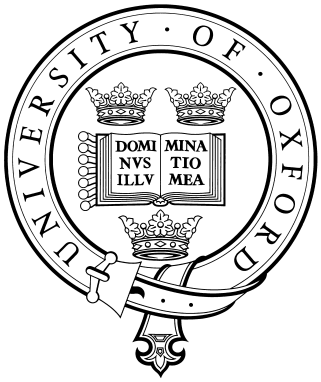
Molecular Modelling of β-barrel Outer Membrane Proteins

Prapasiri Pongprayoon

Structural Bioinformatics and Computational Biochemistry Unit,

Department of Biochemistry and Wolfson College, Oxford



A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

June 2010

Molecular modelling of β-barrel outer membrane proteins

Prapasiri Pongprayoon

Department of Biochemistry and Wolfson College, Oxford

Submitted for the degree of Doctor of Philosophy, Michaelmas 2010

Abstract

In Gram-negative bacteria, the Outer membrane (OM) acts as a first barrier to screen unwanted compounds whilst enabling ions and very small solutes to diffuse into the cell. Most of nutrients and essential ions are effectively transported across a membrane via the outer membrane proteins (OMPs). The water-filled β-barrel OMPs are called porins. These pores are classified into two groups, non-specific and substrate-specific porins. Each of them has different mechanisms to facilitate its substrate translocation. To reveal the process of substrate permeation and selectivity in microscopic detail, molecular dynamics (MD) simulations and applications were performed in this thesis.

The studies in this thesis focus on a series of classical porins. These proteins share similar feature where extracellular loop(s) (generally loop 3 (L3)) is folded into the middle of the pore and act as a constriction site which is important for substrate specificity and selectivity. The studies firstly concentrate on the structural properties and dynamics of the general trimeric porins, OmpC and OmpF whose sequences share 60% identity. OmpC and OmpF are found to have similar mechanism of latching loop (L2) to maintain trimeric stability. The smaller pore size allows OmpC to be more cation-selective than OmpF. Additionally, the major driving force for cation permeation in both porins is not from electrostatic properties. This differs from the phosphate-selective porin, trimeric OprP, where a phosphate diffusion depends on electrostatic interactions with positively charged pore-lining residues. The charge brush-like behavior of interior Arg and Lys residues plays a major role in phosphate selectivity. Also, the free energy profiles (PMF) reveal two key regions that are important for differentiating phosphate from other anions. The brush-like mechanism of OprP were also implanted to the simplified model pores in order to determine the possibility of transferring phosphate-selective properties of OprP to a model which may be useful for future design of nanopores. It is found that the duplication of functional residues and pore cavity can turn a model into the highly phosphate-selective pore. Importantly, the phosphate-binding affinity is dependent on the ability of the pore to interfere and occupy the hydration shell of a translocating phosphate where such ability can be maximized by an increase in sidechain flexibility. In case of uptake of more complex substrates, OpdK also employs a constriction site to select its substrate, aromatic vanillate (VNL) with total charge of -1. Unlike ion-specific porins, the free VNL is attracted by polar and aromatic interactions and sequentially directed through the periplasmic vestibule by charged residues insides the pore. The correct orientation of VNL on arrival is crucial for OpdK to recognize and enable the permeation process.

Acknowledgements

The research work presented in this thesis was conducted under the supervision of Prof. Mark SP Sansom, whom I wish to thank for welcoming me in his research group, as well as for his support and advice throughout my DPhil studies. Mark is more than a supervisor to me. The way of his thinking in both academia and real life acts as a template for my future career. Especially, I love his optimistic ideas. Personally, I am too pessimistic in both academic and personal life, but after working with him for 4 years, his positive actions have gradually changed me. Mark is the best supervisor that I have ever had.

Also, these studies were funded by the Royal Thai government. I am most grateful to those anonymous people that, on behalf of the government, decided to give me such a life-changing opportunity.

Completing this DPhil has been a very significant challenge, not only academically, but also personally. In my point of view, scientific research is not only a matter of knowledge, good ideas, and hard work, but also of integrity. I am very grateful to those that made this difficult journey a bit more bearable. Firstly, I want to thank the group of Thai friends here, for continuous help, encouragement, and also good conversations that making me less homesick, especially, Petchara Pattarakijwanich who dedicated his time to addressing most of my physics questions. Secondly, I also want to thank Syma Khalid, Jayne Wallace, Michelle Sahai and Ranjit Vijayan for being such good friends since I first came here, especially Ranjit who always gives me a hand when I face with some computer problems. Finally, Oliver Beckstein is another person that I want to say “thank you”. My DPhil studies here cannot be completed without his kind help. He is the first person who made me realized how intelligent one could be. Having academic discussions with him always gives me the great motivation to learn more and more in order to know and understand things as good as him (although in practice, it is not possible to be as good as him within a couple of days).

To conclude, I also want to thank my parents for their unconditional support. Although they are not rich and cannot fulfill some of my educational dreams, it is a great pleasure to be raised under their arms.

Oxford, June 2010

Table of Contents

Table of abbreviations

APBS	Adaptive Poisson Boltzmann Solver
COO^-	Carboxylate group
EC	Extracellular
EM	Energy minimization
FP	Free energy perturbation
LPS	Lipopolysaccharide
OM	Outer membrane
OMA	Outer membrane auxiliary
OMP	Outer membrane protein
PB	Poisson-Boltzmann
PDB	Protein data bank
Pi	Phosphate ion
PMF	Potential of mean force
PP	Periplasmic
MD	Molecular dynamics
SMD	Steered Molecular dynamics
TI	Thermodynamic integration
VdW	Van der Waals
VNL	Vanillate

1 INTRODUCTION

1.1 General features of the biological membrane

The biological membrane defines the boundary between the cell and its environment (1). Membranes are responsible for the uptake of all nutrients and exclusion of toxic agents (2). Biological membranes comprise lipids, proteins, and carbohydrates (Figure 1-1A). The lipid molecules constitute about 50% of the mass of the cell membrane (1). All lipids in a membrane are amphiphilic where each lipid molecule contains both a hydrophilic (head group) domain and hydrophobic tails (1). The lipid molecules are arranged as a continuous bilayer where the head groups form the upper and lower surfaces which are exposed to the aqueous environment and the tails are buried in between (Figure 1-1A). The most abundant lipids are the phospholipids (1). These have a phosphate-containing head group and a range of fatty acids as hydrocarbon tails. In the bilayer, 5% of the membrane is composed of the sugar-containing lipids called glycolipids where polar groups are external to the lipid bilayer (Figure 1-1A). Apart from lipids, two types of proteins are also found in the membrane. Proteins that have one or more transmembrane (TM) segment embedded in the bilayer are called integral membrane proteins (Figure 1-2A) (1, 2); these in turn are classified into two groups, α-helical and β-barrel proteins (Figure 1-1B) (3). The α-helix bundle is more abundant and occurs in the great majority of membranes. The β-barrel membrane proteins are restricted to the outer membrane of Gram-negative bacteria (3) and related membranes such as the outer membrane of mitochondria. By contrast, proteins that associate at the membrane surface are known as peripheral membrane proteins (Figure 1-2A) (1, 2). In this study, only the integral outer membrane proteins are discussed.

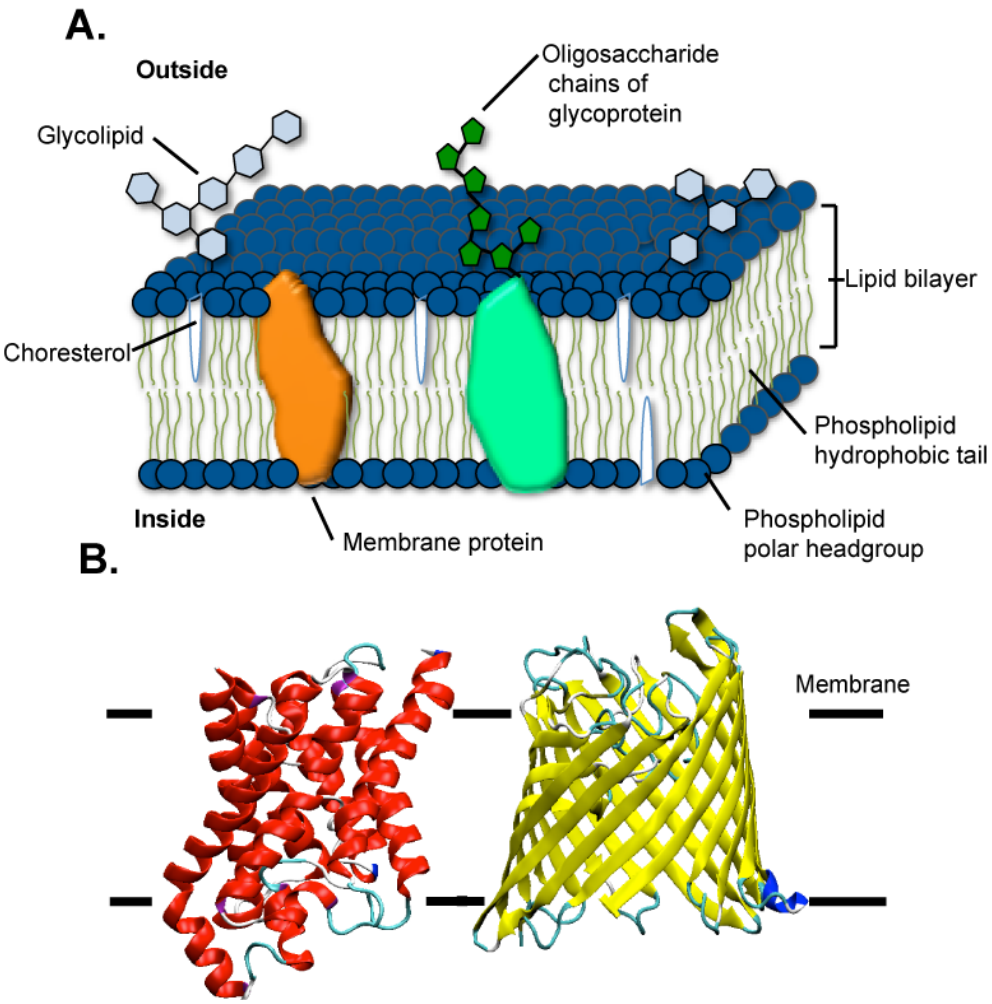


Figure 1-1 (**A**) Membrane components. (**B**) The two classes of membrane proteins: an α-helix bundle (*left*) demonstrated by GlpF (4) and a β-barrel (*right*) demonstrated by OpdK (5). The black line indicates the location of membrane.

Bacteria have an extra cell wall which is a multilayered structure located external to the cytoplasmic membrane (6). The differences in the chemical and physical properties of the cell wall are used to categorise bacterial types. Based on the structure and chemical composition of the cell wall, bacteria are differentiated into two main groups, Gram-negative and Gram-positive (7). The major differences of the Gram-positive from Gram-negative bacteria are the high amount of peptidoglycan in a cell wall and lack of an outer membrane (Figure 1-2B). Gram-negative cells have a complex outer membrane (OM) with the presence of lipopolysaccharide (LPS) in the outer leaflet (Figure 1-2B) (6). Such bacteria also have a periplasmic space between the outer-membrane layer and the inner

membrane which contains a number of water soluble proteins such as periplasmic binding proteins involved in transport from the outer to the inner membrane. The OM acts as a first permeability barrier that enables the membrane to maintain charge and concentration gradients (2). The essential nutrients and ions are transported across OM by the presence of the outer membrane proteins (OMPs).

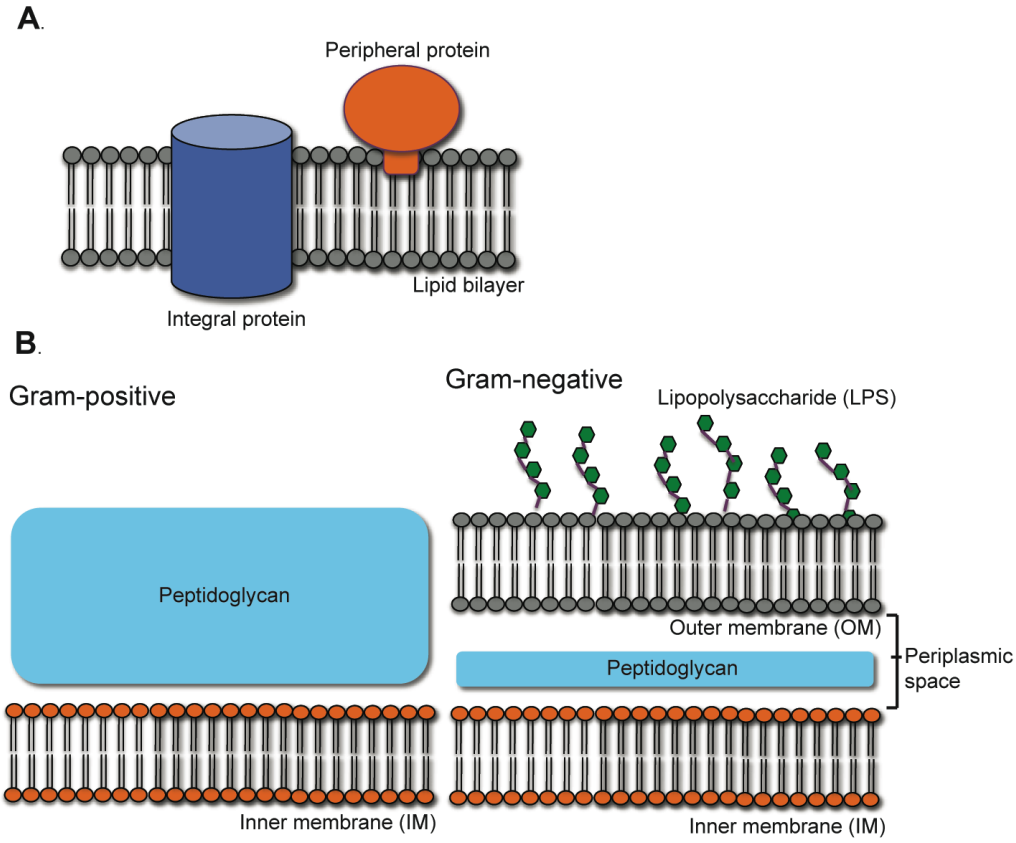


Figure 1-2 (**A**) Two types of membrane protein found in the lipid bilayer. (**B**) Cell wall structures of gram-positive (*left*) and gram-negative (*right*) bacteria. Note that the peptidoglycan in gram-positive bacteria is much thicker than in gram-negative bacteria and only gram-negative bacteria have an outer membrane (containing lipopolysaccharide (LPS)) (6).

1.2 The outer membrane proteins (OMPs)

To date, the majority of OMPs are transmembrane β-barrel proteins. α-helix bundles are rarely found in the OM (3); within the currently known OMP structures, there are only two crystal structures of α-helical barrel outer membrane proteins. Firstly, it is the outer membrane auxiliary (OMA), Wza, which is crucial for the capsule assembly pathways in *Escherichia coli* (8) (Figure 1-3). The second known α-helical barrel is a part of a type IV secretion system (T4SS) whose process is important for the conjugative transfer of DNA (9). At present, this structure is the largest determined for an outer membrane protein (10).

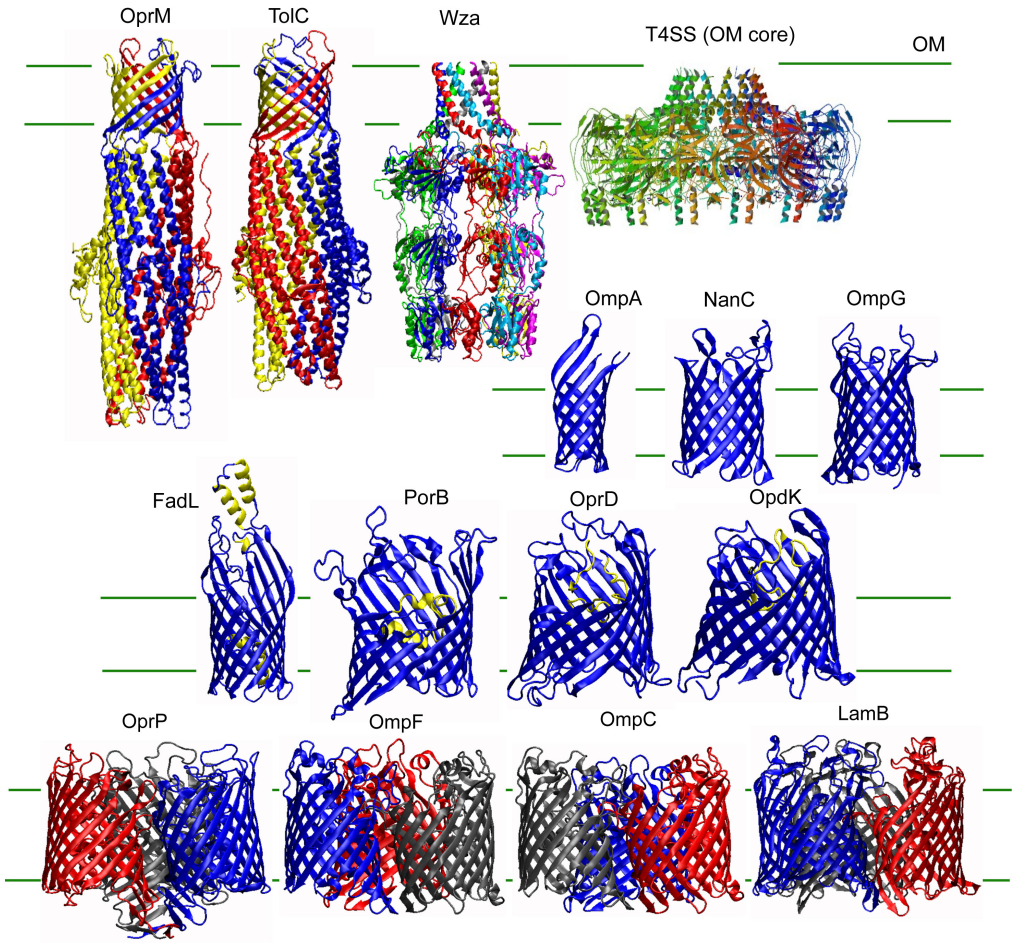


Figure 1-3 Side views of examples of outer membrane proteins (OMPs) found in Gram-negative bacteria. Different colors of each barrel represent a subunit. The yellow helical loop folded into the lumen of the pore in the third row labels the constriction loops that control the substrate translocation. Two green lines show the boundary of the outer membrane. From the top row, OprM (PDB code: 3D5K), TolC (PDB code: 2VDE), Wza (PDB code: 2J58), T4SS (PDB code: 3JQO), OmpA (PDB code: 1QJP), NanC (PDB code: 2WJQ), OmpG (PDB code: 2IWV), FadL (PDB code: 2R88), PorB (PDB code: 3A2R), OprD (PDB code: 2ODJ), OpdK (PDB code: 2QTK), OprP (PDB code: 2O4V), OmpF (PDB code: 2OMF), OmpC (PDB code: 2J1N), and LamB (PDB code: 1AF6).

Many OMPs are water-filled β-barrel proteins called porins. This group is crucial for the uptake of small soluble molecules and ions and the extrusion of toxic agents such as antibiotics from the periplasm to the extracellular environment. Importantly, porins play a key role in screening unwanted molecules entering into a cell; only selected polar molecules or ions less than 600 Daltons are able to permeate to the periplasmic space (1). In general, the passage of small molecules across the OM is mediated by three classes of proteins: (1) non-specific porins, (2) substrate-specific channels, and (3) active transporters. Both non-specific porins and substrate-specific channels do not require energy. Substrate diffusion through these pores is facilitated according to concentration gradients (5). The most common non-specific porins are OmpC and OmpF (Figure 1-3) which are present in a large number of Gram-negative bacteria (5). At low substrate concentration, passive diffusion through general porins is no longer efficient. Substrate-specific channels such as maltose-specific LamB (11), phosphate-specific OprP (12), and the fatty acid transporter FadL (13) which contain binding sites for their substrate are needed (Figure 1-3). Indeed, many Gram-negative bacteria, such as pseudomonads and other soil-dwelling microbes only have substrate-specific porins. All chemical transport in these living cells is based only on substrate-specific channels, which is an efficient defense against toxin agents. Also, this protection barrier allows the human pathogen (i.e. pseudomonads) to be more difficult to treat clinically. Many antibiotics employ porins as an entryway. To fight against them, bacteria develop sequential mechanisms to reduce influx of various drugs by means of reduction of porin expression and porin mutation, together with an efficient efflux system (14). To extrude drugs to the extracellular environment, bacteria utilize a discharge duct that directs drug passage across both inner and outer membrane. The most common ducts are TolC in *E. coli* (15) and OprM in

Pseudomonas aeruginosa (16, 17) in association with other inner membrane proteins (Figure 1-3). However, the complete structures of both complexes are still currently unknown.

Each porin has different characteristics to specify its substrate. They can be generally classified according to their activity (specific and non-specific pores) and their functional structure (monomeric or trimeric) (14). Based on the protein activity, the general porins are expressed at high levels while the more specific ones (e.g. PhoE and LamB) are strongly induced when needed (18). The size of porins is determined by the number of transmembrane β-strands which start at 8 strands (e.g. OmpX) and run up to 22 (e.g. iron transporters, FhuA and FepA) (3).

The first OMPs structurally characterized are referred to as classical porins, OmpK36, OmpF, PhoE, LamB, and ScrY (19). All form highly stable trimers with 16 or 18 transmembrane β-strands in each subunit (19). Recently, more trimeric structures have been solved such as OmpC (20), OprP (12), and PorB (21). In general, the classical porins functionally exist as either monomers or trimers. Examples of recently solved structures of monomeric porins are OprD (22) and OpdK (5). Most classical porins share similar feature where the extracellular loop L3 (or constriction loop) is folded into the lumen serving as the constriction site for substrate recognition (a yellow loop in Figure 1-3).

The remainder have different features to conduct a substrate molecule. Unlike most of classical porins, monomeric OmpA (23), NanC (19), and OmpG (24) composed of 8, 12, and 14 strands, respectively have no long loop folded into the middle of the pore (Figure 1-3). Also, the long-chain fatty acid transporter FadL (13) comprising of 14 β-strands contains a central hatch domain where L3 and L4

are instead located at the extracellular side to facilitate fatty acid translocation (Figure 1-3).

Various patterns of interior structures among porins represent different mechanisms of their substrate translocation and selectivity. To better understand how such porins facilitate their substrate transport, many experimental and computational studies have been employed to date. In the next section, I will give the reader a bird's eye view of previous studies on porins.

1.3 Previous studies on porins

To date many studies of porins have been published. However, the literature review here is restricted to those porins that are studied in this thesis. Overall, I will start with the classical porins and sequentially move on to more specific channels.

1.3.1 Outer membrane protein C and F (OmpC and OmpF)

The classical porins, OmpC and OmpF, in *E. coli* have been widely studied since 1982. The cation-selective properties of both porins were firstly found by Nikaido *et al.* (25), where such properties were observed by measuring the diffusion rate of antibiotics, cephalosporins, through both pores. A batch of simulation studies of OmpF was initiated when its high resolution X-ray structure was solved (26). The electrostatic interactions of OmpF were compared with those of PhoE via continuum electrostatics calculations (27). It was observed that the pore sizes of both OmpF and PhoE were controlled by pH-dependent interactions between two charged clusters at the opposite sides in the constriction area (27). The continuum electrostatic models showed that OmpF had no ion preference, whereas PhoE was anion-preferable (27) where this contradicted the cation-selective properties of OmpF found in an earlier experimental study (28).

Previous studies of OmpF suggested a role of the L3 loop in gating activity and L2 in trimeric stability (29-31). The key residues for gating and structure stability were identified both computationally and experimentally. Further details can be seen in Chapter 3. The first X-ray structure of OmpF motivated many further experimental studies on ion selectivity of its homologue OmpC. Since the OmpC sequence shares 60% identity to OmpF, key residues for ion selectivity and structural stability in OmpF are conserved in OmpC (32). However, OmpC was observed to be slightly more cation-selective than OmpF due to a smaller pore size (32). Additionally, the oppositely charged clusters in the constriction site of both porins similarly generate a strong transversal electric field that facilitates ion translocation (27, 33), especially negatively charged residues that were found to be responsible for the cation preference (34). Later, two well-separated specific ion pathways with screw-like fashion in OmpF were visualized by Brownian dynamics (BD) simulations (35). The BD simulations suggested that charge asymmetry in the constriction site gave rise to separate permeation pathways for cations and anions (36, 37). Cations were found to follow the negatively charged track along the OmpF pore, while a stream of anions was at the opposite side of the pore (35-38). This prediction has recently been verified by X-ray diffraction (39). To better understand ionic selectivity, both single-channel recording and free energy calculations of OmpF were performed (40). The conductance of monovalent metal ions through OmpF was experimentally found to be proportional to the cation mobility in bulk aqueous solution (40). The free energy calculations also showed that the affinity of alkali metal binding was increased with their atomic radii, which agreed very well with experiment (40). To date, not many studies have focussed on the ion selectivity of OmpC. Some studies of OmpC revealed similar pH-dependent gating and selectivity to that found in OmpF (34, 41-44). Moreover, the pH significantly

affects the assignment of ionization states of interior charged residues in both pores. To date, the ionization states of some pore-lining charged residues are still under debate (27, 29, 30, 45, 46). The X-ray structure of OmpF has been solved for 20 years, but aspect of the mechanism of ion permeation remain unclear. The understanding of gating and selective mechanisms of OmpF is a big challenge. Any insight will also serve as a guideline to explore the behaviour of its homologue OmpC. Recently, the X-ray structure of OmpC has been solved (20). This structure does not reveal any features that substantially differ from OmpF, therefore, the question about the key properties that distinguish OmpC from OmpF remains largely open.

Understanding the permeation process of nutrients and antibiotics via porins is also important for biotechnology and medical applications. Therefore, the trend of OmpF and OmpC research is nowadays directed towards medical and technology aspects. The majority of current studies still concentrate on the well-studied OmpF. Within the medical applications, non-specific OmpF was found to be an entryway for a number of antibiotics, especially β-lactam compounds (47). The antibiotics are classified based on mechanism of action and chemical structure, which can mainly be classified into five different groups, β-lactam (penicillins), cephalosporins, polymixins, quinolones, and sulfonamides (6). Single channel experiments showed that both OmpF and OmpC displayed the ability to facilitate the uptake of two classes of antibiotics, fluoroquinolones and cephalosporins (48) where a zwitterionic drug was found to diffuse through the pore more rapidly compared to other charged compounds (14, 32, 49). Such experimental studies have been complemented with analysis of computational approaches (48, 50-53). For example, computational methods have been used to reveal detailed interactions between drug

and pore-lining residues (49, 50, 54-56), studies which are relevant for future drug design.

Apart from being an entryway for drugs, OmpF was also found to serve as both receptor and translocator for antibacterial protein (the colicins) through the OM (57-62). These colicins display a great ability to kill *E. coli* or related bacteria by forming a channel in the inner membrane (63-65). The interface between OmpF and lipopolysaccharide (LPS), which is the major lipid of the outer leaflet interface was found to serve as a receptor for colicins (65). Recently, the crystal structure of OmpF with the N-terminal segment of colicin present in the constriction site was solved (66). This crystal structure will further enable both computational and experimental studies to elucidate the passage process of colicins from the extracellular environment to the inner membrane.

In term of biotechnology, the non-specific properties of OmpF have attracted the attention of many researchers. With the known 3D structure and advances in site-directed mutagenesis, the pore properties of OmpF have been widely engineered to fabricate new selective properties. A pioneering attempt was made to transmutate OmpF into the Ca^{2+}-selective channels by mimicking the fixed charge and volume of the Ca^{2+} channel (67). Besides, the mutation studies showed the potential ability of modified OmpF as a future nanodevice (59, 68, 69). OmpF was also used as a test case to fabricate novel sensing techniques (70-73). Moreover, the insertion of OmpF into different types of artificial membranes has been widely studied in order to achieve cellular surface mimicry and applications in coating bioimplants for enhanced biocompatibility (70, 74-76). To date, the non-specific and well-studied OmpF appears to continuously have wide applications in bionanotechnology.

1.3.2 Outer membrane protein P (OprP)

Under conditions of phosphate limitation, *P. aeruginosa* was found to show induction of a phosphate-specific trimeric OprP (77). Initially, the affinity of OprP to conduct anions especially phosphate ion was compared with that of PhoE from *E. coli*. OprP was found to be phosphate-selective, while the larger PhoE pore has no selectivity for phosphate ions (78). It was shown that OprP has a binding site for phosphate with an affinity at least 100-500-fold higher than the affinity for chloride or other anions (79-81). The ability to select phosphate ion was suggested to be due to the constricted channel containing a phosphate-binding site (79). The phosphate-binding site was thought to comprise a lysine cluster inside the OprP pore (79, 82). Chemical modification of all interior lysines suggested their involvement in this phosphate selectivity (79, 82.). Consequentially, the key lysines for ion selectivity were then identified by site-directed mutagenesis (83). To date, the positively charged architecture and phosphate-binding site have recently been identified by the crystal structure of OprP (12). The 3D structure revealed that the binding site is formed by arginines and polar residues on the L3 constriction loop at the middle of the pore where a series of lysines was located in the periplasmic chamber. From this X-ray structure, the detailed mechanisms of phosphate permeation and the role of the lysine cluster will be studied computationally in Chapter 4.

1.3.3 Outer membrane protein OpdK

OpdK is one of the substrate-specific porins of *P. aeruginosa*. This protein belongs to the OprD family (84). In the OprD family, phylogenetic analysis revealed two subfamilies, the OprD group and the OpdK group (5, 84). Members of the OpdK subfamily have roles in the transport of organic carbon sources, while the OprD group is responsible for passage of amino acids (84). The OpdK pore was found to facilitate permeation of vanillate (4-hydroxy-3-methoxy benzoate) and related aromatic acids (5). Recently, the 3D structures of members of each subfamily, OprD and OpdK, have been solved. They share similar features in their constriction sites. Unlike general porins, the constriction site of both OprD and OpdK is formed by two extracellular loops (5, 22). OprD is selective for basic amino acids, while OpdK with a combination of polar and negatively charged regions prefers an aromatic vanillate molecule (5). To date, there is a lack of biophysical studies into the mechanism of substrate selectivity of OpdK and OprD. The more complete crystal structure of OpdK not only establishes the general features of the OprD family, but also provides a starting point for further studies. In Chapter 6, the translocation of vanillate through OpdK will be observed at atomic level.

2 THEORY AND METHODS

In the previous chapter, I provided the biological background to my thesis studies of Gram-negative outer membrane proteins. The main aim of this thesis is to gain an insight into the uptake process of small molecules through the bacterial cell membrane via porins. Thus, MD simulations are an ideal technique because not only do they include a fully atomistic description of protein, solvent, and membrane, but also they can reveal functionally relevant motions. Furthermore, computational approaches also help us to understand the specificity and selectivity of porins for their substrates. In this chapter I describe the theoretical basis of MD simulations, and their extension to steered MD (SMD) and the calculations of potential of mean force (PMF). I also summarize the use of Poisson-Boltzmann calculation to compute the electrostatic energy profiles of the ion translocation through porins.

2.1 Molecular dynamics (MD) simulations

The MD simulations described in this thesis are based on the classical approximation i.e. using molecular mechanics forcefields. A discussion of these methods can be found in many textbooks e.g. that by Leach, A which is followed in this thesis (85).

In MD simulations, the dynamics of atoms and molecules in a simulation box or system can be explained by the classical mechanics based on Newtonian's mechanics. Interactions within a system simply depend on the processes of stretching bonds, opening and closing angles, rotating bonds, and implicit electrostatics (85). These molecular properties can be represented in term of the potential energy function.

2.1.1 Potential function and parameters

The overview of potential energy and interaction parameters used in MD simulations has been given. In standard MD simulations, each atom is treated as a spherical space with a point charge and the potential energy of the system as a function of positions of all particles in a system. MD simulations are based on the intra- and inter-molecular forces within a system. The set of potential functions exploited in each simulation, as well as their associated parameters, is referred to as the forcefield. The minimal equation to present the potential energy (U) of the system is,

$$U(r) = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{torsions} k_\phi (\cos(n\phi - \delta) + 1)$$
$$+ \sum_{\substack{nonbonded \\ pairs}} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{r_{ij}} \right] \qquad (2.1)$$

The first three summations are over bonds, angles, and torsions where b_0 and θ_0 are the bond and angle at equilibrium, k_b, k_θ, k_ϕ represent the force constant for the bond length, bond angle, and torsion angle terms, respectively. These interactions are pictured in Figure 2-1. The torsion term also includes "improper" torsions, where four atoms defining an angle do not connect to each other by covalent bonds (Figure 2-1). The final sum (over pairs of atoms i and j) describes the combination of VdW and electrostatic terms. The q_i and q_j are the partial atomic charges on atom i and j that interact via Coulomb's law and r_{ij} represents the distance between atom i and j. The VdW term is represented by Lennard-Jones potential function.

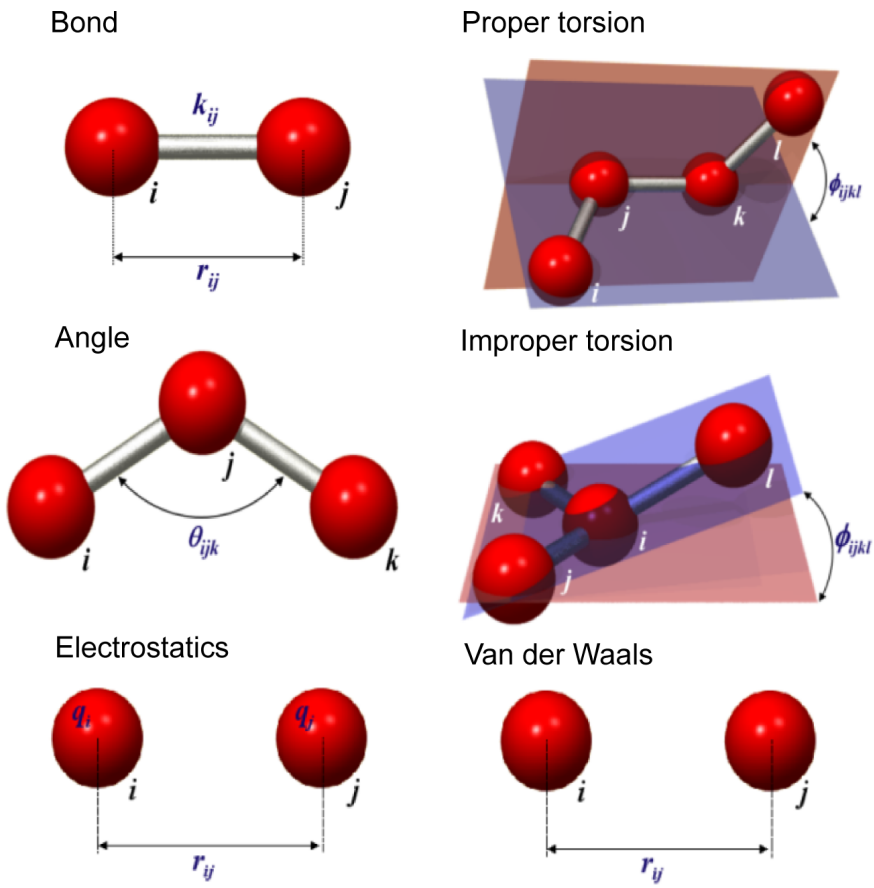


Figure 2-1 View of forcefield interactions. Atoms and bonds are colored in red and grey, respectively. This figure is adapted from www.pumma.nl/index.php/Theory/Potentials.

For the first term, the bond term can simply be defined by Hooke's law where energy varies with the square of displacement from reference bond length b_0 (equation 2.1). This reference bond length can be called as the "equilibrium" bond length, which is the value that is adopted in a minimum energy structure (85). The angle term is frequently described by Hooke's law as found in the bond term.

The bond and angle terms are often regarded as 'hard' degree of freedom (85). Most of the variation in structure and relative energies is due to the interplay between the change in torsional and non-bonded contributions (85). The barrier to rotate chemical bonds is a fundamental key to understand the structural properties and conformational change of molecules. Especially for a flexible molecule, the major conformational change is due to the bond rotations. Therefore, it is essential

to include the torsion term in the forcefield where the torsion potentials are expressed as a cosine series expansion in equation 2.1.

In case of VdW interactions, these interactions are formed by either atomic repulsion owning to overlapping of electronic clouds at short inter-atomic distance (Pauli repulsion) or long-range inter-attraction due to the instantaneous dipole interaction (London dispersion). The VdW interactions are often treated with the Lennard-Jones 12-6 function, which is the combination of an attractive long-range term and a repulsive short-range one. The attractive contribution is referred to as the dispersion forces (London forces) formed by the fluctuation of electromagnetic dipole moment in an atom. This fluctuation can induce a dipole in adjacent atom resulting in a r^{-6} attractive term in the potential (equation 2-1). The short-range repulsion is understood in term of Pauli principle where two electrons in a system are forbidden to have the same set of quantum numbers (85). This short-range replusive r^{-12} term aims to prohibit pairs of electrons to occupy the same region of space (i.e. the internuclear region).

For electrostatics and VdW, both of them are derived differently. The electrostatic interaction is formed between molecular or atomic distributions of charges where these charges are represented as a set of fractional point charges. Each of them is centered on an atom in a molecule. Because the partial charge of a given atom type is fixed throughout the simulations, the polarization effects are not counted here in biological forcefields. Nevertheless, many attempts are made to include the polarizability in the classical MD. Inducible dipoles alongside the point charge and partial charges, which can be varied during the simulations, are some of the methods used in order to observe the polarization effects in classical MD (86).

The non-bonded interactions, especially the electrostatic term, appear to be computationally expensive. Therefore, a cut-off for non-bonded interactions can be applied where only particles in a certain distance of each other are counted. The cut-off distance is given so as to include as many interactions as possible with an affordable computational cost. In term of electrostatic interactions, groups of atoms with an integer charge groups are normally treated together in order to avoid the creation of charges from the cut-off leading to the mal-behaviour of electrostatic function at the distance near a cut-off radius. Also, if the nearest image distances between the centers of two charged groups are within the cut-off radius, all non-bonded interactions between these two groups are then calculated. Also, an error due to the truncation of interactions can be reduced by applying the smoothing functions to the potentials. These smoothing functions aim to replace the truncated forces by other forces that continuous and have continuous derivatives at the cut-off radius. Replacing such forces can give smaller error from the time-step integration. Besides, the computational cost can be further reduced by the combination of cut-off methods and the concept of neighbouring list where this list helps to identify which pairs of atoms should be considered in the calculation. A list contains all interaction pairs located in a cut-off distance that is updated as the simulation progresses. The accuracy can also be increased by twin-range cut-off where interaction pairs that are out of the normal cut-off distance but do fall within larger cut-off can also be counted during neighbouring searching.

To treat electrostatic forces in MD simulations, the most common method used is the so-called the Particle Mesh Ewald algorithm (PME) (87). This method is based on the Ewald summation, which was originally used to compute the long-range electrostatic energy of crystalline system without truncating interactions (85). PME significantly improves the performance of the Ewald scheme and constitutes a

rigorous alternative to cut-off calculations for medium to large simulations which implements periodic boundary conditions.

Both VdW and electrostatic energies are calculated between pairs of atoms. These forcefields vary in their application and method of derivation. Transferring parameters from one forcefield to another one is difficult. Newly constructed forcefields should be extensively tested for accuracy in the same manner as an original or existing set was tested. Parameters in each forcefield should be tested and improved so as to gain accurate results. To date, a number of forcefields in the context of biomolecular systems and membrane proteins are in use such as CHARMM (88), OPLS (89), AMBER (90), and, as in use in this thesis, GROMOS (91, 92).

Advances in quantum mechanics also play a key role in the development of forcefields for classical MD simulations. Both *ab initio* quantum calculations and semi-empirical methods are successfully used to derive and optimize the parameterization of non-bonded potentials. Other parameters are extracted from the experiment. For small molecules, the geometric potentials can be obtained from microwave and electron diffraction whereas the bond potentials for larger molecules can be obtained from the X-ray diffraction. Similarly, vibrational data from infrared or Raman spectroscopy can give information on the rotation barriers leading to the force constants.

In this thesis, the general potential function is used (equation 2-1). The electrostatic term via Coulomb's law and Lennard-Jones potential are satisfactory without additional terms (as in the GROMOS forcefield).

2.1.2 Calculating MD trajectories

In principle the exact trajectories of every atom in the system could be obtained by integrating the equations of motion derived from Newton's laws if initial atomic coordinates, velocities, and interactions in the system were known. Nevertheless, such calculations in many-body system are not possible in practice. It is because a particle simultaneously impacts and depends on the motion of adjacent particles. Apparently, the equations of motion that describe the dynamics of the system are coupled.

To solve this problem, the concept of finite difference techniques is applied to generate the MD trajectories. The main idea is to break down an integral into small steps separated by finite time intervals Δt. The time intervals must be shorter than the time of the most rapid motions occurring in the simulation in order to avoid unrealistic overlap between two atoms. The total force (f) on each particle at time t can be computed by the summation of all interactions with other particles. From the force the accelerations (a), velocities (v), and positions (x) at time t are then determined. The combination of these values at time t is further used to calculate the new set of v and x at time $t+\Delta t$ where the force is assumed to be constant during the time step. To produce the MD trajectories, initial coordinates and velocities of all particles in the system are required. The coordinates are usually obtained from X-ray or NMR structures. The initial velocities are normally assigned according to a Maxwell-Boltzmann distribution where the probability of having an atom of mass m with velocity v at temperature T is given. The k_B is the Boltzmann constant:

$$P(v) = \sqrt{\frac{m}{2\pi k_B T}} \cdot \exp\left(-\frac{mv^2}{2k_B T}\right)$$

To actually obtain a trajectory, the differential forms of Newton's equation of motion are solved with respect to time for every atom.

$$\frac{\delta^2 r}{\delta t^2} = \frac{F}{m} \qquad \frac{\delta r}{\delta t} = at + v_t \qquad r = v_t t + \frac{1}{2} a t^2 + r_t$$

This set corresponds to the truncated Taylor series for displacement. The algorithms used for the numerical integration over finite time intervals Δt can be seen below:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} a(t)\Delta t^2 + \ldots$$
$$v(t + \Delta t) = v(t) + a(t)\Delta t + \frac{1}{2} b(t)\Delta t^2 + \ldots$$
$$a(t + \Delta t) = a(t) + b(t)\Delta t + \frac{1}{2} c(t)\Delta t^2 + \ldots$$

where *b* and *c* are the third and fourth derivative, and so on.

The most widely used method for integrating the equations of motion in MD simulations is the Verlet algorithm (85). This integrator is based on the two Taylor expansions, a forward expansion *(t+Δt)* and a backward expansion *(t-Δt)*:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} a(t)\Delta t^2 + \ldots$$
$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{1}{2} a(t)\Delta t^2 - \ldots$$

The sum of these two expansions yields an equation for propagating the position. Only force-derived acceleration and positions at the current and previous time step are used as an input:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2$$

To improve the velocity evaluation and accuracy, modifications of this algorithm have been introduced. For instance, the leap-frog algorithm, so named because the velocities are computed at the mid-point of the position evaluation, and vice versa. The new acceleration and mid-step velocity can be computed by the current acceleration and previous mid-step velocity:

$$r(t+\Delta t) = r(t) + v(t+\frac{\Delta t}{2})\Delta t$$
$$v(t+\frac{\Delta t}{2}) = v(t-\frac{\Delta t}{2}) + a(t)\Delta t$$
$$v(t) = \frac{1}{2}\left[v(t+\frac{\Delta t}{2}) + v(t-\frac{\Delta t}{2})\right]$$

where velocity at time *t* can also be obtained. Comparing with the standard Verlet algorithm, the leap-frog algorithm has the advantage of explicitly including the velocity, which can be further used to calculate the kinetic properties. All simulations in this thesis, performed by GROMACS, employ this leap-frog algorithm.

2.1.3 Constraints, boundaries, and ensembles

As mentioned before, the time step Δt is an important factor to obtain reasonable MD results. Too large time step will produce errors into the trajectories due to rough approximation of constant acceleration. On the other hand, using infinitesimal Δt generates fewer errors but it is computationally expensive. In general, the time step in MD simulations is dictated by the highest frequency motion (usually bond vibrations) present in the system (85). Typically, for biological systems, to observe all classical phenomenon, the suitable time interval is from 0.5 fs to 2 fs.

Certain methods have been developed to increase this time step and extend the length of trajectories for a given simulation time. These methods are based on a constraint principle where the individual internal coordinates are fixed during the simulation without influencing other internal degree of freedom (85). The most frequently used methods for applying constraints, particularly in simulations of biological systems are LINCS (93), SETTLE (94), and SHAKE (95). SETTLE is an analytical method that constrains the internal degree of freedom of small molecules. SHAKE is more general and based on an iterative procedure. By SHAKE, unconstrained coordinates are set sequentially to correct values, according to a list of distance constraints. The iterative treatment is limited by a short time step (~1 fs). In comparison, LINCS is non-iterative and allows time steps of up to 3 fs which makes LINCS more efficient than SHAKE. Also, LINCS can perform 3-4 times faster and is more easily parallelizable.

Along with the corrections for constraints, methods to maintain a system in a particular thermodynamic ensemble are also implemented. In non-biological system, molecular dynamics are performed in the constant microcanonical ensemble (NVE) (85). This ensemble allows free dynamics without rescaling of velocities and only need is the volume of the system. In this sense, most simulations use periodic boundary conditions to represent the simulation cell in all direction by replica images. Under the periodic boundary condition, particles crossing the cell returns on the opposite side. It means that the number of particles in the unitcell remains constant. Periodic boundary condition can help to remove an error from unwanted boundaries in an isolated system. Importantly, the distance between the box edges and the molecule of interest has to be seriously determined to avoid unwanted periodic correlations. So, the “minimum image convention” is introduced to define

the minimum separation between any two images of these particles. The cut-off must also be adhered in order to avoid duplicated interactions.

For biomolecular systems, different ensembles are used. The most common alternative ensembles are constant *NVT* and constant *NPT* ensembles. The constant temperature simulation is desired to determine the behaviour of the system changes with temperature such as the unfolding of a protein (85). If the specific temperature is known, the constant pressure condition is more appropriate. Many experimental measurements are conducted under the conditions of constant temperature and pressure so the isothermal-isobaric ensemble are the most relevant to experiment data (85). To control either temperature or pressure, the scaling functions are introduced. The desired temperature and pressure can be maintained by the thermostat and barostat developed by Berendsen (96). The system of interest is coupled to the external bath with fixed temperature or pressure (96). The corrections in temperature and pressure are obtained by the scaling factor derived from the difference between the reference (external bath) and the system. When the thermodynamic properties are of interest, the more sophisticated methods are introduced in order to observe changes in pressure or temperature. The most commonly used are the Nosé-Hoover scheme for thermostat (97, 98) and the barostat proposed by Anderson (99, 100). These methods allow the fluctuation of both temperature and pressure on the average value. The fictitious degrees of freedom are introduced to the equation of motion so as to result in the appropriate ensemble.

2.1.4 Energy minimization

Energy minimization (EM) is an essential component of simulations. Based on a potential function, it is used to find the low-energy structure via numerical methods. EM does not provide the time-dependent trajectory, but it can be used to prepare the starting coordinates of the system of interest for further MD simulations. This can adjust protein conformation and remove overlaps between protein and solvent or ligand etc.

EM generally optimizes the atomic positions in the Cartesian coordinates. The most commonly used methods are steepest descents and conjugate gradients. Both of them are based on the first derivative of the potential energy function (85) where the coordinates of atoms are gradually changed and moved towards the minimum point via an iterative approach. The steepest descent uses a line-search procedure to identify the minimum point in the function along a line through the multi-dimensional space where the current step is independent from the previous one. This method is efficient for the initial configuration that is far away from its energy minimum. All EM in this thesis employs the steepest descent methods since it is ideal for removing bad contacts among components in the system. On the other hand, the conjugate gradient method exploits the combination of direction used in the previous step and the newly computed gradient of the potential energy. This method is greatly efficient when the system is in the proximity of the energy minima.

2.2 MD techniques to study solute translocation

Experimentally, the ion permeation can be detected as electrical current flow, where the ion flux is the electric current across the membrane. The magnitude of current is then converted to the rate of passage of ions through one open channel (101). In general, the open channel can pass at least 1 pA of current corresponding to the number of 10^6 monovalent ions per second (101). Although an experiment offers an insight into the ion permeation and also selectivity of channels, the mechanistic details cannot be probed in atomic level. MD simulations therefore play a crucial role in the elucidation of transport mechanism of channels. Add-ons to MD simulations have been applied to a system in order to mimic a membrane environment. One of them is an applied electric field across a membrane, which mimics a voltage clamp experiment. Recently, the long simulations with an applied electric field have been successfully revealed the detailed mechanism of K^+ ion permeation through K channel (102). Not only have the atomic-resolution observations of ion conduction been captured, but also the gating mechanism (102). In addition, an external electric field is widely used to study ion permeation through other protein channels and DNA threading through a nanopore (103-110). I therefore employed this technique to study vanillate transport through the OpdK pore in Chapter 6.

2.2.1 Steered Molecular dynamics (SMD) simulations

Due to the limitations of the MD timescale, the rare events are difficult to be observed during the conventional MD run. For example, suppose we want to observe the full transport of ions through OmpF. In experiment, the estimated permeation rate is ~ 10^8 ions per second (111). Therefore, equilibrium simulations of 1-10 ns duration are too short to obtain information on such transport processes.

The more advanced techniques have to be applied to increase the probability of seeing unlikely configuration by reducing the energy barrier. One of the most commonly used methods is the SMD (112). The idea is to apply an external force to manipulate the particle of interest. With this method, the entire permeation process can be captured by pulling the solute through the channel. SMD allows us not only to focus on the important configuration, but also to minimize the computational cost. SMD has been widely used to understand protein dynamics that influence folding and unfolding mechanism as well as solute permeation (113-116). Especially, SMD can be applied to study the permeation of solute with high degrees of freedom such as DNA (117). In this thesis, SMD technique is then used to understand the solute translocations through the OprP and OpdK pores in Chapter 4 and 6, respectively.

SMD is based on the harmonic oscillation of spring. For the solute transport, the virtual spring is then attached to the solute in order to apply a force to pull it later on. The pulling group is harmonically constrained with a force $F=-k(x-vt)$, where k is the spring constant, x is the coordinate of the pulling group, v is the velocity of the pulling group. The forces exerted on the pulling group are then recorded. Also, the force profile can be quantitatively used to identify the free energy landscape only if the pulling rate is so slow that still adheres the system close to quasi-equilibrium. Since SMD is the non-equilibrium process, to observe equilibrium properties it is important that either force or velocity used is small or slow enough to drive the system not too far away from equilibrium. In general, SMD can be performed under the conditions of either constant force or constant velocity, but only constant-velocity SMD is used here.

2.2.2 Free energy calculations

The free energy of molecular system is often considered to be the most important quantity in thermodynamics since it describes the tendencies of atoms or molecules to associate and react (85, 118). Hence, the ability to predict this quantity would be an important goal. However, the free energy is a difficult value to obtain. Many attempts have been tried to progress towards the goal in recent years. The attention is also drawn to the application of MD simulations. Nonetheless, the free energy cannot be determined by the standard MD due to the inadequate samples in some regions in phase space dominating the free energy. For example, MD always seeks out the lower energy regions and never adequately samples the region with higher energy (85). The advance methods to permit the system to explore all phase space are needed. To date, there are certain methods to calculate the free energy difference based on MD simulations such as thermodynamic integration (TI), and free energy perturbation (FP) which are commonly used to calculate the free energy difference between two states. Moreover, the free energy can also be compute as a function of reaction coordinates, named potential of mean force (PMF), which is very useful for studying the energetic properties of solute transport across a channel. In this thesis only PMF calculations are conducted in order to probe the free energy landscape of the channel interior due to the solute translocation.

2.2.2.1 Potential of Mean Force (PMF)

Calculation of a PMF is one of techniques to calculate the free energy changes as a function of reaction coordinates *R*. The PMF can be determined as

$$W(R) = -kT \ln P(R)$$

where *R* is the coordinate, and *P(R)* is the probability of having system at the coordinate *R, i.e.,*

$$P(R) = Q^{-1} \iint \delta[R'(r) - R] e^{-E(r,p)/kT} dqdp$$

$$P(R) = \frac{Q(R)}{Q}$$

where *Q* is the full partition function, δ is the Dirac delta function. The more simplified explanation is that the probability can be obtained from the ratio of the partition function at *R*, *Q(R)*, to the full one *Q*. The PMF of particular coordinate can be obtained as seen below,

$$W(R') = -kT \ln Q(R') = -kT \ln P(R') - kT \ln Q = -kT \ln P(R') + C$$

where *C* is the integration constant (free energy constant) and the PMF at R' is derived from the natural logarithm of the probability of seeing system at R'. The constant *C* is an important factor to iteratively find $P(R')$ in the next step of smoothing PMF function.

To overcome the sampling problem, one of the most commonly used methods is the umbrella sampling. The idea is to modify the potential function in order to sufficiently sample unfavourable states (85). The system is simulated in the presence of an artificial biasing window potential, $U(R,R')$. The biasing potential is used to confine the variations of the coordinate *R* within a small interval (window)

around some prescribed values (119). This helps to achieve an efficient sampling (119). The modification of the potential function can be written as a perturbation:

$$W^{biased}(R') = -kT\ln P^{biased}(R') + U(R,R') + C$$

where $U(R,R')$ is an umbrella potential in a harmonic function:

$$U(R') = \frac{1}{2}k(R - R')^2$$

To obtain the PMF over the entire reaction coordinates, the multiple biased simulations are needed to sample the different regions of *R*. Ultimately, the results from each window are unbiased and combined together to obtain the final unbiased PMF (119). The unbiased PMF at each particular region are obtained from the adjustment of the free energy constant, *C*. This can be efficiently determined by performing an algorithm called Weighted Histogram Analysis Method (WHAM). In practice, it is achieved through iteration procedure. The initial guess of the free energy constants are made to estimate an unbiased probability $P(R')$. The iteration process is repeated until the new free energy constants match the prescribed value and then the correct $W(R')$ is obtained. Carrying out multiple simulations will give several *W(R)* outputs. WHAM also serves to patch all PMF pieces to generate a single curve.

However, there remains some issues relating to convergence. The proper sampling of the phase space is needed to obtain the correct PMF. Also, the good overlapping partial PMF has to be considered as well because both inadequate sampling and bad overlapping partial PMF can lead to the unreliable PMF.

Free energy calculations, especially of PMFs, are very computationally expensive and time-consuming therefore some alternatives that are less expensive

such as continuum and implicit solvation models have been explored which will be discussed in the next section.

2.2.3 Electrostatic energy based on Poisson-Boltzmann theory

Electrostatics plays a key role in many biological procedures (120) where it often dominates both inter- and intra-molecular interactions (121). Many biological processes including both structural and dynamic properties of biomolecules are mainly driven by the set of electrostatic interactions. Therefore, in the charge-dominating system, only electrostatic properties may be enough to elucidate the whole biological process.

In general, MD simulations treat all solvent molecules explicitly, which costs a lot of computer power. Thus, a model of implicit solvent where solvent and ions are treated as a bulk continuum is proposed (121). This technique is efficient for the study of charge-dominating molecules or system such as DNA, RNA, and ion conductance (122-128). The most widely used method is the Poisson-Boltzmann (PB) model where it is used to determine the electrostatic contribution to solvation energy and binding energy of ligand (121).

2.2.3.1 Electrostatic energy and Poisson-Boltzmann (PB) calculation

For PB model, the ionic solution is treated as a continuous dielectric medium with average distribution of ionic species (121). The electrostatic potential (φ) related to the work of assembling a charge density (ρ) is obeyed the Poisson's law (121, 129):

$$\nabla \bullet \varepsilon(x)\nabla\varphi(x) = -4\pi[\rho^{f}(x) + \rho^{ion}(x)]$$

where ε is the dielectric constant. The charge density is the sum of the fixed charges of the solute (ρ^{f}) and mobile counterions and co-ions (ρ^{ion}) in the solvent. In

general, the dielectric constant of water solution is set to 80 and of biomolecule may vary between 2 to 12 (130, 131). The border between these two dielectric regions can be placed by the VdW surface, but it is more often to assume on the water-excluded molecular surface which is determined by the center of a probe sphere with a radius of 1.4 Å rolled over the VdW surface of the solute or biomolecule (121). The local ionic concentrations in the ion-accessible domain can be determined by the Boltzmann distribution depending only on the electrostatic potential (121).

The PB equation needs to be solved numerically. One of the recently developed programs is the adaptive Poisson-Boltzmann solver (APBS) where the linearized PB equation is solved numerically with the finite difference method (132). The electrostatic potentials and fixed charge densities are represented in a box on 3D grid (121). The values of dielectric constant and ion strength are then allocated to each grid point. The boundary between solute and solvent can be defined by either the molecular surface or the accessible surface as mentioned earlier (85). The grid points outside the surface are assigned the high dielectric constant (water ~ 80) and a value of ionic strength. The grid points within the surface are defined as a solute or macromolecule with low dielectric constant.

Compared with the more sophisticated method PMF, the PB calculation can dramatically reduce the CPU time and computer resources. Moreover, the PB calculation can also estimate the energy as a function of pathway in the term of Born energy. This value is the energy difference of moving a fixed charge from vacuum to the particular environment of interest such as solvent and inside the protein interior,

$$\Delta G_{born} = \frac{q^2}{8\pi a}\left(\frac{1}{\varepsilon_0} - \frac{1}{\varepsilon}\right)$$

where q is the ion charge, a is the ion radius, ε_0 is the dielectric constants in a vacuum, and ε is the dielectric constant of the region of interest (biomolecule), respectively. In this thesis a set of the Born energies is calculated to determine and compare the electrostatic properties of the pore interiors of OmpC and OmpF in Chapter 3.

Although this method is computationally inexpensive and efficient for a charge-dominating system, it still has some limitations and must be applied with care. It is due to the fact that this approach appears to fail under the condition of highly charged system (121). The advanced techniques are needed to account for the improved electrostatic correlation and more precise representation of exclusion volume term (121). Besides, when the system is not solely dependent on the electrostatics, the Born energy may not be sufficient to represent the energetic properties of the entire system. For example, in the system with few charges, only the electrostatic potential cannot be used to explain the system behaviour since the entropic effect is also dominated. Therefore, the free energy is still important.

3 SIMULATIONS OF OMPF AND OMPC FRATERNAL TWIN

3.1 Introduction

Porins are water-filled pore-forming proteins that reside in the outer membrane of Gram-negative bacteria. They control the influx and efflux of nutrients and metabolites (32, 133). Most porins are homotrimers and are extremely stable in a wide range of pH, organic solvents, chaotropic salts, and detergents (31). The most conspicuous structural aspect of the trimeric porins is the presence of the constriction region or 'eyelet' located within the hourglass-shaped channel which is lined with charged residues. This feature is conserved in general diffusion porins and is found to be important physiologically (134). One of the most studied porins is OmpF from *E. coli*. It consists of 16 antiparallel β-strands connected by turns at the periplasmic side and loops at the extracellular side (26, 36, 135-138). One of the loops (L3) is folded into the lumen in each monomer acting as a constriction loop (Figure 3-1A and B). Moreover, there is the so-called latching loop (L2) reaching over to the neighbouring monomers (Figure 3-1A and B). These structural features are also conserved in the recently solved crystal structure of OmpC (20) whose amino acid sequence is 60% identical to that of OmpF (Figure 3-2A).

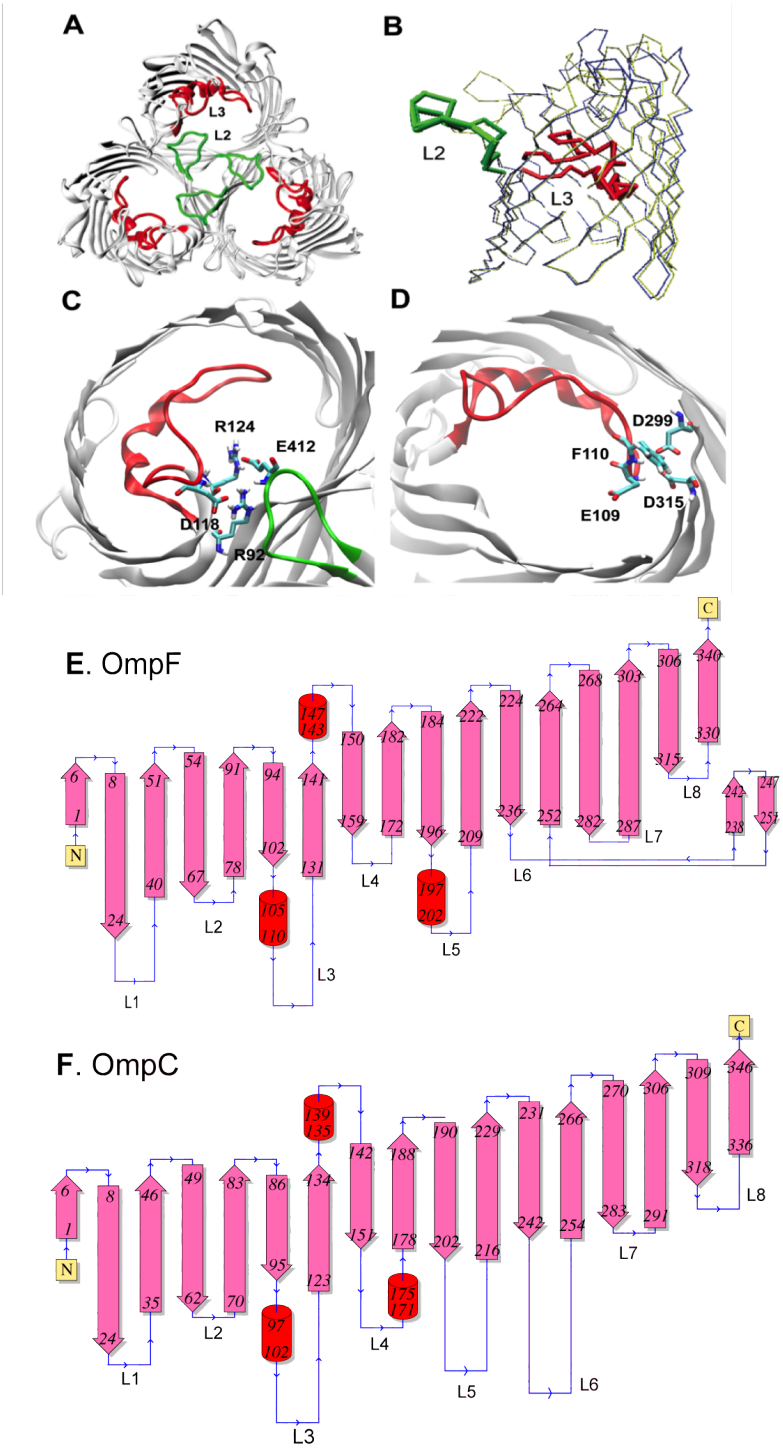
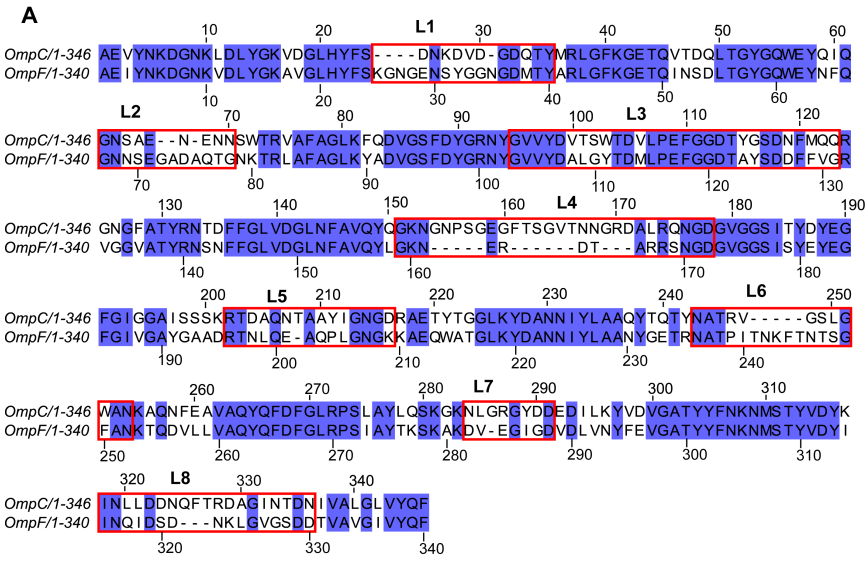


Figure 3-1 Top view of trimeric OmpC showing the latching loop (L2) and constriction loop (L3) (**A**). (**B**) Side-viewed superimposition of OmpC (yellow) and OmpF (blue) monomer. L2 and L3 are colored in green and red, respectively. Detailed views of the amino acid residues participating in the interactions (**C**) between the latching loop (L2) and its neighbouring monomer in OmpC and (**D**) between L3 and the adjacent barrel wall. Topologies of both OmpF and OmpC (adopted from PDBsum (www.pdbsum.com)) are shown in (**E**) and (**F**). The β-barrel and helix are colored in pink and red respectively.

Experimental studies of OmpF (31, 42, 138, 139) showed that the salt-bridge, E411-R100 in L2, was involved in trimer stability (the locations of these residues are shown in Figure 3-1C). Moreover, E411 was found to form hydrogen bonds with R132 on the channel wall of the adjacent subunit (31). Several OmpF experiments also revealed that D312 and E296 on the opposite barrel wall play a role in the control of L3 flexibility (20, 27, 135, 139). The nature of its homologue, OmpC, has also been investigated in comparison with OmpF. Conserved gating residues in OmpC are found to play a similar role as those of OmpF. Studies of the D315A mutant of OmpC (homologous to D312 in OmpF) indicated a considerable change in gating activity from the wild-type (139). Although some previous studies showed that OmpF and OmpC share similar structural and functional properties, the massive increase in the proportion of OmpC in extreme environments (e.g. high temperature or toxic agents) rather than OmpF still attracts researcher's attentions. OmpC is experimentally found to be somewhat more cation selective than OmpF and its pore was known to be smaller (20, 32). Many experiments revealed the functional residues that are homologous to those of OmpF, but what makes OmpC more specific remains obscure. Therefore, I have conducted simulation studies to reveal the structural and functional properties of OmpC that differ from OmpF.



OmpC	R92	E109	F110	D118	R124	D299	D315	E66(412)
OmpF	R100	E117	F118	D126	R132	E296	D312	E71(411)

Figure 3-2 (**A**) Sequence alignment of OmpC and OmpF. The sequence identifiers are given. Each extracellular loop is shown in red boxes. (**B**) Nomenclature of homologous residues in OmpC and OmpF.

Molecular dynamics (MD) simulations have been used in a number of earlier studies to investigate the dynamic properties of OmpF. A 1-ns MD study of the trimeric OmpF revealed deviations of dynamical structure relative to the crystal structure. It also showed that L3 flexibility affected a change in pore cavity (30). Furthermore, MD simulations were used to study the influence of protonation states of titratable residues on the OmpF function (138). It appears that key aspects of the protein behaviour can be captured on the MD timescale. In this study, I therefore employ MD simulations not only to understand the pore activity and the dynamics of OmpC, but also to explore the properties that cause OmpC to be more specific than OmpF. Attention is also drawn to the effect of key residues on conformational stability of the two porins.

3.2 Methods

3.2.1 Simulation setup

The trimeric OmpC (2J1N, resolution 2.0 Å) and OmpF (2OMF, resolution 2.4 Å) crystal structures consisting of 346 and 340 amino acids in each monomer respectively were embedded in a pre-equilibrated dimyristoyl-phosphatidylcholine (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.* (140). This generated a cavity into which proteins were inserted. Water and counter ions were then added into each system. Ions were added by replacing selected water molecules. For each system, a 0.2 ns equilibration of protein restrained dynamics simulation was conducted with a force constant of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$, followed by a 10 ns unrestrained MD production run.

3.2.2 Simulation protocols

All simulations were performed using the GROMACS 3.1 simulation package (www.gromacs.org) (141) with an extended united atom version of the GROMOS96 forcefield (92). 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 (87) 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 (*NPT*) ensemble. The temperature of the protein, DMPC, solvent, and ions were each coupled separately using the Berendsen thermostat (96) 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 analyzed by GROMACS and locally written code. Molecular graphic images were prepared using VMD (142).

3.3 Results and Discussion

3.3.1 General features of OmpC and OmpF

Considering structural stability of both OmpC and OmpF, the time-dependent Cα RMSDs relative to the starting structure for each simulation are calculated. Based on this analysis, for both the intact proteins and for the β-barrels, OmpC appears to be more stable than OmpF (Figure 3-3A and B). A similar conclusion is arrived at if I compare the RMSDs for the backbone atoms of the monomers (Figure 3-3C and D). It is clear that the structures of each OmpC monomer are stable, while the OmpF monomers show a greater degree of structural flexibility (especially in monomer 3) (Figure 3-3D).

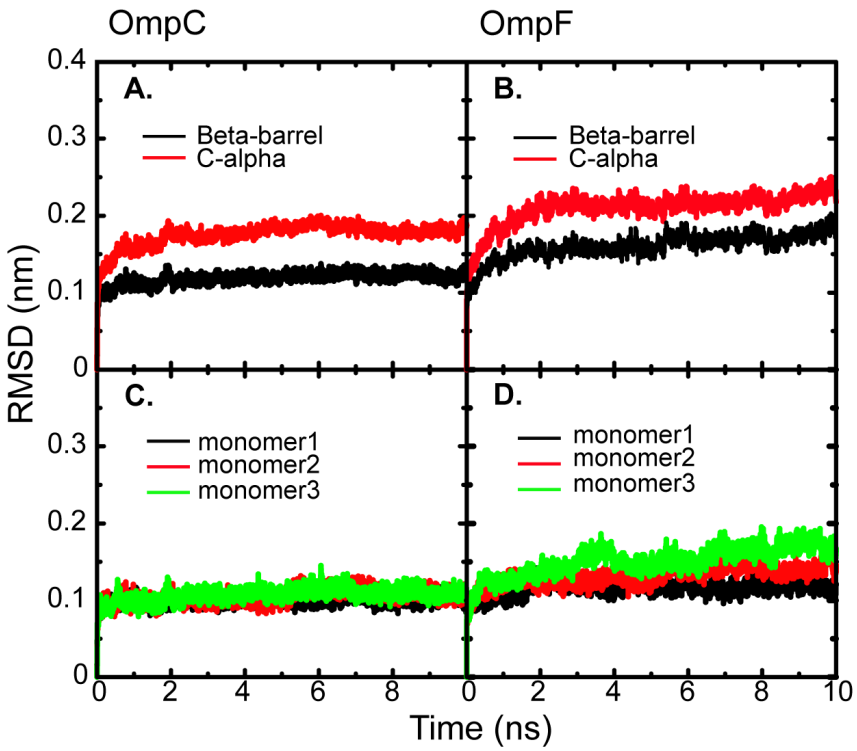


Figure 3-3 RMSDs of C-α atoms of all residues and its β-barrel of trimeric OmpC (**A**) and OmpF (**B**). RMSDs of C-α atoms of each monomer in OmpC and OmpF are shown in (**C**) and (**D**).

From the RMSDs of loop regions (Figure 3-4), it is evident that L2 and L3 do not change their conformations significantly on a 10 ns simulation timescale, which may be related to their functional properties. L1, L4, L6, L7, and L8 in both porins (Figure 3-1E and F and Figure 3-4) show different degrees of structural drift. The longer loops of L1 and L6 in OmpF, and L4 and L7 in OmpC give the structure much more flexibility (Figure 3-1E and F and Figure 3-4). Intriguingly, the L8 in both porins are almost in the same length (OmpC is 3 amino acids longer; see Figure 3-2A), but the RMSDs of L8 are obviously different (Figure 3-4). It was found in a crystal structure of OmpC that the membrane facing side of L8 is a Mg^{2+} binding site (20), which may form a crosslink to a head group of outer membrane. However, neither simulation contained Mg^{2+} ions. There are only Na^{+} and Cl^{-} ions to neutralize the systems. It is interesting that L8 in OmpC remains inflexible. This

restrained L8 may play a role in either structural stability or pore activity. Further studies are needed to elucidate the actual role of L8 in OmpC.

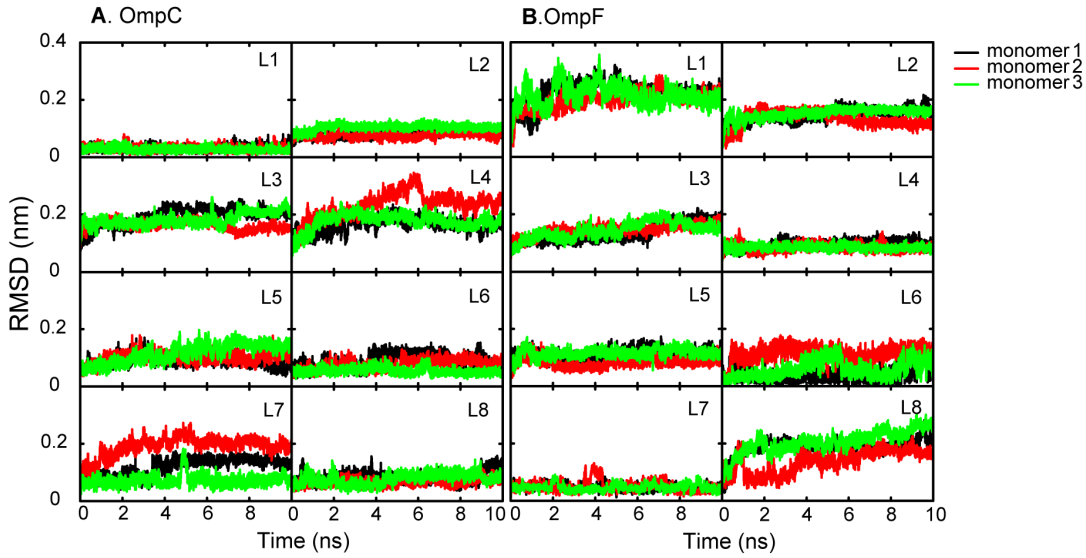


Figure 3-4 RMSD of C-α atoms of all residues in loop L1- L8 of each monomer in OmpC (**A**) and OmpF (**B**).

3.3.2 Interactions at L2 and L3 sites

It is found that one of the monomers in both proteins is in the closed or semi-closed state (Figure 3-5). The pore radii of both monomer 1 in OmpC and monomer 2 in OmpF are fluctuated between ~ 0.1 nm (closed) to 0.3 nm (open) where the rest stay open during the simulations (Figure 3-5). To efficiently compare two data sets, I focus only on monomers that are consistently in the open state (Figure 3-5; monomer 1 and 3 of OmpF and monomer 2 and 3 of OmpC). All numerical results shown in this section are averaged from the two pore-open monomers of each porin. The sequence alignment and the nomenclature of homologous residues in OmpC and OmpF that will be referred in the following sections are shown in Figure 3-2.

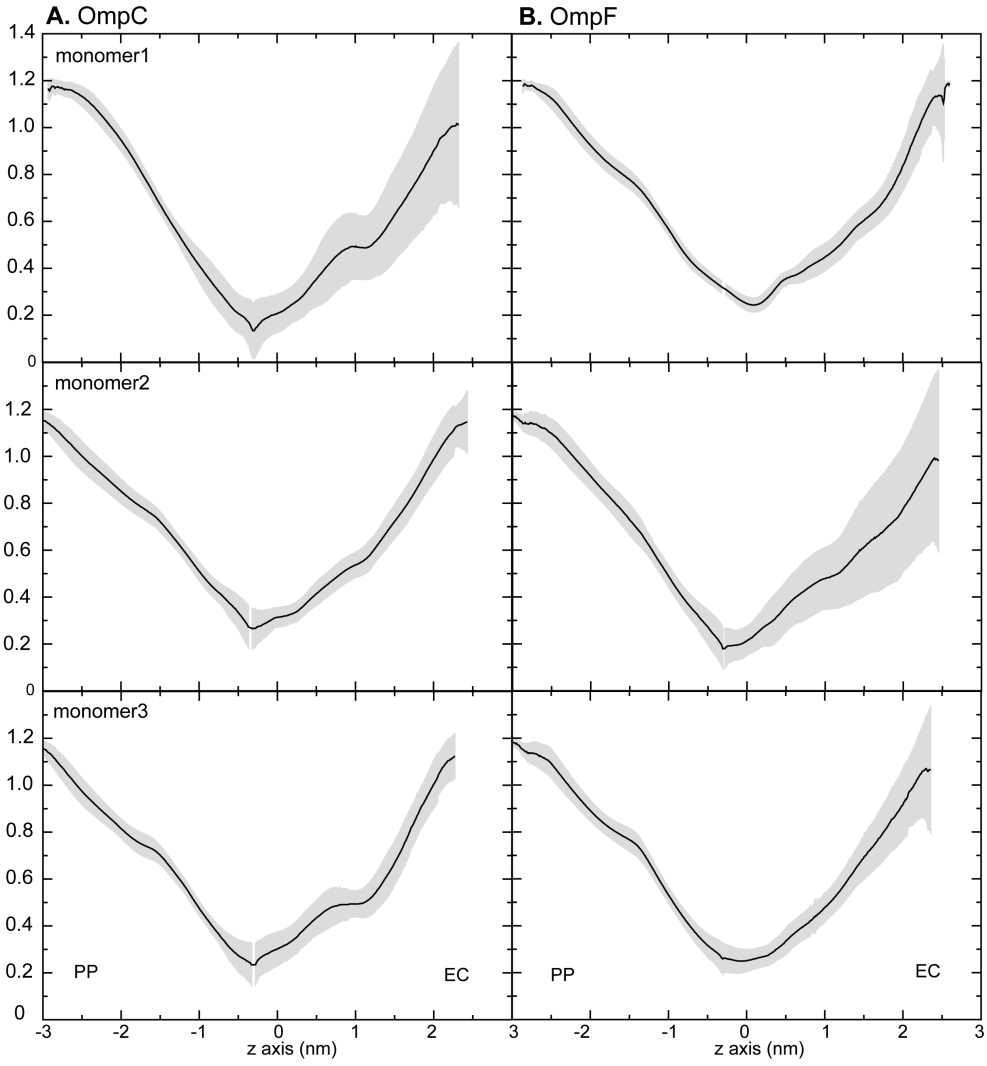


Figure 3-5 Average pore radii over 10 ns simulations of each monomer of OmpF (**A**) and OmpC (**B**)

3.3.2.1 The latching loop L2

The location of L2 and its neighbouring residues are shown in Figure 3-1C. Key interactions between a tip of L2 and an adjacent subunit involved in trimeric stability have been investigated by monitoring the distance between donor and acceptor residues and the possibility of the hydrogen bond formation during the simulations (Figure 3-6 and Table 3-1).

OmpC		OmpF	
Donor-Acceptor	H bond (%)	Donor-Acceptor	H bond (%)
R92-D118	58	R100-D126	71
R92-E412	8	R100-E411	44
R124-E412	64	R132-E411	75
E109-D299	0	E117-E296	0
E109-D315	83	E117-D312	70
F110-D299	0	F118-E296	0
F110-D315	81	F118-D312	86
E412-D118	67	E411-D126	66

Table 3-1 Percentage of one hydrogen bond occurring over 10 ns between two residues in L2 and L3 regions which are proposed to play a key role in the structural stability. Data are averaged out percentage of hydrogen bonds from 2 monomers.

In the OmpF studies, it was suggested that the salt-bridge E411-R100 at the tip of neighbouring L2 contributed to the trimeric stability significantly (31). Besides, this R100 might be involved in two salt bridges: (i) one with D126 of L3 contributing to the L3 flexibility and (ii) one with E411 on L2 of a neighbouring unit (27). Both R100-D126 and R100-E411 interactions can clearly be confirmed by both consistent distances between donor and acceptor and hydrogen bonds with high percentage of 44 and 71 (Figure 3-6C and D and Table 3-1). The homologous bonds, R92-D118 and R92-E412, are also conserved in OmpC although R92-E412 appears to be very transient (Table 3-1). The stronger salt bridge, R100(92)-D126(118), found in both porins implies the dominant contribution of R100(92) to the L3 flexibility over the trimeric stability. This finding is in a good agreement with the previous studies showing the significant role of R100-D126 in controlling the L3 flexibility (31, 139). Although there is no experimental evidence to support a

role of R92-D118 in OmpC, the conserved series of salt bridges found infers the similar role as seen in OmpF.

In addition to the hydrogen bonds with R100, E411(412) can also interact with D126(118) at the base of L3 (Table 3-1). To maintain the trimeric stability, E411(412) is involved in two salt bridges at the base of L3: [1] one with R100(92; weak) and [2] one with D126(118) in both porins.

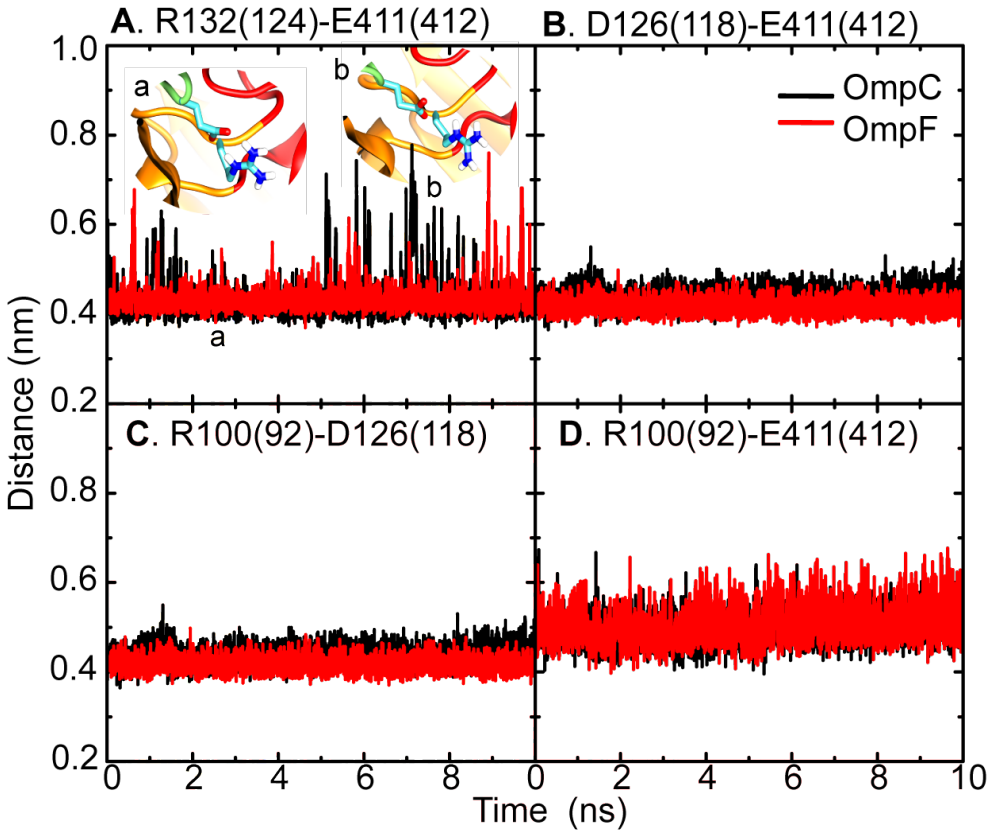


Figure 3-6 Distances between residues surrounding the tip of loop L2 of the adjacent unit. (**A**) Carbon atoms of the functional groups of R132 and E411 in OmpF (124 and 412 in OmpC) and (**B**) carboxylate carbon of D126 and H atom on the backbone of E411 in OmpF (118 and 412 in OmpC). (**C**) Carbon atoms of the functional groups of R100 and D126 in OmpF (R92 and D118 in OmpC) and (**D**) R100 and E411 in OmpF (92 and 412 in OmpC). OmpC is coloured in black and OmpF is in red. (**a**) and (**b**) represent the conformational changes of R132(124) and E411(412) during the simulation.

Apart from residues at the base of L3, the constriction residue, R132(124) surprisingly shows a contribution to the trimeric stability. The salt bridge, R132(124)-E411(412), is consistently observed in both porins (Table 3-1).

R132(124) is reported to be one of the gating residues, but its strong hydrogen bond with E411(412) (Table 3-1) highlights its participation in trimeric stability, especially in the open state. Not only does this finding support the previous studies showing the contribution of R132 to trimer stability in OmpF (31), but also the strong hydrogen bond R132-E411 (75% in Table 3-1) reveals the dominant role of R132 in structural stability rather than gating activity. Although R132 mainly forms a hydrogen bond with E411(412), it is flexible enough to swing back and forth to maintain the gating activity, which can be shown by the large fluctuations in the salt-bridge length in Figure 3-6A (see a and b for the sidechain orientations). Besides, the results reveal the analogous role of R124 in OmpC. The formation of strong R124-E412 bond (64% in Table 3-1) is observed here. Additionally, R124 sidechain flipping can also be seen and it seems to occur more frequently than R132 (Figure 3-6A). The similar change in sidechain orientation and the close contact with E411(412) suggest the role of the gating residue R132 and R124 in both porins are alike.

3.3.2.2 The constriction loop L3 and the adjacent barrel wall

Karshikoff *et al.* (27) proposed that D312 on the barrel wall plays a key role in L3 flexibility that leads to the control of pore size in OmpF. Similarly, OmpC also contains a homologous residue, D315. The analysis is performed in both porins to understand the role of D315 in comparison with known properties of D312. It appears that D312 and D315 share very similar function. Both of them form strong salt bridges with E117(109) and F118(110) at the tip of L3. It is known that D312(315)-E117(109) is crucial for controlling L3 flexibility that affects the change in pore size. The simulation results not only reveal this known interaction, but also reveal a permanent interaction between D312(315) and the backbone of F118(110)

(Table 3-1). Both D312(315)-E117(109) and D312(315)-F118(110) are found to be the major contribution to the L3-barrel wall adhesion in both porins.

On the other hand, neither E296 in OmpF nor its analogue D299 in OmpC is able to interact with its surrounding residues on the tip of L3 (Table 3-1). Also, the amino acid substitution from E to D at position 299 in OmpC does not affect the formation of hydrogen bonds in this case. This finding agrees with the previous experiment showing that E296(D299) is not involved in the interaction between L3 and the adjacent wall (27).

3.3.3 Electrostatic properties of the pore surface

Although both OmpC and OmpF share similar pore-lining architecture and interaction network, OmpC is experimentally found to be slightly more cation-selective than OmpF (32). A number of studies have been studied to understand the ion flow through OmpF, both experimentally and theoretically (36, 40, 143, 144). The free energy calculations showed that the affinity of monovalent cations to the binding site of OmpF increased with their atomic radii, $Li^+ \sim Na^+ < K^+ \sim Rb^+ \sim Cs^+$ (40). This affinity can be determined by the magnitude of the hydration energy, $Li^+ > Na^+ > K^+ > Rb^+ > Cs^+$ (101). Cs^+ easily losses its hydration water and can bind to the binding site with higher affinity (40). To compare the binding affinity, the electrostatic potentials of the monovalent ions along a pore surface of both OmpC and OmpF are computed using APBS (132) (see Chapter 2). The approximation of APBS calculation achieves the high accuracy on a small size of solute therefore only Na+ and Cl^- are discussed here.

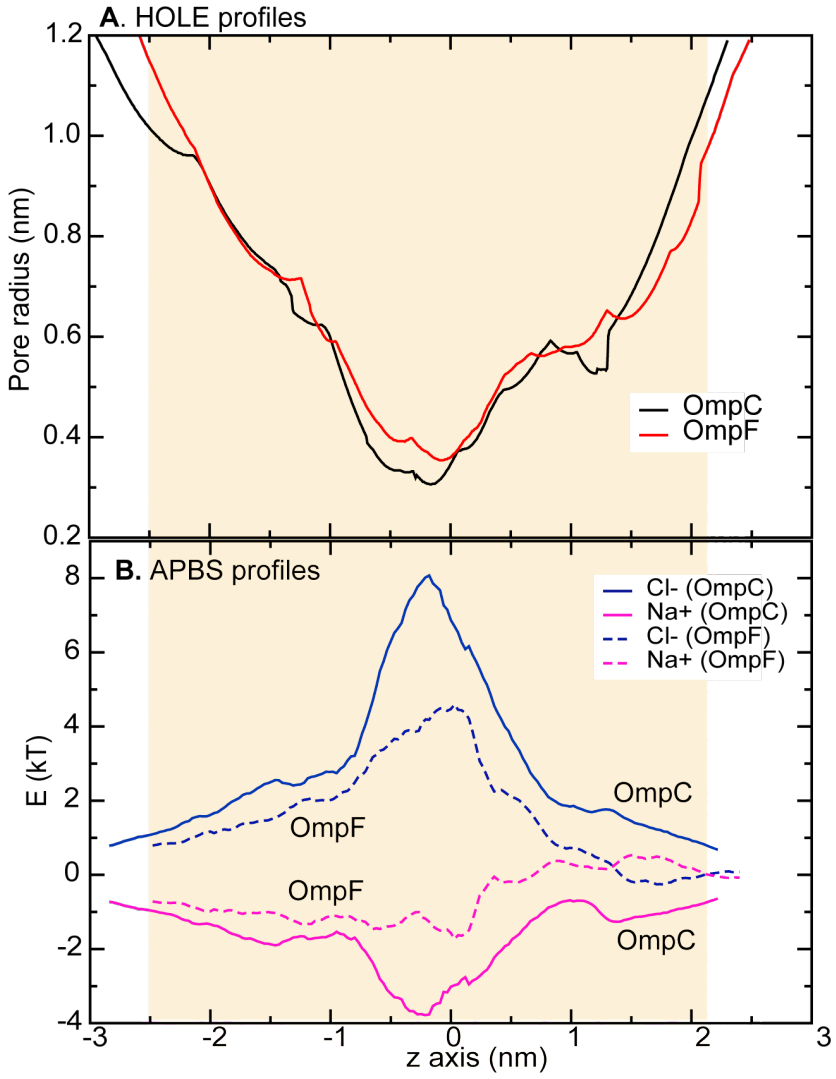


Figure 3-7 (**A**) Pore radii of crystal structures of OmpC (black) and OmpF (red). (**B**) Electrostatic potentials of chloride ion (blue) and Na^+ (pink) along a pore axis of OmpF (dash line) and OmpC (solid line). The profiles were calculated using APBS (132) with dielectric constant (ε) of 78.5 and 4 for water and protein respectively and solvent radius of 0.7. The length of both porins is shown as a yellow box.

The electrostatic potentials of both OmpC and OmpF (Figure 3-7) agree very well with the previous observation (3, 32). The selectivity filter of these two porins is at the constriction site (z ~ 0 nm in Figure 3-7). The negative well depth indicates that both OmpC and OmpF prefer a cation, Na^+ (Figure 3-7B: pink line). Indeed, OmpC appears to show more cation-selective characteristic with the deeper potential well of -4 *kT* (Figure 3-7B). Conversely, the energy barriers of 8 *kT* and 4 *kT* represent the unfavourable permeation of Cl^- along the OmpC and OmpF pore respectively. Even though the pore-ling architectures at the constriction site in both

porins are similar (negatively charged residues at L3 and positively charged patch on an adjacent wall), the smaller pore size permits OmpC to show stronger electrostatic properties (Figure 3-7). It was found that the more cation-selective properties of OmpC play a crucial role in bacterial survival (32). The OM permeability of antibiotics was significantly decreased as a result of increase in OmpC and decrease in OmpF (32). Also, under the acid-stressed condition, the neutralization of the periplasm was achieved by increasing a number of OmpC where OmpC was thought to contribute to the retention of buffer molecules in the periplasm (32). It is conceivable in such extreme conditions retarding and impeding the translocation of noxious solutes is essential for proliferation of the bacterial cell. An increase in the proportion of more selective and smaller OmpC channels appears to minimize the risk.

3.4 Conclusions

OmpC and OmpF are found to share similar functional and structural properties. They have the same set of interaction network to monitor the L3 flexibility and maintain trimeric stability. Apart from the more stable OmpC pore, the major differences that can distinguish OmpC from OmpF are the smaller pore size that increases the selectivity (Figure 3-7). Although both porins prefer cations, OmpC seems to have more cation-selective characteristic as seen by the deeper electrostatic well of 4 kT for Na^+ and higher barrier of 8 kT for Cl^-, while OmpF shows very weak preference for Na^+ with an energy well of 2 kT. However, the electrostatic properties may not the only factor to drive the permeation process in this case. The free energy profiles of the K^+ and Cl^- transport along OmpF pore have been computed using Brownian dynamics simulations (36). Under the condition of 1M KCl solution, significant free energy barriers located in the constriction site opposing the passage of K^+ and Cl^- were found (36). Even though

the K^+ transport is slightly more favourable, a free energy barrier of ~1.7 *kT* was still observed (36). It is implied that the electrostatic attraction is not the major factor to drive the diffusion process in OmpF. The entropic effect rather plays a key role. Despite the fact that the major driving force for ion diffusion in homologous OmpC is still unknown, the deeper electrostatic potential well should have more influence on the shape of the diffusive free energy rather than that of OmpF.

It is interesting that although certain residues proposed to be important for either structural stability or function are highly conserved, the properties of both OmpC and OmpF are not identical. A full structural explanation cannot be offered, but I have revealed the primary differences and similarities between OmpC and OmpF that may aid further studies. Overall properties such as the pore selectivity and the gating mechanism of OmpC also remain to be determined by experimental studies.

4 SIMULATIONS OF ANION TRANSLOCATION THROUGH OPRP REVEAL THE MOLECULAR BASIS FOR HIGH AFFINITY AND SELECTIVITY FOR PHOSPHATE

4.1 Introduction

Under conditions of phosphate starvation, high-affinity uptake of inorganic phosphate (Pi) is required for bacterial growth. In the pathogen *P. aeruginosa*, phosphate limitation leads to the induction of the protein OprP (77), an analogue of the PhoE porin of *E. coli*. Although both OprP and PhoE form anion-selective channels whereby each monomer is formed by a 16-stranded β-barrel, they differ in the nature of their transmembrane pores. OprP has high affinity for Pi uptake via a strong phosphate-binding site (79-81). In contrast, PhoE is weakly anion-selective and does not contain a specific binding site for Pi, but rather is anion-selective due to an overall excess of positively charged residues along the pore surface (145).

The X-ray structure of OprP (12) reveals three positively charged loops (L3, L5, and T7) folded into the lumen (Figure 4-1A). The arginine sidechains of L3 and L5 and lysines of T7 form a 'ladder' along the pore axis which includes the phosphate-binding site of OprP (12). *In vivo* phosphate permeation by OprP is unidirectional, from the external environment to the cell interior, via an association of OprP with a periplasmic phosphate-binding protein (81, 146). OprP is selective for phosphate over other anions, as revealed by conductance studies (77, 79-81) which indicate that phosphate permeation under conditions of phosphate starvation is about 20 fold more efficient than Cl^- transport (81). The primary contributing factor is a 100–500 times higher affinity of OprP for Pi than Cl^- (79-81). This renders OprP one of the few known porins that not only selects anions over cations but also discriminates sensitively between different anions (84).

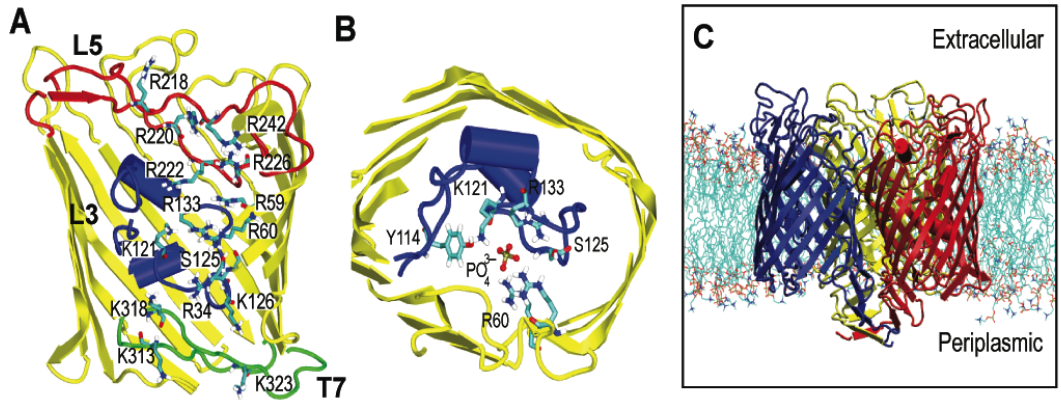


Figure 4-1 (**A**) Structure of the OprP monomer showing the sidechains of the arginine ladder and the lysines at the periplasmic mouth of the pore. The L3 and L5 loops are shown in blue and red respectively, and the T7 turn in green. (**B**) The key residues in the constriction zone surrounding the central phosphate. (**C**) The OprP trimer in a DMPC bilayer. The three protein subunits are in red, blue and yellow. The lipid molecules are in cyan (tails) and red/blue/brown (headgroups). Water and counterions have been omitted for clarity.

Crystallographic (12) and mutational studies (83) have resulted in a plausible model for phosphate transport through OprP, stressing the role of the sidechain of K121 (Figure 4-1B) as a "molecular ratchet" (12). To understand the ion permeation and a role of K121 in atomic level, many advanced computational methods can be used. Recently, continuum electrostatic calculations based on Poisson-Boltzmann equation have been used to probe the electrostatic potential of pore-lining surface (132). However, the static protein conformation and continuum electrostatic approximation are the major limitations of the application of this method to asymmetric and very narrow pores such as OprP. More sophisticated methods are needed. Here I therefore use a multi-technique molecular dynamics (MD) simulation approach to OprP embedded in a phospholipid bilayer (Figure 4-1C) to explore the relationship between structure and function, and to evaluate and refine the proposed transport mechanism for Pi and for Cl^-. In particular, calculations of the potential of mean force (PMF; i.e. the relative free energy profile) for permeant ions and solutes along the pores of K^+ channels (147) and aquaporins (148-150) have been reported. In both of these cases the pore is

relatively narrow (i.e. of comparable radius to a water molecule) and inflexible. In the current study I present the PMF profile for phosphate and Cl^- ions along the OprP pore. In contrast to previous simulation studies of pores and channels I am analyzing a somewhat wider pore, lined by flexible basic sidechains, that nevertheless exhibits significant selectivity. Thus our simulation results provide insights not only into the biology of OprP, but into more general aspects of the design principles of nanopores.

4.2 Methods

4.2.1 System preparation

The OprP trimer (PDB id 2O4V; resolution 1.94 Å) (12) was embedded in a pre-equilibrated dimyristoyl-phosphatidylcholine (DMPC; 290 lipids) bilayer using published methods (140). Water (SPC) molecules (47092 waters) and counter ions (39 ions) were then added. Tribasic phosphate ions were placed at either mouth of the pore for the equilibrium MD simulations. Water molecules overlapping with phosphate ions were removed manually. Energy minimization (steepest descent algorithm) was run for 5000 steps, followed by a 1ns equilibration phase during which protein and phosphate were harmonically restrained with a force constant of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$. Systems for umbrella sampling were prepared in the same manner except that no phosphate was added at this stage.

All simulations were carried out with GROMACS (151) version 3.3 (www.gromacs.org) with the extended united atom GROMOS96 forcefield (91). Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method (87) 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 *NPT* ensemble at constant number of particle, pressure, and temperature. The

temperature of protein, DMPC, solvent, and ions were each coupled separately to a Berendsen thermostat (96) at 310 K with couple constant τ_t=0.1 ps. The pressure was coupled isotropically using the Berendsen algorithm at 1 bar with coupling constant τ_p=1 ps. The time step for integration was 2 fs and all bonds were constrained with the LINCS algorithm (93). The two production runs of equilibrium MD were performed for 18 ns and 17 ns.

Constant-velocity SMD simulations were performed for 10 ns each to pull the phosphate in the ±*z* directions. Phosphate was initially at the centre of the pore. An elastic spring with a force constant of 350 $kJ \cdot mol^{-1} \cdot nm^{-2}$ was attached to the centre of mass of the phosphate and was moved in the direction parallel to the channel axis (*z* axis) at a rate of 0.25 nm/ns. Each SMD run was repeated twice.

Data were analyzed using GROMACS and locally written code. Molecular graphic images were prepared using VMD (142).

4.2.2 Umbrella Sampling to Calculate Potential of Mean Force Profile

The starting system for the umbrella sampling simulations was obtained from a 1 ns equilibrium MD run. The 1D reaction pathway corresponds to the distance along the *z*-axis between the centre of mass of each tribasic phosphate ion and each OprP monomer. The range of interest from –5.6 nm to +5.6 nm was divided into 226 equidistant windows (each of width 0.05 nm). For each window simulation, a phosphate was placed into each of the three pores at a center of an umbrella window, with the *xy*-location of each phosphate determined by the pore radius profiles as calculated using HOLE (152). Water molecules that overlapped with a phosphate were removed. Each window was initially equilibrated for 0.5 ns with an isotropic harmonic restraint of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ on the phosphate ions.

Umbrella sampling calculations were carried out by applying a harmonic biasing potential on the centre of each of the three phosphate ions relative to the center of mass of the corresponding monomer; the force constant was 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ and acted only along z. Data were collected for 1 ns and the last 600 ps of each run were used to construct the PMF. The unbiased PMF was computed with the weighted histogram analysis method (WHAM) (153). The final PMFs were obtained as the average of the PMFs for three ions, one for each pore of the OprP trimer. Errors were estimated from performing the above analysis in three trajectory blocks of length 200 ps and averaging the block-PMFs.

I calculated the dissociation constant (K_d) of phosphate and chloride ions from their PMFs. The standard free energy of binding, ΔG^o, is related to K_d via $K_d = V^0 \exp(-\Delta G^o/RT)$ where $V^0 = 1.661$ nm^3 is the standard volume corresponding to a standard concentration of 1 M. ΔG^o was computed from the 1D PMF, $W(z)$, using the method of Doudou *et al.* (154) under the assumption that confinement of the three sampled ions in the bulk region is provided by the simulation cell.

4.3 Results and Discussion

4.3.1 Three types of simulations

Three classes of simulations have been performed for the OprP trimer embedded in a phospholipid bilayer (Figure 4-2): (i) equilibrium MD simulations (duration ~17 ns) in which a Pi anion is released at either mouth of the pore; (ii) constant-velocity steered MD (SMD; (113)) in which a Pi is pulled through the pore on a timescale of 10 ns; and (iii) umbrella sampling calculations of the PMF of a Pi and of a Cl^- ion as a function of their position along the pore axis. All three approaches provide evidence for an anion-binding site at the centre of the pore, close to the key K121 sidechain. In the equilibrium simulations (Figure 4-2B) a Pi ion released at either mouth of the pore diffuses to the central binding pocket within ~10 ns where it remains for the remainder of the simulation. In the SMD simulations (duration ~10 ns; Figure 4-2C and D) the Pi ions were pulled in either direction from the central binding site. In each case the anion is slowly released from this site (over ~4 ns), interacting more weakly with other sites before exiting through either the extracellular or periplasmic mouth of the pore. The exact locations of the interaction sites within and at the mouths to the pore are revealed in the PMF (Figure 4-2E).

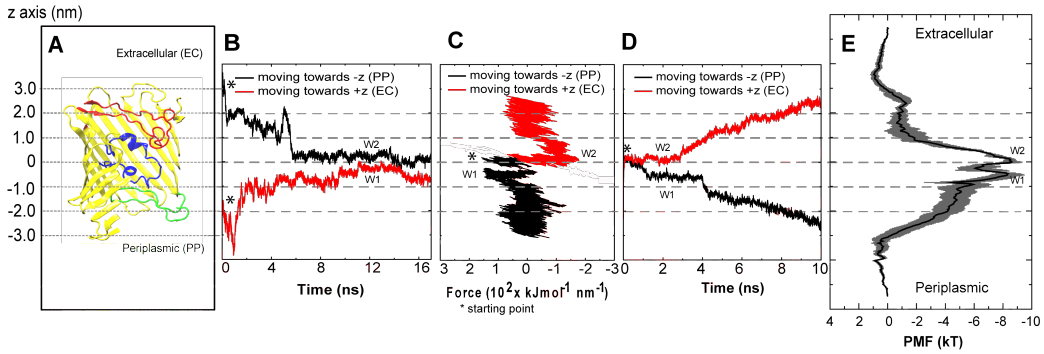


Figure 4-2 **(A)** Cross section of the OprP monomer showing L3 (in blue), L5 (in red), and T7 (in green) and defining the z-axis coordinates used in other figures. **(B)** Phosphate trajectories from equilibrium MD simulations for an ion initially at the extracellular (black; -*z*) and periplasmic (red; +*z*) mouth of the pore. W1 and W2 represent the ion present at the two central energy wells. **(C)-(D)** Force profiles (**C**) and trajectories (**D**) from inward (black; -*z*) and outward (red; +*z*) constant-velocity SMD simulations of phosphate through the pore. Each phosphate was initially placed in the centre of the pore. **(E)** Potential of mean force (PMF) for a tribasic phosphate ion along the permeation pathway. The grey bars indicate one standard deviation on the free energy as calculated from block averaging.

4.3.2 Energetics of the phosphate binding sites.

The PMF profile reveals no significant barrier to permeation anywhere along the pathway, both for Pi and Cl^-. This is in contrast with the corresponding profiles of narrower pores (e.g. KcsA (147) and aquaporins (148-150) which show a complex energy landscapes composed of mixtures of wells and barriers. In the OprP phosphate profile two adjacent energy wells (W1 and W2; Figure 4-3A) are seen in the center of the pore in the vicinity of the L3 loop, forming the most favorable sites for Pi binding with a well depth of ~ –8 *kT*. These favourable Pi binding sites were also detected by equilibrium MD and SMD (W1 and W2 in Figure 4-2C-D). Thus, in SMD pulling a Pi from the extracellular to a periplasmic side (–*z* in Figure 4-2C) or in the reverse direction (+*z*) reveals a sudden increase in the force on the Pi of magnitude 150 $kJ \cdot mol^{-1} \cdot nm^{-1}$ at *z* values from –1 to 1 nm corresponding to the energy wells W1 and W2 found in the PMF. These two energy wells were thus observed in two independent SMD simulations and two equilibrium simulations.

The SMD simulations also revealed two other minor phosphate-binding 'sites' which seem to correspond to shoulders/plateau regions in the PMF.

I computed a dissociation constant K_d = 6±1 μM for Pi from the PMF (see Methods). This is within the range of reported experimental estimates of the K_d of Pi at neutral pH (0.15 μM to 310 μM) (79, 81, 84). A K_d of 6 μM corresponds to a standard free energy of binding ΔG^o= –12.0±0.2 kT. Such tight binding is also consistent with the presence of bound Pi in the X-ray structure of OprP (12).

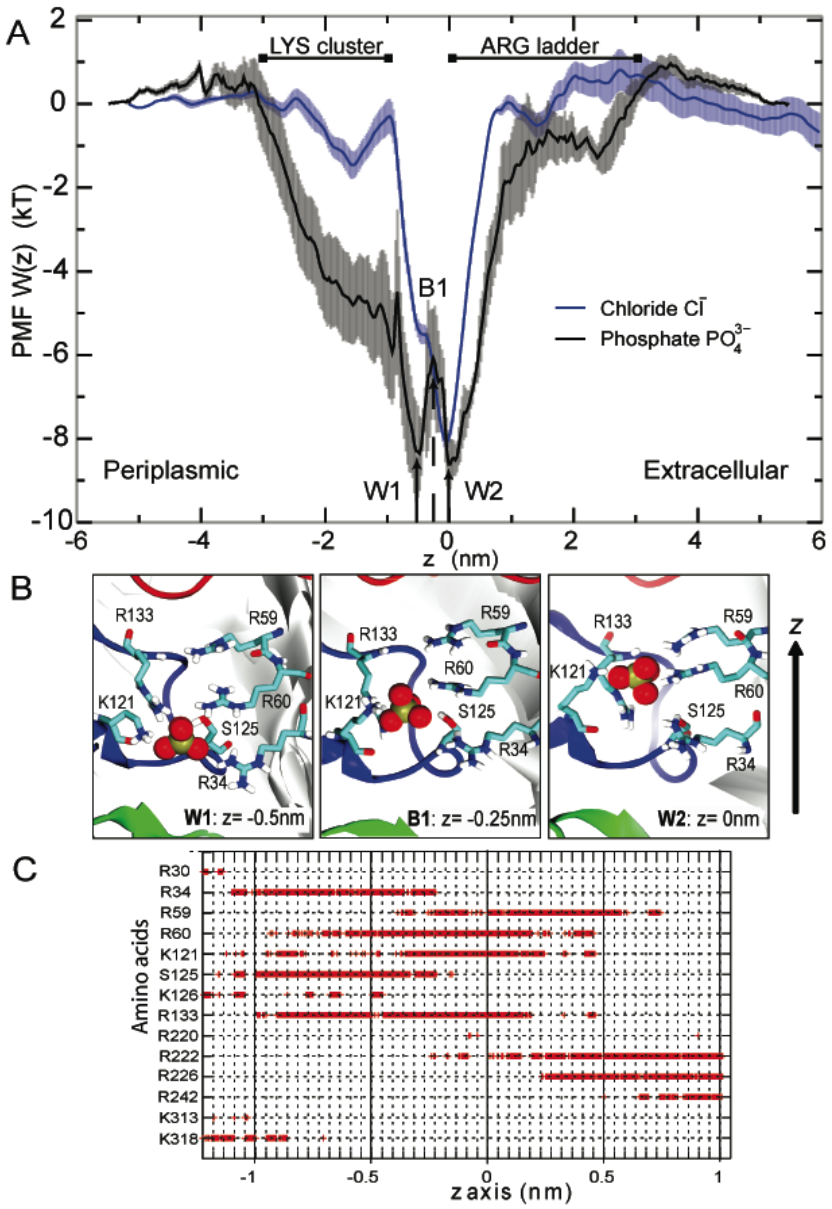


Figure 4-3 **(A)** Relative free energy profile (PMF) of anions along the OprP permeation pathway. Phosphate is stabilized in two wells (W1, W2), separated by a small barrier (B1) in the constriction zone. A Cl^- ion can bind in the W2 site. Error bars indicate one standard deviation of the data as obtained from averaging over three blocks of 200 ps length. The approximate positions of clusters of basic residues along the pore axis are marked by extended lines. **(B)** Conformations of positively charged sidechains with Pi (in spacefilling format) at W1, B1, and W2. (**C**) Hydrogen bonds between Pi and the pore-lining amino acids as a function of position of the ion on the *z* axis.

4.3.3 Basic side chains form the phosphate-binding site.

The Pi binding site in the L3 region is unusual for a porin in that it binds the substrate tightly and selectively. The two subsites W1 and W2 are separated by only a small free energy barrier B1 of ~2 *kT* (Figure 4-3A). W1 and W2 are about 0.5 nm apart. Analysis of the Pi-protein interactions indicates that the double-well structure of the binding site may be due to the movement of a single side chain, K121 (Figure 4-3B). Hydrogen-bonding analysis (Figure 4-3C) indicates that interactions with R59, R60, K121, and R133 are involved in the formation of the energy well at W2, whilst interactions with R34, R60, S125, and R133 form the second Pi binding site at W1. The free energy barrier at B1 appears to result from the loss of favorable electrostatic interactions. At the transition between two subsites, the hydrogen bond to R34 and S125 are being broken and the one to R59 is not yet formed, thus destabilizing the phosphate ion.

The sidechain of K121 has a special role: It makes direct contact with the Pi during the whole transition between W1 and W2, which requires it to switch its rotameric state. Hydrogen bonds with R60 and R133 also form stable interactions with the Pi across the whole central binding site (i.e. W1, B1, and W2), thus contributing to the overall stabilization of the charged Pi at the centre of the membrane. It is found that interactions with other residues are localized at the edges of the free energy wells. K121 is thus a key residue that can interact with Pi across two regions; R60 and R133 can also facilitate Pi movement across the constriction site but to a lesser degree (Figure 4-4A and C).

A positive correlation is observed between the *z* positions of K121 and the Pi (Figure 4-4B), indicating that the dynamics of K121 and Pi are coupled. This in turn suggests a mechanism in which K121 plays a crucial role in directing a Pi ion

from W2 to W1 by “sweeping” it along as a charged brush. This agrees with the suggestion from crystallographic studies that different conformations of K121 could enable the transfer of Pi to the constriction site (12). Site-directed mutagenesis showed that the K121E mutation both reduces Pi conductance and also the ability of Pi to inhibit Cl^- conductance (83).

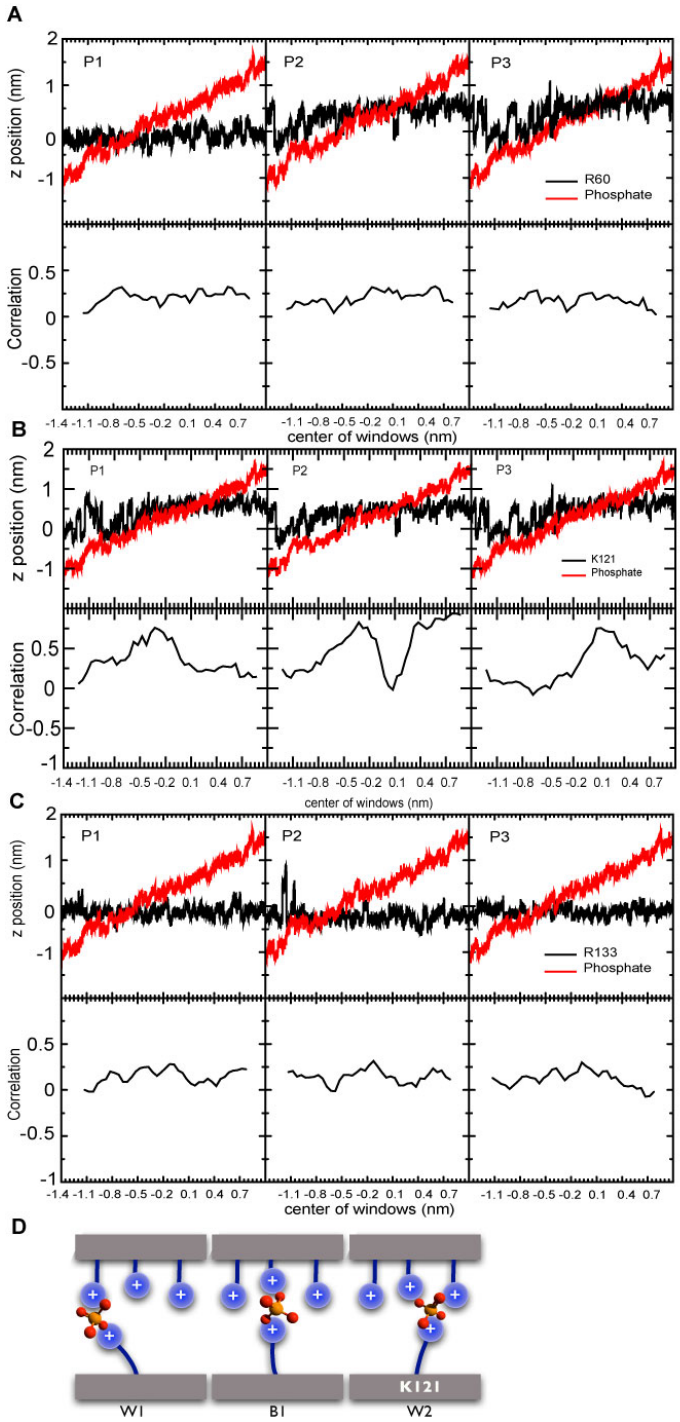


Figure 4-4 (**A)-(C)** *Top panel:* Trajectories of the N_ε atom of the side chain of R60, K121, and R133 and a Pi in a constriction site from one OprP monomer extracted from the umbrella sampling simulations. The *z*-position of the N_ε atom or Pi ion is shown on the vertical axis as a function of the *z*-coordinate of the centre of the window (i.e. of the Pi ion restraint) on the horizontal axis. *Bottom panel:* Correlation coefficient between the *z*-position of the R60, K121, and R133 N_ε and the centre of mass of the Pi as a function of the *z*-coordinate of the centre of the umbrella window. (**D**) Schematic diagram of the mechanism of positively charged sidechains acting as charged brushes to facilitate Pi transport along the pore (by courtesy of Oliver Beckstein).

My simulations reveal that the Pi binding site is not rigid, but instead is flexible adapting its size to the presence of the solute. Analysis of the pore radius profile in the presence of trapped Pi at W1, W2, and B1 demonstrates that the change in conformation of the K121 sidechain correlates with a change in the radius of the pore in the region of L3 (Figure 4-5). The terminal amino group of the K121 sidechain always points towards the narrowest region of the pore. I conclude that K121 is crucial for the formation of Pi-binding site *and* for dynamic changes at the narrowest region of the pore. Thus, the OprP permeation pathway may be thought of as a wide and relatively rigid tunnel provided by the β-barrel scaffold which is lined by sidechains protruding into the lumen that provide a flexible but highly specific binding site. In this picture the pore may be thought of as resembling a polymer brush folded into a tube, forming a brush-like nanopore.

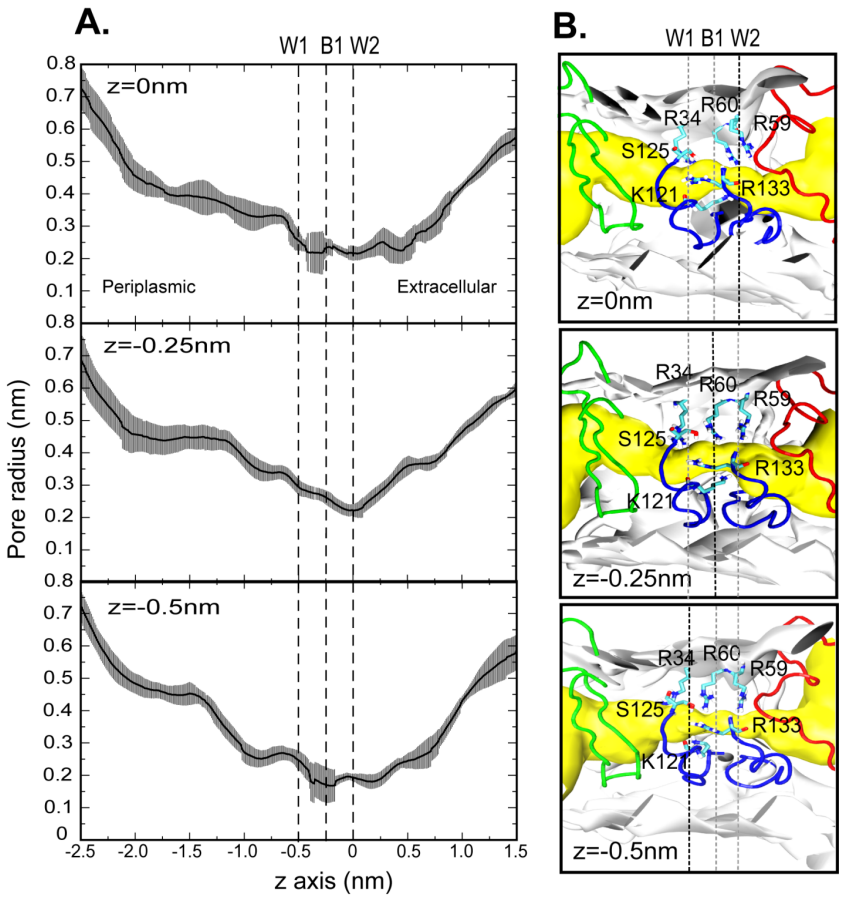


Figure 4-5 (A) Average pore profiles of a OprP monomer with phosphate present at z=0 (W2), -0.25 (B1), and -0.5 (W1) nm, created with HOLE (152). (B) shows pore-lining surfaces which are scaled and aligned to correspond to the pore profile above with a presence of R34, R59, R60, K121, S125, and R133. The black dash line represents a phosphate position.

In addition to the dominant interaction network at the pore constriction site, the other interactions with residues lining the channel seen in the simulations support the *arginine ladder* mechanism proposed on the basis of the X-ray structure (12). Thus, the overlapping hydrogen bonds of R222 and R226 in the L3 and L5 region seem to be important in directing the phosphate to the constriction site (Figure 4-3C). These major interactions found in L3 and L5 regions are also clearly seen in the SMD and equilibrium MD (Figure 4-6).

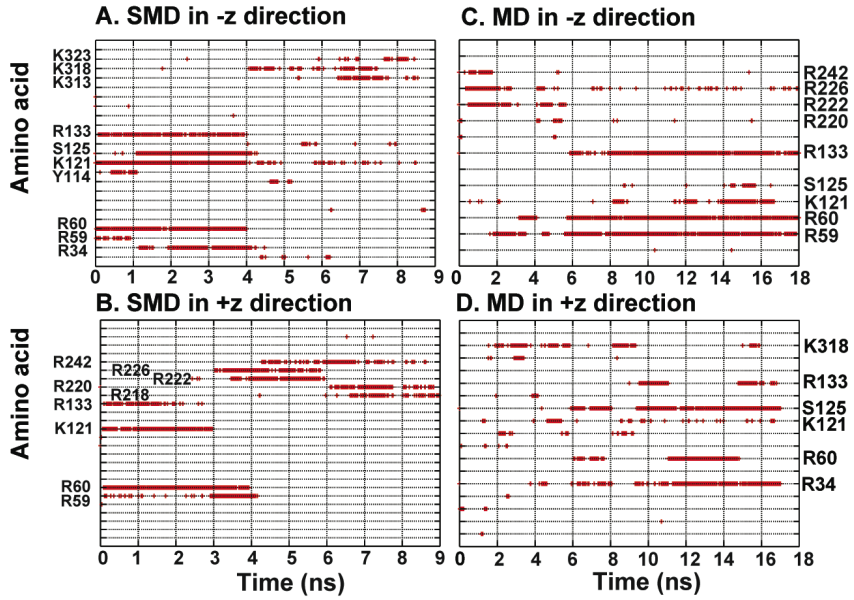


Figure 4-6 Time-resolved hydrogen bond existence between the phosphate and the pore-lining amino acids during MD. (**A**) and (**B**) are the hydrogen-bonding network resulting from SMD, while (**C**) and (**D**) are from unbiased MD simulations in both $-z$ and $+z$ directions.

4.3.4 Comparison of phosphate and chloride

In order to elucidate the selectivity of OprP of Pi over Cl^- ions I also computed the PMF for Cl^- ions (Figure 4-3A). The monovalent anion is stabilized in the W2 site to the same degree as the trivalent phosphate, with a free energy well depth of ~8 *kT*. In contrast to Pi, this is the only binding site for a Cl^- ion. A large portion of the periplasmic side of the pore that is lined by the "Lys-cluster" (12, 77, 80-83) is also attractive for Pi but barely so for Cl^-. The K_d as computed from the PMF is 44±2 μM (equivalent to ΔG^o=–10.0±0.1 *kT*), which corresponds to a seven-fold stronger binding of Pi ions than Cl^-. This is smaller than an experimental estimate of the binding selectivity of ~100–500-fold as derived from the measured K_d values (79, 80). Nevertheless, the calculation of the binding constants as integrals over pore volumes suggests that OprP is selective of Pi over Cl^- because the total attractive pore volume is much larger for Pi than for Cl^- (Figure 4-3A).

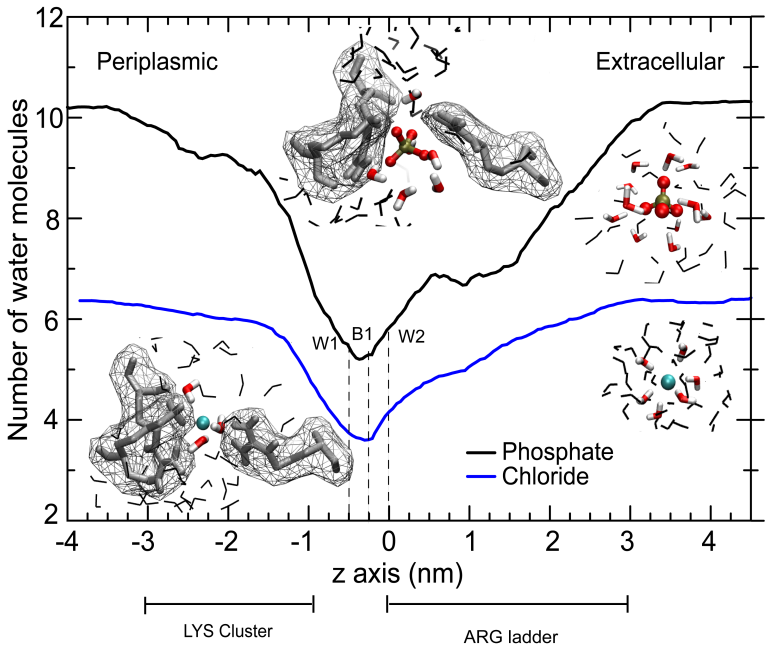


Figure 4-7 Number of water molecules within an ion's first hydration shell was obtained from umbrella sampling simulations of Pi and Cl^-. The approximate position of the "arginine ladder" and the "lysine cluster" are shown along the permeation pathway. Insets depict the ion and water in CPK representation. Water molecules within the first hydration shell (water oxygen–anion distance <0.275 nm for Pi, <0.3 nm for Cl^-) are shown in color, other nearby waters are grey. Pore-lining basic residues are shown as grey sticks together with their solvent accessible surface as a mesh. Images on the left hand side are from the binding site near K121, right hand side images are in bulk water.

The mechanistic basis for selectivity between anions can then be traced to the differential hydration of the two ions (Figure 4-7). The pore at the Lys-cluster is sufficiently wide to allow a Cl^- ion to be fully solvated by a bulk-like hydration shell of roughly six water molecules. Pi is larger and confinement strips about two waters from its hydration shell. Thus the highly charged, partially desolvated Pi anion can interact strongly with the basic Lys residues whereas the Cl^- ion with its smaller charge is screened by its complete hydration shell. Partial desolvation is sufficient to establish direct interactions between the negative charges on the phosphate oxygen and the positive charge of the lysine ε-amino group.

Unlike the Lys-cluster, the Arg ladder motif does not appear to differentiate energetically between Pi and Cl^- (Figure 4-3A). Its role is rather to facilitate

movement of the negatively charged particles towards the center of the low dielectric membrane through a narrowing pore. It accomplishes this by gradual desolvation of the anions (which strips them of roughly half of their bulk hydration shell; Figure 4-7) while stabilizing them through favorable electrostatic interactions.

4.4 Conclusions

4.4.1 Biological function

The use of a combination of simulation approaches has revealed the presence of two well defined and closely spaced energy wells (W1 and W2) along the permeation pathway of a Pi anion through the OprP pore. W1 corresponds to the Pi binding site in the X-ray structure. The transition from W2 to W1 is facilitated by K121. The highly-correlated Pi-K121 motion, the hydrogen-binding analysis, and the change in pore cavity due to the conformation change of K121 support the idea that K121 serves as a 'charged brush' to facilitate Pi passage between the two subsites. From the PMF I am able to compute a standard free energy of binding of phosphate to OprP ($\Delta G^o = -12.0 \pm 0.2\ kT$) equivalent to a dissociation constant of $K_d = 6 \pm 1\ \mu M$. In this sense OprP is not a typical specific porin such as LamB with a dissociation constant only in the mM range (155). This reflects the function of OprP under conditions of phosphate *starvation* when available phosphate must be scavenged from the extracellular environment at concentrations <200 μM (77). The OprP-bound phosphate is passed to a specific periplasmic binding protein at the periplasmic mouth of the pore in order to facilitate unidirectional phosphate uptake (81, 146).

OprP is also remarkable in that it discriminates between anions. The experimentally observed stronger binding of tribasic phosphate over chloride ions could be qualitatively explained from the simulations and traced to the cluster of N-

terminal lysine residues. The periplasmic mouth of the pore is sufficiently wide to admit a fully solvated Cl^- ion but partially desolvates a Pi ion. Thus the monovalent ion is screened from the positive charge of the lysine amino groups by its own hydration shell whereas the unscreened Pi ion is stabilized by strong charge-charge interactions. While the lysine-cluster is the selectivity filter for anions, the arginine ladder primarily facilitates the gradual dehydration of incoming anions from the extracellular side without discriminating between Cl^- and Pi. Both Cl^- and Pi are stabilized at the center of the pore to a similar degree, with roughly half of their hydration shells stripped; it is likely that the binding site itself also acts as a filter for anions over cations and a general size-exclusion filter.

4.4.2 PMFs of biological pores

Free energy profiles (PMFs) have been calculated for a number of permeant species through biological pores, including e.g. K^+ ions through the KcsA channel (147), water and glycerol through aquaporins and aquaglyceroporins (148-150), and ammonia through AmtB (156, 157). It is therefore useful to compare these pores with OprP. Both the KcsA selectivity filter and the Aqp pore are (relatively) inflexible, narrow (~0.2 nm radius) and are lined at least in part by backbone atoms. These pores present barriers to permeation of K^+ ion (KcsA) or water (hAqp1) or glycerol (GlpF) of the order of 5 *kT* (147-150). In contrast, the pore of OprP which is flexible (especially at K121), relative wide (minimum radius >0.2 nm) and lined largely by sidechains has an energy well of depth ~8 *kT*, formed by positively charged sidechains as discussed above. This is closer to the situation for the ammonia channel AmtB which has been estimated to present an interior energy well (formed largely by stabilizing interactions with His sidechains) of overall depth ~17 *kT* for ammonia molecules (157), although in the latter study the authors comment

on the structural invariance of the AmtB pore regardless of the location of the ammonia molecule. Thus it would seem that the flexible, sidechain-lined pore constriction of OprP provides for a relatively high affinity binding site whilst allowing high permeation rates for phosphate and chloride anions (81). It is of particular interest that a sidechain-lined pore seems to be associated with formation of a well rather than a flatter energy landscape with one or more small barriers to permeation. Hence OprP does not share the free energy landscape of permeation with "typical" channels or pores but it rather resembles a protein with a binding site with two openings that bridge the outer membrane. OprP can be understood as a brush-like nanopore in which charged side chains provide a high local concentration of counter charge to stabilize a charged substrate, similar to highly simplified but successful models of sodium and calcium channels (158). In addition it utilizes differential desolvation to probe different anions and can so discriminate between a chloride ion that "hides" under its hydration shell and a larger phosphate ion whose size forces it to make direct contact with the sensing charged brushes. This mechanism depends critically upon the degree of flexibility of the pore lining sidechains. This may provide a new design principle for e.g. derivatization of nanopores (159) to form ion selective biomimetic channels (see Chapter 5). Furthermore, an improved understanding may aid design of selective blockers of OprP to aid in fighting *Pseudomonas* infection.

5 SIMULATIONS OF BRUSH-LIKE NANOPORES

5.1 Introduction

Ion channels and porins are nanopores that in many cases, display high levels of solute specificity. Their selectivity has been exploited in new technologies; for instance synthetic functional nanopores have been fabricated based on the permeation properties of biological channels (104, 105, 160-163). Although such artificial pores have a wide range of possible applications, their selectivity and recognition properties remain inadequate. Certain researchers have attempted to modify biological pores to increase their sensing ability (69, 164). Nevertheless, the relative instability of modified biological pores under industrial conditions remains a major limitation. The interplay between biological- and synthetic or inorganic-based designs may result in the more practical nanopores. Recently, the array of carboxyl backbones in selectivity filter of KcsA channel has been implanted to the interior surface of carbon nanotubes (CNT) in silico in order to duplicate its K^+-selective behavior (165). With an interior architecture similar to that in KcsA, the modified CNTs were able to display K^+ selectivity (165). This is one example showing that the high selectivity of biological channels can in principle be transferred to a synthetic nanopore.

In Chapter 4, OprP was shown to exhibit a high affinity for the uptake of a phosphate (Pi). The PMF indicates a crucial role of basic pore-lining residues which provide a 'charged brush' in determining Pi selectivity. In this chapter, OprP is used as a template for designing a series of computational models of synthetic nanopores that exhibit both OprP's anion-vs-cation selectivity and a preference for Pi over other anions.

A number of studies have used MD simulations of simple models of nanopores to explore e.g. water permeation (166-170) and channel gating (171, 172) in more complex biological systems. In this study, I used MD simulations to understand the permeation of ions through my pore models. Via the calculation of PMFs, I have explored the influence of pore-lining charge mobility on the phosphate permeability. The relative selectivity for phosphate of OprP and the synthetic nanopores is also compared.

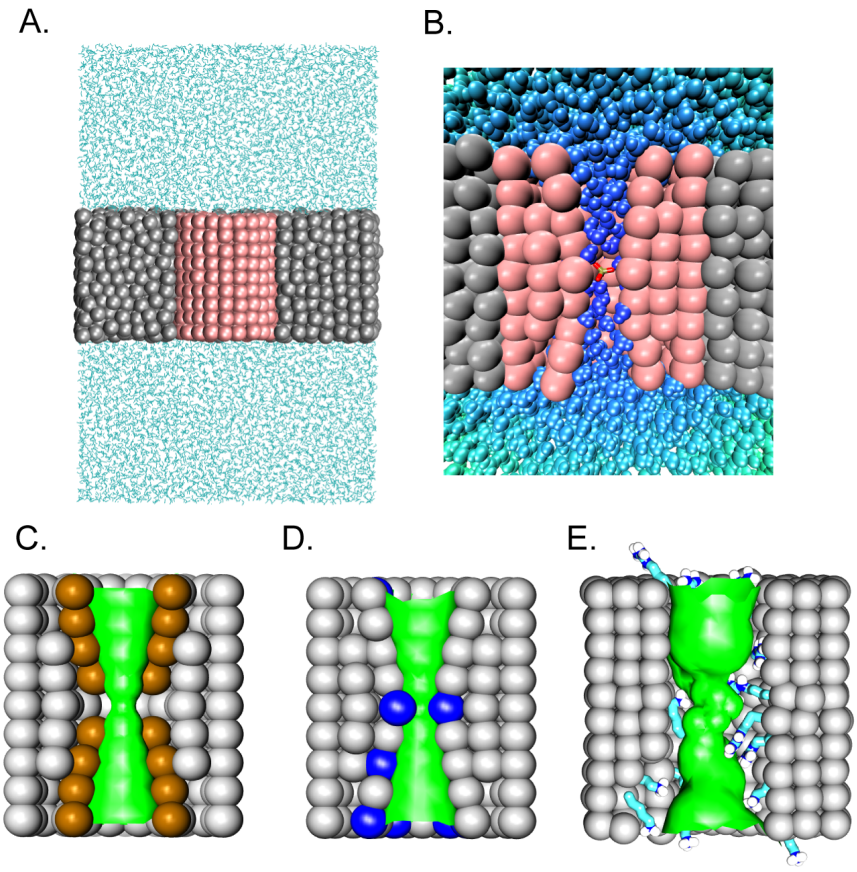


Figure 5-1 Cartoon representation (**A**) Model system. A model pore is colored in pink. Biomimetic membrane slab and water molecules are in gray and blue. (**B**) Cross-sectional view of a pore model with the presence of a phosphate ion. (**C**)-(**E**) represent the interior architecture of each model with the HOLE surface. A pore shape is labelled as yellow beads in (**C**). Blue beads in (**D**) indicate the +1 charged particles.

5.2 Methods

5.2.1 Design of the Nanopores

As in some earlier studies (168, 169, 173), the cylindrical pores were constructed from methane-like hydrophobic particles (Figure 5-1A and B). Functional pore-facing amino acids in OprP were transplanted into the model pores. In addition, to study the benefit of having mobile charged residues on a pore surface, the energetic properties of three classes of pore models (non-charged, static and flexible charged) are investigated (Figure 5-1C-E). Since the brush-like behaviour of OprP plays a vital role in Pi selectivity, it is interesting to understand how different degrees of sidechain flexibility can affect ion selectivity. Thus, models with different lengths of charged sidechains were constructed.

The pore models were designed to mimic the dimensions and characteristics of OprP. The functional arrangement of 20 positively charged residues (Arg and Lys) found in OprP is transplanted to the model pores. Each model comprises concentric rings of methane particles with van der Waals radius of 0.195 nm. A pore is composed of two mouth regions (radius ~0.7 nm, 1.75 nm < length < 2.5 nm) and an inner pore region (radius ~0.3 nm, length ~0.5 nm). All methane-like particles were harmonically restrained to their initial position with the force constant k_0 = 10,000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ in order to conserve the OprP-like pore cavity. Three models with different interior environments were constructed here (Table 5-1). The hydrophobic pore (HP) is used as a reference pore. To construct the static charged pore (SP), 20 charges of +1 were assigned to given particles on an interior wall that match positively charged architecture inside the OprP pore. The mobile charged model (MP1) was created by attaching 9 Arg and 11 Lys sidechains extracted from the OprP crystal structure (12) onto the pore surface. The sidechains were arranged to

represent the OprP signature (i.e. Arg ladder, Lys cluster, and phosphate-binding site). The end of each sidechain was capped with a methyl group, the position of which was restrained by harmonic constraints with force constant k_0. To study the effect of degrees of sidechain flexibility, the other two mobile charged pores, MP2 and MP3 were built with the different lengths of pore-ling sidechains (see Table 5-1 for details). The location of the charged residues and pore cavity of all models were designed to replicate those of OprP so as to focus only on the flexibility (Figure 5-2). Pore radii of each pore type were computed using HOLE (152).

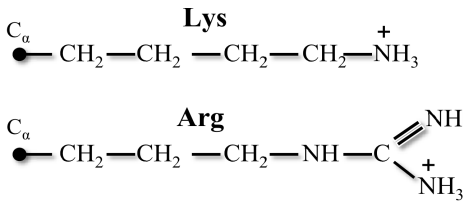
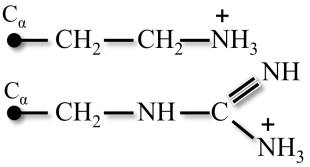
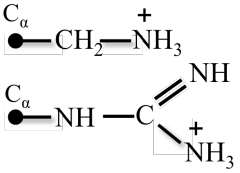
Model	Description	Sidechain
HP	Hydrophobic pore Radius profile matches to OprP	-
SP	Static charged pore Radius profile matches to OprP +20 charges matched to OprP	-
MP1	Mobile charged pore Radius profile matches to OprP 9 Arg + 11 Lys sidechains	
MP2	As for MP1, but sidechains truncated by 2 CH_2 groups Radius profile matches to OprP	
MP3	As for MP1, but sidechains truncated by 3 CH_2 groups Radius profile matches to OprP	

Table 5-1 Details of each model nanopore

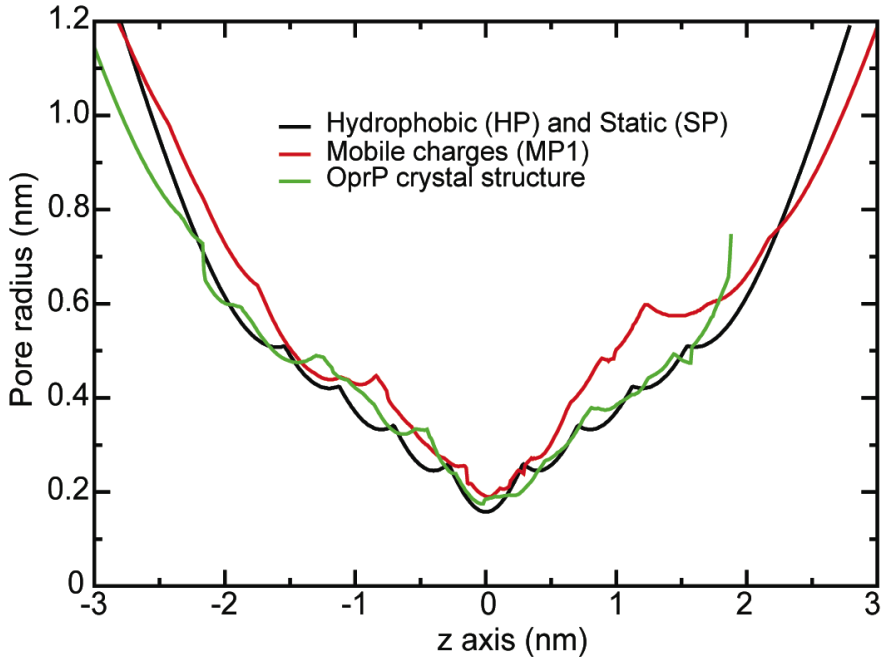


Figure 5-2 Pore radius profiles of all models using HOLE.

Each pore model was embedded in a slab of methane particles held on a cubic lattice with the force constant of 3,000 $kJ \cdot mol^{-1} \cdot nm^{-2}$. This simple model of a membrane has been used in earlier models of nanopores (168, 169, 173). Water (SPC) molecules with 3 Na^+ counterions for HP and 23 Cl^- ions for SP and MP series were then added. Energy minimization (steepest descent algorithm) was run for 600 steps, followed by a 2 ns equilibration phase where the pore framework was harmonically restrained with a force constant of k_0.

All simulations were carried out with GROMACS (151) version 3.3 (www.gromacs.org) with the extended united atom GROMOS96 forcefield (91). Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method (87) 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 *NPT* ensemble at constant number of particles, pressure, and temperature. The temperature of pore, membrane, solvent, and ions were coupled to a Berendsen thermostat (96) at

300 K with coupling constant τ_t=0.1 ps. The pressure was coupled isotropically using the Berendsen algorithm at 1 bar with coupling constant τ_p=1 ps. The time step for integration was 2 fs and all bonds were constrained with the LINCS algorithm (93). Data were analyzed using GROMACS and locally written code. Molecular graphic images were prepared using VMD (142).

5.2.2 Umbrella Sampling to Calculate a Potential of Mean Force Profile

The 2 ns-equilibrated pore systems were used as initial structures for the umbrella sampling simulations. The 1D reaction coordinate corresponds to the distance between the centre of mass of a tribasic phosphate ion and the center of mass of a pore model. The pathway (-4.6 nm < z < 4.6 nm) was divided into 190 equidistant windows (each of width 0.05 nm). The detailed procedures and parameters used to setup each umbrella window can be seen in the "Methods" section in Chapter 4. PMF Data were collected for 1 ns for HP and SP models and 1.5 ns for MP1 and the last 700 ps of each run were used to construct the PMF. In case of MP2 and MP3, each window was run for 2.5 ns and the last 1.5 ns of each simulation was used to calculate the PMF. Errors were estimated from performing a block analysis (174) in the trajectory blocks of length 300 ps and averaging the block-PMFs.

5.3 Results and Discussions

5.3.1 Are the phosphate-selective properties transferable?

The dimensions of the model pores measured via HOLE are given in Figure 5-2. The pore radius profiles indicate the almost identical sizes of the constriction site in the models to that of OprP pore (radius ~ 0.2 nm) although MP1 has a slightly wider vestibule.

The influence of pore interior on the ion passage can be understood from the PMF profiles in Figure 5-3. Phosphate translocation through the fully hydrophobic HP model is opposed by a barrier of ~ +120kT (Figure 5-3A) similar to that seen in a study of purely hydrophobic nanopores (173). The pore with static charges lining the interior (SP) does not exhibit a barrier but instead has a pronounced energy well of depth ~ -160kT (Figure 5-3B). This illustrates the important role of the charged components in reducing the energy barrier across membranes (175). Although the energy well is a key factor for attracting anions, such a deep well (Figure 5-3B), will lead to phosphate binding which is too strong for fast translocation to occur. A more 'pore-like' energy profile can be obtained by mimicking the brush-like behaviour in MP1 (Figure 5-1E and Figure 5-3C). The Arg and Lys sidechains stabilize a Pi and allow binding of Pi anion with an energy well depth of ~ -20kT (Figure 5-3C). Thus, it appears that the 'brush-like' properties are essential to obtain a high permeability anion-selective nanopore.

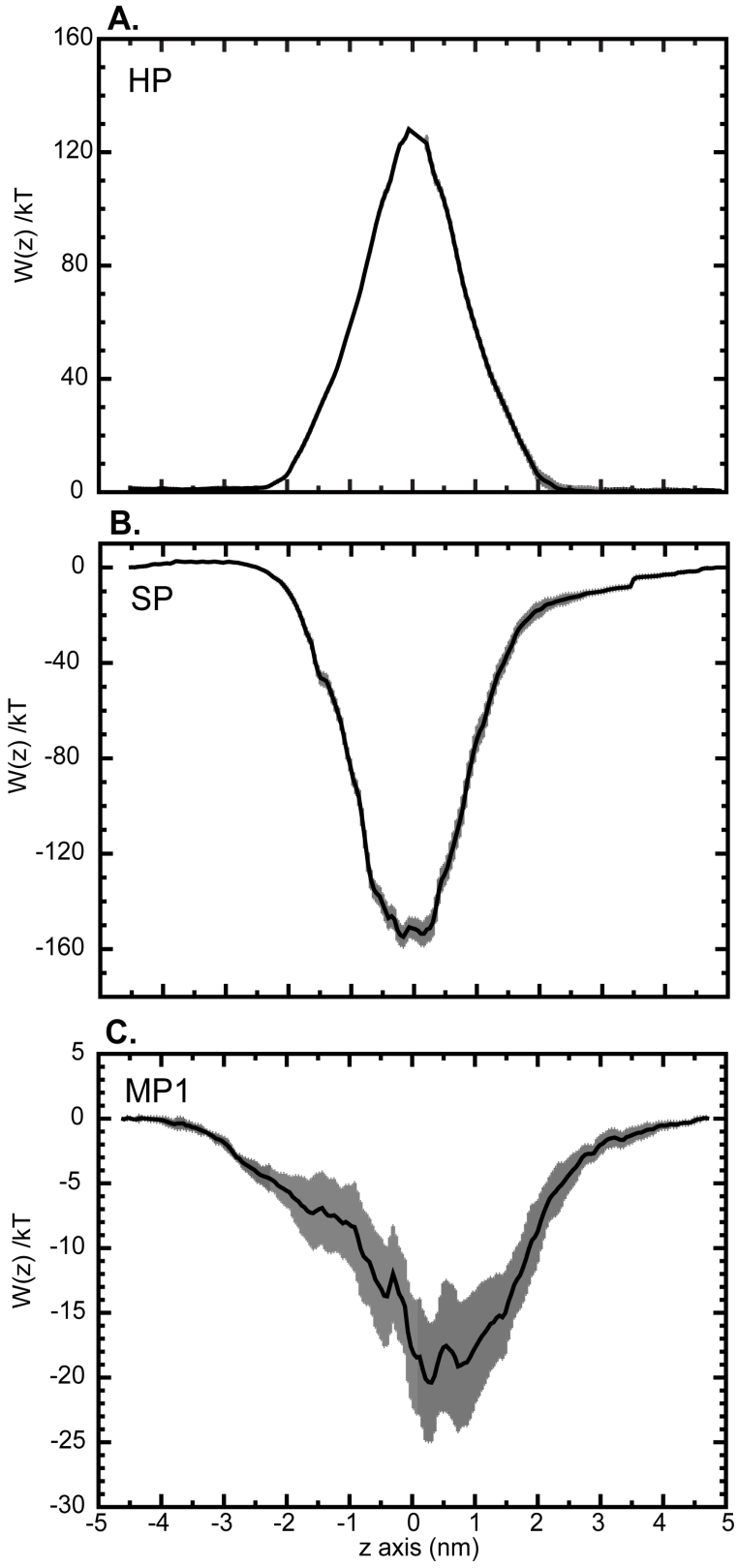


Figure 5-3 Phosphate PMF profiles with error bars calculated from the block average. (**A**) Hydrophobic pore (HP). (**B**) Static charged pore (SP). (**C**) Mobile charged pore (MP1).

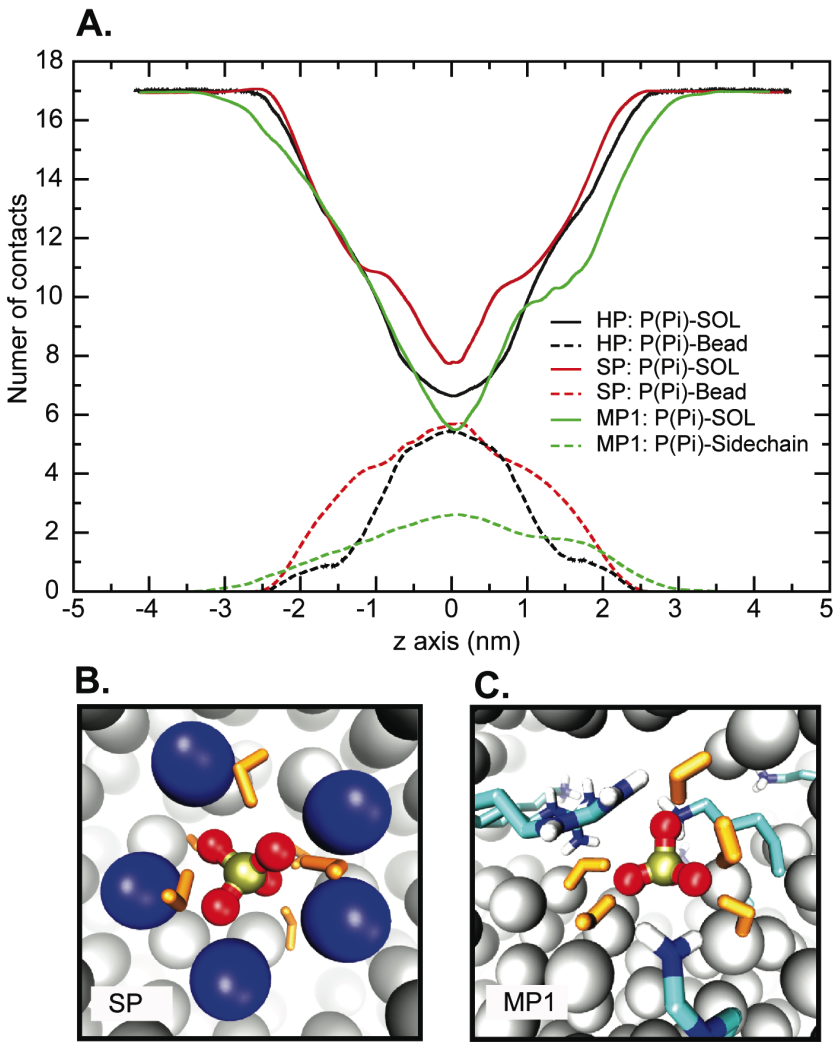


Figure 5-4 (**A**) Average contact number of Pi with water molecules (solid line), the center of mass (COM) of pore-lining methane-like particle in SP and, HP and Arg and Lys sidechains in MP1 (dash line) along the reaction coordinates. All contacts are calculated using a cut-off of 0.5 nm. (**B**) and (**C**) Bird's eye views of phosphate and its neighbours in the constriction site of SP and MP1 nanopores. Pi is represented in spacefilling format. Waters in the first hydration shell are in orange. The blue beads in SP shows +1 charged particles in a constriction site.

The Pi-binding affinity among these 3 model pores can be investigated further by tracking the contacts between Pi and its neighbours. Travelling along the narrow pore, dehydration is necessary for Pi to fit the pore cavity. As a rule, the incomplete hydration shell of Pi in a pore can be sequentially filled up by interactions with pore-lining residues. Therefore, the change in hydration number of Pi allows us to follow and understand the process of Pi passage. In the model with static charges (SP), it shows the highest hydration number of ~8 in the constriction

site. However, SP also has a high number of pore contacts of ~6 (Figure 5-4A) so that the total number of contacts (~14) provides a nearly bulk-like coordination shell (~17 hydration number in bulk) for Pi at the constriction point. This high number of pore contacts and the bulk-like coordination shell in the constriction side of SP give the most stable Pi-binding sequentially leading to a very deep energy well. In contrast for MP1, Pi loses ~11 of water molecules from its hydration shell at the constriction site (Figure 5-4A), but only ~ 3 pore-lining sidechains of MP1 are able to occupy the hydration shell. The lesser total number of contacts with Pi in MP1 (~9 contacts compared to ~14 total contacts in SP) helps to explain the reduction in energy well depth. The 'smear' of flexible sidechains in MP1 appears to provide more 'pore-like' energy profile (Figure 5-5). The trade-off between substrate-binding affinity and diffusion rate are crucial to take into account.

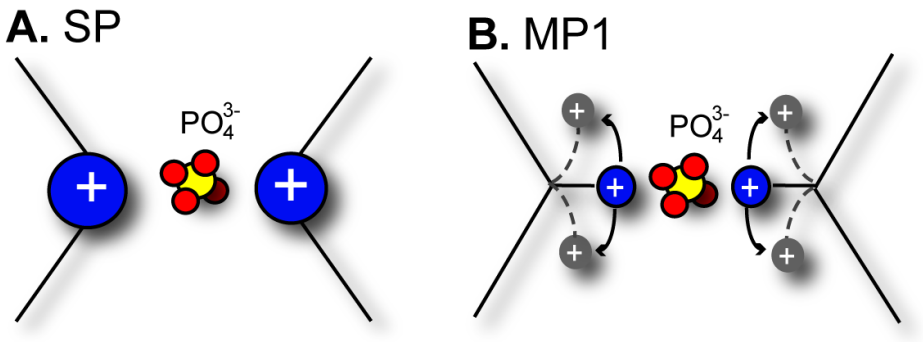


Figure 5-5 Cartoon views of pore-lining surface of SP (**A**) and MP1 (**B**) in the presence of Pi in the constriction site. The possible orientations of flexible sidechains in the constriction site of MP1 are shown in grey.

In case of the fully hydrophobic pore (HP), waters can also be observed inside HP. A similar phenomenon was seen in previous studies of hydrophobic nanopores. Beckstein *et al*. (168, 169) show that water can be conducted through narrow hydrophobic pores (radius ≤ 0.4 nm) with an increase in polarity. Therefore, the phosphate inside a pore can induce a number of water molecules to enter the

pore. Despite the fact that it is unfavourable for water to enter HP, ~ 7 water molecules are found to solvate Pi in the constriction site in association with ~ 6 pore contacts (Figure 5-4A). However, the loss of half hydration shell cannot be compensated by hydrophobic interactions with beads on the wall surface of HP resulting in the formation of a massive energy barrier to translocation through the pore.

It is also found that the replication of charged architecture of OprP in MP1 produces a similar PMF profile to that of OprP, but with > 2× deeper wells (Figure 5-6A). The deeper energy well found in MP1 may be due to the presence of only positively charged residues inside a pore whereas the OprP pore also contains ~ 16 of pore-lining acidic residues. Apart from the constriction zone, the other phosphate-favourable site in OprP, the "Lys sink" at the periplasmic mouth, can also be captured by the MP1 model (Figure 5-6A). On the contrary, it seems that the *fixed* charge architecture inside SP is insufficient to preserve the energetic signature of OprP. The bulk-like environment inside SP is important to stabilize Pi, but the high selectivity can be arrived when the Pi-dehydrating ability is at maximum. The differences in PMFs and hydration numbers of SP and MP1 demonstrate that the flexibility of pore-lining residues does matter. The mobile Arg- and Lys-mimicking sidechains show greater ability to shed the hydration shell of Pi yielding the more 'pore-like' energy profile and sequentially the conservation of PMF profile in MP1.

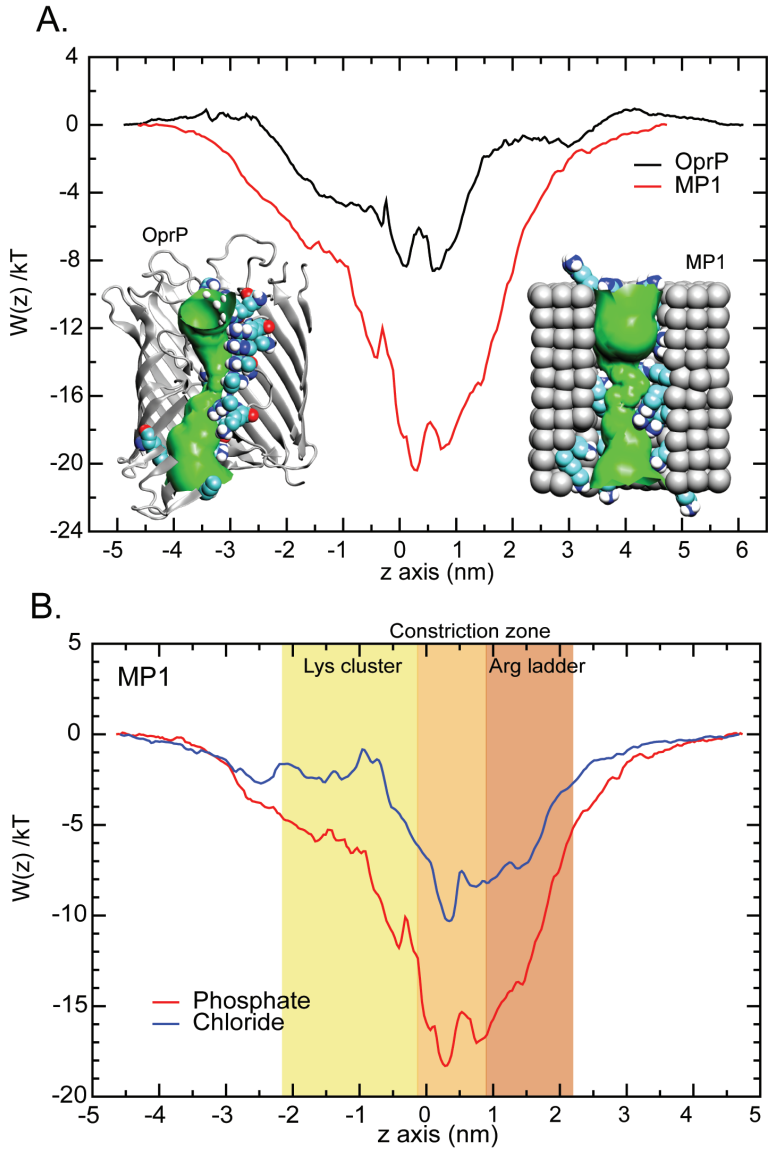


Figure 5-6 Comparison of phosphate PMFs of (**A**) MP1 (red) and the template OprP (black). The OprP PMF was aligned to the z coordinates of MP1. The PMF profiles of both phosphate (red) and chloride (blue) ion through MP1 model is shown in (**B**). The key regions are colored in yellow, orange, and brown, respectively.

To further probe selectivity properties of these nanopores, a Cl^- PMF was also computed. Comparision of the Cl^- and Pi PMFs for MP1 PMFs show that Pi binds more strongly in the constriction site than chloride due to the energy well of ~ 7kT deeper in the Pi PMF (Figure 5-6B). The exit vestibule, which mimics the Lys sink of OprP, is attractive for Pi but barely so for Cl^-. In OprP, the energy difference between Pi an Cl^- at the region of Lys cluster is one of the features that can differentiate these two ions. Hence, the energy gap at the same region in MP1

implies the ability of MP1 to differentiate between Pi and Cl^-. The K_d values as computed from the PMFs are 0.45 nM for Pi and 0.17 μM for Cl^- corresponding to ~ 400-fold stronger binding of Pi than Cl^- (Table 5-2). It appears that positively charged interior of MP1 facilitates a Pi transport. Furthermore, the presence of a solely positively charged interior increases the binding affinity of Pi dramatically relative to OprP. The simulation results of MP1 importantly reveal that relevant aspects of the function and selectivity of OprP can be transferred to a non-biological model by the duplication of functional architecture and pore geometry using a small number of key sidechains.

Model	K_d (μM)	Relative selectivity ($[K_d]_{Cl^-}/[K_d]_{Pi}$)
MP1 (Cl^-)	0.17±0.16	-
MP1 (Pi)	$4.5\times10^{-4}\pm1.9\times10^{-4}$	380

Table 5-2 Series of K_d as computed from the PMFs and relative selectivity of each model.

5.3.2 The effect of degrees of sidechain flexibility on the phosphate selectivity

As seen in the previous section, flexible sidechains showed a greater ability to desolvate a Pi anion than the nanopore with static charges. It therefore seems that the sidechain flexibility may be crucial for controlling the local environment of Pi, resulting in Pi selectivity. The MP1 nanopore was designed to energetically capture the major features for phosphate selectivity as seen in the natural template OprP. In order to study the effect of the degree of sidechain flexibility on Pi-binding affinity

and selectivity, additional models (MP2 and MP3) were constructed in which the length of sidechains were truncated compared to MP1 (Table 5-1 and Figure 5-7).

The pore sizes of MP2 and MP3 are similar to that of the OprP pore, although the cavity at the mouth region appears to be slightly larger than the OprP pore in all models (Figure 5-7C).

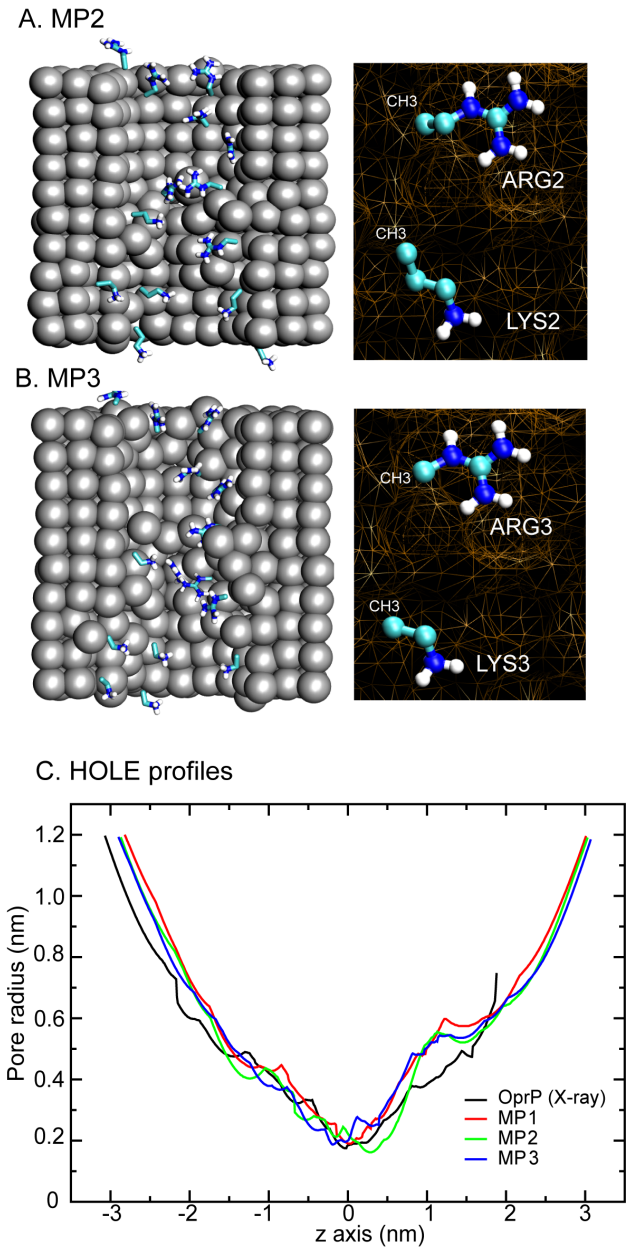


Figure 5-7 Cross sections of model MP2 and MP3 with modified amino acids are represented in (**A**) and (**B**) respectively. The chemical structures of all pore-lining residues in each model are shown on the right of (**A**) and (**B**). (**C**) The pore radii of all models in comparison with the original OprP.

Considering the PMFs of each model, MP2 and MP3 seem to slightly destabilize the Pi ion in the constriction site (Figure 5-8A). Apart from the more shallow energy well, the only obvious difference among these models is at the Lys sink (Figure 5-8A). Since interacting with a Lys sink is a part of the phosphate-selective mechanism, the loss of these interactions thus reduces phosphate-binding affinity. The most rigid MP2 and MP3 sidechains decrease the affinity to bind Pi. The computed K_d of MP2 and MP3 are 1.1 nM and 2.4 nM which are both higher than for the more flexible MP1 (0.45 nM) (Table 5-3). Consequently, the abilities to bind Pi of MP2 and MP3 are reduced by 2-fold and 5-fold from that of MP1, respectively.

Model	K_d (μM)	Relative selectivity $([K_d]_{Pi}^{MPn}/[K_d]_{Pi}^{MP1})$
MP2	$1.1\times10^{-3}\pm9.8\times10^{-4}$	2
MP3	$2.4\times10^{-3}\pm1.9\times10^{-3}$	5

Table 5-3 Series of K_d as computed from the PMFs and relative selectivity of each model.

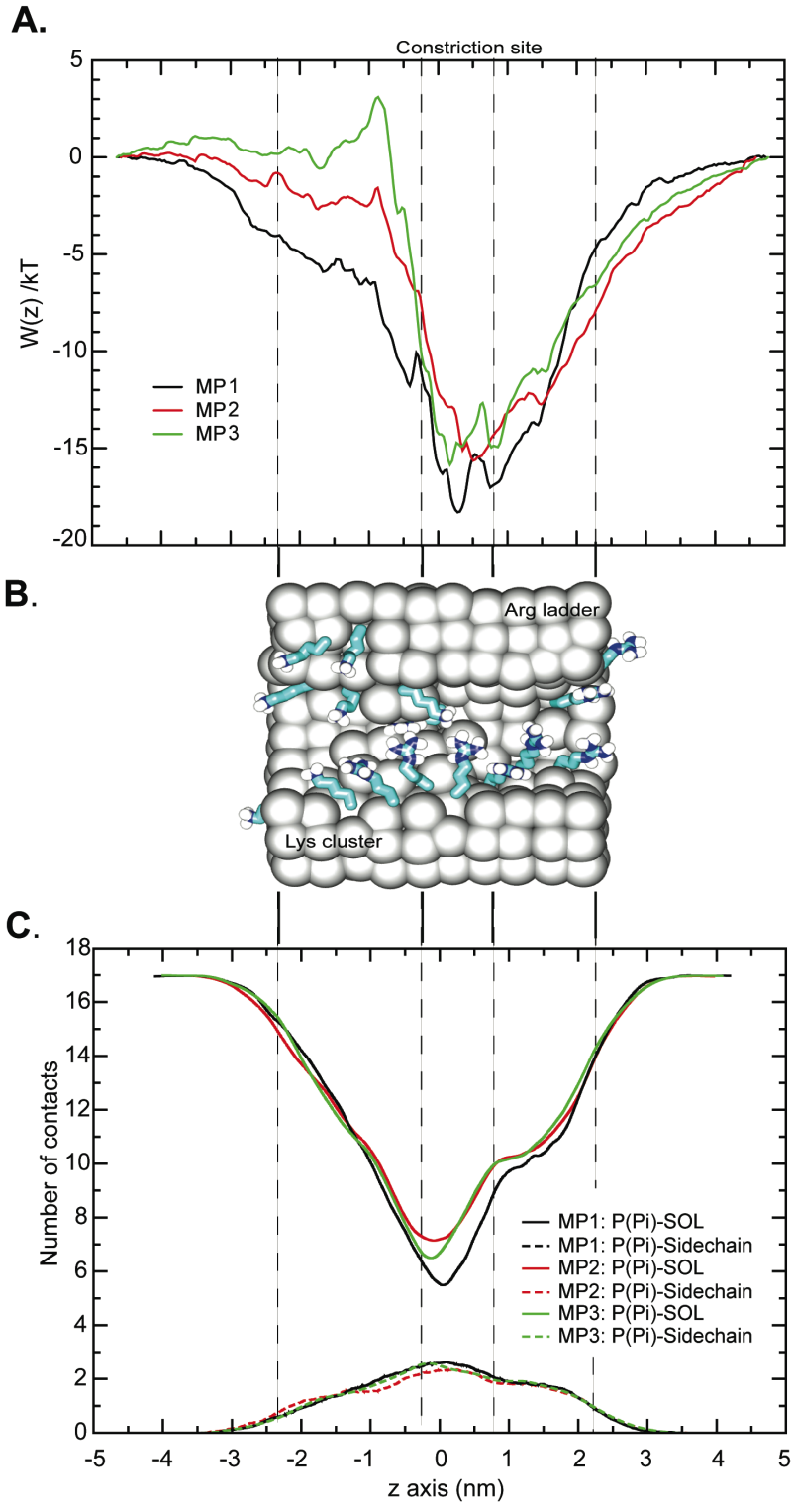


Figure 5-8 (**A**) Phosphate PMFs of pore models. (**B**) Cross-section of MP1 model. (**C**) Average coordination number of phosphate molecule in MP1, MP2, and MP3 with water molecules (SOL: solid line) and pore-lining sidechain (dash line) calculated from umbrella trajectories within the cut-off of 0.5 nm. MP1, MP2, and MP3 are colored in black, red, and green, respectively.

Since the environment of a translocating Pi is crucial to understand the permeation mechanisms, the number of contacts between Pi and its neighbours are investigated. The number of water contacts is gradually reduced in all models until the phosphate reaches the constriction site. In the constriction site, the hydration number is decreased to roughly 5 in MP1 and 7 in MP2 and MP3 (Figure 5-8B). This loss of water contacts has to be compensated by the interactions with the pore in order to stabilize Pi. In each model, 3 sidechains appear to coordinate Pi in the constriction site. Nevertheless, the major difference among nanopores here is the ability to dehydrate Pi. Overall, the hydration number of both MP2 and MP3 are similar to that of MP1 except in the constriction area. In MP1, the high degree of sidechain flexibility appears to show the greatest ability of Pi dehydration in the constriction site (Figure 5-8B). This suggests that the local environment of translocating Pi in a confined region can be controlled by the flexibility of surrounding sidechains. The calculated K_d values indicate that the Pi selectivity is proportional to the Pi-dehydrating ability.

5.4 Conclusions

The extent to which simplified model nanopores may capture the energetics of anion translocation through a biological channel has been examined. Applied to the OprP pore, the simplified models successfully describe the selectivity properties. The charged brush-like model (MP1) preserves the essential phosphate-selective characteristics of OprP. The SP model with static charges is able to show the anion-favourable feature, but it fails to represent Pi selectivity. In an analogous approach, experimental studies have mimicked a biological channel using synthetic nanopores with fixed charges (105, 162, 176). Those models show ionic selectivity, favouring the passage of counterions and avoiding the passage of coions (162), but the specificity for an ion species as found in biological channels (i.e. K^+ channel, Ca^{2+} channel, and OprP) are still absent. The results show that the existence of brush-like characteristics may be one of choices to increase selectivity. Moreover, it is found that the highly selective properties of biological channels can be transferred when their functional architecture and pore geometry are duplicated.

To obtain highly selective, the flexibility of pore-lining sidechains in a confined geometry is crucial. Inside the pore, the sidechains and confined water molecules compete to solvate a phosphate molecule. Especially in the constriction site, the narrow pore cavity with the flexible sidechains takes a great advantage over water molecules to dehydrate Pi and ultimately increase in Pi-binding affinity. The more rigid sidechains of MP2 and MP3 have less ability to desolvate Pi yielding the reduction in Pi selectivity. Apparently, the selectivity is considerably dependent on the ability of the pore to perturb the hydration shell of Pi where such process can be acquired by the increase in sidechain flexibility.

This finding allows us to elucidate the role of sidechain flexibility in ion selectivity, which may be useful to understand the selective mechanisms of other biological pores with brush-like behavior. Such analysis of the relationship between structure and function of pore models may reveal more general design principles for nanopores which could be exploitable from a technological perspective.

6 SIMULATIONS OF VANILATE PERMEATION THROUGH OPDK

6.1 Introduction

To conduct nutrients into the cell, gram-negative bacteria generally employ the water-filled and non-specific porins as an entryway. Atypically *P. aeruginosa* has no general porins (compatible to e.g. OmpC and OmpF of *E coli.*) (84). Instead, passage across the OM is mediated by *substrate-specific* porins (5). This class of proteins includes both trimeric and monomeric channels. They contain a binding site for their substrate that facilitates a diffusion process at low substrate concentration (5). To date, the OprD family is the largest family of substrate-specific channels currently known (5). In *P. aeruginosa*, this family contains 19 members all of which share ~ 45% - 60% sequence similarity (84). These 19 members cluster into two distinct subfamilies (177). The OprD subfamily containing 8 members has been characterized in *P. aeruginosa* to date. This includes the glucose-specific porin OprB, the phosphate- and polyphosphate-specific porins OprP and OprO, respectively, and the basic amino acid–specific porin OprD, which was also found to be an entryway for antibiotics (177, 178). The other group, the OpdK subfamily, contains 11 members, most of which have not been characterized (5). It was discovered that the OpdK subfamily is involved in the uptake of small aromatic compounds (177). Though the OprD family is known to be important for the proliferation of pseudomonads, unfortunately there is still a lack of information regarding structure and function of these proteins. Recently, the first structures of OprD and OpdK were determined (5, 22). The X-ray structures of both proteins revealed an interior architecture that may aid substrate translocation. However, the structural basis of substrate selectivity cannot be elucidated solely from the *static*

protein structures. Thus, MD simulations have been employed to study dynamic aspects of the structure-function relationship.

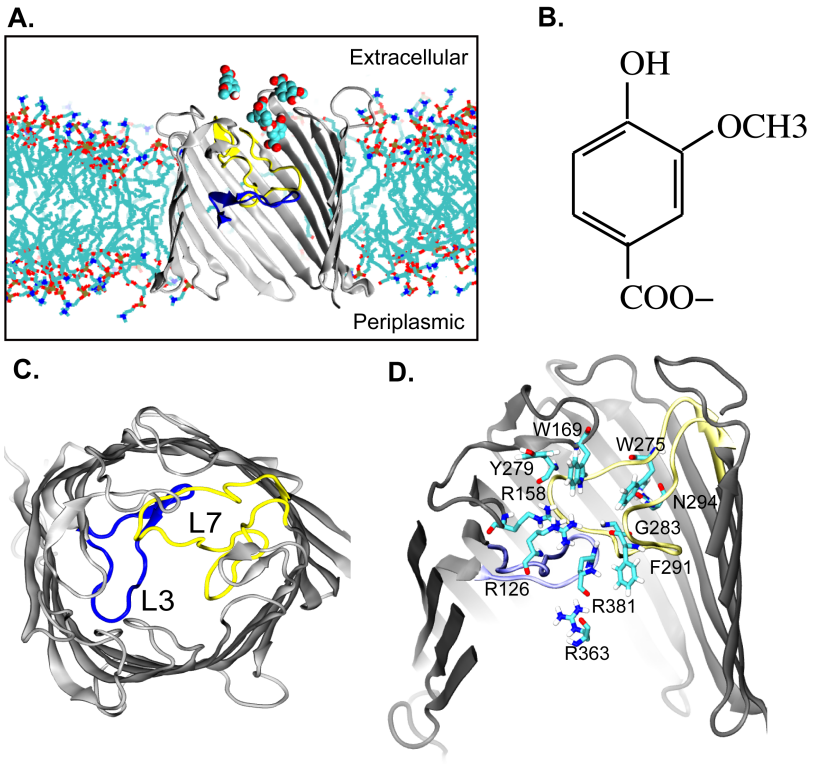


Figure 6-1 (**A**) Cross section of monomeric OpdK embedded in a DMPC bilayer in the presence of four VNL molecules at the extracellular mouth. The lipid molecules are in cyan (tails) and red/blue/brown (headgroups). Water and counterions have been omitted for clarity. The covalent structure of VNL is shown in (**B**). (**C**) Top view of a OpdK pore. Two loops (L3 and L7) are folded into the lumen to form the constriction site. L3 and L7 are colored in blue and yellow, respectively. Some key residues for VNL permeation are shown in (**D**).

Both OprD and OpdK form an 18-stranded β-barrel with a narrow pore in a constriction site (5). Each contains two long loops L3 and L7 folded inward to form a constriction site at the middle of the pore (Figure 6-1). Although OpdK and OprD share a similar basic architecture, the larger pore with a different pore-lining shape of OpdK provides the differences in substrate specificity (5, 22). OpdK was found to be involved in the uptake of vanillate (VNL) and small aromatic compounds (177). A VNL molecule (Methyl 4-hydroxy-3- methoxybenzoate) comprises an aromatic ring

and a carboxylate group with a total charge of -1 (Figure 6-1B). To conduct VNL, the OpdK pore contains a string of positively charged residues (R126, R158, R363, and R381) on one side and aromatic and polar residues (W169, W275, Y279, and N294) on the other site at the L7 region (Figure 6-1C and D). In this study, I focus only on OpdK due to the availability of a more complete X-ray structure. The only published X-ray structure of OprD has missing residues at the constriction site and it is not clear if homology modelling would yield a reliable structure. Thus, I use OpdK as a representative of this family to study the mechanism of substrate translocation.

As a starting point to understand the relationship between substrate translocation and protein dynamics, I performed multiple MD simulation approaches: (i) Steered MD (SMD) and (ii) an applied external electric field. In both of these techniques, an external force is applied to the system to drive the permeation process to completion within the simulation timescale. Both of these methods will be employed to explore the mechanism of VNL translocation through the OpdK pore.

6.2 Methods

The OpdK (2QTK, resolution 2.8 Å) crystal structure consisting of 390 amino acids contains 9 extracellular loops from which there are some missing residues in L1 and L2. The missing loop regions (residue 1-3, 25-30, 70-75, and 84-87) were modelled using Modeller (179). The OpdK monomer was embedded in a pre-equilibrated dimyristoyl-phosphatidylcholine (DMPC) bilayer (pre-equilibrated by running a 2 ns simulation). The solvent-accessible molecular surface of OpdK was used as a template to remove lipids and perform short steered MD simulations of the solvated proteins as described in detail by Faraldo-Gomez *et al.* (140). This generated a cavity into which the protein was inserted. Water and neutralizing counter ions were then added to each system (see Table 6-1). Ions were added by

randomly replacing water molecules. A 0.2 ns equilibration of protein-restrained dynamics was conducted. The initial systems were generated by randomly adding 4 VNL molecules (parameters of VNL have been modified from Tyr) near the extracellular mouth (Figure 6-1A) and followed by 600 steps of energy minimization and a 1 ns equilibration run with restrained VNLs. The 20 ns production run (MD0) was primarily conducted in order to study the possibility of observing VNL translocation within typical equilibrium MD timescale. However, a 20 ns simulation was not long enough to observe full VNL permeation through the OpdK pore. To accelerate the transport process, two other methods were subsequently used: (i) SMD and (ii) applying a constant electric field across the membrane. With an electric field, the same starting coordinates as for the MD0 system were used. A constant electric field of 0.15 V/m was applied perpendicular to the membrane plane in the +z direction. In practise, this fixed electric field drives a charged molecule (VNL in this case) to penetrate through the protein pore. A 1M NaCl solution was also added to the system. The applied electric field simulations were performed 10 times, but full VNL permeation was observed in only 4 of these. Only 2 simulations are analyzed here, as pore distortion was observed in the remaining 8 simulations. A strong electric field caused not only severe distortion of the barrel wall, but also shifted the positions of the L3 and L7 loops (Figure 6-2).

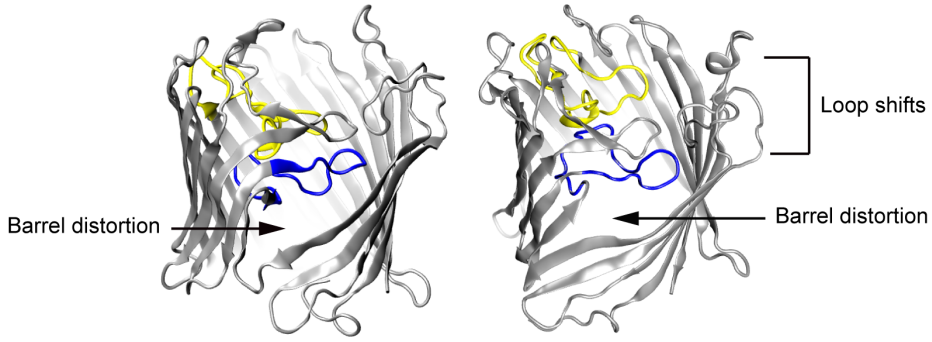


Figure 6-2 Distortions of protein structure due to applying high strength of electric field (0.15 V/m).

In the case of SMD simulations, constant-velocity SMD simulations were performed for 10 ns each to pull the VNL in the ±z directions. A VNL molecule was initially placed at the center of the pore. An elastic spring with a force constant of 500 $kJ \cdot mol^{-1} \cdot nm^{-2}$ was attached to the centre of mass of the VNL and was moved in the direction parallel to the channel axis (*z* axis) at a rate of 0.25 nm/ns.

Additionally, to reveal the orientation and behaviour of VNL in the constriction site, 3 MD runs (MD1, MD2, and MD3; see Table 6-1 for details) were setup by placing a VNL molecule at the center of OpdK pore with the different orientations followed by a 600 step energy minimization and a 1 ns equilibration run with a restrained VNL position.

Simulation	No. of VNLs	No. of counterions	Duration (ns)	E field (V/m)
MD0	4	9	20	-
MD1	1	6	20	-
MD2	1	6	20	-
MD3	1	6	10	-
V1,V2	4	9	10, 20	0.15
SMD	1	6	10	-

Table 6-1 Summary of MD simulations

All simulations were performed using GROMACS 3.3 (www.gromacs.org) (141) with an extended united atom version of the GROMOS96 forcefield (91). To relax steric conflicts generated during setup, energy minimizations used up to 1000 steps of steepest descent. Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method (87) 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 (*NPT*) ensemble. The temperature of the protein, DMPC, solvent, and ions were each coupled separately using the Berendsen thermostat (96) 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 analyzed by GROMACS and locally written code. Molecular graphic images were prepared using VMD (142).

6.3 Results and Discussions

6.3.1 VNL behavior at the extracellular mouth

Unlike the highly charged phosphate ion in Chapters 4 and 5, VNL, which contains an aromatic ring with the total molecular charge of -1, struggled to reach the OpdK vestibule on the 20 ns MD timescale. The MD0 simulation shows that a VNL molecule experiences difficulty in penetrating into the constriction site, even though the set of VNL molecules (VNL1, VNL2, VNL3, and VNL4) was initially placed at the extracellular mouth adjacent to the constriction area (Figure 6-1 and Figure 6-3A and C). It appears that only VNL1 can sit in the extracellular mouth at z ~ 1 nm for 20 ns (Figure 6-3A; black line). The initial orientation of VNL1 backbone is parallel to the z axis (Figure 6-3C and D), but within just 1 ns, the VNL backbone reoriented such that it is now aligned above the L3-L7 interface (z ~ 1 nm), where its carboxylate group points towards the positively charged patch (R126 and R158) and the aromatic ring faces the polar and aromatic residues (W169, W275, and N294) in the opposite direction (Figure 6-3E). VNL1 remains in this position for the rest of the simulation. Since the initial orientation of VNL1 backbone is parallel to the z-axis, where the carboxylate group of VNL points towards the pore axis, it is possible that the R126 and R158 sidechains underneath interact with the COO^- group and then drag the entire molecule into the upper L3-L7 region. Additionally, a decrease in pore size is observed in the presence of VNL at z = 1 nm (Figure 6-3A and B). Both the reorientation of VNL and pore reduction possibly reveal the first step of the mechanism of selection. The ‘correct’ orientation of VNL is essential to stabilize the VNL-protein complex when it approaches the L3-L7 region. The most favourable orientation of VNL at the upper L3-L7 interface is where the carboxylate group points towards the Arg patch while the rest of a molecule is aligned with the opposite

side. Also, it is observed that binding of VNL results in the reduction of pore size (Figure 6-3B). The final snapshot of MD0 (Figure 6-3E) shows the role of the tip of the L7 loop in controlling the pore size. At 20 ns, the tip of L7 appears to point towards the pore surface causing the narrowest pore size (Figure 6-3B and E). The movement of L7 is further highlighted by an increase in RMSDs of L7 (Figure 6-3F and G). Although the overall difference between the RMSDs of L3 and L7 is small (~ 0.1 nm), the sudden increase in RMSDs of L7 is sufficient to indicate the movement of the tip of L7 when compared with the constant RMSDs of L3 (Figure 6-3F and G).

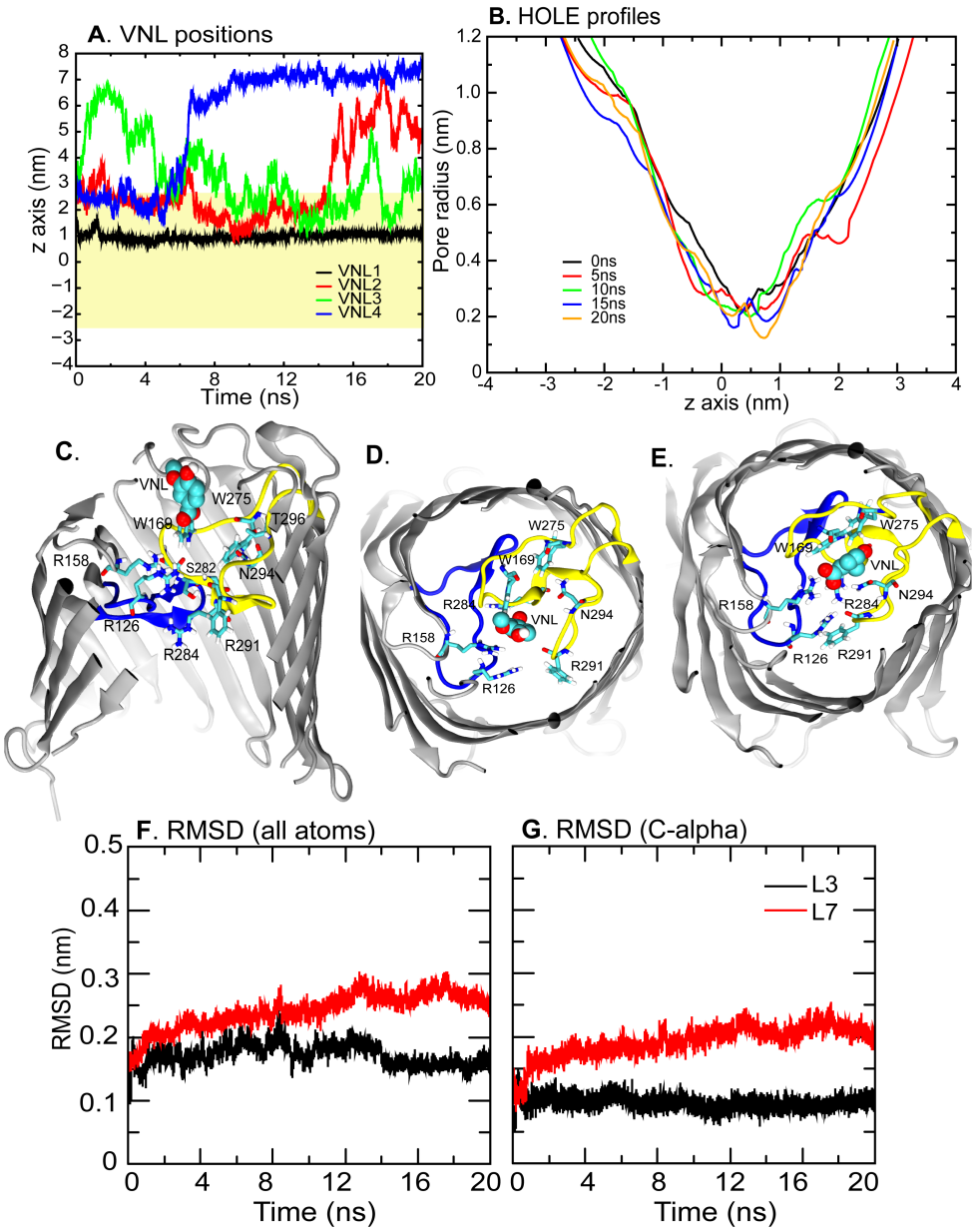


Figure 6-3 (**A**) VNL trajectories in the z direction from a 20 ns equilibrium MD simulation (MD0). The protein length is shown as a yellow box. (**B**) Pore radii of OpdK during the simulations. (**C**)-(**E**) Conformations of key residues with VNL (in spacefilling format). An initial orientation of VNL is shown as side and top views in (**C**) and (**D**) respectively. The final VNL conformation is shown in (**E**). (**F**) and (**G**) represent RMSDs of all atoms (**F**) and Cα atoms (**G**) of L3(black) and L7(red) loops.

Residue	Interaction Energies (kJ/mol)							
	Equilibrium MD				V1 (VNL1)	V2 (VNL2)	SMD	
	MD0	MD1	MD2	MD3			To EC	To PP
W169	**-27.94** ±8.79	**-2.35** ±3.49	**-6.18** ±2.92	-0.48 ±0.46	-0.98 ±3.01	-0.03 ±0.01	**-6.46** ±7.47	0 ±0
W275	**-6.27** ±3.23	-0.06 ±0.13	-0.85 ±0.75	-0.18 ±0.14	-0.01 ±0.01	0 ±0	-0.12 ±0.22	0 ±0
Y279	-1.54 ±1.30	-0.21 ±1.17	0 ±0.01	-0.01 ±0.03	-0.62 ±2.08	-0.07 ±0.56	**-6.60** ±11.96	0 ±0
F291	-0.30 ±0.59	**-4.33** ±6.47	**-21.65** ±6.70	**-10.27** ±3.40	**-4.32** ±6.32	-0.12 ±0	**-4.66** ±5.62	-0.59 ±2.18

Table 6-2 Average interaction energies between VNL and the key aromatic residues in the constriction site from three types of simulations with standard deviations. The strong interactions with VNL are in bold.

Apart from the change in the loop orientation, the key interactions between VNL and neighbouring residues above the L3-L7 interface are also investigated. The presence of hydrogen bonds between VNL and its adjacent residues is calculated as a function of time (Figure 6-4A). The orientation of VNL above the L3-L7 region is influenced by strong hydrogen bonds with R158 and W275 (Figure 6-4A). A COO^- group of VNL firmly interacts with R158, while a OH group on the opposite site favours an interaction with W275 (Figure 6-3E). This analysis also shows the close contact between VNL and other residues on L7, especially a strong interaction with N294 (*see* Hydrogen bonds of residue ~ 280 to 294 in Figure 6-4A). Not only do charged sidechains but also aromatic residues play a key role in the formation of a tight VNL-protein complex. Apart from a hydrogen bond with W275 leading to interaction energy of -6.27 kJ/mol (Figure 6-4B and Table 6-2; MD0), some of the major attractive forces are from W169 and F291. Even though the π-π interaction is undefined in classical MD, the parallel orientation of aromatic rings of VNL and

W169 and F291 are still observed. This face-to-face conformation leads to the maximum favourable interaction energy of ~ -20 kJ/mol in both cases (Figure 6-4B and Table 6-2).

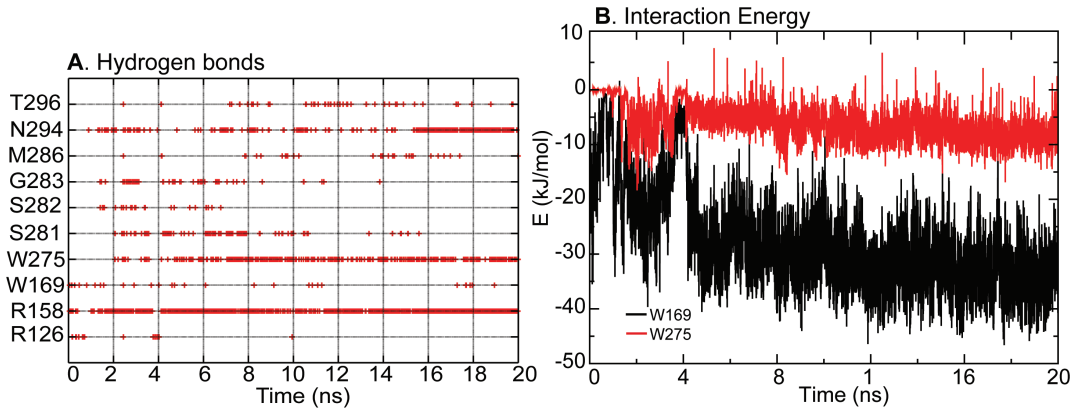


Figure 6-4 (**A**) Hydrogen bonds as a function of time. (**B**) The interaction energy between key residues and VNL.

It appears that the upper L3-L7 interface is the first confined region for VNL. The Arg patch and some aromatic and polar residues at the opposite end may play a role in substrate recognition. Admittedly, when a region above L3-L7 is occupied by VNL, the interactions between VNL and residues on L7 introduce the movement of the tip of L7 that leads to a decrease in pore size. Reducing the pore size may be one of the mechanisms employed by the protein to prevent unwanted molecules travelling though the pore without verification.

6.3.2 The orientation of VNL in the constriction site

The MD0 simulation has revealed the possible orientation of VNL above L3-L7 region where a COO^- group points towards an Arg patch (R158 and R126) and an aromatic ring is aligned to an opposite end. Considering the architecture of the Arg patch, R158 is adjacent to R126 and R381 underneath (Figure 6-1D). The location of these arginines suggests that their interactions may play a major role in guiding the

whole molecule to a constriction or L3-L7 region. One possibility is that a COO^- group of VNL is directed by these arginines to a constriction site. Therefore, the orientation of VNL in a constriction site should be similar to that found in above region. To support this hypothesis, further equilibrium MD simulations (MD1, MD2, and MD3) with different initial VNL orientations were performed. The initial VNL orientation in MD1 and MD2 was set to the opposite direction to the one observed in the previous section (Figure 6-5C and D), such that the aromatic ring is placed next to the Arg patch and the COO^- group points towards the aromatic region. A VNL molecule in MD1 and MD2 is placed at $z \sim 0.6$ nm and 1 nm respectively (Figure 6-5B). In case of MD3, the VNL orientation is preserved from the previous section and placed at $z \sim 0$ nm (Figure 6-5B and E).

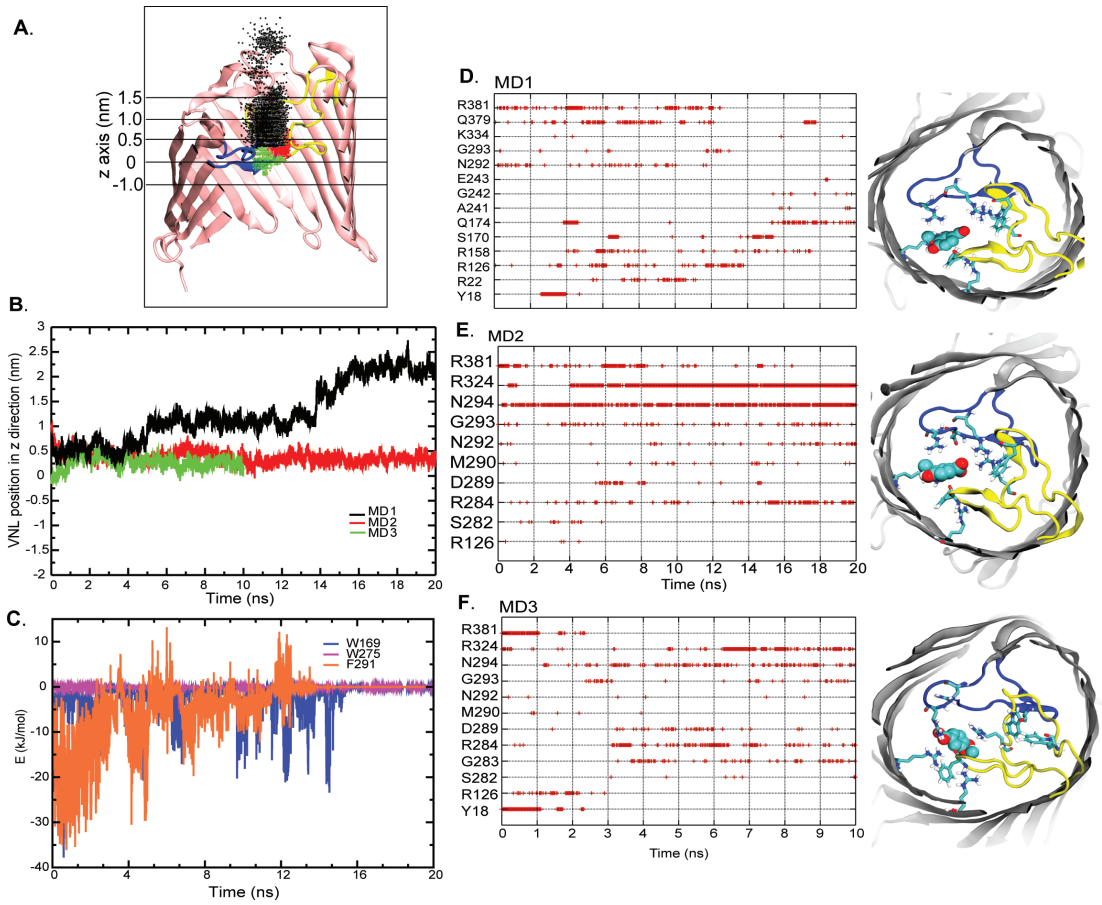


Figure 6-5 (**A**) Cross section of monomeric OpdK showing L3 (in blue) and L7 (in yellow) and defining the z-axis coordinates. The positions of VNL in each simulation are also shown as a cloud where MD1, MD2, and MD3 are colored in black, red and green respectively. (**B**) VNL trajectories from MD1, MD2, and MD3 simulations for a VNL at the constriction site. (**C**) Interaction energies between VNL and aromatic residues in the constriction site. (**D**)-(**F**) Hydrogen bonds between VNL and pore-ling residues in the constriction site (*left*) and the starting conformation of a VNL molecule from each simulations (*right*).

The simulations show that the VNL molecule in MD1 is left for the extracellular vestibule within 4 ns while a similarly oriented VNL in MD2 is remained in the constriction region for 20 ns (Figure 6-5A). It is because that a COO^- group of MD2 immediately reorients itself towards the Arg patch, which is in the same orientation as found in the region above L3-L7 (Figure 6-3E). The reorientation of MD2 confirms the stable conformation of VNL at the constriction site. In MD1 the VNL fails to adapt its conformation to fit into the pore cavity, therefore it is displaced to the extracellular vestibule. In the case of MD3, VNL still sits in the

constriction site although it is driven upwards from z~ 0 nm to z~ 0.25 nm (Figure 6-5A). It appears that in both MD2 and MD3 the VNLs are shifted to the same region at z ~ 0.25 nm (Figure 6-5A) and remain there for the rest of the simulations. At z = 0.25 nm the tip of L3 meets that of L7.

Hydrogen bond analyses of residues that are in close contact with VNL during the simulations are also performed (Figure 6-5C, D, and E). The list of neighbouring residues here also reflects the environment of VNL. It is found that almost all residues adjacent to VNL in MD2 and MD3 are from L7. It appears that L7 is the major contributor to the formation of VNL-protein complex in the constriction zone. The VNL molecules in MD2 and MD3 are found to interact with R284, R324, N294, and R381, especially in MD2 where strong interactions are formed with N294 and R324 (Figure 6-5D and E). It requires a VNL to reorient so as to place its COO^- group adjacent to the Arg patch. MD1 fails to do so, hence the loss of these interactions results in VNL extrusion. Unlike the interactions at the upper L3-L7 region where aromatic residues play a key role in VNL binding, I observe that charge-charge interactions dominate VNL binding in the constriction region. This is indicated by low interaction energies between VNL and aromatic residues in Table 6-2 (MD2 and MD3).

The COO^- group of VNL is found to oscillate within the constriction zone. Three VNL rearrangements are captured from MD2 and MD3. The backbone of VNL is obviously found to align with angles of 60°, 120°, and 180° to the z-axis (Figure 6-6). In MD3, VNL remains constant at an angle of 180° after 10 ns (Figure 6-6). However, these backbone rearrangements can be categorised into two major orientations, almost parallel and perpendicular to a z-axis. It appears that the

oscillation is driven by the interactions of the COO^- group with various arginines in the constriction site.

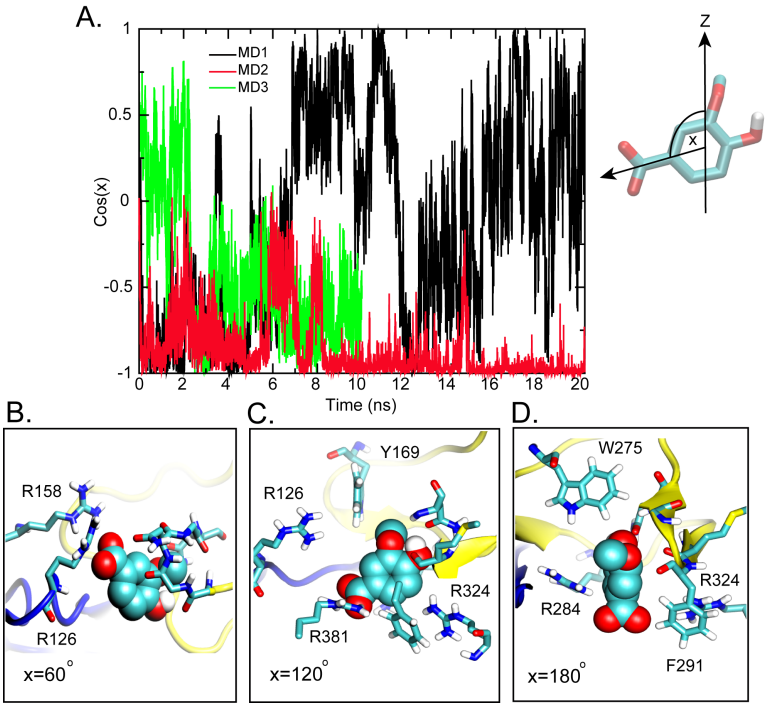


Figure 6-6 Orientation of VNL in OpdK constriction site. (**A**) Time-evolution of the angle between a VNL spine and z axis from MD1 (black), MD2 (red), and MD3 (green) simulations respectively. (**B**)-(**D**) Dominant orientation of VNL (in spacefilling format) in the constriction area (0.5 nm < z < -0.2 nm).

The simulation results here reveal not only the dynamics of VNL but also its orientation in the constriction site. The VNL molecule is stabilized by the charged interactions between its carboxylate group and the Arg patch. Also, its aromatic ring interacts with polar amino acids mostly on L7. The hydrogen bond analysis demonstrates that loop L7 is important for VNL binding. The orientation of the COO^- group pointing towards the Arg patch appears to be the most favourable here. Although these simulations cannot capture the VNL translocation from the upper L3-L7 region to the constriction site, the stable orientation of VNL found in both regions perhaps provides one step along the translocation pathway. The oscillation of the COO^- group of VNL during the simulations suggests the role of the Arg patch in

directing a COO^- group to the desired direction. However, further studies are needed to observe this transition.

The full VNL translocation cannot be observed by conventional equilibrium MD simulations. In the next section, I have introduced another technique where an external fixed electric field is applied to the system to drive VNL molecules through the OpdK pore. Also, the results from an applied electric field are compared to those from SMD.

6.3.3 VNL permeation through a OpdK pore

6.3.3.1 An externally applied electric field

An external electric field of 0.15 V/m, corresponding to a voltage of 450 mV across the membrane, was applied to the OpdK system in the +z direction to accelerate the VNL permeation to the periplasmic side. Although a voltage of 450 mV is a little high when compared with that used in experiments, this magnitude is computationally at a minimum that permits VNL transport in reasonable timescales. The simulation was repeated to give a total of 10 simulations. The VNL translocation is observed in four simulations, but two of them have severe protein distortion (Figure 6-2). Therefore, only two runs are investigated here (V1 and V2). Only one VNL is able to reach the periplasmic side in V1 while 3 VNLs succeed in reaching the periplasmic end in V2 (Figure 6-7A and B). Though the VNL translocation can be observed by applying such a strong electric field here, the pore distortion and structural change that may affect the transport process must be considered. Not only is the VNL molecule, but also the orientation of all charged sidechains are controlled by the strong electric field across the membrane. Under the electric field, all positively charged sidechains point to the extracellular side, while the negative

sidechains point towards the opposite direction. This sidechain rearrangement interrupts not only the structural stability, leading to protein distortion, but the protein function as well. The electric field significantly drives all charged residues at the constriction zone to reorient and leads to a dramatic pore expansion in V1 and V2 (Figure 6-7C and D). The loss of interactions between translocating VNL and pore-lining residues in V1 and V2 indicates that the major driving force for VNL transport is from the electric field, not the OpdK pore itself (Figure 6-7E-H). Especially, in a wider pore VNL1 in V1 and VNL3 in V2 are driven though the pore with only few protein contacts (Figure 6-7E and G). Besides, VNL2 and VNL3 in V2 translocate though the pore while VNL1 is trapped at z = 1 nm and released from the constriction site after both VNL2 and VNL3 reach the periplasmic mouth. It is uncertain if this is a genuine biological phenomenon or an artifact of the method used.

Despite the fact that applying an electric field perturbs the pore geometry, a quantitative interaction network can still be extracted from these simulations. Only interactions in the constriction site are captured here. The hydrogen bond analyses highlight the role of L7 in VNL binding. Apart form R126 and R158 in the Arg patch, VNL interacts with residues located in L7 region (Figure 6-7E-H). Additionally, an interaction with R121 in both V1 and V2 is also observed when VNL leaves the constriction site. The interactions with aromatic residues are also significant. W169, W275, and Y279 located in an upper L3-L7 region favourably contribute to acquire a free VNL molecule, especially W169. F291 at the tip of L7 is also found to interact with a VNL in the constriction area (Table 6-2: V1 and V2).

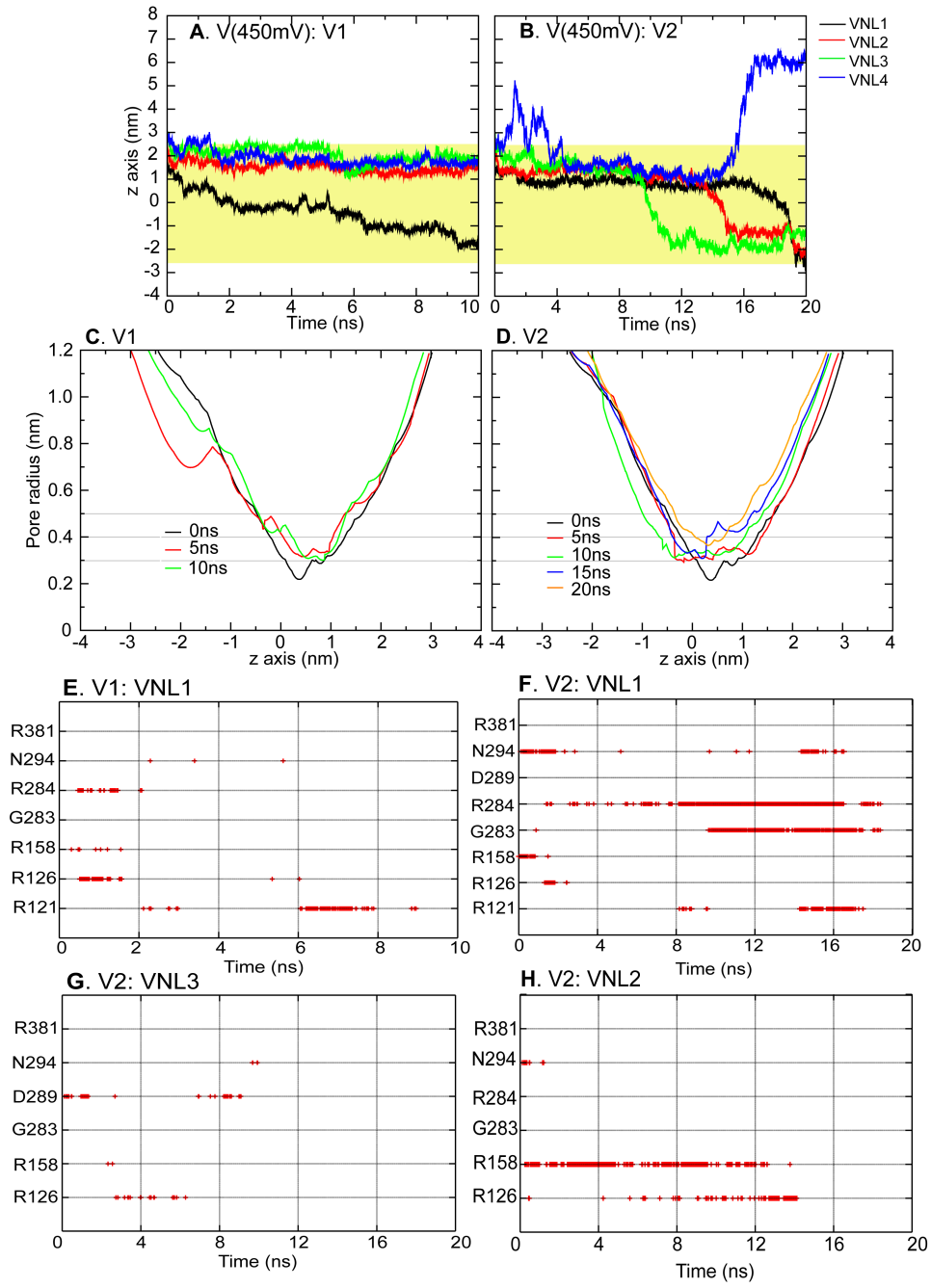


Figure 6-7 (**A**) and (**B**) VNL trajectories under the condition of an externally applied electric field (V1 and V2). The protein region is shown as a yellow box. The time-evolution pore radii of V1 and V2 are in (**C**) and (**D**). The hydrogen bond analyses between each translocating VNL and its neighbouring residues are shown in (**E**)-(**H**)

The study indicates that an electric field is not suitable for short simulations with a low-charged and large solute such as the VNL molecule. To drive a large molecule with small charges though a narrow pore, a higher electric field is needed to gain enough forces to push a solute though the desired direction. However, this strong electric field not only advances a VNL transport, but also distorts the protein

secondary structure. To better understand the translocation process without possible artifacts, I therefore performed SMD in the next section.

6.3.3.2 Steered Molecular Dynamics simulations (SMD)

In equilibrium simulations (Figure 6-3 and Figure 6-5) VNL molecules cannot penetrate to the periplasmic end within 20 ns. In the SMD simulations (duration 10 ns; Figure 6-8) VNL molecules are pulled in either direction from the central binding site. In each case the VNL molecule is slowly released from this site (~1 ns), interacting with other sites before exiting through either the extracellular or periplasmic mouth of the pore. Pulling VNL to the periplasmic side (red line in Figure 6-8B) or in the reverse direction (black line in Figure 6-8B) reveals an increase in the force profiles on the VNL molecule of magnitude 100 $kJ \cdot mol^{-1} \cdot nm^{-1}$ at $-1 < z < 0.5$ nm (Figure 6-8B). This indicates the presence of the major VNL-binding site at the constriction region.

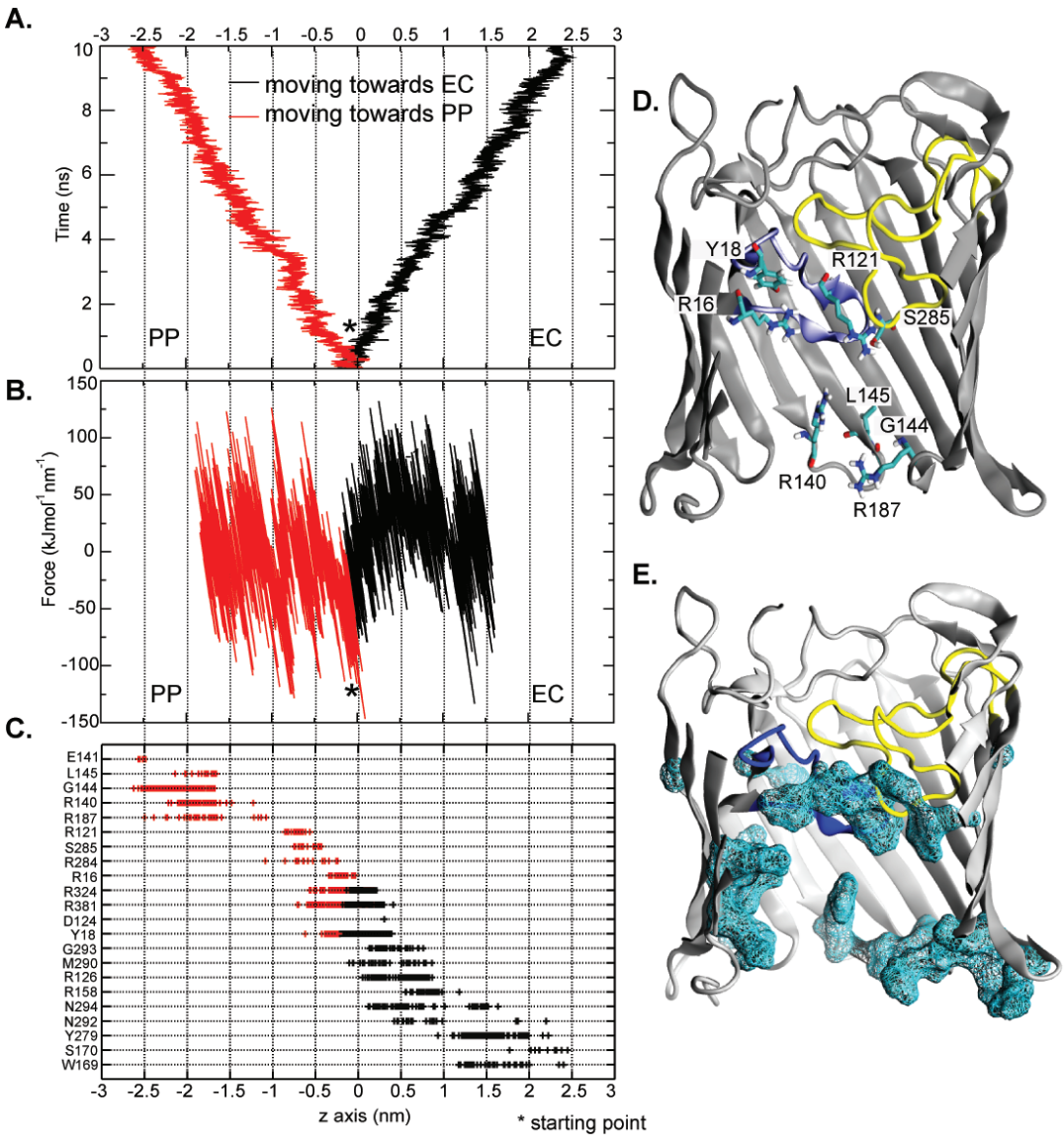


Figure 6-8 (**A**) Trajectories from SMD for a VNL initially at the center of the pore move towards extracellular (EC; black) periplasmic (PP; red) mouth of the pore. (**B**) Force profiles from constant-velocity SMD simulations of VNL through the OpdK pore. (**C**). Hydrogen bonds between VNL and the pore-lining amino acids along the pathway as a function of position of the VNL on the z-axis. The length of periplasmic area is colored in gray. (**D**) Cross section of the OpdK pore reveals pore-lining residues at a periplasmic end. (**E**) Positively charged residues at the periplasmic vestibule in a cyan wireframe format.

The SMDs reveal not only the major VNL-binding site, but also the possible interaction pathway along the OpdK pore. When a free VNL enters the extracellular mouth, it interacts strongly with two aromatic residues, W169 and Y279, with interaction energies of -6.46 and -6.60 kJ/mol (Figure 6-8C and Table 6-2; SMD). This set of interactions indicates the major role of aromatic residues in attracting a

free VNL, where this agrees well with the equilibrium MD results in the previous section. The electrostatic interactions become dominant when a VNL molecule diffuses close to the constriction site. R16, R126, R158, R324, and R381 are found to form strong interactions with a translocating VNL (Figure 6-8C). In particular, R158 appears to interact with VNL during the transition from the extracellular region to the constriction zone (Figure 6-8C). Moreover, a set of interactions from residues located in L7 (residue 290 to 294 in Figure 6-8C) reveals a key role of L7 in VNL binding at the center of the pore, which is in a good agreement with my equilibrium MD results. Besides, R16 and Y18 located on the barrel wall under the tip of L3 also contribute to the VNL binding at a lower L3-L7 area (Figure 6-8C and D). It is also found that the VNL interacts with R121, R284, and S285 before leaving the constriction chamber (Figure 6-8C and D). At the periplasmic end, VNL interacts with R140, R187, and backbones of G144 and L145. However, this may not be the only possible route to reach the periplasmic end of the pore. A positively charged cluster along the periplasmic chamber suggests multiple routes to exit the pore (Figure 6-8E). Further studies are needed to elucidate the behaviour of VNL in the periplasmic outlet.

6.3.3.3 SMD vs. An constant electric field

The results indicate that the entire process of VNL translocation cannot be captured by equilibrium MD on a timescale of ~20 ns. In this chapter, SMD or an application of an external electric field were used to observe this process. In SMD, the VNL molecule is pulled with the maximum exerted force of ~ 100 $kJ \cdot mol^{-1} \cdot nm^{-1}$ (Figure 6-8B). Based on the equation of electric force, F=qE, where q is the charge of molecule of interest (C) and E is the electric field (V/m), the electric field used in this study (0.15 V/m) can be converted to the force of 144 $kJ \cdot mol^{-1} \cdot nm^{-1}$. Despite the fact that both SMD and an applied electric field generate a force of comparable

magnitude on the VNL molecule, the external electric field fails to reveal the detailed mechanism of VNL transport. The high strength of voltage used in this study plays a major role in protein distortions (Figure 6-2). Under the condition of a constant electric field, a voltage difference is generated over the whole system, which duplicates an applied voltage in experiments (solid line in Figure 6-9). With the constant electric field in the +z direction, all negatively charged particles in the system experience electrical forces exerted on them in the –z direction and in the reverse direction for an opposite charge. Consequently, the mobile VNL molecule is then driven through the OpdK pore (Figure 6-9), where the possible permeation pathway can be obtained by continuous forces exerted on the carboxylate group of VNL (dash line in Figure 6-9)

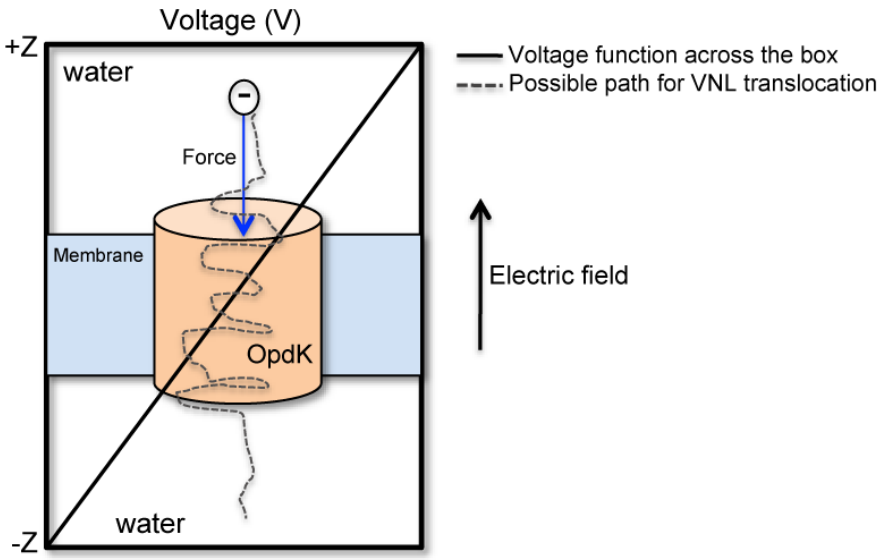


Figure 6-9 Cartoon representation of VNL trajectories under the condition of a constant electric field.

Since the magnitude of electrical force is proportional to an applied electric field, the higher electric field applied, the larger the force exerted on charged particles becomes. In simulations with a high-strength electric field, VNL molecules can translocate through the OpdK pore with few protein contacts (Figure 6-7). This

implies that the electric field of 0.15 V/m is too high that the VNL can be linearly driven through the Opdk pore by the sole electric force. Besides, the orientation and movement of all charged sidechains are strictly dictated by a large electric field which leads to the protein distortions found in this study. In previous study, the electric field was applied to the simple system with a fragment of voltage sensor of KvAP channel embedded in a membrane (180). They illustrated that the protein structure and function could be maintained, even up to 5 V of an electric field. My studies appear to give the opposite outcome. Considering the system in a previous study (180), almost all charged residues located in helices of a voltage sensor are buried in a membrane. The membrane insulator gives adequate protection against the high-strength electric field where this allows charged residues to function. On the other hand, all functional charged residues lining along water-filled pore surface in my system encounter the same constant field as in a bulk region leading to the confined sidechain orientations and leading to misleading results. In a short MD timescale (~10 ns), SMD appears to give more meaningful results although both SMD and an applied electric field are non-equilibrium techniques. With the limit of simulation time, using such a high-strength electric field significantly interrupts the native orientations of all charged particles in my system, which drives the system further away from equilibrium. This will then hide the occurrence of the actual phenomenon.

6.4 Conclusions

The MD0 simulations show that the upper L3-L7 region serves as a first dock for VNL. The orientation of VNL plays a key role in this docking. To avoid an extrusion, the VNL conformation has to fit the Arg patch and a series of aromatic and polar residues at the opposite end in the upper L3-L7 region (Figure 6-3E). Both Arg patch and aromatic residues are crucial for VNL recognition. A COO^- group on VNL is stabilized by the Arg patch while the rest of VNL molecule is in contact with aromatic and polar residues on L7 at the other end. Both L3 and L7 in the constriction site contribute to VNL binding, but only L7 has a significant role in both stabilizing an aromatic ring of VNL and controlling the dimension of the pore cavity. It is found that the motion of the tip of L7 contributes to a change in pore size. Furthermore, hydrogen bond analysis from SMDs reveals a role of positively charged patch in threading a VNL molecule to the constriction site. Besides, the MD results suggest the dominant role of aromatic (W169 and W275) and polar (N294) residues in attracting a free VNL from an extracellular environment. This finding is supported by the electronegative extracellular periphery of OpdK (5). This electronegative rim prevents anions and negatively charged molecules approaching the vestibule. However, VNL molecules or other small aromatic molecules with negative charge can be tracked by interactions with aromatic and polar residues at the extracellular mouth instead. When a VNL approaches the upper L3-L7 region, the Arg patch can then facilitate translocation to the constriction site. Subsequently, the highly positive charges on the periplasmic interior surface lead the VNL to the outlet.

Comparing the interior architecture of OpdK with the known crystal structures of other members in the OprD family, OprD and OprP (12, 22), not only does OpdK contain a major substrate-binding site at the middle of the pore like OprD

and OprP, but all of them also show the similar behaviour of trapping their substrates. OprD and OprP are basic amino acid- and phosphate-selective, respectively. The opposite sign of the electrostatic potential on the extracellular pore surface seems to play a key role in attracting anions and positively charged amino acids from the environment (12, 22). Similarly in the OpdK pore, the positively charged patch on one side and a mixed aromatic and polar region at the opposite site are located on the pore surface in order to attract a free VNL and for substrate recognition. However, it appears that the trick of acquiring either highly charged small amino acids or anions is dominated by the electrostatic attraction, while the larger aromatic molecules with small charges are tracked by polar interactions. Although the diverse types of attractive forces are used to procure free substrates, the key factor to thread substrates to the exit hall is the electrostatic forces. The sink of either negatively or positively charges at the periplasmic mouth existing in all OpdK, OprD, and OprP appears to take a major responsibility in streaming a solute to the periplasmic end, especially that of OprP also found to contribute to a phosphate selectivity (181). Overall, according to the findings here and previous studies both unique substrate-binding site and a charged interior at the periplasmic end acts as a part of the signatures for the OrpD family.

Finally, elucidating the OpdK behaviour responding to its substrate (VNL) not only serves as a representative to better understand the structural and functional basis of the other members in the OprD family, but also may yield fundamental information for the lead design of new drugs against *P. aeruginosa*. Also, the mechanism of VNL translocation shown here may play a role in screening for compounds with better permeability properties.

7 CONCLUSIONS

The work described in this thesis addresses several aspects of OMPs, such as protein dynamics, uptake of substrates, selectivity, as well as the OMP-based models for design of synthetic nanopores. Initially, the general porin, OmpC, was studied in comparison with the well-studied OmpF (Chapter 3). Understanding the nature of both general and prevalent porins acts as a primary source to advance our knowledge to better understand the behaviour of more complex porins in the uptake processes, as seen in the Chapter 4 and 6. The majority of this thesis concentrates on probing translocation processes of solute via substrate-specific porins. Also, the knowledge of selectivity properties among these porins serves as a template for nanopore design in Chapter 5.

Firstly, the major porins, OmpC and OmpF, in *E. coli* have been investigated in the Chapter 3 where these general pores serve as models to understand protein dynamics and function via the conventional MD. The nature of well-studied OmpF was investigated parallel to the newly crystallized OmpC. Despite the fact that both porins share similar structural features the simulation results show that OmpC is not only more stable than OmpF on the 10 ns MD timescale, but also more cation-selective. Both porins are prevalent in *E. coli*, but the expression of OmpF is reduced and reversed in the production of OmpC under the extreme condition (32). This is one of the survival mechanisms of bacteria. The more stable and more selective OmpC is therefore important to minimize the risk. The MD results also illustrate the role of the constriction loops (L3) in binding to the substrate, which is the unique feature of porins. The constriction loop L3 not only serves as the most favourable site for solute but controls the pore cavity as well. Importantly, for such non-specific porins, the electrostatic potentials of both porins in an association with the previous

study (36) reveal the key role of entropic effects in the solute diffusion. It can be initially learnt from this study that electrostatic factor is not always the major driving force for solute uptake, especially for less charged or bulky solutes. This initial finding plays a key role in the Chapter 6 where the translocation of bulky charged solute is determined.

Secondly, the more complex substrate-specific channel, OprP, was studied in Chapter 4. Unlike the general porins in Chapter 3, the more complicated features such as the Arg ladder and Lys sink make the OprP pore more specific to anions, especially a phosphate ion. In this project, the behaviour of OprP responding to the substrate binding is highlighted. However, the entire process cannot practically be observed on the standard MD timescale. More advanced techniques are then applied to capture such unlikely processes. Beneficially, the phosphate-selective properties of OprP over a common chloride ion were also drawn to light. In OmpC and OmpF, the excessive negatively charged residues inside a pore permit both porins to favour cations, but this cluttered charged interior does provide neither specificity nor selectivity among cationic species. On the contrary, OprP contains an array of positively charged residues along the pore surface to attract free anions. Even though the organized Arg ladder located in the extracellular vestibule serves as a snare for all anionic species, key charged residues on the constriction loop L3 at the middle of the pore efficiently construct the energetic barriers to characterize anions (in this case are phosphate and chloride ion) where the phosphate can be recognized by the double-well energy landscape. Besides, the larger size and higher charges of phosphate ion permitting stronger interactions allow a phosphate to deeply sink into the constriction site and also be continuously glided to a Lys sink at the periplasmic chamber. Unlike phosphate, although both chloride and phosphate share the most

favourable binding site at the constriction zone, a chloride ion experiences more difficulties to translocate to the periplasmic chamber. The key factor of the phosphate translocation is the "sweeping" mechanism. All arginines and lysines generate the brush-like environment where Pi is swept along the pore axis until it reaches the periplasmic exit. On the other hand, the chloride ion appears to be fully hydrated which prevent interactions with pore-lining residues leading to the loss of favourable free energy at the periplasmic vestibule. The differences in the energetic profiles and hydration shells allow both phosphate and chloride to be differentiated. From Chapter 3 and 4, it is clear that the array of flexible and charged residues in the confined environment plays a nontrivial role in the substrate selectivity. Based on this finding, an OprP-like synthetic nanopore is modelled in Chapter 5 in order to study the possibility of implanting phosphate-selective properties into a pore model for possible applications in bionanotechnology.

Thirdly, the pore-lining charge architecture of OprP was transferred to the modelled pore to mimic the selective behaviour for phosphate ion as found in OprP (Chapter 5). Since my simulation results reveal the important role of charged brush-like behaviour in phosphate selectivity, the models built in Chapter 5 duplicated both interior architecture of positively charged residues and pore cavity. The sidechains of all pore-lining Arg and Lys residues were planted to the pore surface of the modelled nanopore. Both phosphate and chloride PMFs were computed to compare with those from the biological OprP pore. The phosphate PMF profile of the modelled nanopore gives the similar energy landscape as found in the original OprP pore, which successfully implies the high possibility of the pore model to express phosphate-selective behaviour. With only an array of positively charged residues, the ability to become phosphate-selective is dramatically increased to ~ 400 times greater than a

chloride one. Moreover, to emphasize the effects of charged brush-like characteristics on the phosphate translocation, the set of pore models with the different length of charged sidechains were investigated. With the narrowest pore diameter of ~ 0.6 nm, the most flexible model (MP1) seems to give the most favourable phosphate-binding site. The sidechain flexibility is found to play a role in dehydration of Pi. The longer sidechains allow an increase in the degree of freedom to tightly bind to Pi, especially in the narrowest constriction zone. The Pi selectivity of models with shorter sidechains, MP2 and MP3, is reduced 2-fold and 5-fold from the fully flexible MP1. It appears that the affinity of Pi binding depends on the ability of the pore to dehydrate and interact with Pi where such ability can be obtained by an increase in the flexibility of pore-lining sidechains.

Fourthly, in the past two chapters, the translocations of common ions with high charges and small degrees of freedom were investigated. Many simulation techniques were applied in order to probe such transport processes. All experiences gaining form previous chapters were bought to the last 'Result' chapter. The translocation of more complex solute, aromatic VNL with a charge of -1 (see Chapter 6), via the OpdK pore were studied. For ion channels, the major attraction for solute is the electrostatics where the selectivity can be determined by the confined geometry inside the constriction zone. On the other hand, the more complex and less polar solute like VNL appears to act differently. The molecular orientation and polar interactions seem to dominate both selectivity and permeation. My finding suggests that the translocation can be driven when the VNL molecule reorients to fit the subsite above the constriction area where the carboxylate group of VNL points towards the positively charged patch and the rest of the molecule is aligned to the opposite direction which is surrounded by the aromatic and polar amino acids

(Chapter 6). The simulation results show that the success VNL has to reorient before passing through the periplasmic chamber. When the orientation of VNL matches with the pocket above the constriction site. The hydrogen bonds between the carboxylate group of the VNL molecule and the Arg patch are then drag the entire molecule into the constriction site. Unlike the OprP pore where the free anion is trapped and translocated through the pore by electrostatic interactions along the pore axis, the OpdK pore scavenges for free VNL by firstly tracking aromatic and polar interactions before utilizing electrostatic interactions to drive VNL to the periplasmic hall. Although both OprP and OpdK employ different strategies to trap their substrate from the environment, both of them share the same feature of the charged sink at the periplasmic end. The charged clusters in both OprP and OpdK are found to be responsible for the substrate translocation from the constriction site to the periplasmic end. Interestingly, the crowd of charged residues at the periplasmic end seems to be one of unique features of the currently known structures in the OprD family (5, 12, 22).

Finally, the thesis has demonstrated the potential exploitation of computational techniques, especially MD simulations to gain novel insight into the biological function and dynamics of OMPs. Also, the novel knowledge from my thesis may help to supplement and inspire further structural and functional experiments that aid to fight *E. coli* and *P.aeruginosa*. Moreover, the highly phosphate-selective properties of OprP revealed in this thesis can also serve as a base for the future design of robust nanopores in the bionanotechnology aspect.

Bibliography

1. Alberts B, Bray D, Lewis J, Raff M, Roberts K, & Watson JD (1994) *Molecular biology of the cell* (Garland Publishing, Inc., New York & London) 3 Ed.
2. Luckey M (2008) *Membrane structural biology: with biochemical and biophysical foundatations* (Cambridge University Press, New York).
3. Schulz GE (2002) The structure of bacterial outer membrane proteins. *Biochimica et Biophysica Acta* 1565(2):308-317.
4. Lee JK, Khademi S, Harries W, Savage D, Miercke L, & Stroud RM (2004) Water and glycerol permeation through the glycerol channel GlpF and the aquaporin family. *Journal of Synchrotron Radiation* 11(Pt 1):86-88.
5. 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.
6. Levinson W (2008) *Review of Medical Microbiology and Immunology* (The McGraw-Hill Companies, Inc.) 10 Ed.
7. David B, Holt JG, Krieg NR, & Sneath PHA (1994) *Bergey's Manual of Determinative Bacteriology* (Lippincott Williams & Wilkins) 9 Ed p 787.
8. Dong CJ, Beis K, Nesper J, Brunkan-LaMontagne AL, Clarke BR, Whitfield C, & Naismith JH (2006) Wza the translocon for E-coli capsular polysaccharides defines a new class of membrane protein. *Nature* 444(7116):226-229.
9. Fronzes R, Christie PJ, & Waksman G (2009) The structural biology of type IV secretion systems. *Nature Reviews Microbiology* 7(10):703-714.
10. Chandran V, Fronzes R, Duquerroy S, Cronin N, Navaza J, & Waksman G (2009) Structure of the outer membrane complex of a type IV secretion system. *Nature* 462(7276):1011-1015.
11. Dutzler R, Wang YF, Rizkallah PJ, Rosenbusch JP, & Schirmer T (1996) Crystal structures of various maltooligosaccharides bound to maltoporin reveal a specific sugar translocation pathway. *Structure* 4(2):127-134.
12. Moraes TF, Bains M, Hancock RE, & Strynadka NC (2007) An arginine ladder in OprP mediates phosphate-specific transfer across the outer membrane. *Nature Structural & Molecular Biology* 14(1):85-87.
13. van den Berg B, Black PN, Clemons WM, Jr., & Rapoport TA (2004) Crystal structure of the long-chain fatty acid transporter FadL. *Science* 304(5676):1506-1509.
14. Pages JM, James CE, & Winterhalter M (2008) The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nature Reviews Microbiology* 6(12):893-903.
15. Koronakis V, Sharff A, Koronakis E, Luisi B, & Hughes C (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405(6789):914-919.

16. Akama H, Kanemaki M, Yoshimura M, Tsukihara T, Kashiwagi T, Yoneyama H, Narita S, Nakagawa A, & Nakae T (2004) Crystal structure of the drug discharge outer membrane protein, OprM, of Pseudomonas aeruginosa: dual modes of membrane anchoring and occluded cavity end. *Journal of Biological Chemistry* 279(51):52816-52819.

17. Phan G, Benabdelhak H, Lascombe M, Benas P, Rety S, RPicard M, Ducruix A, Etchebest C, & Broutin I (2010) Structural and Dynamical Insights into the Opening Mechanism of P. aeruginosa OprM Channel. *Structure* 18(4):507-517.

18. Koebnik R, Locher KP, & Van Gelder P (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Molecular Microbiology* 37(2):239-253.

19. Wirth C, Condemine G, Boiteux C, Berneche S, Schirmer T, & Peneff CM (2009) NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family. *Journal of Molecular Biology* 394(4):718-731.

20. Basle A, Rummel G, Storici P, Rosenbusch JP, & Schirmer T (2006) Crystal structure of osmoporin OmpC from E. coli at 2.0 A. *Journal of Molecular Biology* 362(5):933-942.

21. Tanabe M, Nimigean CM, & Iverson TM (2010) Structural basis for solute transport, nucleotide regulation, and immunological recognition of Neisseria meningitidis PorB. *Proceedings of the National Academy of Sciences of the United States of America*

22. Biswas S, Mohammad MM, Patel DR, Movileanu L, & van den Berg B (2007) Structural insight into OprD substrate specificity. *Nature Structural & Molecular Biology* 14(11):1108-1109.

23. Pautsch A & Schulz GE (2000) High-resolution structure of the OmpA membrane domain. *Journal of Molecular Biology* 298(2):273-282.

24. Subbarao GV & van den Berg B (2006) Crystal structure of the monomeric porin OmpG. *Journal of Molecular Biology* 360(4):750-759.

25. Nikaido H, Rosenberg EY, & Foulds J (1983) Porin channels in Escherichia coli: studies with beta-lactams in intact cells. *Journal of Bacteriology* 153(1):232-240.

26. Cowan SW, Garavito RM, Jansonius JN, Jenkins JA, Karlsson R, Konig N, Pai EF, Pauptit RA, Rizkallah PJ, Rosenbusch JP, & et al. (1995) The structure of OmpF porin in a tetragonal crystal form. *Structure* 3(10):1041-1050.

27. Karshikoff A, Spassov V, Cowan SW, Ladenstein R, & Schirmer T (1994) Electrostatic properties of two porin channels from Escherichia coli. *Journal of Molecular Biology* 240(4):372-384.

28. Benz R, Schmid A, & Hancock RE (1985) Ion selectivity of gram-negative bacterial porins. *Journal of Bacteriology* 162(2):722-727.

29. Watanabe M, Rosenbusch J, Schirmer T, & Karplus M (1997) Computer simulations of the OmpF porin from the outer membrane of Escherichia coli. *Biophysical Journal* 72(5):2094-2102.

30. Tieleman DP & Berendsen HJ (1998) A molecular dynamics study of the pores formed by Escherichia coli OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. *Biophysical Journal* 74(6):2786-2801.

31. Phale PS, Philippsen A, Kiefhaber T, Koebnik R, Phale VP, Schirmer T, & Rosenbusch JP (1998) Stability of trimeric OmpF porin: the contributions of the latching loop L2. *Biochemistry* 37(45):15663-15670.

32. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews* 67:593-656.

33. Schirmer T (1998) General and Specific Porins from Bacterial Outer Membrane. *Journal of Structural Biology* 121:101-109.

34. Saint N, Lou KL, Widmer C, Luckey M, Schirmer T, & Rosenbusch JP (1996) Structural and functional characterization of OmpF porin mutants selected for larger pore size. II. Functional characterization. *Journal of Biological Chemistry* 271(34):20676-20680.

35. Schirmer T & Phale PS (1999) Brownian dynamics simulation of ion flow through porin channels. *Journal of Molecular Biology* 294(5):1159-1167.

36. Im W & Roux B (2002) Ion permeation and selectivity of OmpF porin: a theoretical study based on molecular dynamics, Brownian dynamics, and continuum electrodiffusion theory. *Journal of Molecular Biology* 322(4):851-869.

37. Im W & Roux B (2002) 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. *Journal of Molecular Biology* 319(5):1177-1197.

38. Phale PS, Philippsen A, Widmer C, Phale VP, Rosenbusch JP, & Schirmer T (2001) Role of charged residues at the OmpF porin channel constriction probed by mutagenesis and simulation. *Biochemistry* 40(21):6319-6325.

39. Dhakshnamoorthy B, Raychaudhury S, Blachowicz L, & Roux B (2009) Cation-selective Pathway of OmpF Porin Revealed by Anomalous X-ray Diffraction. *Journal of Molecular Biology*

40. Danelon C, Suenaga A, Winterhalter M, & Yamato I (2003) Molecular origin of the cation selectivity in OmpF porin: single channel conductances vs. free energy calculation. *Biophysical Chemistry* 104(3):591-603.

41. Liu N & Delcour AH (1998) Inhibitory effect of acidic pH on OmpC porin: wild-type and mutant studies. *FEBS Letters* 434(1-2):160-164.

42. Liu N, Samartzidou H, Lee KW, Briggs JM, & Delcour AH (2000) Effects of pore mutations and permeant ion concentration on the spontaneous gating activity of OmpC porin. *Protein Engineering* 13(7):491-500.

43. Nestorovich EM, Rostovtseva TK, & Bezrukov SM (2003) Residue ionization and ion transport through OmpF channels. *Biophysical Journal* 85(6):3718-3729.

44. Alcaraz A, Nestorovich EM, Aguilella-Arzo M, Aguilella VM, & Bezrukov SM (2004) Salting out the ionic selectivity of a wide channel: the asymmetry of OmpF. *Biophysical Journal* 87(2):943-957.

45. van der Straaten TA, Tang JM, Ravaioli U, Eisenberg RS, & Aluru NR (2003) Simulating Ion Permeation Through the ompF Porin Ion Channel Using Three-Dimensional Drift-Diffusion Theory. *Journal of Computational Electronics* 2:29-47.

46. Varma S & Jakobsson E (2004) Ionization states of residues in OmpF and mutants: effects of dielectric constant and interactions between residues. *Biophysical Journal* 86(2):690-704.

47. Jaffe A, Chabbert YA, & Semonin O (1982) Role of porin proteins OmpF and OmpC in the permeation of beta-lactams. *Antimicrobial Agents and Chemotherapy* 22(6):942-948.

48. Mahendran KR, Kreir M, Weingart H, Fertig N, & Winterhalter M (2010) Permeation of Antibiotics through Escherichia coli OmpF and OmpC Porins: Screening for Influx on a Single-Molecule Level. *Journal of Biomolecular Screening*

49. Danelon C, Nestorovich EM, Winterhalter M, Ceccarelli M, & Bezrukov SM (2006) Interaction of zwitterionic penicillins with the OmpF channel facilitates their translocation. *Biophysical Journal* 90(5):1617-1627.

50. Mach T, Neves P, Spiga E, Weingart H, Winterhalter M, Ruggerone P, Ceccarelli M, & Gameiro P (2008) Facilitated permeation of antibiotics across membrane channels--interaction of the quinolone moxifloxacin with the OmpF channel. *Journal of the American Chemical Society* 130(40):13301-13309.

51. Mahendran KR, Chimerel C, Mach T, & Winterhalter M (2009) Antibiotic translocation through membrane channels: temperature-dependent ion current fluctuation for catching the fast events. *European Biophysics Journal with Biophysics Letters* 38(8):1141-1145.

52. Neves P, Sousa I, Winterhalter M, & Gameiro P (2009) Fluorescence Quenching as a Tool to Investigate Quinolone Antibiotic Interactions with Bacterial Protein OmpF. *Journal of Membrane Biology* 227(3):133-140.

53. 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. *Journal of Physical Chemistry B* 114(15):5170-5179.

54. Ceccarelli M, Danelon C, Laio A, & Parrinello M (2004) Microscopic Mechanism of Antibiotics Translocation through a Porin. *Biophysical Journal* 87(1):58-64.

55. Neves P, Berkane E, Gameiro P, Winterhalter M, & de Castro B (2005) Interaction between quinolones antibiotics and bacterial outer membrane porin OmpF. *Biophysical Chemistry* 113(2):123-128.

56. Vidal S, Bredin J, Pages JM, & Barbe J (2005) Beta-lactam screening by specific residues of the OmpF eyelet. *Journal of Medicinal Chemistry* 48(5):1395-1400.

57. Chai T, Wu V, & Foulds J (1982) Colicin A receptor: role of two Escherichia coli outer membrane proteins (OmpF protein and btuB gene product) and lipopolysaccharide. *Journal of Bacteriology* 151(2):983-988.

58. Bourdineaud JP, Fierobe HP, Lazdunski C, & Pages JM (1990) Involvement of OmpF during reception and translocation steps of colicin N entry. *Molecular Microbiology* 4(10):1737-1743.

59. Fourel D, Hikita C, Bolla JM, Mizushima S, & Pages JM (1990) Characterization of ompF domains involved in Escherichia coli K-12 sensitivity to colicins A and N. *Journal of Bacteriology* 172(7):3675-3680.

60. Bourdineaud JP, Fierobe HP, Lazdunski C, & Pages JM (1990) Involvement of Ompf during Reception and Translocation Steps of Colicin-N Entry. *Molecular Microbiology* 4(10):1737-1743.

61. ElKouhen R & Pages JM (1996) Dynamic aspects of colicin N translocation through the Escherichia coli outer membrane. *Journal of Bacteriology* 178(17):5316-5319.

62. Zakharov SD, Eroukova VY, Rokitskaya TI, Zhalnina MV, Sharma O, Loll PJ, Zgurskaya HI, Antonenko YN, & Cramer WA (2004) Colicin occlusion of OmpF and TolC channels: outer membrane translocons for colicin import. *Biophysical Journal* 87(6):3901-3911.

63. Bredin J, Simonet V, Iyer R, Delcour AH, & Pages JM (2003) Colicins, spermine and cephalosporins: a competitive interaction with the OmpF eyelet. *Biochemical Journal* 376(Pt 1):245-252.

64. Sobko AA, Kotova EA, Antonenko YN, Zakharov SD, & Cramer WA (2006) Lipid dependence of the channel properties of a colicin E1-lipid toroidal pore. *Journal of Biological Chemistry* 281(20):14408-14416.

65. Baboolal TG, Conroy MJ, Gill K, Ridley H, Visudtiphole V, Bullough PA, & Lakey JH (2008) Colicin N binds to the periphery of its receptor and translocator, outer membrane protein F. *Structure* 16(3):371-379.

66. Yamashita E, Zhalnina MV, Zakharov SD, Sharma O, & Cramer WA (2008) Crystal structures of the OmpF porin: function in a colicin translocon. *EMBO Journal* 27(15):2171-2180.

67. Miedema H, Meter-Arkema A, Wierenga J, Tang J, Eisenberg B, Nonner W, Hektor H, Gillespie D, & Meijberg W (2004) Permeation properties of an engineered bacterial OmpF porin containing the EEEE-locus of Ca2+ channels. *Biophysical Journal* 87(5):3137-3147.

68. Vrouenraets M, Wierenga J, Meijberg W, & Miedema H (2006) Chemical modification of the bacterial porin OmpF: gain of selectivity by volume reduction. *Biophysical Journal* 90(4):1202-1211.

69. Miedema H, Vrouenraets M, Wierenga J, Meijberg W, Robillard G, & Eisenberg B (2007) A biological porin engineered into a molecular, nanofluidic diode. *Nano Letters* 7(9):2886-2891.

70. Ho D, Chang S, & Montemagno CD (2006) Fabrication of biofunctional nanomaterials via Escherichia coli OmpF protein air/water interface insertion/integration with copolymeric amphiphiles. *Nanomedicine* 2(2):103-112.

71. Schmitt EK, Vrouenraets M, & Steinem C (2006) Channel activity of OmpF monitored in nano-BLMs. *Biophysical Journal* 91(6):2163-2171.

72. Shah DS, Thomas MB, Phillips S, Cisneros DA, Le Brun AP, Holt SA, & Lakey JH (2007) Self-assembling layers created by membrane proteins on gold. *Biochemical Society Transactions* 35:522-526.

73. Wilk SJ, Petrossian L, Goryll M, Thornton TJ, Goodnick SM, Tang JM, & Eisenberg RS (2007) Integrated electrodes on a silicon based ion channel measurement platform. *Biosensors & Bioelectronics* 23(2):183-190.

74. Ranquin A, Versees W, Meier W, Steyaert J, & Van Gelder P (2005) Therapeutic nanoreactors: Combining chemistry and biology in a novel triblock copolymer drug delivery system. *Nano Letters* 5(11):2220-2224.

75. Broz P, Driamov S, Ziegler J, Ben-Haim N, Marsch S, Meier W, & Hunziker P (2006) Toward intelligent nanosize bioreactors: A pH-switchable, channel-equipped, functional polymer nanocontainer. *Nano Letters* 6(10):2349-2353.

76. Dudia A, Kocer A, Subramaniam V, & Kanger JS (2008) Biofunctionalized lipid-polymer hybrid nanocontainers with controlled permeability. *Nano Letters* 8(4):1105-1110.

77. Hancock RE, Poole K, & Benz R (1982) Outer membrane protein P of Pseudomonas aeruginosa: regulation by phosphate deficiency and formation of small anion-specific channels in lipid bilayer membranes. *Journal of Bacteriology* 150(2):730-738.

78. Dargent B, Hofmann W, Pattus F, & Rosenbusch JP (1986) The selectivity filter of voltage-dependent channels formed by phosphoporin (PhoE protein) from E. coli. *EMBO Journal* 5(4):773-778.

79. Hancock RE & Benz R (1986) Demonstration and chemical modification of a specific phosphate binding site in the phosphate-starvation-inducible outer membrane porin protein P of Pseudomonas aeruginosa. *Biochimica et Biophysica Acta* 860(3):699-707.

80. Benz R & Hancock RE (1987) Mechanism of ion transport through the anion-selective channel of the Pseudomonas aeruginosa outer membrane. *Journal of General Physiology* 89(2):275-295.

81. Benz R, Egli C, & Hancock REW (1993) 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. *Biochimica et Biophysica Acta* 1149(2):224-230.

82. Hancock RE, Schmidt A, Bauer K, & Benz R (1986) Role of lysines in ion selectivity of bacterial outer membrane porins. *Biochimica et Biophysica Acta* 860(2):263-267.

83. Sukhan A & Hancock RE (1996) The role of specific lysine residues in the passage of anions through the Pseudomonas aeruginosa porin OprP. *Journal of Biological Chemistry* 271(35):21239-21242.

84. Hancock RE & Brinkman FS (2002) Function of pseudomonas porins in uptake and efflux. *Annual Review of Microbiology* 56:17-38.

85. Leach AR (2001) *Molecular Modelling: Principles and Applications* (Addison Wesley Longman Limited, Essex).

86. Halgren TA & Damm W (2001) Polarizable force fields. *Current Opinion in Structural Biology* 11(2):236-242.

87. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics* 98(12):10089-10092.

88. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, & Karplus M (2009) CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry* 30(10):1545-1614.

89. Jorgensen WL, Maxwell DS, & TiradoRives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* 118(45):11225-11236.

90. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong GM, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang JM, & Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry* 24(16):1999-2012.

91. van Gunsteren WF, Billeter SR, Elsing AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, & Tironi IG (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide.* (Biomos and Hochschulverlag AG an der ETH Zürich, Zürich).

92. Bonvin AMJJ, Mark AE, & van Gunsteren WF (2000) The GROMOS96 benchmarks for molecular simulation. *Computer Physics Communications* 128(3):550-557.

93. Hess B, Bekker H, Berendsen H, & Fraaije J (1997) LINCS: A Linear Constraint Solver for Molecular Simulations. *Journal of Computational Chemistry* 18:1463-1472.

94. Miyamoto S & Kollman PA (1992) Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *Journal of Computational Chemistry* 13(8):952-962.

95. Ryckaert JP, Ciccotti G, & Berendsen HJC (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* 23(3):327-341.

96. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, & Haak JR (1984) Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics* 81(8):3684-3690.

97. Nosé S (1984) A Molecular-Dynamics Method for Simulations in the Canonical Ensemble. *Molecular Physics* 52(2):255-268.

98. Hoover WG (1985) Canonical Dynamics - Equilibrium Phase-Space Distributions. *Physical Review A* 31(3):1695-1697.

99. Anderson HC (1980) molecular dynamics simulations at constant pressure and/or temperature. *Journal of Chemical Physics* 72:2384-2393.

100. Parrinello M & Rahman A (1981) polymeric transitions in single crystal: a new moleculear dynamics method. *Journal of Applied Physics* 52:7182-7190.

101. Hille B (2001) *Ion channels of Ecitable membranes* (Sinauer Associates Inc.) 3 Ed.

102. Jensen MO, Borhani DW, Lindorff-Larsen K, Maragakis P, Jogini V, Eastwood MP, Dror RO, & Shaw DE (2010) Principles of conduction and hydrophobic gating in K+ channels. *Proceedings of the National Academy of Sciences of the United States of America* 107(13):5833-5838.

103. Aksimentiev A & Schulten K (2005) Imaging alpha-hemolysin with molecular dynamics: ionic conductance, osmotic permeability, and the electrostatic potential map. *Biophysical Journal* 88(6):3745-3761.

104. Cruz-Chu ER, Aksimentiev A, & Schulten K (2009) Ionic Current Rectification through Silica Nanopores. *Journal of Physical Chemistry C* 113(5):1850-1862.

105. Cruz-Chu ER, Ritz T, Siwy ZS, & Schulten K (2009) Molecular control of ionic conduction in polymer nanopores. *Faraday Discussions* 143:47-62.

106. Heng JB, Aksimentiev A, Ho C, Marks P, Grinkova YV, Sligar S, Schulten K, & Timp G (2006) The electromechanics of DNA in a synthetic nanopore. *Biophysical Journal* 90(3):1098-1106.

107. Siu SWI & Bockmann RA (2007) Electric field effects on membranes: Gramicidin A as a test ground. *Journal of Structural Biology* 157(3):545-556.

108. Sigalov G, Comer J, Timp G, & Aksimentiev A (2008) Detection of DNA sequences using an alternating electric field in a nanopore capacitor. *Nano Letters* 8(1):56-63.

109. Mirsaidov U, Timp W, Zou X, Dimitrov V, Schulten K, Feinberg AP, & Timp G (2009) Nanoelectromechanics of methylated DNA in a synthetic nanopore. *Biophysical Journal* 96(4):L32-34.

110. Schulz R & Kleinekathofer U (2009) Transitions between Closed and Open Conformations of ToIC: The Effects of Ions in Simulations. *Biophysical Journal* 96(8):3116-3125.

111. Suenaga A, Komeiji Y, Uebayasi M, Meguro T, Saito M, & Yamato I (1998) Computational observation of an ion permeation through a channel protein. *Bioscience Reports* 18(1):39-48.

112. Deuflhard P, Hermans J, Leimkuhler B, Mark AE, & Reich S (1998) *Steered molecular dynamics* (Springer-Verlag, Berlin).

113. Isralewitz B, Gao M, & Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology* 11(2):224-230.

114. Gao M, Wilmanns M, & Schulten K (2002) Steered molecular dynamics studies of titin I1 domain unfolding. *Biophysical Journal* 83(6):3435-3445.

115. Gullingsrud J & Schulten K (2003) Gating of MscL studied by steered molecular dynamics. *Biophysical Journal* 85(4):2087-2099.

116. Jensen MO, Yin Y, Tajkhorshid E, & Schulten K (2007) Sugar transport across lactose permease probed by steered molecular dynamics. *Biophysical Journal* 93(1):92-102.

117. Wells DB, Abramkina V, & Aksimentiev A (2007) Exploring transmembrane transport through α-hemolysin with grid-steered molecular dynamic. *Journal of Chemical Physics* 127:125101.

118. Kollman P (1993) Free energy calculations: Applications to chemical and bio- chemical phenomena. *Chemical Reviews* 93:2395-2417.

119. Roux B (1995) The Calculation of the Potential of Mean Force Using Computer-Simulations. *Computer Physics Communications* 91(1-3):275-282.

120. Che J, Dzubiella J, Li B, & McCammon JA (2008) Electrostatic free energy and its variations in implicit solvent models. *Journal of Physical Chemistry B* 112(10):3058-3069.

121. Grochowski P & Trylska J (2008) Continuum molecular electrostatics, salt effects, and counterion binding--a review of the Poisson-Boltzmann theory and its modifications. *Biopolymers* 89(2):93-113.

122. Edwards S, Corry B, Kuyucak S, & Chung SH (2002) Continuum electrostatics fails to describe ion permeation in the gramicidin channel. *Biophysical Journal* 83(3):1348-1360.

123. Faraldo-Gomez JD & Roux B (2004) Electrostatics of ion stabilization in a ClC chloride channel homologue from Escherichia coli. *Journal of Molecular Biology* 339(4):981-1000.

124. Jogini V & Roux B (2005) Electrostatics of the intracellular vestibule of K+ channels. *Journal of Molecular Biology* 354(2):272-288.

125. Peter C & Hummer G (2005) Ion transport through membrane-spanning nanopores studied by molecular dynamics simulations and continuum electrostatics calculations. *Biophysical Journal* 89(4):2222-2234.

126. Sotomayor M, van der Straaten TA, Ravaioli U, & Schulten K (2006) Electrostatic properties of the mechanosensitive channel of small conductance MscS. *Biophysical Journal* 90(10):3496-3510.

127. Aguilella-Arzo M, Garcia-Celma JJ, Cervera J, Alcaraz A, & Aguilella VM (2007) Electrostatic properties and macroscopic electrodiffusion in OmpF porin and mutants. *Bioelectrochemistry* 70(2):320-327.

128. Berashevich J & Chakraborty T (2009) Thermodynamics of G x A mispairs in DNA: continuum electrostatic model. *Journal of Chemical Physics* 130(1):015101.

129. Koehl P (2006) Electrostatics calculations: latest methodological advances. *Current Opinion in Structural Biology* 16(2):142-151.

130. Grycuk T (2002) Revision of the model system concept for the prediction of pK(a)'s in proteins. *Journal of Physical Chemistry B* 106(6):1434-1445.

131. Miyashita O, Onuchic JN, & Okamura MY (2003) Continuum electrostatic model for the binding of cytochrome c2 to the photosynthetic reaction center from Rhodobacter sphaeroides. *Biochemistry* 42(40):11651-11660.

132. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* 98(18):10037-10041.

133. Malek K & Maghari A (2007) Translocation and interactions of L-arabinose in OmpF porin: A molecular dynamics study. *Biochemical and Biophysical Research Communications* 352(1):104-110.

134. Kindal M, Tieleman R, & Tieleman DP (2002) Orientation and interactions of dipolar molecules during transport through OmpF porin. *FEBS Letters* 528(1-3):53-57.

135. Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit RA, Jansonius JN, & Rosenbusch JP (1992) Crystal structures explain functional properties of two E. coli porins. *Nature* 358(6389):727-733.

136. Hoenger A, Pages JM, Fourel D, & Engel A (1993) The orientation of porin OmpF in the outer membrane of Escherichia coli. *Journal of Molecular Biology* 233(3):400-413.

137. Miedema H, Vrouenraets M, Wierenga J, Eisenberg B, Schirmer T, Basle A, & Meijberg W (2006) Conductance and selectivity fluctuations in D127 mutants of the bacterial porin OmpF. *European Biophysics Journal* 36(1):13-22.

138. Varma S, Chiu SW, & Jakobsson E (2006) The influence of amino acid protonation states on molecular dynamics simulations of the bacterial porin OmpF. *Biophysical Journal* 90(1):112-123.

139. Liu N & Delcour AH (1998) The spontaneous gating activity of OmpC porin is affected by mutations of a putative hydrogen bond network or of a salt bridge between the L3 loop and the barrel. *Protein Engineering* 11(9):797-802.

140. Faraldo-Gomez JD, Smith GR, & Sansom MS (2002) Setting up and optimization of membrane protein simulations. *European Biophysics Journal* 31(3):217-227.

141. 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.

142. Humphrey W, Dalke A, & Schulten K (1996) VMD - Visual Molecular Dynamics. *Journal of Molecular Graphics* 14:33-38.

143. Kobayashi Y & Nakae T (1985) The mechanism of ion selectivity of OmpF-porin pores of Escherichia coli. *European Journal of Biochemistry* 151(2):231-236.

144. Benz R, Schmid A, Van der Ley P, & Tommassen J (1989) Molecular basis of porin selectivity: membrane experiments with OmpC-PhoE and OmpF-PhoE hybrid proteins of Escherichia coli K-12. *Biochimica et Biophysica Acta* 981(1):8-14.

145. Bauer K, van der Ley P, Benz R, & Tommassen J (1988) The pho-controlled outer membrane porin PhoE does not contain specific binding sites for

phosphate or polyphosphates. *Journal of Biological Chemistry* 263(26):13046-13053.

146. Poole K & Hancock RE (1984) Phosphate transport in Pseudomonas aeruginosa. Involvement of a periplasmic phosphate-binding protein. *European Journal of Biochemistry* 144(3):607-612.

147. Berneche S & Roux B (2001) Energetics of ion conduction through the K+ channel. *Nature* 414(6859):73-77.

148. Vidossich P, Cascella M, & Carloni P (2004) Dynamics and energetics of water permeation through the aquaporin channel. *Proteins: Structure Function and Bioinfomatics* 55:924-931.

149. Hub JS & de Groot BL (2008) Mechanism of selectivity in aquaporins and aquaglyceroporins. *Proceedings of the National Academy of Sciences of the United States of America* 105(4):1198-1203.

150. Henin J, Tajkhorshid E, Schulten K, & Chipot C (2008) Diffusion of glycerol through Escherichia coli aquaglyceroporin GlpF. *Biophysical Journal* 94(3):832-839.

151. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, & Berendsen HJ (2005) GROMACS: fast, flexible, and free. *Journal of Computational Chemistry* 26(16):1701-1718.

152. Smart OS, Neduvelil JG, Wang X, Wallace BA, & Sansom MS (1996) HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics* 14(6):354-360, 376.

153. Kumar S, Bouzida D, Swendsen RH, Kollman PA, & Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* 13 1011-1021

154. Doudou S, Burton NA, & Henchman RH (2009) Standard Free Energy of Binding from a One-Dimensional Potential of Mean Force. *Journal of Chemical Theory and Computation* 5(4):909-918.

155. Klebba PE & Newton SM (1998) Mechanisms of solute transport through outer membrane porins: burning down the house. *Current Opinion in Microbiology* 1(2):238-247.

156. Lin Y, Cao Z, & Mo Y (2006) Molecular dynamics simulations on the *Escherichia coli* ammonia channel protein AmtB: mechanism of ammonia/ammnium transport. *Journal of American Chemical Societies* 128:10876-10884.

157. Bostick DL & Brooks CL, 3rd (2007) Deprotonation by dehydration: the origin of ammonium sensing in the AmtB channel. *PLoS Computational Biology* 3(2):e22.

158. Boda D, Nonner W, Valisko M, Henderson D, Eisenberg B, & Gillespie D (2007) Steric selectivity in Na channels arising from protein polarization and mobile side chains. *Biophysical Journal* 93(6):1960-1980.

159. Liu B, Li XY, Li BL, Xu BQ, & Zhao YL (2009) Carbon nanotube based artificial water channel protein: membrane perturbation and water transportation. *Nano Letters* 9(4):1386-1394.

160. Ramirez P, Aguilella-Arzo M, Alcaraz A, Cervera J, & Aguilella VM (2006) Theoretical description of the ion transport across nanopores with titratable fixed charges. *Cell Biochemistry and Biophysics* 44(2):287-312.

161. Dehez F, Tarek M, & Chipot C (2007) Energetics of ion transport in a peptide nanotube. *Journal of Physical Chemistry B* 111(36):10633-10635.

162. Ramirez P, Gomez V, Cervera J, Schiedt B, & Mafe S (2007) Ion transport and selectivity in nanopores with spatially inhomogeneous fixed charge distributions. *Journal of Chemical Physics* 126(19):194703.

163. Jovanovic-Talisman T, Tetenbaum-Novatt J, McKenney AS, Zilman A, Peters R, Rout MP, & Chait BT (2009) Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. *Nature* 457(7232):1023-1027.

164. Garcia-Gimenez E, Alcaraz A, Aguilella VM, & Ramirez P (2009) Directional ion selectivity in a biological nanopore with bipolar structure. *Journal of Membrane Science* 331(1-2):137-142.

165. Gong X, Li J, Xu K, Wang J, & Yang H (2010) A controllable molecular sieve for Na+ and K+ ions. *Journal of the American Chemical Society* 132(6):1873-1877.

166. Allen TW, Kuyucak S, & Chung SH (1999) The effect of hydrophobic and hydrophilic channel walls on the structure and diffusion of water and ions. *Journal of Chemical Physics* 111(17):7985-7999.

167. Hummer G, Rasaiah JC, & Noworyta JP (2001) Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* 414(6860):188-190.

168. Beckstein O & Sansom MS (2003) Liquid-vapor oscillations of water in hydrophobic nanopores. *Proceedings of the National Academy of Sciences of the United States of America* 100(12):7063-7068.

169. Beckstein O & Sansom MS (2004) The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Physical Biology* 1(1-2):42-52.

170. Beckstein O (2009) Teaching old coefficients new tricks: new insight into the meaning of the osmotic and diffusive permeation coefficients. *Biophysical Journal* 96(3):763-764.

171. Gullingsrud J, Kosztin D, & Schulten K (2001) Structural determinants of MscL gating studied by molecular dynamics simulations. *Biophysical Journal* 80(5):2074-2081.

172. Beckstein O, Biggin PC, Bond P, Bright JN, Domene C, Grottesi A, Holyoake J, & Sansom MSP (2003) Ion channel gating: insights via molecular simulations. *FEBS Letters* 555(1):85-90.

173. Beckstein O, Tai K, & Sansom MS (2004) Not ions alone: barriers to ion permeation in nanopores and channels. *Journal of the American Chemical Society* 126(45):14694-14695.

174. Flyvbjerg H & Petersen HG (1989) Error-Estimates on Averages of Correlated Data. *Journal of Chemical Physics* 91(1):461-466.

175. Parsegian.A (1969) Energy of an Ion Crossing a Low Dielectric Membrane - Solutions to 4 Relevant Electrostatic Problems. *Nature* 221(5183):844-&.

176. Ramirez P, Mafe S, Alcaraz A, & Cervera J (2003) Modeling of pH-switchable ion transport and selectivity in nanopore membranes with fixed charges. *Journal of Physical Chemistry B* 107(47):13178-13187.

177. Tamber S, Ochs MM, & Hancock REW (2006) Role of the novel OprD family of porins in nutrient uptake in Pseudomonas aeruginosa. *Journal of Bacteriology* 188(1):45-54.

178. Trias J & Nikaido H (1990) Protein-D2 Channel of the Pseudomonas-Aeruginosa Outer-Membrane Has a Binding-Site for Basic-Amino-Acids and Peptides. *Journal of Biological Chemistry* 265(26):15680-15684.

179. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, & Sali A (2000) Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure* 29:291-325.

180. Roux B (2008) The membrane potential and its representation by a constant electric field in computer simulations. *Biophysical Journal* 95(9):4205-4216.

181. Pongprayoon P, Beckstein O, Wee CL, & Sansom MSP (2009) Simulations of anion transport through OprP reveal the molecular basis for high affinity and selectivity for phosphate. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21614-21618.

Appendix: Publication arising from this Thesis

A research article has resulted from the simulation work presented in this thesis. Chapter 4 is largely based on publication, whilst Chapter 5 and 6 can be considered as “manuscripts in preparation”.

Research article

Pongprayoon P, Beckstein O, Wee CL, & Sansom MSP (2009) Simulations of anion transport through OprP reveal the molecular basis for high affinity and selectivity for phosphate. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21614-21618.