhibits the Na^+ permeation. Apparently, sCO is more chloride-selective than lCO (Figure 6). This finding clearly shows the entrance effect on transport properties and ion selectivity is not only a property of the system itself but is also connected to the idea of directionality.

4. Conclusions

This work demonstrates the geometrical effects of pores on transport properties in both micro- and nanoscale. In microchannels, even though CY provides a slow water flow, it can conduct the largest portion of water due to its largest and constant cross-sectional area. This common property is also inherited in CY nanopore. Furthermore, FE models show that a presence of narrow region in HG and CO can significantly increase the water velocities. This finding highlights a key role of the orifice on controlling water flow rate. Both hydrophobic micro- and nanopores show similar effects of pore geometries on a water passage. The fast water transport can be reached using pores with a constriction (HG and CO), while CY is suitable for mass transport. In case of selectivity, it can be explained by our atomistic MD work. The MD results reveal that HG and CO nanopores promote more ion selectivity than CY. We also highlight a role of non-uniform surface charges on enhanced transport properties. The presence of non-uniform surface charges also helps to improve ion selectivity. Furthermore, the water flux is reduced in the presence of conical entrance/exit. HG and sCO successfully promote the rejection of cation translocation. HG and sCO show similar degrees of water and ion flux as well as ion selectivity. The hourglass geometry is commonly found in pore-forming membrane proteins. The high selectivity of HG observed here help to explain why most of membrane proteins have the hourglass pore geometry. CY appears to be improper for cellular transport activity due to massive fluxes of water and ions and importantly less selectivity. Besides, our results reveal a key role of the entrance on transport properties and ion selectivity. Interestingly, sCO and lCO give us clear evidence that the flow direction through an asymmetric pore determines the selectivity of nanopores. The tip-to-base flow (sCO) does not only show very strong ion selectivity, but also promote asymmetric ion flow which can lead to an enhancement of ICR. In general, both HG and sCO serve as an excellent molecular sieve, but the fabrication of HG solid-state nanopore with desired size appears to be difficult and complex. A high-resolution drilling technique is required. Generally, focused ion beam and focused electron beam drilling methods are used to fabricate nanopores, but the limited beam diameter is a major bottleneck to make a sub-nanometre pore [52]. Due to a reduction of beam strength when passing through a media, the drilling process commonly generates a conical pore with diameter >10 nm [53]. To obtain sub-5-nm pore, various assistant approaches (ie the adoption of a dedicated focused ion beam nanowriter [54]) are required. Currently, making a complex geometry like HG in atomic scale appears to be too difficult. Therefore, many studies focus on utilising conical nanopore. Our results obviously show that HG and CO with the same diameter at mouth and end provide similar ion selectivity and solution fluxes, but the degree of solute selectivity of CO is flow direction-dependent. However, we only investigated pore geometries with the same pore diameter here. None of other dimensional factors is involved. Many previous studies are devoted to explore effects of cone angle, pore length and diameter size on transport properties, but no work can define the best pore dimension on solution transport. Our study explains effects of pore shapes on transport properties at a fundamental level. More advanced insight into pore properties on the transport ability is still needed.

Acknowledgement

We would like to thank for financial supports from Kasetsart University Research and Development Institute (KURDI) (code: S-K1.56 and S-K161.59), Thailand Research Fund (code: TRG5880230), National Research Council of Thailand (NRCT), Health Systems Research Institute (HSRI) and Royal Society of Chemistry (RSC) for financial supports. We also thank the reviewers for useful suggestions and comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Kasetsart University Research and Development Institute (KURDI) [code: S-K1.56 and S-K161.59], Thailand Research Fund [code: TRG5880230], National Research Council of Thailand (NRCT), Health Systems Research Institute (HSRI) and Royal Society of Chemistry (RSC) for financial supports.

References

[1] Roux B. The membrane potential and its representation by a constant electric field in computer simulations. Biophys J. 2008;95:4205–4216. Epub 2008/07/22.
[2] Jensen MO, Borhani DW, Lindorff-Larsen K, et al. Principles of conduction and hydrophobic gating in K+ channels. Proc Nat Acad Sci. 2010;107:5833–5838. Epub 2010/03/17.
[3] Tieleman DP, Leontiadou H, Mark AE, et al. Simulation of pore formation in lipid bilayers by mechanical stress and electric fields. J Am Chem Soc. 2003;125:6382–6383. Epub 2003/06/06.
[4] Singh A, Orsat V, Raghavan V. Soybean hydrophobic protein response to external electric field: a molecular modeling approach. Biomolecules. 2013;3:168–179.
[5] Sokhan VP, Nicholson D, Quirke N. Fluid flow in nanopores: accurate boundary conditions for carbon nanotubes. J Chem Phys. 2002;117:8531–8539.

[6] Kalra A, Garde S, Hummer G. Osmotic water transport through carbon nanotube membranes. Proc Nat Acad Sci. 2003;100:10175–10180.
[7] Siwy Z, Heins E, Harrell CC, et al. Conical-nanotube ion-current rectifiers: the role of surface charge. J Am Chem Soc. 2004;126:10850–10851.
[8] Siwy Z, Kosinska ID, Fulinski A, et al. Asymmetric diffusion through synthetic nanopores. Phys Rev Lett. 2005;94:048102.
[9] Vlassiouk I, Smirnov S, Siwy Z. Nanofluidic ionic diodes. comparison of analytical and numerical solutions. ACS Nano. 2008;2:1589–1602.
[10] Portella G, de Groot BL. Determinants of water permeability through nanoscopic hydrophilic channels. Biophys J. 2009;96:925–938.
[11] Zhu CQ, Li H, Meng S. Transport behavior of water molecules through two-dimensional nanopores. J Chem Phys. 2014;141:18C528.
[12] Gravelle S, Joly L, Ybert C. Large permeabilities of hourglass nanopores: from hydrodynamics to single file transport. J Chem Phys. 2014;141:18C526.
[13] Tang D, Yoo YE, Kim D. Molecular dynamics simulations on water permeation through hourglass-shaped nanopores with varying pore geometry. Chem Phys. 2015;453:13–19.
[14] Daiguji H, Yang PD, Majumdar A. Ion transport in nanofluidic channels. Nano Lett. 2004;4:137–142.
[15] Siwy ZS. Ion-current rectification in nanopores and nanotubes with broken symmetry. Adv Func Mater. 2006;16:735–746.
[16] Qian SZ, Joo SW, Ai Y, et al. Effect of linear surface-charge non-uniformities on the electrokinetic ionic-current rectification in conical nanopores. J Colloid Interface Sci. 2009;329:376–383.
[17] Lan WJ, Edwards MA, Luo L, et al. Voltage-rectified current and fluid flow in conical nanopores. Acc Chem Res. 2016;49:2605–2613.
[18] Tang D, Li L, Shahbabaei M, et al. Molecular dynamics simulation of the effect of angle variation on water permeability through hourglass-shaped nanopores. Materials. 2015;8:7257–7268.
[19] Tang D, Kim D. Study on the transport of water molecules under the geometry confinement of aquaporin-like nanopores. Appl Therm Eng. 2014;72:120–125.
[20] Laohakunakorn N, Keyser UF. Electroosmotic flow rectification in conical nanopores. Nanotechnology. 2015;26:275202.
[21] Li H, Francisco JS, Zeng XC. Unraveling the mechanism of selective ion transport in hydrophobic subnanometer channels. Proc Nat Acad Sci. 2015;112:10851–10856.
[22] Mochizuki K, Koga K. Solid–liquid critical behavior of a cylindrically confined Lennard–Jones fluid. Phys Chem Chem Phys. 2015;28:18437–18442.
[23] Cervera J, Schiedt B, Neumann R, et al. Ionic conduction, rectification, and selectivity in single conical nanopores. J Chem Phys. 2006;124:104706.
[24] Ramirez P, Gomez V, Cervera J, et al. Ion transport and selectivity in nanopores with spatially inhomogeneous fixed charge distributions. J Chem Phys. 2007;126:194703.
[25] Jiang XK, Liu Y, Qiao R. Current rectification for transport of room-temperature ionic liquids through conical nanopores. J Phys Chem C. 2016;120:4629–4637.
[26] Ileri N, Létant SE, Palazoglu A, et al. Mesoscale simulations of biomolecular transport through nanofilters with tapered and cylindrical geometries. Phys Chem Chem Phys. 2012;14:15066–15077.
[27] Hao L, Su JY, Guo HX. Water permeation through a charged channel. J Phys Chem B. 2013;117:7685–7694.
[28] Suk ME, Aluru NR. Molecular and continuum hydrodynamics in graphene nanopores. RSC Adv. 2013;3:9365–9372.
[29] Berezhkovskii A, Hummer G. Single-file transport of water molecules through a carbon nanotube. Phys Rev Lett. 2002;89:4629–4637.
[30] Tan SW, Wang L, Liu H, et al. Single nanoparticle translocation through chemically modified solid nanopore. Nanoscale Res Lett. 2016;11:50.
[31] Li PY, Cheng KY, Zheng XC, et al. Facile synthesis of water-soluble graphene-based composite: non-covalently functionalized with chitosan-ionic liquid conjugation. Funct Mater Lett. 2016;9:1650045.
[32] Hwang JY, Eltohamy M, Kim HW, et al. Self assembly of positively charged carbon nanotubes with oppositely charged metallic surface. Appl Surf Sci. 2012;258:6455–6459.
[33] Shan CS, Yang HF, Han DX, et al. Water-soluble graphene covalently functionalized by biocompatible poly- l-lysine. Langmuir. 2009;25:12030–12033.
[34] Beckstein O, Sansom MS. Liquid-vapor oscillations of water in hydrophobic nanopores. Proc Nat Acad Sci. 2003;100:7063–7068. Epub 2003/05/13.
[35] Beckstein O, Sansom MS. The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. Phys Biol. 2004;1:42–52. Epub 2005/10/06.
[36] Beckstein O, Tai K, Sansom MS. Not ions alone: barriers to ion permeation in nanopores and channels. J Am Chem Soc. 2004;126:14694–14695. Epub 2004/11/13.
[37] Beckstein O, Sansom MS. A hydrophobic gate in an ion channel: the closed state of the nicotinic acetylcholine receptor. Phys Biol. 2006;3:147–159. Epub 2006/07/11.
[38] Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model. 2001;7:306–317.
[39] Bonvin AMJJ, Mark AE, van Gunsteren WF. The GROMOS96 benchmarks for molecular simulation. Comput Phys Commun. 2000;128:550–557.
[40] Darden T, York D, Pedersen L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys. 1993;98:10089–10092.
[41] Berendsen HJC, Postma JPM, van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. J Chem Phys. 1984;81:3684–3690.
[42] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996;14:33–38.
[43] Goldsmith J, Martens CC. Pressure-induced water flow through model nanopores. Phys Chem Chem Phys. 2009;11:528–533.
[44] Suk ME, Aluru NR. Ion transport in sub-5-nm graphene nanopores. J Chem Phys. 2014;140:084707.
[45] Farimani AB, Aluru NR, Tajkhorshid E. Thermodynamic insight into spontaneous hydration and rapid water permeation in aquaporins. Appl Phys Lett. 2014;105:083702.
[46] Hou X, Guo W, Jiang L. Biomimetic smart nanopores and nanochannels. Chem Soc Rev. 2011;40:2385–2401.
[47] Bocquet L, Tabeling P. Physics and technological aspects of nanofluidics. Lab Chip. 2014;14:3143–3158.
[48] Hummer G, Rasaiah JC, Noworyta JP. Water conduction through the hydrophobic channel of a carbon nanotube. Nature. 2001;414:188–190.
[49] Picaud F, Paris G, Gharbi T, et al. Biomimetic solution against dewetting in a highly hydrophobic nanopore. Soft Matter. 2016;12:4903–4911.
[50] Balme S, Picaud F, Manghi M, et al. Ionic transport through sub-10 nm diameter hydrophobic high-aspect ratio nanopores: experiment, theory and simulation. Sci Rep. 2015;5.
[51] Samoylova ON, Calixte EI, Shuford KL. Molecular dynamics simulations of ion transport in carbon nanotube channels. J Phys Chem C. 2015;119:1659–1666.
[52] Patterson N, Adams DP, Hodges VC, et al. Controlled fabrication of nanopores using a direct focused ion beam approach with back face particle detection. Nanotechnology. 2008;19:235304.
[53] Liu ZW, Wang YF, Deng T, et al. Solid-state nanopore-based DNA sequencing technology. J Nanomater. 2016.
[54] Gierak J, Madouri A, Biance AL, et al. Sub-5 nm FIB direct patterning of nanodevices. Microelectron Eng. 2007;84:779–783.

Journal of Molecular Graphics and Modelling 65 (2016) 1–7

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Why do the outer membrane proteins OmpF from *E. coli* and OprP from *P. aeruginosa* prefer trimers? Simulation studies

Jitti Niramitranon[a], Mark SP Sansom[b], Prapasiri Pongprayoon[c,*]

[a] Department of Computer engineering, Faculty of Engineering, Kasetsart Univeristy, Chatuchak, Bangkok 10900, Thailand
[b] Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
[c] Department of Chemistry, Faculty of Science, Kasetsart University, Chatuchak, Bangkok 10900, Thailand

ARTICLE INFO

Article history:
Received 15 October 2015
Received in revised form
18 December 2015
Accepted 6 February 2016
Available online 10 February 2016

Keywords:
OprP
OmpF
Outer membrane protein
Porin
Trimer
Molecular dynamics simulations

ABSTRACT

Porins are water-filled protein channels across the outer membrane of gram-negative bacteria. They facilitate the uptake of nutrients and essential ions. Solutes are filtered by a constriction loop L3 at the mid of a pore. Porins are heat-stable and resistant to toxic agents and detergents. Most porins are trimer, but no clear explanation why trimeric form is preferable. In this work, we thus studied effects of oligomerization on porin structure and function in microscopic detail. A well-studied OmpF (general porin from *Escherichia coli*) and well-characterised OprP (phosphate-specific pore from *Pseudomonas aeruginosa*) are used as samples from 2 types of porins found in gram-negative bacteria. MD simulations of trimeric and monomeric pores in pure water and 1 M NaCl solution were performed. With a salt solution, the external electric field was applied to mimic a transmembrane potential. Expectedly, OprP is more stable than OmpF. Interestingly, being a monomer turns OmpF into an anion-selective pore. The dislocation of D113's side chain on L3 in OmpF causes the disruption of cation pathway resulting in the reduction of cation influx. In contrast, OprP's structure and function are less dependent on oligomeric states. Both monomeric and trimeric OprP can maintain their anion selectivity. Our findings suggest that trimerization is crucial for both structure and function of general porin OmpF, whereas being trimer in substrate-specific channel OprP supports a pore function.

1. Introduction

Porins are water-filled, pore-forming protein channels across the outer membrane of gram-negative bacteria. They allow diffusion of nutrients and metabolites across outer membrane with various grades of selectivity [1–3]. Especially, porins also serve as an entryway for many antibiotics [4–7]. Porins can roughly be classified as general and substrate-specific pores. General porins (e.g. OmpC and OmpF) filter solutes based on their molecular size, while substrate-specific porins, such as OprP, have specific binding site for certain molecules [8]. Generally, porins have a β-barrel structure connected with extracellular loops and intracellular turns. Most of them have the extracellular loop L3 (constriction loop) folds back into the β-barrel lumen creating a constriction region.

Most porins are homotrimers (e.g. OmpF [9], OmpC [10], OprP [11], and PhoE [12]) and some exist as a monomer such as OccD1 [13], OmpG [14], and NanC [15]. The previous studies found that the multimerization was important for function and structure in soluble proteins. The oligomerization gives shape to active sites, increases affinity of protein complexes for ligand binding, and promotes protein stability [16,17]. However, present understanding of protein oligomerization is mostly originated from studies of soluble proteins. It is unclear whether the same principles can be applied to membrane proteins, especially porins. In a single porin, the oligomerization was found to be not a requirement for stability and membrane insertion [18], whereas the role of multimerization in a triplet-pore porin remains unclear. To better understand the significance of being trimer, in this study, an example of general porins, OmpF, and substrate-specific pores, OprP, were studied in comparison. Both are from different groups of porins and adequate details on structure and function are available. The study of both porins will help us better understand the nature of most trimeric porins.

OmpF is a well-studied porin prevalently found in *Escherichia coli* (*E. coli*). It consists of 16 antiparallel β-strands connected by

Abbreviations: OprP, outer membrane protein P; OmpF, outer membrane protein F; MD, molecular dynamics.

* Corresponding author. Fax: +66 2579 3955.
E-mail addresses: fsciprpo@ku.ac.th, ppongprayoon@yahoo.com (P. Pongprayoon).

http://dx.doi.org/10.1016/j.jmgm.2016.02.002
1093-3263/

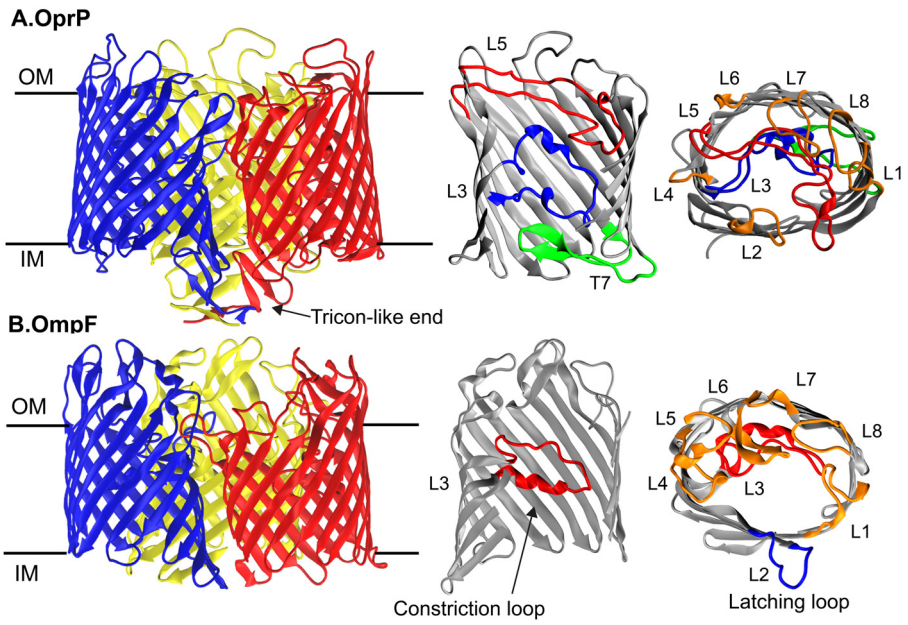


Fig. 1. Trimeric porins (A) OprP (B) OmpF. The locations of key loops are shown in cross-section and top views.

turns at the periplasmic side and loops at the extracellular side [9,19–23]. Like other OMPs, it has an extracellular loop (L3) folded into the lumen in each monomer acting as a molecular filter (Fig. 1). Besides, there is the so-called latching loop (L2) reaching over to the neighbouring monomers to stabilize a trimeric form (Fig. 1B). OmpF is defined as a general or non-specific porin even though it is slightly cation-selective. Despite the observed trimeric form of OmpF, dimeric and monomeric states have also been observed experimentally [8,24–26]. Dimer was found to act as an intermediate in trimerization [24]. Moreover, key residues that are crucial for protein-protein interaction interface were also identified [8].

OprP from *Pseudomonas aeruginosa* (*P. aeruginosa*) is a well-characterised phosphate-selective homotrimeric pore. OprP plays a key role in high-affinity phosphate uptake under the condition of phosphate starvation. In an absence of phosphate, OprP can also conduct common anions. Each pore adopts 16 antiparallel β-strands lined by three positively charged loops (L3, L5, and T7) folded into its lumen (Fig. 1A). These folded loops create an arginine ladder and lysine cluster which are the key features for phosphate selectivity. Unlike other porins, OprP has an extended periplasmic 'tricon-like' end that is involved in stabilizing trimer [11]. Due to its high selectivity for phosphate, almost all OprP studies are devoted to phosphate selectivity and translocation mechanism [27–31]. No previous work emphasizes the role of oligomerization on structural and functional properties.

In this study, Molecular dynamics (MD) simulations were performed to explore the effect of oligomerization state on structure and function in microscopic level. MD simulations have been widely used in many earlier studies to investigate the dynamic properties of OmpF. Studies of trimeric OmpF revealed deviations of dynamical structure relative to the crystal structure. They also showed that L3 flexibility affected a change in pore cavity [32,33]. MD simulations were also successfully used to observe behaviour and solute (such as antibiotics) passage through OmpF in comparison with experiments [2,33–37]. MD simulations were also conducted to reveal a mechanism of phosphate transport by OprP [29–31]. In this study, we then used MD simulations to reveal the importance of being trimer in OmpF and OprP. To generate an ion flow, we apply the external electric field across a membrane. Recently, MD simulations with applied external electric field become a popular tool for studying ion channels [38–40]. Despite concerns about a degree of artificiality, recent studies have been shown that a constant external electric field is a valid representation of the influence of an electromotive force, exerted by a voltage difference [38,41]. Despite the fact that OprP is phosphate-specific, only a common salt (NaCl) was used in this study so as to compare results with general OmpF pores. Understanding structural and functional properties affected by different oligomeric states here can facilitate further studies on the structural biology of outer membrane proteins (OMPs) and the development of nanopore technology.

Table 1
8 systems set up in this study.

Name	Condition		Time (ns)
	Pure water (MD)	1 M NaCl (1 M)	
TF	Trimeric OmpF	Trimeric OmpF	30
MF	Monomeric OmpF	Monomeric OmpF	30
TP	Trimeric OprP	Trimeric OprP	30
MP	Monomeric OprP	Monomeric OprP	30

2. Method

2.1. Molecular dynamics simulations

The trimeric OprP (PDB ID: 2O4V) and OmpF (PDB ID: 2OMF) crystal structures consisting of 411 and 340 amino acids in each monomer respectively were downloaded from the Protein Data Bank (www.rcsb.org). The protonation states of all charged amino acids were set at physiological pH. To study the effects of oligomeric states on structure and function, trimeric and monomeric systems of both OmpF and OprP in electrolyte solution (1 M NaCl) and pure water were set. There were 8 simulations performed as seen in Table 1 TF stands for a trimeric OmpF and TP is a trimeric OprP. TF1–TF3 represent OmpF monomer 1–3 and TP1–TP3 are OprP monomer 1–3. MP and MF are for stand-alone OprP and OmpF, respectively.

Each system was embedded in a pre-equilibrated dimyristoylphosphatidylcholine (DMPC) bilayer (pre-equilibrated by running a 2 ns simulation). The solvent-accessible molecular surface of both

OmpC and OmpF were used as templates to remove lipids and perform short steered MD simulations of the solvated proteins as described in detail by Faraldo-Gomez et al.[42]. This generated a cavity into which proteins were inserted. Water and counter ions were then added into each system. Ions were randomly added by replacing water molecules. For each system, a 3 ns equilibration of protein restrained dynamics simulation was conducted with a force constant of 1000 kJ mol^{-1} nm^{-2}. To mimic a transmembrane potential, a constant external electric field of 0.05 V/nm (equivalent to a potential of 150 mV across a membrane) perpendicular to the membrane plane (z axis) was applied in all salt solution systems. The 30 ns unrestrained MD production run were conducted. All monomeric simulations were run with a repeat (each system has a different randomized velocity from a Maxwell distribution at the beginning of a simulation). After a production run, the total energy of each simulation was investigated to ensure a convergence of all simulations. All results from trimeric systems are an average among three monomers, while those of monomeric system are an average between 2 replicas.

All simulations were performed using the GROMACS 4.5 simulation package (www.gromacs.org) [43] with an extended united atom version of the GROMOS96 forcefield [44]. To relax steric conflicts generated during setup, all energy minimizations used up to 1000 steps of steepest descent. Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method [45] with a short range cutoff of 1 nm, a Fourier spacing of 0.12 nm, and fourth-order spline interpolation. All simulations were performed in the constant number of particles, pressure, and temperature (NPT) ensemble. The temperature of the protein, DMPC, solvent, and ions were each coupled separately using the Berendsen thermostat [46] at 300 K with a coupling constant $\tau_t = 0.1$ ps. The pressure was coupled using the Berendsen algorithm at 1 bar with a coupling constant $\tau_p = 1$ ps. The time step for integration was 2 fs. Coordinates were saved every 2 ps for subsequent analysis.

The results were analysed by GROMACS and locally written code. Molecular graphic images were prepared using VMD [47].

3. Results and discussions

Overall, the RMSDs, calculated on alpha carbon, demonstrate that trimeric OMPs are apparently more stable than stand-alone pores (Fig. 2). Considering monomeric OmpF (MF), a protein shows high structural flexibility in both pure water and salt solution (all-atom RMSDs ~ 0.4 nm). This implies the instability of MF framework. For loop regions, most loops behave similarly to those of trimer. L1 appears to be the most fluctuated loop, while the motion of L8 is quite random. However, the constriction loop L3 of MF becomes more mobile in salt solution. Like MF, the monomeric OprP (MP) are less stable in pure water especially barrel structure (Fig. 2C). Most loops are quiet inside a pore, however the high flexible L5 and L8 are observed. On the other hand, the flexibility of MP is dramatically reduced in the presence of salt solution and applied potential (Fig. 2D). L5, L7, and L8 are quite flexible. Under the nature-like environment, OprP appears to be tough enough to survive solely in membrane. Furthermore, Principle Component Analysis (PCA) is also applied to probe a major motion of all proteins. The results from the first eigenvector confirm that stand-alone pores are more mobile than triplet pores (Fig. S1 in Supplementary information), but a single pore becomes less fluctuated in salt water.

Considering interactions inside proteins, a number of hydrogen bonds were calculated in Table 2. Overall, within a monomer, each TF pore can roughly form ~280 hydrogen bonds and TP has ~345 hydrogen bonds. OprP in both forms (TP and MP) are stabilized by similar number of hydrogen bonds, while a salt solution causes more hydrogen bonds in OmpF. Overall, the trimerization creates extra ~40 hydrogen bonds in TF and ~60 hydrogen bonds in TP at protein interfaces (Table 2). Interestingly, one third of interface hydrogen bonds (~22 hydrogen bonds) in TP are from a tricon-like region. The low RMSDs of this region in TP also suggest that being trimer can stabilise and make a rigid tricon-like region (low RMSDs of ~0.1 nm in Fig. 5A). On the contrary, a tricon-like motif is very fluctuated in MP (Fig. 5A). The instability of a tricon-like end observed in monomeric OprP may serve as one of the factors that drive an oligomerization in OprP. The results here reflect not only the tighter packing and more stable structure of OprP, but also an important role of tricon-like end in trimeric assembly. Considering a pore interior, L3 appears to stay clam inside a pore (~20 L3-protein hydrogen bonds in TF and ~25 hydrogen bonds in TP) although a number of hydrogen bonds found in a salt solution are slightly deviated. With 28–29 hydrogen bonds, L3 seems to bind to a protein wall more tightly in MP, while L3 in MF seems to be varied (16 hydrogen bonds in MF (MD) and 20 bonds in MF (1 M)). This confirms the more stable structure of OprP than OmpF. Apparently, the interaction network in OmpF seems to be sensitive to a salt solution, whereas OprP is more resistant.

Table 2
Number of hydrogen bonds occurring between L3 and protein, within each monomer, and tricon-like region. * Hydrogen bonds are computed from a triplet system. The data in a bracket are differences in a number of hydrogen bonds between a system of whole triplet protein (TF and TP) and a sum of 3 single monomers in a trimeric system (TF1 + TF2 + TF3 and TP1 + TP2 + TP3).

No. of hydrogen bonds	OmpF		OprP	
	MD	1M	MD	1M
L3-protein (MF/MP)	16	20	28	29
L3-protein (TF1/TP1)	20	18	25	24
L3-protein (TF2/TP2)	20	23	26	25
L3-protein (TF3/TP3)	20	19	26	26
Monomer(MF/MP)	277	281	346	347
Monomer1(TF1/TP1)	277	281	339	340
Monomer2(TF2/TP2)	284	283	344	348
Monomer3(TF3/TP3)	275	282	348	346
3 Monomers*	879(+43)	884(+38)	1093(+62)	1091(+58)
Tricon-like end			21	24

Furthermore, the average pore cavities of both porins are measured in Fig. 3. Both OmpF and OprP have small pore sizes ranging from ~0.2–0.3 nm (Fig. 3). For OmpF, TF in both pure water and salt solution show similar pattern where one of them gives the smallest pore and the rest have comparable pore sizes. TF3 in pure water and TF1 in salt solution provide the smallest pore size (~0.2 nm), while the rest have similar pore cavities of ~0.25 nm. This can confirm that each monomer behaves independently as found in previous study [33]. For all OprP, MP promotes the smallest pore in both salt and pure water system (~0.2 nm). A pore cavity appears to be varied at the periplasmic end, especially in pure water. Besides, both proteins have the hourglass pore shapes, but OprP provides much narrower pore cavity along a protein axis. OmpF gives quite large chambers at both extracellular and periplasmic mouths. Too large room in OmpF can blind solutes. Moving diffusively inside those areas results in low pore selectivity. In contrast, a narrow region in OprP permits the direct contacts with pore-lining residues so the efficient mechanism of solute selectivity can take place. The pore shape and size can primarily tell different degrees of solute selectivity between both porins. Due to large entrance and exit in OmpF, the constriction site appears to be the only barrier that screens incoming guests. Although the size of an extracellular mouth in OprP is similar to that of OmpF (Fig. 3), a pore-lining arginine ladder allows an accumulation of anionic solutes at an entrance. The narrow tunnel from a constriction site to a periplasmic end permit OprP to efficiently monitor and screen solute permeation.

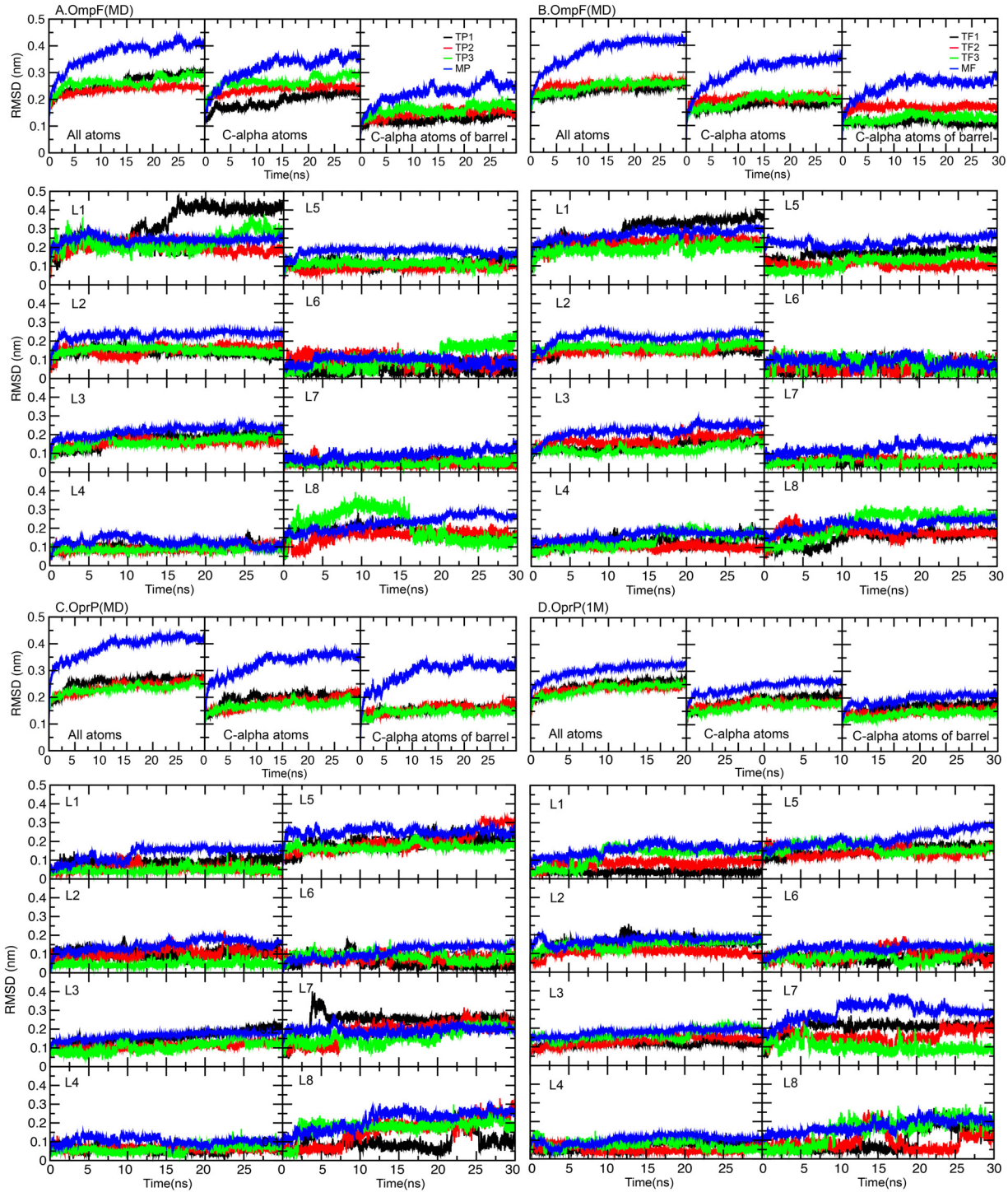


Fig. 2. RMSDs of OmpF and OprP.

To make a clearer picture, a set of cumulative ion fluxes is computed (Fig. 4). Fig. 4 clearly displays the highly anion-selective properties of both TP and MP. The sodium conductance is completely blocked (Fig. 4B). MP appears to accept more anion influx. Comparing with TP, MP has narrower constriction site, but the wider entrance permits an accumulation of anions (Fig. 3). Consequently, this anion accumulation drives more chlorides to dive down a concentration gradient. Our finding demonstrates that the anion selectivity in OprP is oligomeric-state independent. Both TP and MP can maintain their strong electropositive properties inside a pore (see Fig. S2 in Supplementary information for contour maps of electrostatic potentials). Unlike OprP, OmpF allows both cation and anion flows, but a little larger cation influx causes TF become slightly selective for cations. This finding also agrees well with previous studies [21,49–51]. Surprisingly, Na^+ flux is dramatically reduced in monomeric OmpF (Fig. 4A). Being a OmpF monomer significantly alters the electrostatic potential of a pore (Fig. S2 in Supplementary information). MF unexpectedly loses its innate selectivity by turning into an anion-selective pore. The trimeric assembly appears to play a significant role in structure and function of OmpF. To observe a change in ion selectivity, the locations of key residues in a constriction area of OmpF are investigated.

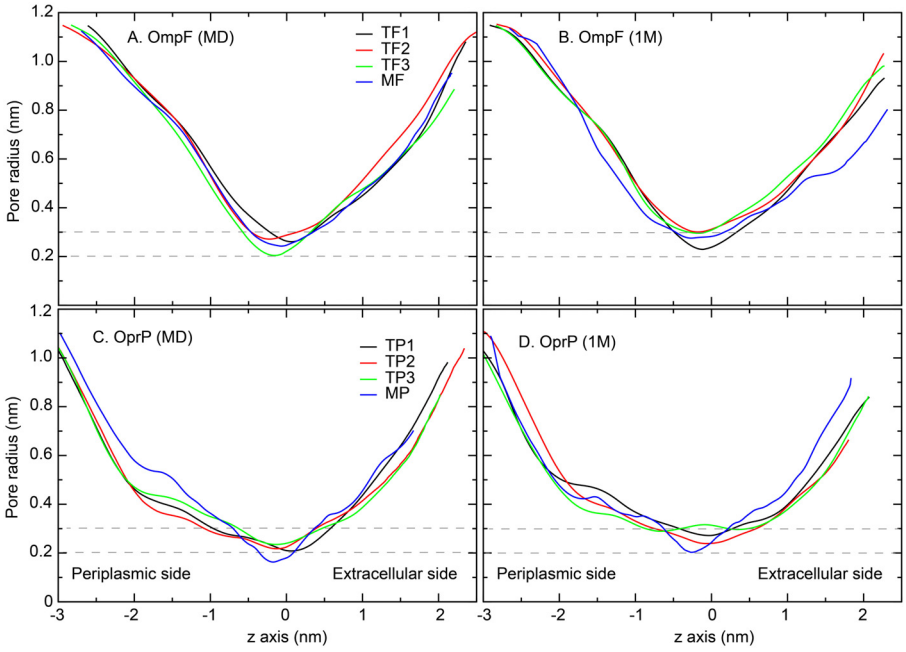


Fig. 3. Average pore radii calculated by HOLE program [48].

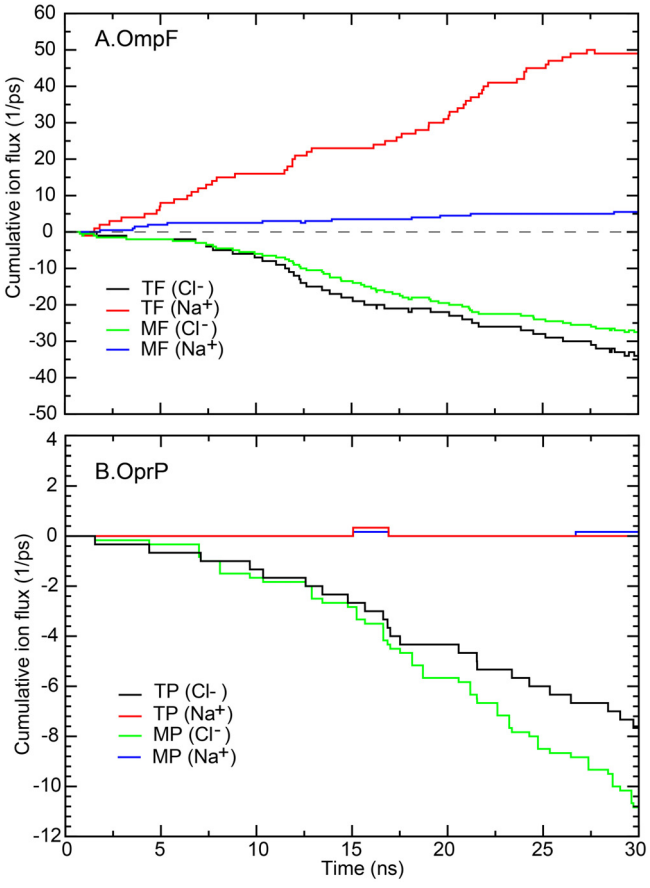


Fig. 4. Average cumulative fluxes recording the indices of translocating sodium and chloride ions inside a pore.

In a constriction site, a set of arginine patch (R42, R82, R132) and 2 negatively charged residues (D113 and E117) are found to generate an electrostatic field that control a solute passage (Fig. 5G). Two well-separated pathways for cations and anions were reported in previous studies [21,52]. An arginine patch was found to facilitate an anion passage, while an opposite site is for cations. For both trimeric and monomeric OmpF, the *z* positions of all arginine's side chains are preserved. The quanidinium groups of all arginines point upward to an extracellular side (Fig. 5B–E). With these preserved positions, both MF and TF can maintain similar degree of anion uptake (Fig. 4A). E117 at the tip of L3 also behaves in the same manner as arginines. In contrast, D113 of MF instead points down to a periplasmic chamber. This severely interrupts an electrostatic field at a constriction site resulting in the serious reduction in cation influx found in Fig. 4A. Being a monomer causes OmpF more structural fluctuation and importantly the malfunction. This also confirms a vital role of trimeric state on structure and function of OmpF.

4. Conclusions

Overall, our findings clearly indicate that a general porin, OmpF, is much less stable than substrate-specific OprP. Even though being a monomer results in more overall structural fluctuation in both porins, OprP can efficiently preserve its anion selectivity. Especially, MP's structure becomes as stable as TF under a salt condition. This implies the possibility of MP existence in reality. Oppositely, OmpF appears to be dependent on oligomeric states. With the increased structural flexibility and malfunction, the stand-alone OmpF does not seem to benefit a bacterial cell. Having such pores can harm a cellular function. Clearly, the multimerization promotes more structural stability in both porins. Being a triplet pore definitely helps OmpF enhance a structural stability and importantly support a pore function. In case of OprP, since OprP seems to be strong enough to survive as a single pore, improving structural stability may not be a major goal for trimerization. Based on our findings, substrate-specific porin like OprP definitely more stable and tougher than general porins (OmpF) and importantly a substrate-specific pore can function efficiently even in a stand-alone form. Therefore, containing a lot of substrate-specific porins allow bacteria to thrive in nutrient-poor and extreme environment. Nonetheless, this study only focuses on a transport of common ions. The effects of oligomerization on phosphate specificity and

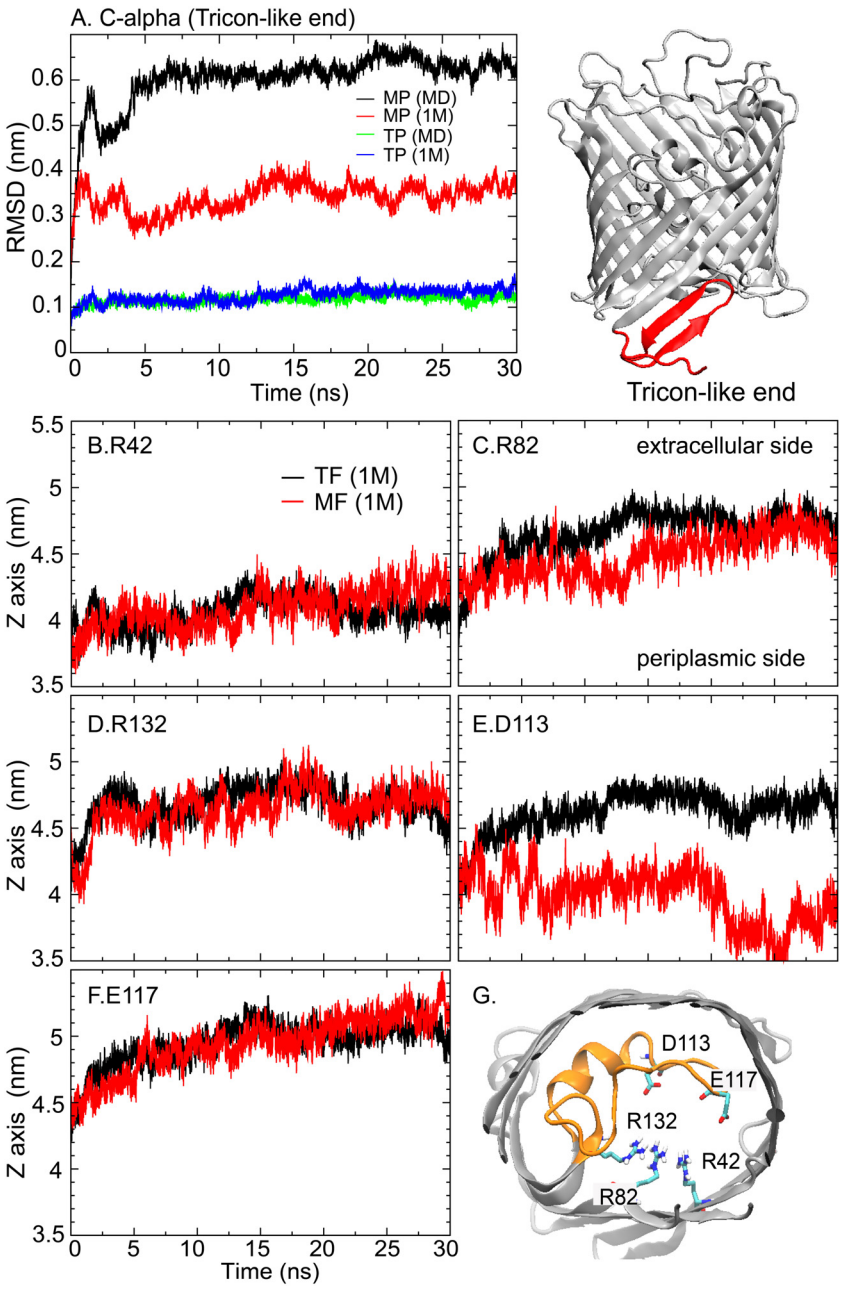


Fig. 5. (A) C-alpha RMSDs of tricon-like region (left) and a cartoon view of OprP showing a tricon-like region (right). (B)–(F) are z positions of functional groups of key residues (guandininium carbons of R42, R82, R132 and carboxylic carbons of D113 and E117) lining a constriction site. (G) is a cartoon view showing locations of key residues in a constriction area.

selectivity will be further work. This study is just a primary step to better understand the nature of porins and gram-negative bacteria. Other major topics such as how the regulatory network allows these porins to be expressed when needed or how they adapt themselves under the condition of drug resistance remain widely open.

Acknowledgements

We would like to thank the Kasetsart University Research and Development Institute (KURDI) (code: VT86.56 and MV3.58) and the Thailand Research Fund (TRF) (code: TRG5880230) for financial supports. Also, we thank Syma Khalid for good discussion.

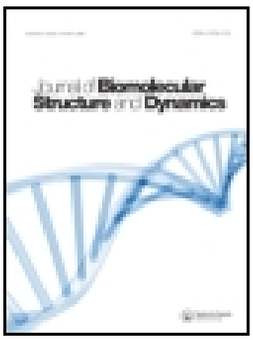
Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2016.02.002.

References

[1] H. Nikaido, Molecular basis of bacterial outer membrane permeability revisited, Microbiol. Mol. Biol. Rev. 67 (2003) 593–656.
[2] K. Malek, A. Maghari, Translocation and interactions of L-arabinose in OmpF porin: a molecular dynamics study, Biochem. Biophys. Res. Commun. 352 (2007) 104–110.
[3] S. Galdiero, A. Falanga, M. Cantisani, R. Tarallo, M.E. Della Pepa, V. D'Oriano, et al., Microbe-host interactions: structure and role of Gram-negative bacterial porins, Curr. Protein Pept. Sci. 13 (2012) 843–854.

[4] E.M. Nestorovich, T.K. Rostovtseva, S.M. Bezrukov, Residue ionization and ion transport through OmpF channels, Biophys. J. 85 (2003) 3718–3729.
[5] A. Basle, R. Qutub, M. Mehrazin, J. Wibbenmeyer, A.H. Delcour, Deletions of single extracellular loops affect pH sensitivity, but not voltage dependence, of the *Escherichia coli* porin OmpF, Protein Eng. Des. Sel. 17 (2004) 665–672.
[6] J.C. Todt, W.J. Rocque, E.J. Mcgroarty, Effects of pH on bacterial porin function, Biochemistry 31 (1992) 10471–10478.
[7] S. Kojima, H. Nikaido, Permeation rates of penicillins indicate that *Escherichia coli* porins function principally as nonspecific channels, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) E2629–34.
[8] H. Naveed, D. Jimenez-Morales, J. Tian, V. Pasupuleti, L.J. Kenney, J. Liang, Engineered oligomerization state of OmpF protein through computational design decouples oligomer dissociation from unfolding, J. Mol. Biol. 419 (2012) 89–101.
[9] S.W. Cowan, T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R.A. Pauptit, et al., Crystal structures explain functional properties of two *E. coli* porins, Nature 358 (1992) 727–733.
[10] A. Basle, G. Rummel, P. Storici, J.P. Rosenbusch, T. Schirmer, Crystal structure of osmoporin OmpC from *E. coli* at 2.0 A, J. Mol. Biol. 362 (2006) 933–942.
[11] T.F. Moraes, M. Bains, R.E. Hancock, N.C. Strynadka, An arginine ladder in OprP mediates phosphate-specific transfer across the outer membrane, Nat. Struct. Mol. Biol. 14 (2007) 85–87.
[12] B.K. Jap, K.H. Downing, P.J. Walian, Structure of PhoE porin in projection at 3.5 A resolution, J. Struct. Biol. 103 (1990) 57–63.
[13] S. Biswas, M.M. Mohammad, D.R. Patel, L. Movileanu, B. van den Berg, Structural insight into OprD substrate specificity, Nat. Struct. Mol. Biol. 14 (2007) 1108–1109.
[14] G.V. Subbarao, B. van den Berg, Crystal structure of the monomeric porin OmpG, J. Mol. Biol. 360 (2006) 750–759.
[15] C. Wirth, G. Condemine, C. Boiteux, S. Berneche, T. Schirmer, C.M. Peneff, NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family, J. Mol. Biol. 394 (2009) 718–731.
[16] M.H. Ali, B. Imperiali, Protein oligomerization: how and why, Bioorg. Med. Chem. 13 (2005) 5013–5020.
[17] J.B. Pereira-Leal, E.D. Levy, C. Kamp, S.A. Teichmann, Evolution of protein complexes by duplication of homomeric interactions, Genome Biol. 8 (2007) R51.
[18] P. Van Gelder, J. Tommassen, Demonstration of a folded monomeric form of porin PhoE of Escherichia coli in vivo, J. Bacteriol. 178 (1996) 5320–5322.
[19] A. Hoenger, J.M. Pages, D. Fourel, A. Engel, The orientation of porin OmpF in the outer membrane of *Escherichia coli*, J. Mol. Biol. 233 (1993) 400–413.
[20] S.W. Cowan, R.M. Garavito, J.N. Jansonius, J.A. Jenkins, R. Karlsson, N. Konig, et al., The structure of OmpF porin in a tetragonal crystal form, Structure 3 (1995) 1041–1050.
[21] W. Im, B. Roux, Ion permeation and selectivity of OmpF porin: a theoretical study based on molecular dynamics, Brownian dynamics, and continuum electrodiffusion theory, J. Mol. Biol. 322 (2002) 851–869.
[22] H. Miedema, M. Vrouenraets, J. Wierenga, B. Eisenberg, T. Schirmer, A. Basle, et al., Conductance and selectivity fluctuations in D127 mutants of the bacterial porin OmpF, Eur. Biophys. J. 36 (2006) 13–22.
[23] S. Varma, S.W. Chiu, E. Jakobsson, The influence of amino acid protonation states on molecular dynamics simulations of the bacterial porin OmpF, Biophys. J. 90 (2006) 112–123.
[24] V. Visudtiphole, M.B. Thomas, D.A. Chalton, J.H. Lakey, Refolding of *Escherichia coli* outer membrane protein F in detergent creates LPS-free trimers and asymmetric dimers, Biochem. J. 392 (2005) 375–381.
[25] Y. Watanabe, Effect of various mild surfactants on the reassembly of an oligomeric integral membrane protein OmpF porin, J. Protein Chem. 21 (2002) 169–175.
[26] T. Surrey, A. Schmid, F. Jahnig, Folding and membrane insertion of the trimeric beta-barrel protein OmpF, Biochemistry 35 (1996) 2283–2288.
[27] R. Benz, C. Egli, R.E.W. Hancock, Anion transport through the phosphate-specific oprp-channel of the pseudomonas-aeruginosa outer-membrane—effects of phosphate, di-basic and tribasic anions and of negatively-charged lipids, Biochim. Biophys. Acta 1149 (1993) 224–230.
[28] A. Sukhan, R.E. Hancock, The role of specific lysine residues in the passage of anions through the *Pseudomonas aeruginosa* porin OprP, J. Biol. Chem. 271 (1996) 21239–21242.
[29] P. Pongprayoon, O. Beckstein, C.L. Wee, M.S.P. Sansom, Simulations of anion transport through OprP reveal the molecular basis for high affinity and selectivity for phosphate, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 21614–21618.
[30] N. Modi, I. Barcena Uribarri, M. Bains, R. Benz, R.E. Hancock, U. Kleinekathoefer, The role of the central arginine R133 towards the ion selectivity of the phosphate specific channel OprP: effects of charge and solvation, Biochemistry (2013).
[31] N. Modi, I. Barcena-Uribarri, M. Bains, R. Benz, R.E. Hancock, U. Kleinekathofer, Tuning the affinity of anion binding sites in porin channels with negatively charged residues: molecular details for OprP, ACS Chem. Biol. (2015).
[32] D.P. Tieleman, H.J. Berendsen, A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer, Biophys. J. 74 (1998) 2786–2801.
[33] P. Pongprayoon, How do the protonation states of E296 and D312 in OmpF and D299 and D315 in homologous OmpC affect protein structure and dynamics? Simulation studies, Comput. Biol. Chem. 53PB (2014) 226–234.
[34] A. Kumar, E. Hajjar, P. Ruggerone, M. Ceccarelli, Structural and dynamical properties of the porins OmpF and OmpC: insights from molecular simulations, J. Phys.-Condens. Matter. (2010) 22.
[35] P.R. Singh, M. Ceccarelli, M. Lovelle, M. Winterhalter, K.R. Mahendran, Antibiotic permeation across the OmpF channel: modulation of the affinity site in the presence of magnesium, J. Phys. Chem. B. 116 (2013) 4433–4438.
[36] B.K. Ziervogel, B. Roux, The binding of antibiotics in OmpF porin, Structure 21 (2013) 76–87.
[37] K.R. Mahendran, E. Hajjar, T. Mach, M. Lovelle, A. Kumar, I. Sousa, et al., Molecular basis of enrofloxacin translocation through ompf, an outer membrane channel of *Escherichia coli*—when binding does not imply translocation, J. Phys. Chem. B 114 (2010) 5170–5179.
[38] B. Roux, The membrane potential and its representation by a constant electric field in computer simulations, Biophys. J. 95 (2008) 4205–4216.
[39] M.O. Jensen, D.W. Borhani, K. Lindorff-Larsen, P. Maragakis, V. Jogini, M.P. Eastwood, et al., Principles of conduction and hydrophobic gating in K+ channels, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 5833–5838.
[40] D.P. Tieleman, H. Leontiadou, A.E. Mark, S.J. Marrink, Simulation of pore formation in lipid bilayers by mechanical stress and electric fields, J. Am. Chem. Soc. 125 (2003) 6382–6383.
[41] A. Singh, V. Orsat, V. Raghavan, Soybean hydrophobic protein response to external electric field: a molecular modeling approach, Biomolecules 3 (2013) 168–179.
[42] J.D. Faraldo-Gomez, G.R. Smith, M.S. Sansom, Setting up and optimization of membrane protein simulations, Eur. Biophys. J. 31 (2002) 217–227.
[43] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, J. Mol. Model. 7 (2001) 306–317.
[44] A.M.J.J. Bonvin, A.E. Mark, W.F. van Gunsteren, The GROMOS96 benchmarks for molecular simulation, Comput. Phys. Commun. 128 (2000) 550–557.
[45] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[46] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[47] W. Humphrey, A. Dalke, K. Schulten, VMD—visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[48] O.S. Smart, J.G. Neduvelil, X. Wang, B.A. Wallace, M.S. Sansom, HOLE: a program for the analysis of the pore dimensions of ion channel structural models, J. Mol. Graph. 14 (1996) 354–360, 76.
[49] M. Aguilella-Arzo, J.J. Garcia-Celma, J. Cervera, A. Alcaraz, V.M. Aguilella, Electrostatic properties and macroscopic electrodiffusion in OmpF porin and mutants, Bioelectrochemistry 70 (2007) 320–327.
[50] B. Dhakshnamoorthy, S. Raychaudhury, L. Blachowicz, B. Roux, Cation-selective pathway of OmpF porin revealed by anomalous X-ray diffraction, J. Mol. Biol. (2009).
[51] S. Pezeshki, C. Chimerel, A.N. Bessonov, M. Winterhalter, U. Kleinekathofer, Understanding ion conductance on a molecular level: an all-atom modeling of the bacterial porin OmpF, Biophys. J. 97 (2009) 1898–1906.
[52] W. Im, B. Roux, Ions and counterions in a biological channel: a molecular dynamics simulation of OmpF porin from *Escherichia coli* in an explicit membrane with 1 M KCl aqueous salt solution, J. Mol. Biol. 319 (2002) 1177–1197.



Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: http://www.tandfonline.com/loi/tbsd20

Exploring the interactions of a DNA aptamer with human serum albumins: simulation studies

Wanwisa Panman, Deanpen Japrung & Prapasiri Pongprayoon

To cite this article: Wanwisa Panman, Deanpen Japrung & Prapasiri Pongprayoon (2016): Exploring the interactions of a DNA aptamer with human serum albumins: simulation studies, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2016.1224733

To link to this article: http://dx.doi.org/10.1080/07391102.2016.1224733

Journal of Biomolecular Structure and Dynamics, 2016
http://dx.doi.org/10.1080/07391102.2016.1224733

Taylor & Francis
Taylor & Francis Group

LETTER TO THE EDITOR

Exploring the interactions of a DNA aptamer with human serum albumins: simulation studies

Wanwisa Panman[a], Deanpen Japrung[b*] and Prapasiri Pongprayoon[c,d,e*]

[a]*School of Science, Walailak University, Nakhon Si Thammarat 80161, Thailand;* [b]*National Nanotechnology Center, National Science and Technology Development Agency, Thailand Science Park, Pathumthani 12120, Thailand;* [c]*Faculty of Science, Department of Chemistry, Kasetsart University, Chatuchak, Bangkok 10900, Thailand;* [d]*Center for Advanced Studies in Nanotechnology and Its Applications in Chemical, Food and Agricultural Industries, Kasetsart University, Bangkok 10900, Thailand;* [e]*Computational Biomodelling Laboratory for Agricultural Science and Technology (CBLAST), Kasetsart University, Bangkok 10900, Thailand*

Communicated by Ramaswamy H. Sarma

(*Received 10 February 2016; accepted 9 August 2016*)

Introduction

The level of blood glucose is a traditional indicator for diabetes diagnosis and treatment. To track diabetes progression, the levels of blood sugar and the percentage of glycated hemoglobin (HbA1c) are commonly measured. However, accurate analysis of HbA1c is costly, requires sizeable equipment, and is not suitable for short-term data analysis due to its long lifespan (120 days). Besides, the HbA1C percentage is unreliable in patients with aberrant red blood cells or kidney/liver disease (Inaba et al., 2007). To improve the diagnostic accuracy and precision, new protein indicators based on early stage glycation products are needed. Recently, human serum albumin (HSA) has been receiving increasing attention because of its shorter half-life (12–21 days). In addition, the concentration of glycated human serum albumin (GHSA) can be measured directly by a number of methods. In particular, GHSA was found to be a better glycemic indicator than HbA1c in diabetic patients with a kidney disorder (Inaba et al., 2007).

HSA is a protein with a heart-like shape and is composed of 585 amino acid residues and 3 domains (I, II, and III); each domain is subdivided into 2 subdomains (A and B) (Figure 1(A)). It is the most abundant protein in blood plasma and serves as a nutrient and drug carrier. HSA contains two drug-binding sites [Sudlow sites I (warfarin-azapropazone-binding site) and II (indole-benzodiazepine-binding site)] as shown in Figure 1(A). In earlier studies, various researchers suggested many glycation sites on HSA (K12, K51, K93, K159, K199, K205, K233, K276, K281, K286, K378, K414, K439, K525, K538, and K545) (Barnaby, Cerny, Clarke, & Hage, 2011; Wa, Cerny, Clarke, & Hage, 2007). GHSA was found to have low ligand-binding capacity and properties that differ from those of normal HSA (Lim, Cheng, & Yang, 2007). The level of GHSA in patients with diabetes was found to be two- to five-fold higher than that in healthy people (Paroni et al., 2007). Thus, GHSA can serve as a marker of diabetes. Many methods [e.g. enzymatic assays, high-performance liquid chromatography, and enzyme-linked immunoassays (ELISAs)] were tested to precisely quantify GHSA. One of the promising and advanced methods is an aptamer-based sensor.

An aptamer is a single-stranded oligonucleotide (DNA or RNA) with the length < 100 nucleotides. It has strong binding affinity and selectivity for specific targets. On the basis of their high selectivity, aptamers are widely used in many applications as a target-specific marker or aptasensor (Kaida et al., 2013; Yunn et al., 2015). These site-selective aptamers can serve as a diagnostic or theranostic tool. Recently, we experimentally identified some DNA aptamers that can bind apo HSA and GHSA using systematic evolution of ligands by exponential enrichment (SELEX) (Apiwat et al., 2016). This protocol has been widely used for screening of various single-stranded DNAs or RNAs against a target of interest. By means of SELEX, 17 aptamers were selected and characterized in our recent study (Apiwat et al., 2016). Due to the lowest free energy (–4.08 kcal/mol) and K_d of 5.78 μM, a 23-nucleotide aptamer was chosen in that study for further binding analysis. Nevertheless, this aptamer was experimentally found to bind both apo and glycated HSA, but GHSA was more preferable. No clear atomic-scale evidence was obtained regarding why the above oligonucleotides prefer apo HSA and GHSA. Knowledge about the orientation and location of the aptamer and albumin during binding at the atomic level should advance the understanding of the HSA–aptamer binding mechanism. This insight will help to identify

*Corresponding authors. Email: deanpen@nanotec.or.th (D. Japrung); fsciprpo@ku.ac.th (P. Pongprayoon)

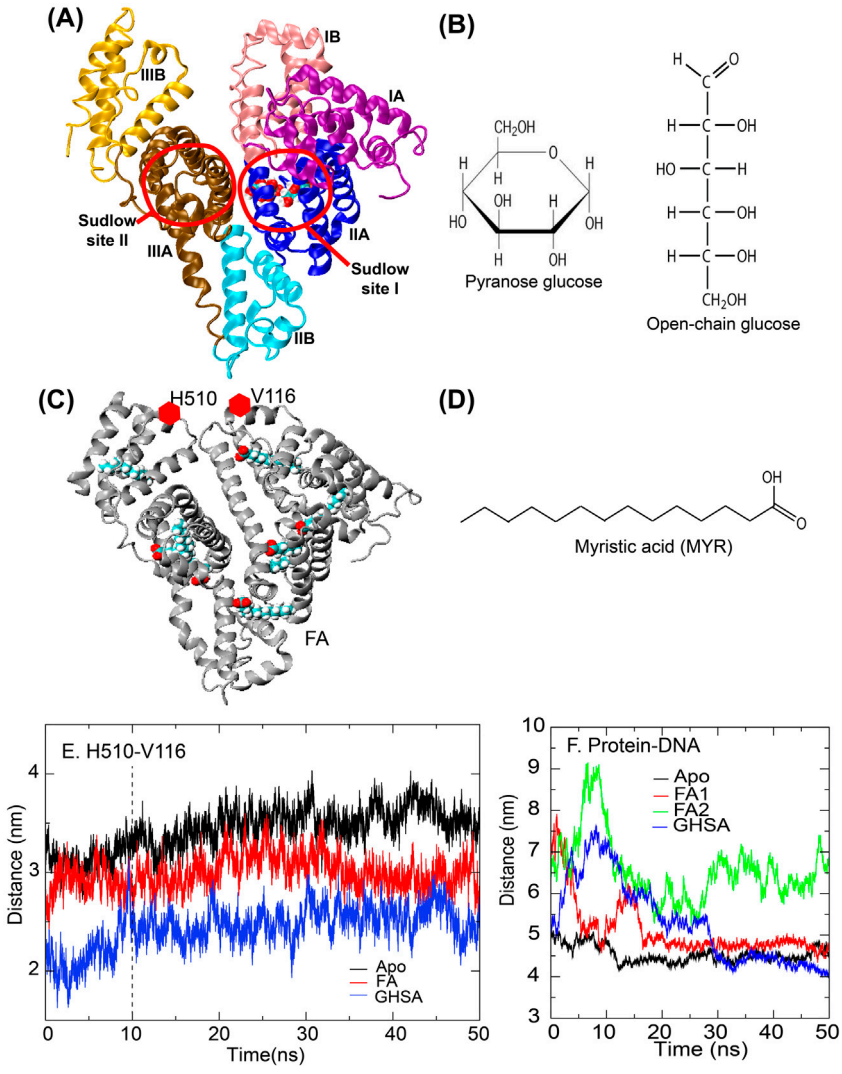


Figure 1. (A) HSA structure and (B) Two glucose conformations (left is a pyranose (closed) form (GLC) and open-chain glucose (GLO) is shown on the right). (C) Represents myristic acid (FA)-bound HSA where the structure of myristic acid is shown in (D). The sets of distances between H510 and V116 (their locations are shown as red dots in (C)) and protein and DNA are shown in (E) and (F), respectively.

key features for future aptasensor development (especially for diabetes screening and monitoring).

To obtain an atomic-scale picture, we employed Molecular Dynamics (MD) simulations to visualize progression of the aptamer–protein binding. Recently, the crystal structure of GHSA was solved (Wang et al., 2013). To determine the protein-binding affinity of an aptamer, we conducted MD simulations of an aptamer (the 23-nucleotide aptamer from our previous work) with 3 possible forms of HSA: normal HSA (Apo), fully fatty acid (myristic acid)-bound HSA (FA), and GHSA and then compared these data (Figure 1). We aimed to understand the binding mechanism of the aptamer and to reveal the key interactions and aptamer characteristics that are important for specific GHSA binding. Not only does this study provide a primary tool for elucidation of a mechanism of aptamer binding but also offers some guidelines for future development of more GHSA-specific aptamers.

Methods

Preparation of DNA structure

The three-dimensional structure of the 23-nucleotide aptamer (5′-TGCGGTTGTAGTACTCGTGGCCG-3′) was created by means of 3D-DART. This nucleotide chain contains two adenines (A), five cytosines (C), nine guanines (G), and seven thymines (T). The structure was parameterized using AMBER99SB force field. The DNA aptamer was solvated in water with counterions. This DNA system was equilibrated for 10 ns, followed by a 50-ns production run (details of the conditions used are

shown in the *Simulation protocols* section below). Root mean square deviations (RMSDs) of P atoms in the DNA backbone were calculated to track changes in DNA conformation. After 25 ns, a reduction in RMSDs was observed. This result is indicative of a less mobile and more packed DNA structure (Figure S1 in Supplementary Information). The final snapshot was used for our further analysis of DNA–protein binding due to its less fluctuating conformation and the presence of a hairpin-like loop at the 5′ end. This loop is similar to a recently published conformation observed experimentally (Apiwat et al., 2016).

Preparation of a DNA–HSA complex

The three crystal structures – Apo, FA, and GHSA (Protein Data Bank [PDB] IDs: 1E78, 1E7G, and 4IW2) – were downloaded from www.rcsb.org. The protonation states of all charged amino acid residues were set at physiological pH. The Adaptive Poisson–Boltzmann Solver (APBS) was used to calculate an overall electrostatic potential of HSA. The calculations were run locally for all molecules. Charges on a protein that were generated by PDB2PQR were used as input for APBS. The dielectric constants of 2 and 78.5 were chosen for the protein and solvent, respectively. The contact surface selection was mapped by means of a radius of 0.14 nm. The ion accessibility surface was defined with the probe radius of 0.2 nm. To examine DNA binding for all molecules, we had three systems: (1) Apo HSA–DNA (Apo) (2) fatty-acid-bound HSA–DNA (FA), and (3) glucose-bound HSA–DNA (GHSA). For each system, the aptamer with the same orientation (the 50-ns snapshot) was manually placed at the back of an HSA structure (at least 5 nm away from the center of mass of a protein to prevent the aptamer from seeing a protein at the beginning). The location of the aptamer was assigned based on the position of a large electropositive region on the protein surface obtained from APBS calculations (Figure 3, but the overall views of electrostatic potentials can be seen in Figure S2 in Supplementary Information).

Simulation protocols

All simulations were performed in the GROMACS 4.5.4 simulation package (www.gromacs.org) with AMBER99SB force field. Each structure was placed in a cubic simulation box solvating with water and counterions. To relax the steric conflicts generated during the setup, all energy minimizations involved up to 1000 steps of the steepest descent. Long-range electrostatic interactions were analyzed using the particle mesh Ewald method with a short range cutoff of 1 nm, Fourier spacing of 0.12 nm, and fourth-order spline interpolation. All simulations were carried out with the constant number of particles, pressure, and temperature (NPT ensemble). The temperatures of the protein, solvent, and ions were each coupled separately using the v-rescale thermostat at 300 K with the coupling constant $\tau_t = 0.1$ ps. The pressure was coupled using the Berendsen algorithm at 1 bar with the coupling constant $\tau_p = 1$ ps. The time step for integration was 2 fs. Coordinates were saved every 2 ps for subsequent analysis. The 10-ns equilibration runs were performed and followed by 50-ns production runs. Each system was run twice (1 and 2 represent simulations 1 and 2). All the results provided here represent average values from the two simulations. The data were analyzed by means of GROMACS and our own custom code. Molecular graphic images were prepared in VMD. C-alpha RMSD and RMSF calculations were computed using an initial structure from each production run as a reference. Principle component analysis (PCA) was conducted with default ‘g_covar; and ‘g_anaeig; options in GROMACS. Only the first eigenvector was used to analyze the major protein motions in all cases.

Results and discussion

At the beginning, the convergence of all systems was verified. The total MD energies were measured. In addition, other studies showed that the insignificant change in binding energies can reflect the convergence of systems (Su & Johnson, 2016; Su, Tsai, Mehboob, Hevener, & Johnson, 2015). We therefore measured a change in the binding energy using APBS calculations. The constant total energies and insignificantly fluctuating APBS binding energies observed in all the systems can ensure convergence of our simulations (see Figure S3 in Supplementary Information for energy plots). Moreover, other studies on DNA binding involved a ≤50-ns timescale to successfully explore DNA–protein binding events of even more complex DNA structures (Babin, Wang, Rose, & Sagui, 2013; Wheatley, Pieniazek, Mukerji, & Beveridge, 2012). This observation indicates that the simulation time of 50 ns is still sufficient to examine biological phenomena even though longer simulation time can make investigators more confident in the data. This work is an extension of our recent empirical study (Apiwat et al., 2016). Major findings here agree well with those experiments. The aim of the present study was to explain an aptamer–HSA binding event at the atomic level to improve aptamer specificity for future aptasensor design.

Overall, the C-alpha RMSDs and RMSFs of all three proteins showed a similar degree of structural flexibility, but the GHSA structure had somewhat higher flexibility due to slightly higher RMSDs and RMSFs (Figure S4 in Supplementary Information). To examine a domain motion, the distances between H510 and V116 at the interface between domains I and III were measured in all

cases (Figure 1(E)). The different distances between H510 and V116 in each simulation confirmed the previous finding that the presence of a bound ligand causes displacement of domains I and III (Pongprayoon & Gleeson, 2014; Awang et al., 2016). Especially, two glucose molecules in Sudlow site I make domains I and II very close to each other (Figure 1(E)).

Additionally, to obtain a dynamic picture, PCA was performed on C-alpha atoms in all cases. This analysis revealed that the first eigenvector accounts for the majority of motions in most systems, and there are some indications that a second eigenvector also plays a role (see Figure S5 in Supplementary Information). To make a comparison, only the dynamic motions obtained for the first eigenvector were studied here. In Figure 2, the RWB color format represents time-dependent structure motions for each system, where an initial conformation is in red and the final one is colored in blue. PCA indicates that domain I and subdomain IIIB are the most mobile regions (Figure 2). To facilitate visualization, only highly flexible regions are shown in Figure 2, but the conformational changes in the whole protein can be seen in Figure S6 in the Supplementary Information. If we examine the domain motion, subdomain IIIB and domain I in Apo move apart, and a scissor-like motion in FA is visible well (Figure 2). These motions are in good agreement with the results of other studies. If we compare motions of DNA-bound proteins with movements of the unbound versions from other studies, the directions of domain motion observed here are similar to those of proteins without bound DNA (Pongprayoon & Gleeson, 2014). This finding implies that the presence of bound DNA does not interfere with overall protein dynamics in all cases. Additionally, for GHSA, the inward movement of both domain I and subdomain IIIB in Figure 2 results in the shortest H510–V116 distance observed in Figure 1(E), whereas the longest observed distance in Apo is explained by the separation of domains I and III (Figures 1(E) and 2). The inward motion of GHSA observed here is in good agreement with the data from one study (Apiwat et al., 2016).

The two bound sugars packed inside Sudlow site I cause domain I and subdomain IIIB to move toward each other; this phenomenon was not observed in either Apo or FA (Figure 2).

As for aptamers, they show a conformational fluctuation during simulations, but DNA chains in all systems seem to be packed steadily within 25 ns (see RMSDs of the DNA backbone in Figure S1 in Supplementary Information). At the beginning, aptamers share a similar starting conformation with a mobile 3′ end and a hairpin-like loop at the 5′ tail (see Figure S1 in Supplementary Information regarding DNA conformation). The protein–DNA simulations were performed for 50 ns as in simulations in other studies where key dynamic properties could be successfully captured within this time scale (Babin et al., 2013). The protein–aptamer distances between the centers of mass (COMs) of individual molecules were measured in order to examine the DNA binding (Figure 1(F)). The distances between DNA and protein COMs decrease and then become constant (ranging from approximately 4.5–5 nm) in all cases except for FA2 (Figure 1(F)). When a short protein–DNA distance is observed (~4.5 nm), GHSA and Apo can bind to DNA in all simulations, whereas only FA1 (not FA2) can interact with DNA (distance of ~4.7 nm). The aptamer in FA2

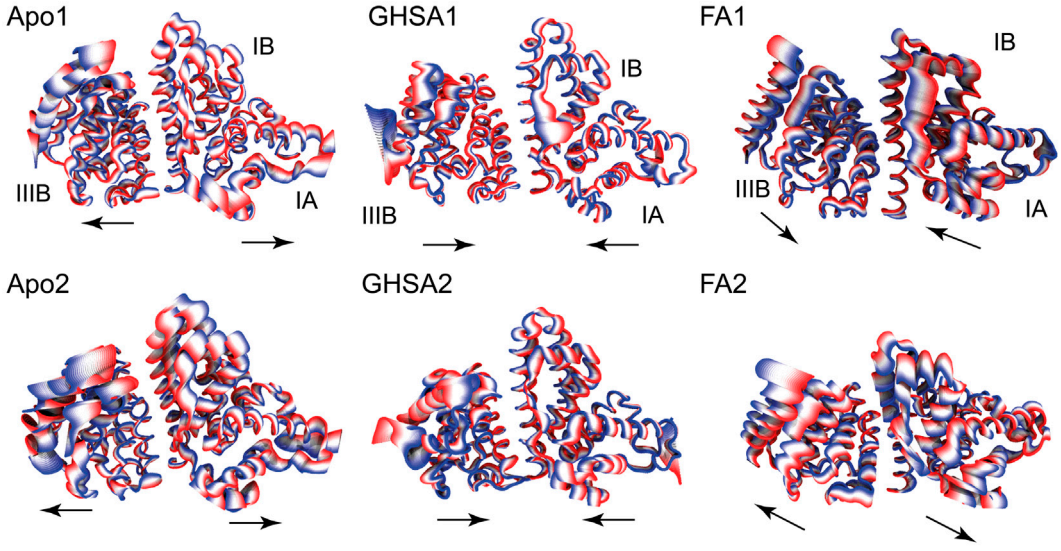


Figure 2. Cartoon showing time-dependent motion calculated from the first principle component 1 (PC1) for all systems. The colors are in RWB format where the initial position is in red and the blue represents the final position.

seems to be blind and fails to stick to HSA (at a distance of ~6 nm from HSA). The above pieces of evidence show the weakest DNA binding for FA; this result is in line with a recent empirical study (Apiwat et al., 2016).

All successful aptamers were found to bind to the positively charged patch on the back of a protein. This area is located at the center of the heart-like structure and mostly covers subdomain IIIB (blue color in Figure 3). Such large electropositive area enables different DNA-binding orientations. To better understand DNA–protein binding, electrostatic potentials of all the proteins were computed. In Apo, both DNA chains (in Apo1 and Apo2) bind to the area near the center of the protein because of a large positively charged environment although they adopt different bound conformations (Figure 3(A)). For FA1, a DNA chain appears to align near subdomain IIIB. So do aptamers in GHSA. Our APBS results suggest that different ligands promote variation in surface electrostatic potentials, but the presence of glucose molecules and myristic acid molecules activates DNA binding in subdomain IIIB only. Aptamers in GHSA1 and GHSA2 similarly align parallel to the

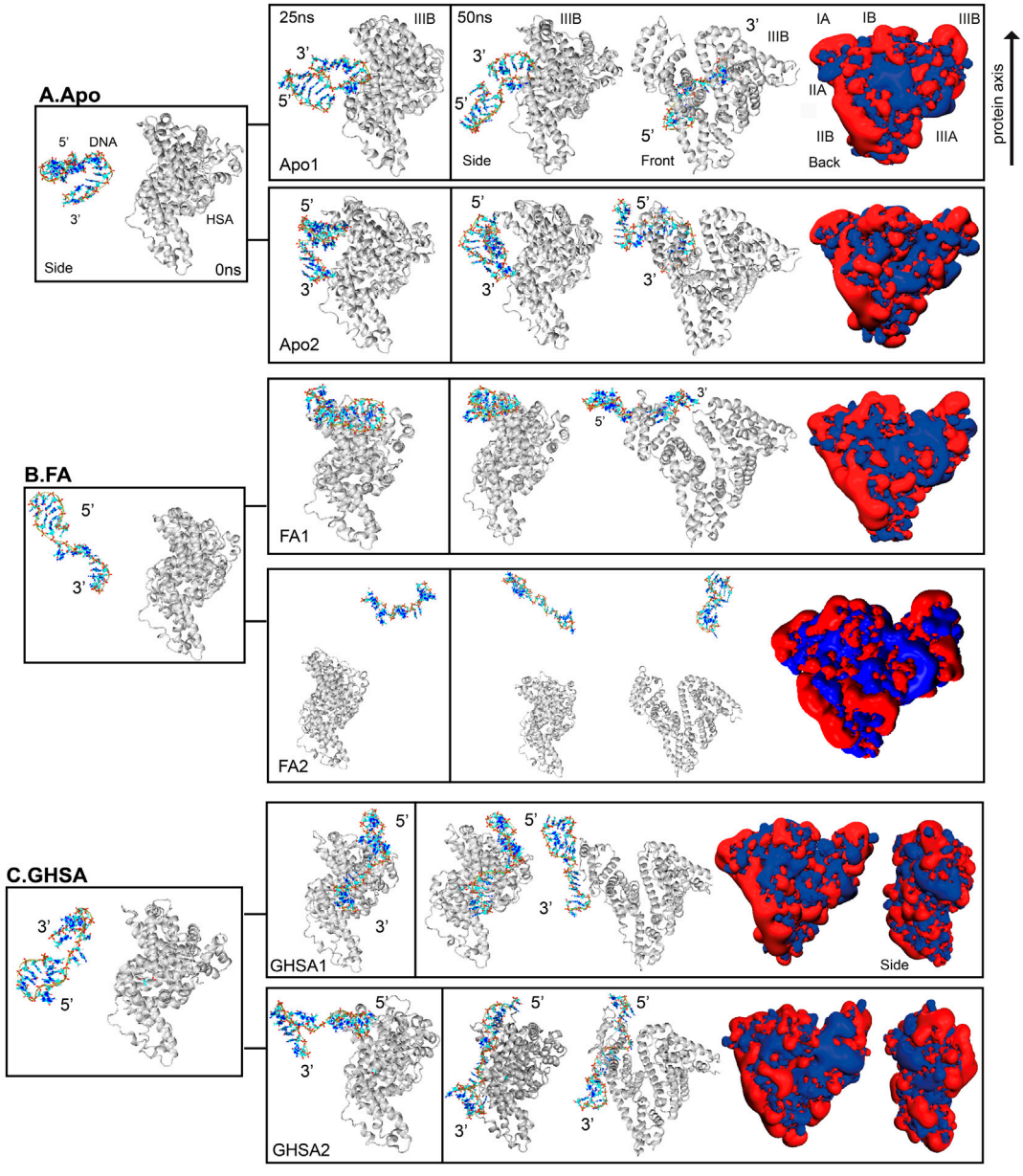


Figure 3. Cartoon views of Apo (A), FA (B), and GHSA (C) in the presence of DNA aptamer at 0, 25, and 50 ns, respectively. DNA aptamer is shown in licorice format. The electrostatic potentials of the last snapshot (50 ns) for all system are also shown on the right. The electropositive potential is shown in blue, while red represent the electronegative area.

protein axis (Figure 3). Both DNA chains in all GHSA simulations bind to the same subdomain (IIIB; Figure 3(C)). Thus, subdomain IIIB seems to be site-specific for aptamer binding in GHSA.

The capacity for aptamer binding depends on the stability and fixation of the electropositive surface. Lysines and arginines are important for formation of such an electropositive environment. Lysines are distributed evenly over a protein surface, whereas arginine residues appear to be in the middle of the protein (Figure S7 in Supplementary Information). HSA contains 60 lysines and 24 arginines; therefore, lysines seem to play a dominant role in the electropositive properties on the protein surface. Nonetheless, HSA also has some negatively charged amino acid residues to balance the protein charge (Figure S7 in Supplementary Information). According to APBS results in Figure 3, the facial electrostatic potentials are dynamic. They depend on delocalization of charged side chains on the protein surface. Hence, the magnitude and frequency of domain moving and shaking become a key factor that determines the coulombic environment for aptamer binding. Generally, the more flexible the domains, the more liberated become the amino acid side chains. Too mobile charged side chains cannot sustain the static and stable electropositive environment, which is essential for aptamer binding. This notion helps to explain why the flexible binding domains in FA are a bad host for aptamer binding. To confirm this idea, the protein–DNA electrostatic binding energies were computed using APBS. Although other nonelectrostatic terms are not strongly involved, these energies can be used to roughly estimate binding affinity of each protein–DNA complex. The trend of these binding energies is well consistent with our findings. All HSA forms can bind to an aptamer but with different affinity. FA shows the lowest coulombic binding energy among the three, whereas GHSA is the best host for DNA binding because of the highest binding energy (see Table S1 in Supplementary Information). The conformational entropy calculation is not included here because of a computational cost, and it is difficult to calculate conformational entropy accurately due to the wild fluctuation of entropy values.

The formation of a binding complex is further supported by a number of hydrogen bonds. As a rule, the 7 self-formed hydrogen bonds and 250 hydrogen bonds with adjacent water molecules are observed in the unbound aptamer (257 hydrogen bonds in total; Table 1). During the binding, the aptamer loses some water contacts, but they are compensated by protein interactions. The aptamer can form five hydrogen bonds with Apo and six hydrogen bonds with FA or GHSA. Both the protein and aptamer seem to be more packed in FA because of 10 internal hydrogen bonds in the aptamer and 463 in the protein (Table 1). In comparison with the unbound aptamer, binding to FA yields 3 more internal hydrogen bonds with a loss of 14 hydrogen bonds with water (Table 1). These constitute 252 hydrogen bonds in total for FA, while Apo and GHSA have 257 hydrogen bonds just as in the stand-alone aptamer in solution (Table 1). Ten internal hydrogen bonds and six protein–aptamer interactions may be insufficient for the aptamer to compensate for the loss of 14 hydrogen bonds with water in FA. When compared with the other HSA forms, the reduction in the total number of hydrogen bonds in FA can reflect the less favorable environment for aptamer binding.

Table 1. The average number of hydrogen bonds during simulations.

No. of hydrogen bonds	Apo	FA	GHSA
Protein (within molecule)	460	463	458
Aptamer (within molecule)	8	10	7
Protein–aptamer	5	6 (FA1)	6
Aptamer (unbound)*		7	
Aptamer–H_2O (bound)	244	236	244
Aptamer–H_2O (unbound)*		250	

Note: The asterisk represents the results from the simulation of a stand-alone aptamer in water.

To better understand the binding mechanism, a detailed hydrogen bond analysis was carried out. At the beginning, there were no protein–aptamer hydrogen bonds because the aptamer was located far away from the protein (Figure 4(A)–(E)). The aptamer discovers the protein approximately within 10 ns (Figure 4(A)–(E)). Except for GHSA2, the aptamer spends half the time in bulk water before seeing the protein after 25 ns (Figure 4(D)). The hydrogen bond analysis clearly shows that the aptamer can bind to all forms of HSA but with different affinity as suggested earlier. This finding was also made in our previous empirical study (Apiwat et al., 2016). All aptamers attach to the back of a protein in subdomain IIIB (except for Apo1 where the aptamer tends to interact with the middle of the protein structure) for various aptamer conformations and locations. If we review hydrogen bonds for an individual complex, although there are divergent binding behaviors observed here, they still share a similar pattern: lysines are the main players in aptamer binding. At least two lysines are involved in a binding event, and these residues can form more than one hydrogen bond with the aptamer. In Apo1, nucleotide sequences A13–G17 in the middle of the aptamer chain (Figure 4(F)) first interact with a protein in the back and stay there for the rest of simulations (Figure 3(A)). Those nucleic-acid sequences form strong hydrogen bonds with two arginines (R160 and R445), two lysines (K439 and K444), H440, and E442. The aptamer aligns from domain I to domain III (see Figures 3A and 4G for final location). Just as in Apo1, the aptamer in Apo2 utilizes nucleic acid sequences in the

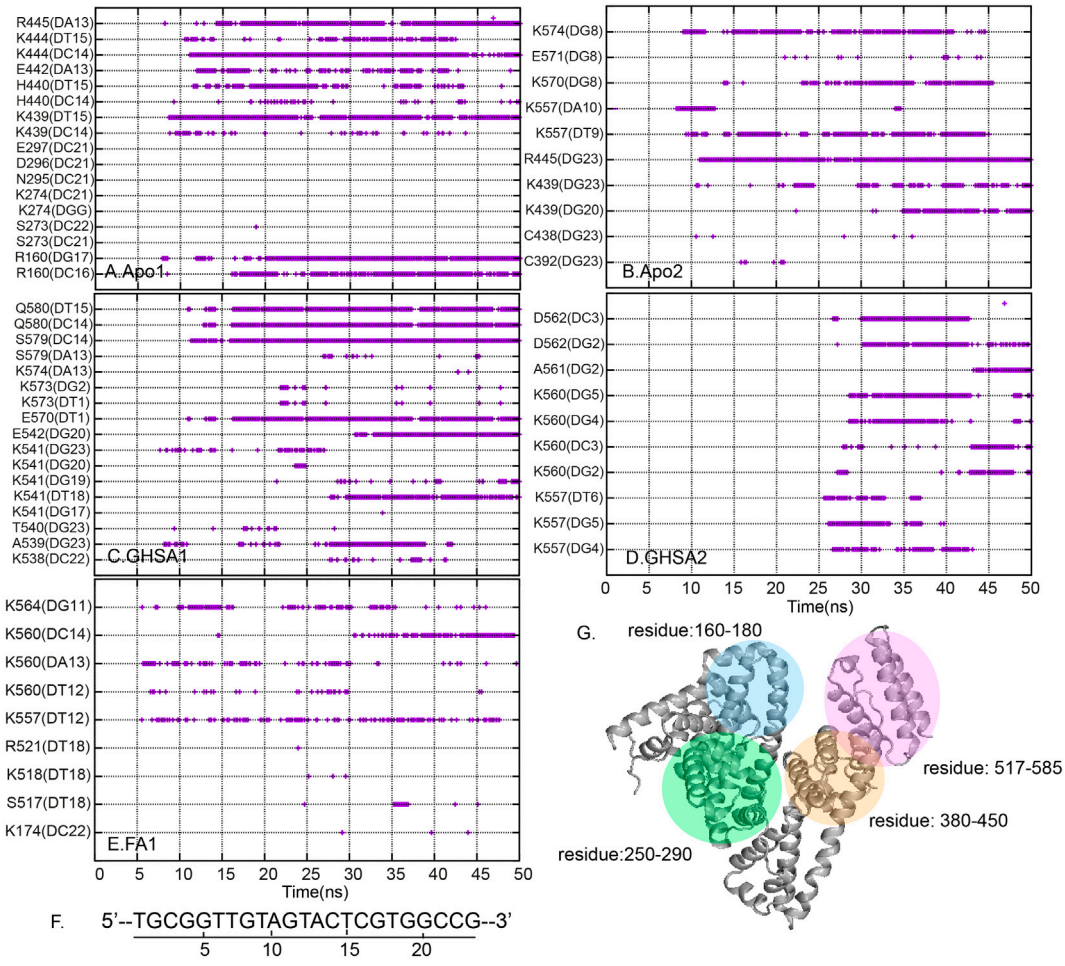


Figure 4. A number of hydrogen bonds as a function of time for all cases. The rough locations of involving residues on the back of protein are colored and shown at the bottom right corner.

middle of the chain [sequences G8–A10 (Figure 4(F))] to interact with C392, C438, R445, and four lysines (K439, K557, K570, and K574; Figure 4(B)–(G)). This middle of the aptamer chain initially binds to subdomain IIIB before the 3′ end lies down on subdomain IIIB and leaves the 5′ end unbound (see Figures 3(A) and 4(G) for locations). This binding pattern (the middle part of the aptamer chain interacts with the protein) in both Apo1 and Apo2 is similar to that of FA1 in spite of different binding regions. The aptamer in FA1 uses sequences G11–C14 (see Figure 4(F) for a sequence list) to attract K557, K560, and K564 in subdomain IIIB before the whole chain sits down on the back of subdomain IIIB (Figure 3(B)). In case of GHSA, the aptamer in GHSA1 first attaches the 3′ end [sequences G20–G23 (see Figure 4(F) for a sequence list)] to subdomain IIIB before the rest of the chain lies beside subdomain IIIB (Figure 3(C)). On the other hand, the aptamer sticks to the GHSA2 surface via the 5′ end. Sequences G2–T6 (see Figure 4(F) for a sequence list) in the 5′ tail initially interact with the back of subdomain IIIB. The 5′ end forms hydrogen bonds with K557, K560, A561, and D562 (Figure 4(D)). This set of interactions causes the whole chain to align on the back of subdomain IIIB throughout the simulation.

It is noteworthy that the binding of 5′ end to GHSA observed here is in good agreement with an empirical study where the 5′ end was found to be the main player in GHSA binding (Apiwat et al., 2016).

In general, although there is no precise binding site for the aptamer, the aptamers in all simulations still have similar binding and structural features. All aptamers employ the middle of their chain to bind to subdomain IIIB. After the binding, the 5′ ends in all cases are left unbound in solution. This end folds back to form a loop that protects this end from outside interactions. An aptamer instead utilizes the 3′ tail to stick to the protein surface by interacting with lysines and other polar residues. As suggested earlier, lysine is the major determinant of aptamer attachment. Nevertheless, the

contribution of other residues (especially arginines) is also crucial. FA1 appears to have the smallest number of contributors. The major protein–aptamer interactions are derived from three lysines, whereas aptamers in Apo and GHSA are supported by many polar and charged residues. The different numbers of interactions and different electrostatic properties cause the variation in binding affinity.

Conclusions

MD simulations can be an effective method for analysis of a complex event such as biomarker–aptamer binding. According to this study, MD simulations can be used to study the binding mechanisms of protein–aptamer complexes. Lysine residues play an important role in this binding. Additional polar and charged residues also help to maintain the binding to an aptamer. Although no specific binding mechanism was observed here, it is obvious that subdomain IIIB is the target for aptamer binding. Furthermore, the dynamic properties of domains I and III, causing a change in electropositive potentials, are also important for this binding event. The more mobile domains change electrostatic potentials, thus creating a poor environment for aptamer binding. Even though all results here are consistent with the notion that subdomain IIIB serves as a DNA-binding site, exploring other possible DNA-binding sites in HSA would still be interesting for aptamer development in the future. Apparently, the aptamer binds to GHSA with the highest affinity, but the specificity and selectivity for GHSA remain insufficient. A more specific and selective aptamer is required for the development of future aptasensors. An atomic-scale insight obtained here may facilitate SELEX and consequently can be applied to aptasensors for control and monitoring of diabetes and other metabolic diseases.

Abbreviations

HSA	Human serum albumin
GHSA	Glycated human serum albumin
MD	Molecular dynamics

Supplementary material

The supplementary material for this article is available online at http://dx.doi.org/10.1080/07391102.2016.1224733.

Acknowledgments

We would like to thank Kasetsart University Research and Development Institute (code:VT(D)-87.59), Thailand Research Fund (TRF), National Research Council of Thailand (NRCT), Health Systems Research Institute (HSRI), and Royal Society of Chemistry (RSC) for financial support. We are also grateful to the anonymous reviewers for good comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Kasetsart University Research and Development Institute [code: VT(D)-87.59]; Thailand Research Fund (TRF) [TRG5880230]; National Research Council of Thailand (NRCT); Health Systems Research Institute (HSRI); and Royal Society of Chemistry (RSC).

References

Apiwat, C., Luksirikul, P., Kankla, P., Pongprayoon, P., Treerattrakoon, K., Paiboonsukwong, K., … Japrung, D. (2016). Graphene based aptasensor for glycated albumin in diabetes mellitus diagnosis and monitoring. *Biosensors and Bioelectronics, 82*, 140–145. doi:10.1016/j.bios.2016.04.015

Awang, T., Wiriyatanakorn, N., Saparpakorn, P., Japrung, D., & Pongprayoon, P. (2016). Understanding the effects of two bound glucose in Sudlow site I on structure and function of human serum albumin: Theoretical studies. *Journal of Biomolecular Structure and Dynamics,* 1–10. doi:10.1080/07391102.2016.1160841

Babin, V., Wang, D. L., Rose, R. B., & Sagui, C. (2013). Binding polymorphism in the DNA bound state of the Pdx1 homeodomain. *PLoS Computational Biology, 9*, e1003160. doi:10.1371/journal.pcbi.1003160

Barnaby, O. S., Cerny, R. L., Clarke, W., & Hage, D. S. (2011). Comparison of modification sites formed on human serum albumin at various stages of glycation. *Clinica Chimica Acta, 412*, 277–285. doi:10.1016/j.cca.2010.10.018

Inaba, M., Okuno, S., Kumeda, Y., Yamada, S., Imanishi, Y., Tabata, T., … Grp, O. C. E. R. (2007). Glycated albumin is a better glycemic indicator than glycated hemoglobin values in hemodialysis patients with diabetes: Effect of anemia and erythropoietin injection. *Journal of the American Society of Nephrology, 18*, 896–903. doi:10.1681/ASN.2006070772

Kaida, Y., Fukami, K., Matsui, T., Higashimoto, Y., Nishino, Y., Obara, N., … Yamagishi, S. (2013). DNA aptamer raised against AGEs blocks the progression of experimental diabetic nephropathy. *Diabetes, 62*, 3241–3250. doi:10.2337/db12-1608

Lim, P. S., Cheng, Y. M., & Yang, S. M. (2007). Impairments of the biological properties of serum albumin in patients on haemodialysis. *Nephrology, 12*, 18–24. doi:10.1111/j.1440-1797.2006.00745.x

Paroni, R., Ceriotti, F., Galanello, R., Battista Leoni, G., Panico, A., Scurati, E., … Mosca, A. (2007). Performance characteristics and clinical utility of an enzymatic method for the measurement of glycated albumin in plasma. *Clinical Biochemistry, 40*, 1398–1405. doi:10.1016/j.clinbiochem.2007.08.001

Pongprayoon, P., & Gleeson, M. P. (2014). Probing the binding site characteristics of HSA: A combined molecular dynamics and cheminformatics investigation. *Journal of Molecular Graphics and Modelling, 54*, 164–173. doi:10.1016/j.jmgm.2014.10.007

Su, P. C., & Johnson, M. E. (2016). Evaluating thermodynamic integration performance of the new amber molecular dynamics package and assess potential halogen bonds of enoyl-ACP reductase (FabI) benzimidazole inhibitors. *Journal of Computational Chemistry, 37*, 836–847. doi:10.1002/jcc.24274

Su, P. C., Tsai, C. C., Mehboob, S., Hevener, K. E., & Johnson, M. E. (2015). Comparison of radii sets, entropy, QM methods, and sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA ligand binding energies of F-tularensis enoyl-ACP reductase (FabI). *Journal of Computational Chemistry, 36*, 1859–1873. doi:10.1002/jcc.24011

Wa, C., Cernry, R. L., Clarke, W. A., & Hage, D. S. (2007). Characterization of glycation adducts on human serum albumin by matirx-assisted laser desorption/ionzation time-of-flight mass spectroscopy. *Clinica Chimica Acta, 385*, 48–60. doi:10.1016/j.cca.2007.06.011

Wang, Y., Yu, H., Shi, X., Luo, Z., Lin, D., & Huang, M. (2013). Structural mechanism of ring-opening reaction of glucose by human serum albumin. *Journal of Biological Chemistry, 288*, 15980–15987. doi:10.1074/jbc.M113.467027

Wheatley, E. G., Pieniazek, S. N., Mukerji, I., & Beveridge, D. L. (2012). Molecular dynamics of a DNA holliday junction: The inverted repeat sequence d(CCGGTACCGG)4. *Biophysical Journal, 102*, 552–560. doi:10.1016/j.bpj.2011.11.4023

Yunn, N. O., Koh, A., Han, S., Lim, J. H., Park, S., Lee, J., … Ryu, S. H. (2015). Agonistic aptamer to the insulin receptor leads to biased signaling and functional selectivity through allosteric modulation. *Nucleic Acids Research, 43*, 7688–7701. doi:10.1093/nar/gkv767