

**Comparison of H<sub>2</sub>O diffusion mechanism of *Klebsiella pneumoniae* outer membrane lectin with its *E. coli* counterpart**

Shivangi Sachdeva  
BO14MTECH11007

A Dissertation Submitted to  
Indian Institute of Technology Hyderabad  
In Partial Fulfillment of the Requirements for  
The Degree of Master of Technology



भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad

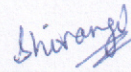
Department of Biotechnology  
Indian Institute of Technology, Hyderabad

June, 2016



## Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.



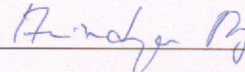
---

Shivangi Sachdeva  
BO14MTECH11007



## Approval Sheet

This thesis entitled "Comparison of water diffusion mechanism of outer membrane lectin, Wzi in *Klebsiella pneumonia* with that of *E. coli*" by Shivangi Sachdeva is approved for the degree of Master of Technology from IIT Hyderabad.



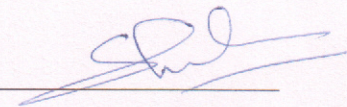
-Dr. Anindya Roy

Associate Professor

(Examiner)

Department of Biotechnology

IIT Hyderabad



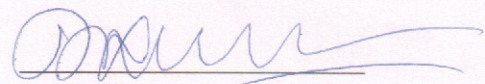
-Dr. D. S. Sharda

Assistant Professor

(Examiner)

Department of Biotechnology

IIT Hyderabad



-Dr. Thenmalarchelvi Rathinavelan

Assistant Professor

(Adviser)

Department of Biotechnology

IIT Hyderabad



## Acknowledgements

I would like to extend my gratitude to my Principal Investigator, Dr Thenmalarchelvi Rathinavelan, Assistant professor, Department of Biotechnology, IIT Hyderabad for her guidance and valuable advice. She has put faith in me from the very first day and helped me come up with thesis project and guided me at every step. She has helped me brainstorm for devising the calculation methodologies. The passion she has for research is very intriguing. She is a major source of inspiration to me.

I thank all my lab members for teaching me basics and helping me throughout the project. Thank you for your advices and providing me strength to carry on. Special thanks to Narendar Kolimi and Sanjana Anil Kumar Nair for their contribution in *Escherichai coli* part of the project.



## Abstract

Emergence of multi drug resistant bacteria is a major cause of infections worldwide. In Gram-negative bacteria, pathogenicity is mainly due to bacterial toxins, cell surface proteins, cell surface polysaccharides and hydrolytic enzymes. Different types of surface polysaccharides are lipopolysaccharide, capsular polysaccharide and exopolysaccharides. The mechanism behind capsular polysaccharide (CPS) translocation to the exterior of bacterium and its association with outer membrane is unknown. Inhibition of CPS biogenesis or exportation to the outer membrane may be an attractive strategy to reduce bacterial virulence. Anchoring of CPS to the outer membrane is done via Wzi, an outer membrane protein. This protein plays dual role as a lectin and porin. This study focuses on understanding water diffusion mechanisms of Wzi through molecular dynamics simulations. Our results show that total five water specific diffusion points have been found, in which three are at the extracellular side and two are from periplasmic side. Interestingly, “YQF” triad, one of the major entry points in Wzi, is structurally similar to sugar-binding site of sodium-galactose cotransporters, implicating its involvement in CPS surface anchorage. This is strengthened by the presence of an ion binding pocket just below YQF triad. In addition, extracellular loop 5 may also have importance in membrane anchorage of Wzi. Thus, our results may help to facilitate the targeting of ligand or drug molecules via disturbing water diffusion or ion binding, and interrupting the loop 5 insertion into the membrane may help in reducing the bacterial virulence.



# Nomenclature

## Abbreviations

LPS-	Lipopolysaccharide
CPS-	Capsular polysaccharide
ABC-	ATP-binding cassette
PCP-	Polysaccharide co-polymerase
OM-	Outer membrane
PG-	Phosphatidylglycerol
UDP-	Uridine diphosphate
OPX protein-	Outer membrane polysaccharide export protein
Man-	Mannose
Gal-1-P-	Galactose-1-phosphate
GlcUA-	Glucuronic acid
Und-pp-	Undecaprenyl diphosphate
MD-	Molecular Dynamics
POPE-	Palmitoyloleoylphosphatidylethanolamine
POPG-	Palmitoyloleoylphosphatidylglycerol
CHARMM-	Chemistry at HARvard Molecular Mechanics



# Contents

Declaration .....	ii
Approval Sheet .....	iii
Acknowledgements .....	iv
Abstract.....	v
Nomenclature.....	vi
<b>1 Introduction</b>	
1.1 Antimicrobial Resistance.....	1
1.2 Gram-negative bacteria .....	1
1.3 Cell-surface glycoconjugates .....	2
1.4 <i>E. coli</i> capsule .....	3
1.5 Biosynthesis of <i>E. coli</i> capsule.....	4
1.5.1 ABC transporter dependent pathway.....	4
1.5.2 Wzy-dependent pathway.....	5
1.6 Wzi structure.....	7
1.7 Scope of the study.....	8
1.8 References.....	8
<b>2 Methodology- Molecular Dynamics Simulations</b> .....	11
2.1 Introduction.....	11
2.2 Force field and simulation engines.....	11
2.3 Energy expression.....	12
2.3.1 Valence interactions.....	12
2.3.2 Non-bonded interactions.....	13
2.4 Constraints and Restraints.....	13
2.5 Integration algorithms.....	14
2.5.1 Verlet algorithm.....	14
2.6 Statistical ensemble.....	15
2.7 References.....	15



<b>3 Water diffusion mechanisms in Wzi of <i>Escherchia coli</i></b> .....	16
3.1 Introduction.....	16
3.2 System setup for MD simulations.....	17
3.3 MD simulation.....	18
3.4 Methodology: Analysis of the trajectory.....	18
3.4.1 Quantification of number of water molecules in the barrel.....	18
3.4.2 Quantification of barrel permeation events.....	18
3.4.3 Diffusion permeability.....	20
3.4.4 Entry and exit calculation.....	20
3.4.5 Net flux calculation.....	21
3.4.6 Normalized frequency.....	21
3.4.7 Ion coordination number.....	21
3.5 Results and discussion.....	21
3.5.1 Dual functional role of Wzi as a porin.....	21
3.5.2 Diffusion points.....	23
3.5.3 Bidirectional water permeation.....	23
3.5.4 Diffusion permeability.....	23
3.5.5 Diffusion mechanisms.....	24
3.5.5.1 Entry 1: YQF.....	24
3.5.5.2 Entry 2: Notch region.....	24
3.5.5.3 Entry 3: Threonine flipping.....	25
3.5.5.4 Entry 4: Bifurcated salt bridge.....	25
3.5.5.5 Entry 5: entry between helixes and barrel.....	26
3.5.6 Entry & Exit and Flux calculation.....	27
3.5.7 Transient water wires.....	27
3.5.8 Pore opening.....	29

3.5.9 Ion binding pocket.....	29
3.5.10 Role of extracellular loop 5.....	30
3.6 Conclusion.....	33
3.7 References.....	33
<b>4 Comparision of water diffusion mechanisms of Wzi, <i>Klebsiella pneuemonia</i> with that of <i>E. coli</i> .....</b>	<b>35</b>
4.1 Introduction.....	35
4.2 System setup.....	36
4.3 MD Simulation.....	37
4.4 Analysis of trajectory.....	37
4.4.1 Quantification of water inside the barrel.....	37
4.5 Results and Discussion.....	37
4.5.1 Wzi as a porin.....	37
4.5.2 Diffusion mechanisms.....	38
4.5.2.1 Entry point 1.....	38
4.5.2.2 Entry point 2.....	38
4.5.2.3 Entry point 3.....	39
4.5.2.4 Entry point 4.....	40
4.5.2.5 Entry point 5.....	41
4.5.3 Ion binding on the extracellular side.....	42
4.5.4 Role of extracellular loop 5.....	42
4.6 Conclusion.....	43
4.7 References.....	44

# Chapter 1

## Introduction

### 1.1 Antimicrobial Resistance

Antimicrobial resistance (AMR) is one of the major global threats that human kind is facing today. Microorganisms evolve to render resistance against antimicrobials. Antibiotics have been an important medical advancement of the last century but their misuse and overuse have led to this alarming situation. A recent report prepared by economist Jim O'Neill warns that failure to tackle AMR would claim 10 million lives each year by 2050 [1]. He also mentions that negligent use of antimicrobials, especially antibiotics, for treatment or for crops may be the reason of worldwide threat we are facing today. There is huge amount of antimicrobials sent into the environment each year.

Gram-negative bacteria such as *Escherichia coli* and *Klebsiella* have gained resistance to third generation cephalosporins [2], the last resort antibiotic, which is quiet alarming. Many different strategies can be used to tackle AMR: judicious use of antibiotics, finding new antibiotic resistant mechanisms, reducing or eliminating the use of antibiotics in agriculture and hygienic practices.

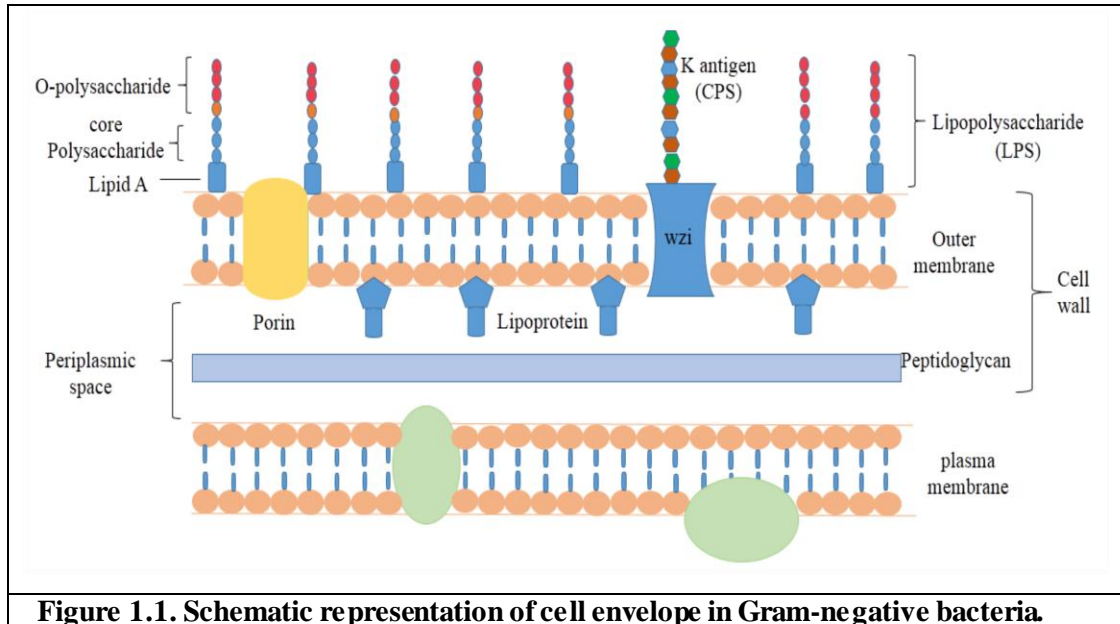
### 1.2 Gram-negative bacteria

Gram-negative bacteria are named so because in Gram staining, the destaining step washes away the primary stain (crystal violet) [3]. This is due to thin cell wall of Gram-negative bacteria that is composed of single layer of peptidoglycan layer and lipid-rich outer membrane (Fig 1.1). While in Gram-positive bacteria peptidoglycan layer is thick and more resilient thus doesn't allow crystal violet to be washed off. This method was devised by the Danish scientist, Hans Christian Gram in 1884.

Gram-negative bacteria causes many diseases in humans. For instance, *Shigella* and *Salmonella* are major cause of foodborne illness. *Citrobacter* is responsible for urinary tract infections, blood stream infections. *Klebsiella pneumonia* also causes blood stream infections, pneumonia and meningitis. *E. coli* is responsible for urinary tract infections, blood stream infections and neonatal meningitis. Some of the virulent factors present in Gram-negative bacteria are toxins, for example pertussis toxin from



*Bordetella pertussis*, effectors such as *Yersinia* outer-membrane proteins (Yops) and adhesins. Adhesins include pili, fimbriae and cell surface glycoconjugates [4]. Gram-negative bacteria have three types of cell-surface glycoconjugates: lipopolysaccharide, capsular polysaccharides and exopolysaccharides (explained in section 1.3).



**Figure 1.1. Schematic representation of cell envelope in Gram-negative bacteria.**

### 1.3 Cell-surface glycoconjugates

Typically three types of surface polysaccharides are found on Gram-negative bacteria. They are lipopolysaccharide (LPS), capsular polysaccharide (CPS) and exopolysaccharide. LPS molecule possess a core-oligosaccharide and an O-polysaccharide (O-PS) or O-antigen associated with the lipid A [5] (Fig 1.1).

Exopolysaccharides are loosely associated to the bacterium and helps in biofilm formation. CPS or K antigen is associated with outer membrane and forms a capsule around the bacterium. In *E. coli*, there are approximately 170 O antigen, 80 K antigen and 56 H antigen. These in combination gives rise to 50,000-100,000 different serotypes [6]. Number of known O, K and H antigen of some Gram negative bacteria is given in Table 1.1. Some other polysaccharides are also present in *E. coli* but these are not serotype specific. Under specific growth conditions, many *E. coli* produce colanic acid (M antigen). Also most strains produce enterobacterial common antigen [7]. Capsular polysaccharides of *E. coli* have been discussed in more detail in next section.

**Table 1.1. Number of known O, K and H antigens in different Gram-negative bacteria. (Symbol ‘-’ means unknown)**

Species	O	K	H
<i>Edwardsiella</i>	17	-	11
<i>Citrobacter(freundii)</i> [8]	42	a few	90
<i>Citrobacter(koseri)</i> [8]	14	none	7
<i>Enterobacter</i> [8]	53	-	57
<i>Hafina</i> [8]	68	-	34
<i>Serrita</i> [9]	11	-	134
<i>Proteus inconstans</i>	62	few	34
<i>Escherichia coli</i> [7]	170	80	56
<i>Klebsiella pneumoniae</i> [10]	9	77	-

#### 1.4 *E. coli* capsule

The high-molecular weight polysaccharide protective structure surrounding some bacteria is known as capsule. They are one of the major virulent determinants in *E. coli*. Some of the diseases have been found in association with a particular K antigen, eg. Neonatal meningitis is associated with *E. coli* K1 [11]. In *E. coli*, there are 80 different capsular serotypes that are classified into 4 groups on basis of genetic and biochemical criteria (Table 1.2).

**Table 1.2. *E. coli* K antigens classified in four groups [12].**

Group I	Group II	Group III	Group IV
<b>K26, K27, K28, K29, K30, K31, K33, K34, K35, K36, K37, K39, K42, K55, K102, K103</b>	<b>K1, K2a, K2ab, K4, K5, K6, K7, K12, K13, K14, K15, K16, K18, K19, K20, K22, K23, K24, K51, K52, K53, K74, K92, K93, K95, K97, K100</b>	<b>K3, K10, K11, K54, K96, K98</b>	<b>K8, ,K9, K38, K40, K44, K46, K47, K48, K49, K50, K57, K83, K85, K87, K101</b>

Capsule acts as physical barrier and keeps cell hydrated. It not only helps in adherence to host cells but also helps in evading human immune system in many ways:

- i. It prevents opsonization by antibodies and resist recognition by macrophages [13].
- ii. It mimics host molecules and thus is recognized as self.
- iii. It inhibits complement cascade by binding to serum factor H, inactivating c3b.

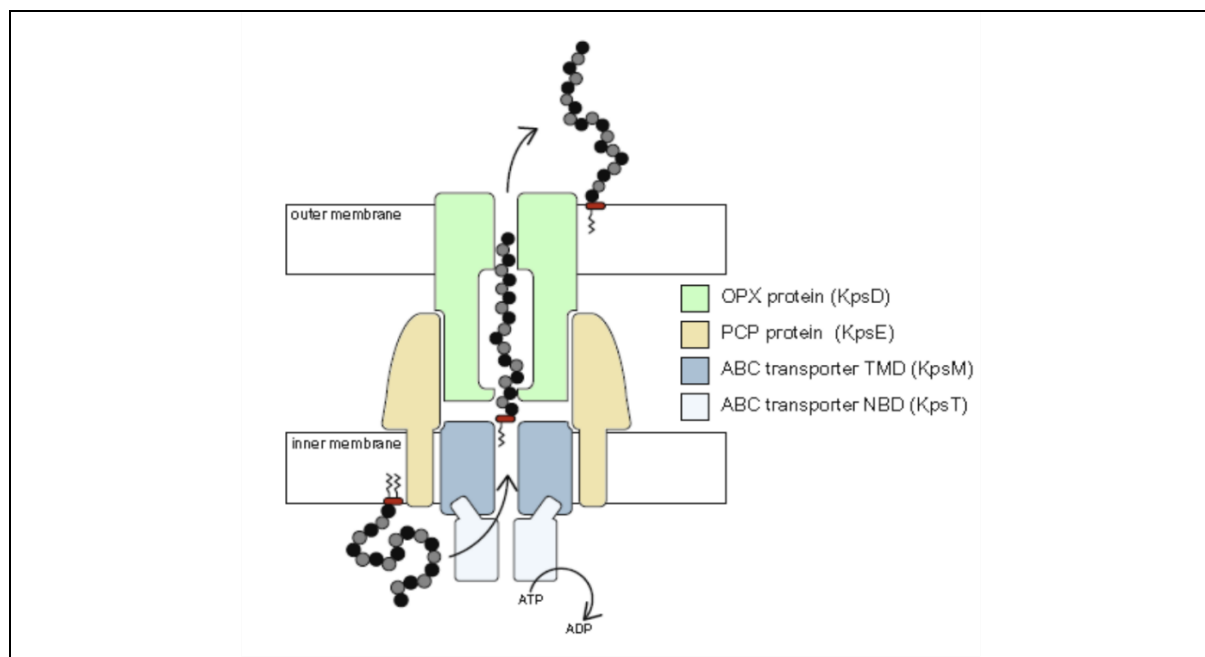
Moreover capsular polysaccharides can also bind to cationic antibiotics. Exposure to antibiotics increases the protection of CPS. Thus, thickening the capsule and reducing antibiotic penetration into the cell [14, 15].

## 1.5 Biosynthesis and exportation of *E. coli* capsule

Biosynthesis and exportation of capsular polysaccharides is a complex process. Group I and IV capsules share a common process that is Wzy-dependent pathway, while group II and III use ATP-binding cassette (ABC) transporter dependent assembly. Both pathways are explained in detail in the following sections.

### 1.5.1 ABC transporter dependent pathway

ABC-transported dependent pathway is present in group II and group III *E. coli*. These bacteria are mainly responsible for diseases like septicemia, meningitis, urinary tract infections, gastrointestinal infections, and otitis in humans. The enzyme, glycosyl transferase catalyzes the transfer of specific sugar to non-reducing end of growing CPS glycan. Synthesized CPS is transported to cell surface by inner membrane ABC transporter, polysaccharide co-polymerase (PCP) protein and outer membrane polysaccharide export (OPX) protein.



**Figure 1.2. ATP-transporter assembly involved in CPS synthesis. [Adopted from Lisa M. Willis, 2013.]**

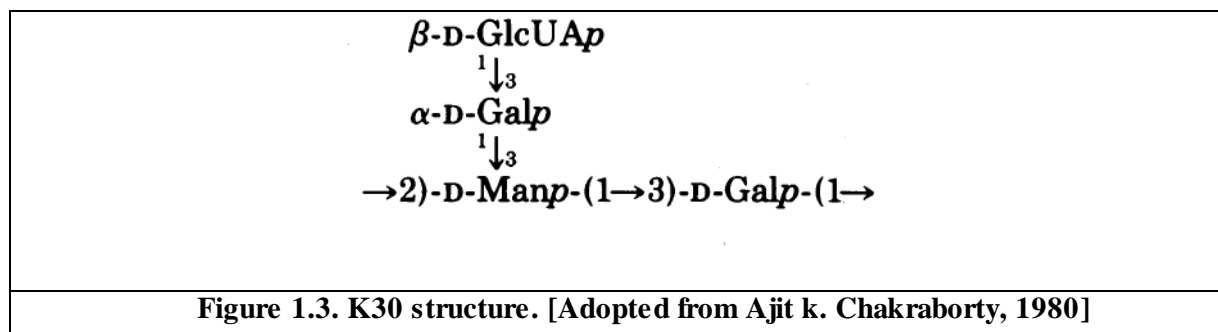


ABC transporter includes two nucleotide binding domains (NBD), KpsT and two integral membrane domains, KpsM (Fig 1.2). NBDs are responsible for ATP hydrolysis, while the transmembrane domains facilitate transportation by conformational changes. KpsE from polysaccharide copolymerase (PCP-3) family, transports CPS through the periplasmic region. Outer membrane polysaccharide export protein, KpsD, forms octameric outer membrane channel for exportation of nascent polysaccharide.

CPS assembled by ATP-transporter assembly has phospholipid, lyso-phosphatidylglycerol (PG), at the reducing end of the polysaccharide chain. PG is attached to CPS via a novel  $\beta$ -linked poly-3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) linker [9]. This lipid may be involved in attachment of CPS to cell surface [10].

### 1.5.2 Wzy- dependent pathway

Wzy-dependent pathway is present in group I and IV capsules, which are associated with gastrointestinal diseases. K30 CPS is used as a prototype for group I capsular polysaccharide. Repeating unit of K30 CPS [16] is shown in Fig. 1.3. All group K antigens are negatively charged due to the presence of uronic acid or pyruvate [12].



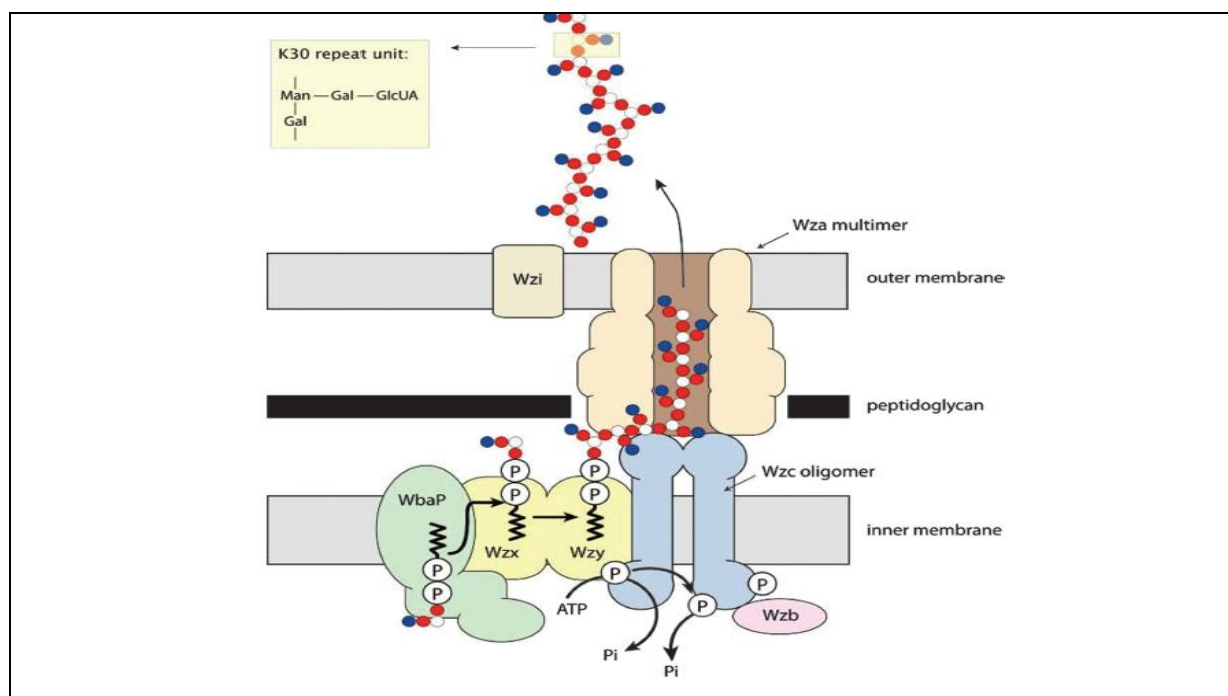
Initiation of group I *E. coli* capsular polysaccharide synthesis begins with WbaP enzyme (Fig. 1.4) from polyisoprenyl-phosphate hexose-1-phosphate transferases family [17]. WbaP has two functions i) uridine diphosphate (UDP)-galactose: Undecaprenol-P galactose-1-phosphate transferase reaction and ii) release of Und-PP-galactose from WbaP [18].

Undecaprenol-PP-linked tetrasaccharide is then flipped to periplasmic space by Wzx flippase (Fig. 1.4). It belongs to the family of polysaccharide-specific transport (PST) proteins and is designated as PST-1 due to involvement in capsule transport assembly. Wzx (*Pseudomonas aeruginosa*) has a cationic vestibule and mediates acidification of proteoliposome due to presence of outward proton transport. Since the protein undergoes conformational changes on binding and release of protons. It

was proposed that proton dependent antiport mechanism may be responsible for translocation of nascent tetrasaccharide across inner membrane [19].

Subsequently, Wzy polymerase (Fig. 1.4) present in the inner membrane polymerizes the tetrasaccharide. It has approximately 12 transmembrane segments and a large periplasmic loop. Wzy has the capability of releasing the K antigen tetrasaccharide from its lipid carrier [7] and thus polymerize.

Wzc, tyrosine (BY) autokinase, (Fig. 1.4) participates in high level polymerization of K antigen by cycling of its phosphorylation state. It has two domains, an N-terminal transmembrane loop and a cytosolic domain. Cytosolic domain has Walker A and B motifs which are involved in ATP binding and hydrolysis [20]. A Y-cluster is also present in this domain which constitutes of seven tyrosine residues within last 17 amino acids. It also forms part of capsule translocation complex along with Wza. Wzb, cognate cytoplasmic phosphotyrosine phosphatase, is responsible for dephosphorylation of tyrosine residues of Wzc and thus making phosphorylation a cyclic event [20]. This cycle of phosphorylation and dephosphorylation regulates the polysaccharide chain length. Wzb (Fig. 1.4) belongs to cysteine-dependent tyrosine phosphatase superfamily [21].



**Figure 1.4. Schematic representation of Wzy dependent pathway. [Adopted from Changjiang Dong, 2006]**

Wza (Fig 4.), an outer membrane translocon, forms SDS-stable octamer spans that across periplasmic region and outer membrane to form exportation channel [22],[23]. The bulk of protein that is located

in periplasmic space comprises of three rings [24], while the transmembrane region is  $\alpha$ -helical barrel. Appearance of octamer is best defined as an amphora without handles [25]. Tyr110 located at the tip of each loop has flexible side chain that is responsible for blocking the pore. On interaction with Wzc the pore opens and the polysaccharide is translocated [26]. Knockout studies of Wza have resulted in *E. coli* without capsule [27].

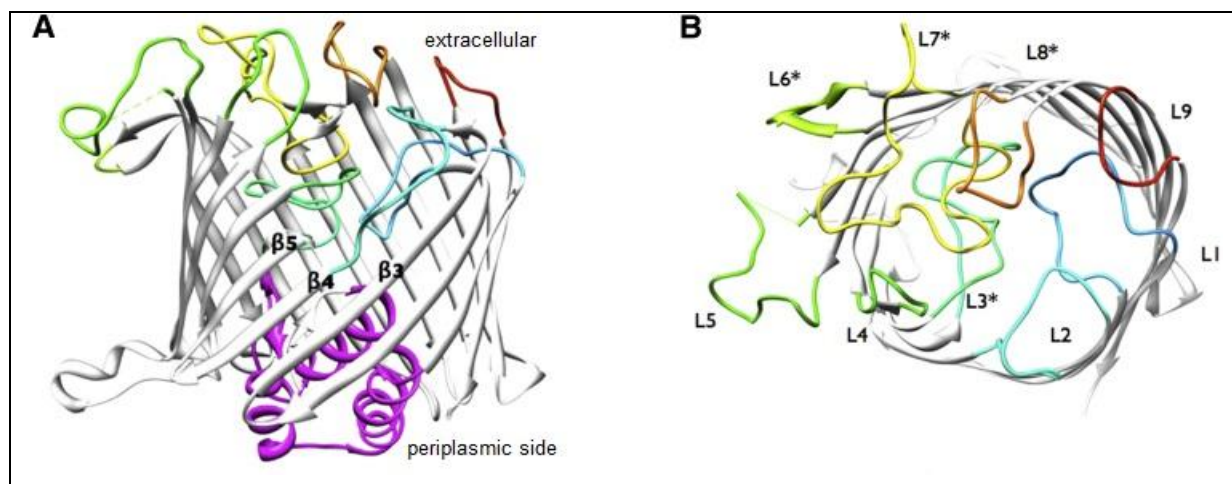
After translocation, K antigen binds covalently or noncovalently onto Wzi (Fig. 1.4), an outer membrane protein. Wzi is the critical factor in initial anchoring of polymeric carbohydrate to the hydrophobic outer membrane and thereby creates a template for capsule formation [28]. The exact binding site of capsular polysaccharide is, however, unknown. Through molecular dynamics (MD) simulation, it has been shown that Wzi also acts as porin and regulate water flow through outer membrane, suggesting that it may play a role in osmoregulation [29]. There are five diffusion points in Wzi, three on extracellular side and two on periplasmic side. Deletion of wzi results in no capsule formation even though polysaccharide is secreted and translocated [30].

Affecting the biogenesis and exportation of capsule of encapsulated bacteria will make the bacterium susceptible to the host immune system. As Wzi and Wza are exposed to the cell exterior, they can be easily targeted. Concomitant to this fact, an effective blocker of Wza has been discovered, Otakis(6-deoxy-6-amino)-cyclomaltooctaose (am8 $\gamma$ CD 13) [31]. It blocks the Wza pore and thus prevents exportation of K antigens. Hampering the nucleation of capsule, i.e binding of capsule onto Wzi, can be another adept target. Since, Wzi also acts as an osmoregulator, affecting the water flow can also lead to cell rupture and death of the bacterium. This study focuses on water conducting property of Wzi and throws light on other potential targets in Wzi. In next section, Wzi structure is explained in detail.

## 1.6 Wzi structure

The 2.6 Å crystal structure of Wzi reveals that it is 18-stranded antiparallel  $\beta$ -barrel structure spanning the outer membrane with a diameter of 36 Å (Fig 1.5 A). Wzi has 9 loops on the extracellular side and 8 turns and 3 amphipathic  $\alpha$ -helices on the periplasmic side (Fig 1.5 B). Three long extracellular loops—L3, L5, and L7—adopt distinct structures, parts of which fold into the barrel, whereas loops L5, L6, and L7 extend beyond the circumference of the barrel wall [32] [28]. Most of the loops fold into the barrel and occlude the pore. The length of  $\beta$ -strands is irregular, smallest one ( $\beta$ 4) that results in a triangular “notch” at the extracellular side.  $\alpha$ -helices are amphipathic in nature and are thus stacked against each other.





**Figure 1.5. Wzi structure** A) Wzi X-ray crystal structure is oriented with the extracellular face at the top. The periplasmic N-terminal helical bundle is colored magenta. Sheets  $\beta 3$ ,  $\beta 4$ , and  $\beta 5$  are labeled. B) Extracellular face is shown with loops colored blue to red (L1-L9). [Adopted from Simon R. Bushell, 2013.]

### 1.7 Scope of the study

Antimicrobial resistance (AMR) is a major growing global threat. Microbes are capable of gaining resistance to drugs quickly on their frequent use. This maintains the need for development of a mechanism that would affect the basic function of bacteria in such a way that it will not acquire antibacterial drug resistance. One such mechanism that may be to affect the basic function of capsules of encapsulated Gram-negative bacteria. Capsular polysaccharides (CPS) are one of major virulence factors in Gram-negative bacteria. For inhibiting capsule formation we can target the capsule formation and assembly mechanism. There are two biosynthesis pathways used by many Gram-negative bacteria for CPS synthesis and translocation- Wzy dependent pathway and ABC-transporter dependent pathway. In Wzy dependent pathway, it has been found in earlier studies that Wzi, an outer membrane lectin, plays a key role in CPS association to outer membrane thus initiating capsule formation. Also, it conducts water that may be essential to keep the capsule hydrated. In this context, here we aim to understand the water conducting property of Wzi in *E. coli* using molecular dynamics simulations. We also wish to explore and compare Wzi water diffusion in *Klebsiella pneumoniae* with that of *E. coli*. Thus, the outcome of this study would aid in the understanding this new function of Wzi protein and also will throw light on CPS hydration and distribution on the membrane. This would facilitate the exploration of new targets for designing antibiotics.

### 1.8 References

1. O'Neill, J., *Tackling Drug-Resistant Infections Globally: final report and recommendations*. 2016.
2. WHO, *Antimicrobial resistance: Global report on surveillance*. 2014.

3. Beveridge, T.J., *Structures of gram-negative cell walls and their derived membrane vesicles*. J Bacteriol, 1999. **181**(16): p. 4725-33.
4. Rasko, D.A. and V. Sperandio, *Anti-virulence strategies to combat bacteria-mediated disease*. Nat Rev Drug Discov, 2010. **9**(2): p. 117-28.
5. Greenfield, L.K. and C. Whitfield, *Synthesis of lipopolysaccharide O-antigens by ABC transporter-dependent pathways*. Carbohydr Res, 2012. **356**: p. 12-24.
6. Ørskov, F. and I. Ørskov, *Escherichia coli serotyping and disease in man and animals*. Canadian Journal of Microbiology, 1992. **38**(7): p. 699-704.
7. Whitfield, C., *Biosynthesis and assembly of capsular polysaccharides in Escherichia coli*. Annu Rev Biochem, 2006. **75**: p. 39-68.
8. *Methods in Microbiology*. 1985.
9. Aucken, H.M., S.G. Wilkinson, and T.L. Pitt, *Re-evaluation of the serotypes of Serratia marcescens and separation into two schemes based on lipopolysaccharide (O) and capsular polysaccharide (K) antigens*. Microbiology, 1998. **144** ( Pt 3): p. 639-53.
10. Qureshi, S. *Klebsiella Infections*. 2015 Oct 06, 2015.
11. Glode, M.P., et al., *Neonatal Meningitis Due to Escherichia coli K1*. The Journal of Infectious Diseases, 1977. **136**: p. s93-s97.
12. Kunduru, B.R., S.A. Nair, and T. Rathinavelan, *EK3D: an E. coli K antigen 3-dimensional structure database*. Nucleic Acids Res, 2016. **44**(D1): p. D675-81.
13. Simoons-Smit, A.M., A.M. Verweij-van Vught, and D.M. MacLaren, *The role of K antigens as virulence factors in Klebsiella*. J Med Microbiol, 1986. **21**(2): p. 133-7.
14. Ganal, S., et al., *Effects of streptomycin and kanamycin on the production of capsular polysaccharides in Escherichia coli B23 cells*. J. Exp. Microbiol. Immunol, 2007. **11**: p. 54-59.
15. Lu, E., et al., *Effect of growth in sublethal levels of kanamycin and streptomycin on capsular polysaccharide production and antibiotic resistance in Escherichia coli B23*. 2008.
16. Chakraborty, A.K., H. Friebolin, and S. Stirm, *Primary structure of the Escherichia coli serotype K30 capsular polysaccharide*. J Bacteriol, 1980. **141**(2): p. 971-2.
17. Valvano, M.A., *Export of O-specific lipopolysaccharide*. Front Biosci, 2003. **8**: p. s452-71.
18. Whitfield, C. and A. Paiment, *Biosynthesis and assembly of Group 1 capsular polysaccharides in Escherichia coli and related extracellular polysaccharides in other bacteria*. Carbohydr Res, 2003. **338**(23): p. 2491-502.
19. Islam, S.T., et al., *Proton-dependent gating and proton uptake by Wzx support O-antigen-subunit antiport across the bacterial inner membrane*. MBio, 2013. **4**(5): p. e00678-13.
20. Wugeditsch, T., et al., *Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in Escherichia coli*. J Biol Chem, 2001. **276**(4): p. 2361-71.
21. Hagelueken, G., et al., *Crystal structures of Wzb of Escherichia coli and CpsB of Streptococcus pneumoniae, representatives of two families of tyrosine phosphatases that regulate capsule assembly*. J Mol Biol, 2009. **392**(3): p. 678-88.
22. Drummel-Smith, J. and C. Whitfield, *Translocation of group 1 capsular polysaccharide to the surface of Escherichia coli requires a multimeric complex in the outer membrane*. EMBO J, 2000. **19**(1): p. 57-66.
23. Ford, R.C., et al., *Structure-function relationships of the outer membrane translocon Wza investigated by cryo-electron microscopy and mutagenesis*. J Struct Biol, 2009. **166**(2): p. 172-82.

24. Collins, R.F. and J.P. Derrick, *Wza: a new structural paradigm for outer membrane secretory proteins?* Trends Microbiol, 2007. **15**(3): p. 96-100.
25. Dong, C., et al., *Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein.* Nature, 2006. **444**(7116): p. 226-9.
26. Reid, A.N. and C. Whitfield, *functional analysis of conserved gene products involved in assembly of Escherichia coli capsules and exopolysaccharides: evidence for molecular recognition between Wza and Wzc for colanic acid biosynthesis.* J Bacteriol, 2005. **187**(15): p. 5470-81.
27. Nesper, J., et al., *Translocation of group 1 capsular polysaccharide in Escherichia coli serotype K30. Structural and functional analysis of the outer membrane lipoprotein Wza.* J Biol Chem, 2003. **278**(50): p. 49763-72.
28. Bushell, S.R., et al., *Wzi is an outer membrane lectin that underpins group 1 capsule assembly in Escherichia coli.* Structure, 2013. **21**(5): p. 844-53.
29. Nair, S.A., *Structure and dynamics of outer membrane lectin involved in the surface expression of Group 1 capsular polysaccharides.* . Masters thesis, Indian Institute of Technology, Hyderabad, 2014.
30. Rahn, A., et al., *A novel outer membrane protein, Wzi, is involved in surface assembly of the Escherichia coli K30 group 1 capsule.* J Bacteriol, 2003. **185**(19): p. 5882-90.
31. Kong, L., et al., *Single-molecule interrogation of a bacterial sugar transporter allows the discovery of an extracellular inhibitor.* Nat Chem, 2013. **5**(8): p. 651-659.
32. Bushell, S.R., et al., *Crystallization and preliminary diffraction analysis of Wzi, a member of the capsule export and assembly pathway in Escherichia coli.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. **66**(Pt 12): p. 1621-5.

# Chapter 2

## Methodology- Molecular Dynamics Simulations

### 2.1 Introduction

Molecular dynamics (MD) simulation is one of the principal tools in the theoretical study of biological molecules. In MD simulations, we calculate time dependent property of a system. The molecular dynamics method was first introduced by Alder and Wainwright in 1957. Through MD simulations one can study structure, dynamics and thermodynamics of biological molecules. It provides detailed information on the fluctuations and conformational changes of the molecules and their complexes which cannot be studied through experimental techniques. MD simulation is also used in the determination of structures from x-ray crystallography and NMR experiments. It also provides means for drug discovery and development. This chapter aims to explain briefly about force field, potential energy equation and how trajectories are generated in MD simulations.

### 2.2 Force field and simulation engines

In MD simulations, the interactions of particles in a system are defined by a potential energy function. The potential energy function gives an idea of the force acting on each particle in the system. These forces are used to determine the dynamics of the system. The backbone of MD simulations, force field, performs these tasks. Force field includes the list of atom types, atomic charges, atom-typing rules, functional forms for the components of energy expression, parameters for the function terms. Goal of force field is to describe entire class of molecules with reasonable accuracy. So, the quality of the force field, its applicability to the model under study, and its ability to predict the particular properties measured in the simulation directly determine the validity of the results. Thus, it necessary to choose appropriate force field.

Simulation engines are the computational packages that handle the application of force fields. They carry out energy minimization, molecular dynamics and force field-based calculations. Few examples

are CHARMM, AMBER. These two packages have their own force field and does not support any other force field. While softwares like GROMACS and NAMD support multiple force fields.

## 2.3 Energy expression

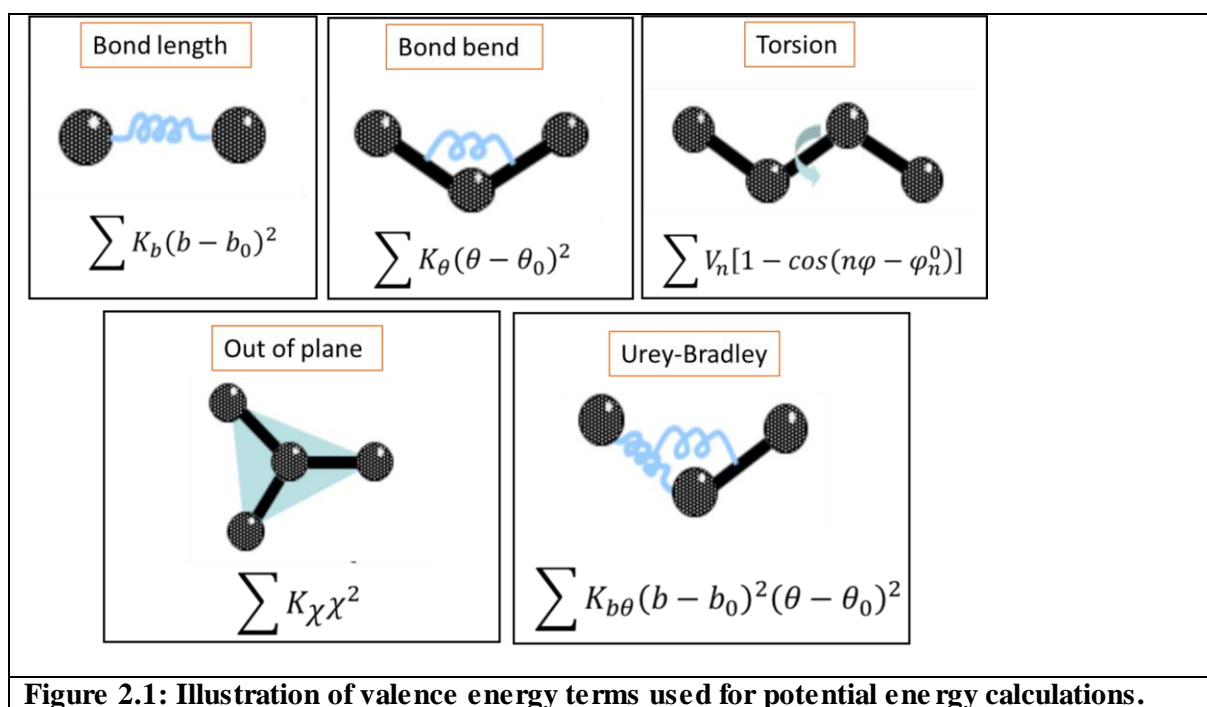
Force field describes the energy expression that drives the simulation. Typical potential energy expression system accounts for two types of interaction, bonded (valence) and nonbonded interactions.

$$E_{\text{total}} = E_{\text{valence}} + E_{\text{nonbonded}}$$

### 2.3.1 Valence interactions

Potential energy for valence interaction involves potential energy due to bond length, bond angles, torsion angle, out of plane and urey-bradley term. Valence energy terms are explained below:

- Bond length: accounts for changes in bond length ( $b$ ) from the equilibrium value ( $b_0$ ), where  $K_b$  is the bond force constant.
- Bond angle: represents change in bond angle ( $\theta$ ), from the equilibrium value ( $\theta_0$ ), where  $K_\theta$  is the angle force constant
- Torsion angle: for change in torsion angle ( $\varphi$ ), where  $\varphi_n^0$  represents the phase shift,  $n$  represent periodicity of dihedral angle and  $V_n$  is the corresponding force constant.
- Out-of-plane bending: where  $K_\chi$  is the force constant and  $\chi^2$  is the change in improper angle.
- Urey-Bradley term: this cross term accounts for angle bending using 1,3 non-bonded interactions.

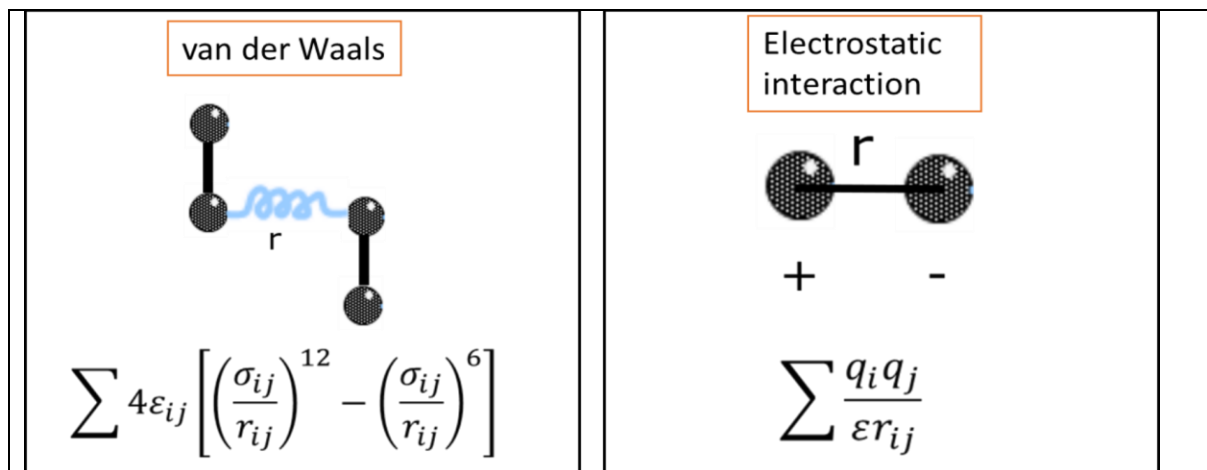


**Figure 2.1: Illustration of valence energy terms used for potential energy calculations.**



### 2.3.2 Non-bonded interactions

Non-bonded interactions include van der Waals and electrostatic interactions (shown in Fig 2.2). These interactions are defined on atom pairs and require number of atoms squared calculations. Non-bonded interactions constitute computationally demanding part of energy calculations.



**Figure 2.2: Snapshots illustrating non-bonded interactions along with the equations used for their energy calculation.**

CHARMM potential energy equation [1]:

$$\begin{aligned}
 E_{pot} &= \sum k_b (r - r_0)^2 + \sum K_\theta (\theta - \theta_0)^2 + \sum |k_\phi| - k_\phi \cos(n\phi) + \sum k_\chi (\chi - \chi_0)^2 \\
 &+ \sum_{excl(i,j)=1} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{excl(i,j)=1} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) sw(r_{ij}^2, r_{on}^2, r_{off}^2) + E_{constraint} + E_{user}
 \end{aligned}$$

### 2.4 Constraints and Restraints

Constraint is an absolute restriction imposed on a calculation while restraint is an energetic bias that tends to force the calculation towards a certain condition. Few basic constraint/restraints are explained below.

- i. *Fixed atom constraint*: atoms are fixed to fixed points in space so that they do not move during simulation run. This saves computation time by calculating only those energy terms that involve at least one atom that is free to move.
- ii. *Harmonic atom restraint*: atoms are given some freedom while constraining them to be near a particular position in space.
- iii. *Dihedral restraint*: the motion of a dihedral angle is restricted usually to a value close to a reference position.

- iv. *Distance restraint*: places a potential between two or more atoms to restrain the distance between these atoms to a reference value.
- v. *Internal coordinate constraint*: all internal coordinates of a part of system are kept fixed. This is useful to study dynamics of structure if whole structure is not well understood or to study interaction.
- vi. *SHAKE*: uses an algorithm that restores specified internal coordinates to their reference values after a minimization or dynamics step.
- vii. *Quadratic droplet constraint*: puts the entire molecule in a cage by constructing a constraining sphere around a molecular system.

## 2.5 Integration algorithms

MD simulations are based on numerical integration of Newton's second law of motion, which states that acceleration ( $a_i$ ) of a particle  $i$  depends directly on the net force ( $F_i$ ) and inversely on mass ( $m_i$ ) of the particle,

$$F_i(t) = m_i a_i(t)$$

The acceleration is given as the derivative of the potential energy with respect to the position,  $r$ ,

$$a = -\frac{1}{m} \frac{\partial E}{\partial r}$$

Newton's equation can then relate the derivative of the potential energy to the changes in position as a function of time,

$$-\frac{\partial E}{\partial r_i} = m_i \frac{\partial^2 r_i}{\partial t_i^2}$$

The initial coordinates are determined in the input file or from a previous operation such as minimization, the initial velocities are randomly generated at the beginning of a dynamics run, according to the desired temperature. Once the initial coordinates and velocities are known, the coordinates and velocities at a later time can be determined. The coordinates and velocities for a complete dynamics run are called the trajectory.

All the integration algorithms assume that the positions, velocities and accelerations can be approximated by a Taylor series expansion. Good integrator speeds up the energy evaluation. One of the algorithms is explained in section 2.5.1.

### 2.5.1 Verlet algorithm

Given a set of position  $r$ , the position after a small time step  $\delta t$  is given by Taylor's series,

$$r(t + \delta t) = r(t) + v(t)\delta t + \frac{1}{2} a(t)\delta t^2 + \frac{1}{6} b(t)\delta t^3 + \frac{1}{24} c(t)\delta t^4 + \dots$$

Similarly, positions from the previous step  $r(t-\delta t)$  is defined as

$$r(t - \delta t) = r(t) - v(t)\delta t + \frac{1}{2}a(t)\delta t^2 - \frac{1}{6}b(t)\delta t^3 + \dots$$

Adding these two equations

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

where  $r$  is the position,  $v$  is the velocity and  $a$  is the acceleration.

## 2.6 Statistical Ensemble

An ensemble is a collection of all microstates of a system, consistent with the constraints with which we characterize a system macroscopically. Different types of ensembles are listed in Table 2.1

Table 2.1 List of ensembles with the constants.

Ensemble	Constants
NVT (canonical ensemble)	No. of atoms, Volume, Temperature
NPT	No. of atoms, Pressure, Temperature
NST	No. of atoms, Stress, Temperature
NVE (microcanonical)	No. of atoms, Volume, Energy
NPH	No. of atoms, Pressure, Enthalpy
NSH	No. of atoms, Stress, Enthalpy

## 2.7 References

1. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program*. J Comput Chem, 2009. **30**(10): p. 1545-614.

# Chapter 3

## Water diffusion mechanism in Wzi of *Escherichia coli*

### 3.1 Introduction

*Escherichia coli* belongs to Enterobacteriaceae family of Gram-negative bacteria. It is a causative agent for many diseases such as urinary tract infections, bloodstream infections, neonatal meningitis, etc. Capsular polysaccharide is one of the major virulent factors in Gram-negative bacteria. *E. coli* K30 antigen is a commonly used prototype for understanding biosynthesis and exportation of Group I CPS. This chapter focuses on outer membrane protein, Wzi, which is known to be a lectin and porin. Herein, water conduction property of Wzi is explored through molecular dynamics simulations.

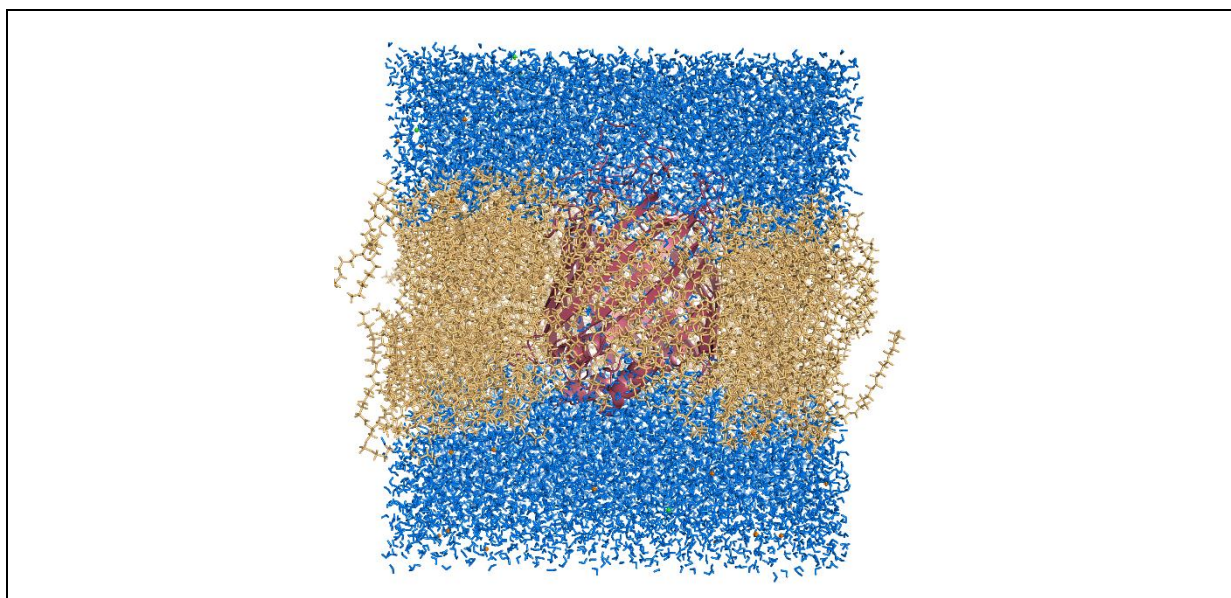
Results show that there are five different diffusion points, three at the extracellular and two at periplasmic face of the barrel which constitutes of conserved residues. There is also a conserved luminal hydrophobic plug that plays a key role in regulating Wzi water diffusion. At each entry point, water selective hydrogen bonding network regulates both water entry and exit. At the periplasmic region, entry 4b acts as a major diffusion point. At the extracellular side, the major contribution for entry is through the notch region (entry 2), while notch region and ‘YQF’ motif contributes equally to exit of water molecules from the barrel. Intriguingly, such an ‘YQF’ motif is observed to be the binding site for sugar in sodium-galactose cotransporters, implicating that it may be the binding site of K30CPS in Wzi. Wzi is predicted to conduct water so as to keep the CPS in a hydrated condition. Diffusion of water molecules in and out of Wzi may be driven by the extracellular concentration of CPS.

Another major finding is the ion binding site present on the extracellular side just below the YQF motif, proposed binding site for K antigen. A single potassium ion moves into the binding site and is strongly held there by coordination with E180 and Q107. Also, it was observed that extracellular loop 5 (L5) splays out and get embedded onto the membrane surface during simulation, elucidating its role in membrane anchorage of Wzi.

To explain the system properly herein two terms are defined. Diffusion entry is accounted when a water molecule that enters into the barrel through anyone of the extracellular entry points and exits the barrel through anyone of the periplasmic entry points. Vice-a-versa is counted as diffusion exit.

### 3.2 System setup for MD simulations

The starting model for the simulation of Wzi in *E. coli* was built from crystal structure (PDB ID: 2YNK). As 15 amino acids of extracellular loop 5 are missing from the crystal structure, the loop was modeled using ModLoop [1-3] to obtain the starting model. System for MD simulations were generated using CHARMM-GUI web interface [4-6] using heterogeneous membrane builder. To mimic the outer membrane of *E. coli*, heterogeneous lipid bilayer is used with POPG and POPE in a 1:3 [7].



**Figure 3.1 Snapshot of the system used for simulation. Protein is shown in red cartoon representation. Bilayer (sand) and water (blue) are shown in stick representation.**

The protein is embedded in the lipid bilayer using replacement method (Fig 3.1). The orientation of the protein in membrane environment is chosen based on the orientations of proteins in membranes (OPM) database [8]. Preequilibrated TIP3P [9] was used to solvate the system by adding a rectangular water box to a thickness of 17.5 Å on top and bottom of the system (~14,000 water molecules). 150 mM KCl (73 K<sup>+</sup> ions and 13 Cl<sup>-</sup>) is added to the solvent, mimic the physiological conditions using Monte-Carlo method [10]. Subsequently, each system is subjected to 0.5 μs simulation. MD is performed using GROMACS 4.6.4[11] software with CHARMM [12,13] 36 force field using the parameters explained in the next section.

### 3.3 MD simulation



After the system is assembled, it was first subjected to 375 ps equilibration run. Prior to equilibration, the system was minimized by running 5000 cycles of steepest descent minimization. Equilibration was done in six steps. First two steps were 25 ps of MD simulation (NVT ensemble) using leapfrog algorithm with 1 ps time interval, wherein, Berendsen thermostat is used to control the simulation temperature with 1 ps time coupling. Subsequently, 25 ps of MD simulation with NPT ensemble with 1 fs time interval was performed. This was followed by three steps of 100 ps of MD simulation with 2 fs time step. Berendsen barostat [14] was used to maintain the semiisotropic pressure (1 bar) with a time coupling 5 ps. Using leapfrog integrator, production run was extended up to 0.5  $\mu$ s for each system along with Nose-Hoover thermostat [15] (303 K) and Parrinello-Rahman [16] barostat (1bar) with 1 ps and 5 ps coupling respectively. Pressure in the system is maintained in a semiisotropic manner i.e. pressure in the X-Y plane of the membrane is independent of Z-direction.

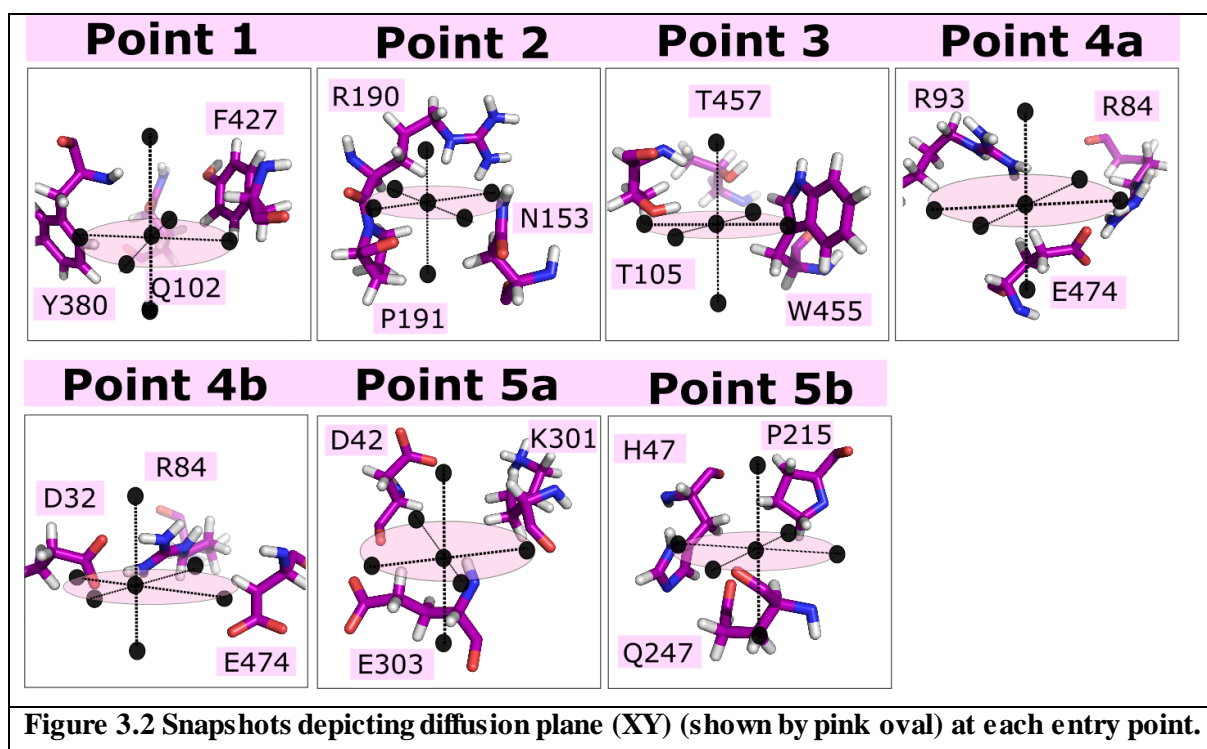
### **3.4 Methodology: Analysis of the trajectory**

#### **3.4.1 Quantification of number of water molecules in the barrel**

Wzi protein is aligned along Z-axis and the center of the barrel is around 50 Å. To calculate number of water molecules inside the barrel per ns, water molecules that fall within the minimum and maximum limits of X & Y coordinates of the  $\beta$ -barrel and between  $-10$  Å &  $10$  Å from the center of the barrel along the Z-axis were averaged over each ns.

#### **3.4.2 Quantification of barrel permeation events**

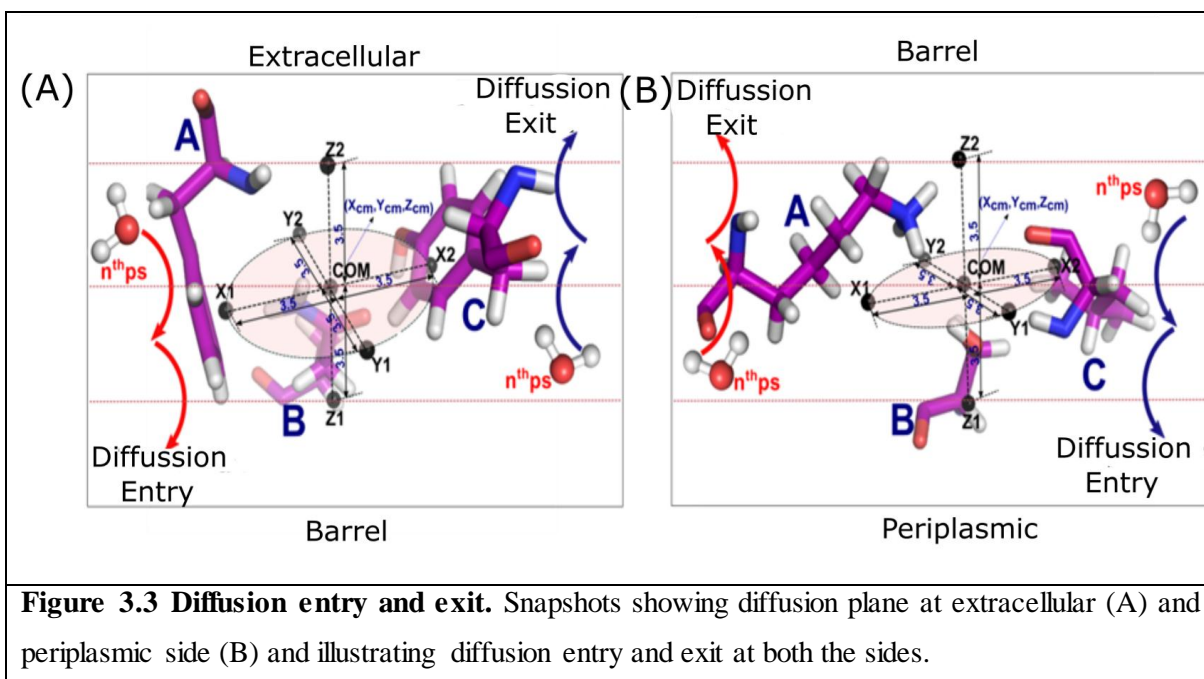
Barrel permeation events are defined as the complete crossing of a water molecule from one side of the barrel to the other. Permeation events were calculated from all the five entry points for last 200ns. The water molecules residing inside the barrel at any given point of time are only considered. A diffusion plane is defined here, through which, a water molecule must pass to be considered for diffusion calculation. At each entry point, the center of mass ( $X_{cm}$ ,  $Y_{cm}$ ,  $Z_{cm}$ ) is calculated by considering the atoms corresponding to the X, Y & Z coordinates of three amino acids that are involved in the regulation of water diffusion. Subsequently, a local diffusion plane (XY-plane) is defined with the upper and lower bounds of ( $X_{cm}+3.5$  Å,  $Y_{cm}+3.5$  Å) and ( $X_{cm}-3.5$  Å,  $Y_{cm}-3.5$  Å) respectively. Diffusion plane of all the entry points with corresponding amino acids are shown in Fig 3.2.



**Figure 3.2 Snapshots depicting diffusion plane (XY) (shown by pink oval) at each entry point.**

Diffusion entry calculation: A water entry into the barrel from the extracellular side is considered only when a water molecule that has Z coordinate greater than  $Z_{cm}$  at nth ps subsequently falls within the defined XY-plane (for the respective entry point) with Z-coordinates between  $Z_{cm}$  and  $(Z_{cm} - 3.5 \text{ \AA})$  ( $Z1$ ) and soon after it crosses  $(Z_{cm} - 3.5 \text{ \AA})$  (Fig 3.3 A). Subsequently, if the water molecule that has Z coordinate between  $Z_{cm}$  and  $(Z_{cm} + 3.5 \text{ \AA})$  ( $Z2$ ) at nth ps at the periplasmic side, falls within the defined XY-plane of the diffusion point with Z-coordinates between  $Z_{cm}$  and  $(Z_{cm} - 3.5 \text{ \AA})$  ( $Z1$ ) and soon after it crosses  $(Z_{cm} - 3.5 \text{ \AA})$  (Fig.3.3 B).

Diffusion exit calculation: Calculation of diffusion exit is swapped in such a way that a water molecule is considered when it that has Z-coordinates lesser than  $Z_{cm}$  at nth ps, subsequently falls within the defined XY-plane with Z-coordinates between  $Z_{cm}$  and  $(Z_{cm} + 3.5 \text{ \AA})$  ( $Z2$ ) and soon after it crosses  $(Z_{cm} + 3.5 \text{ \AA})$  (Fig. 3.3 B) at the periplasmic side. Subsequently, diffusion exit is counted only when the water molecule that has Z-coordinates between  $Z_{cm}$  and  $(Z_{cm} - 3.5 \text{ \AA})$  ( $Z1$ ) at nth ps, falls within the defined XY-plane of one of extracellular entry point with Z-coordinates between  $Z_{cm}$  and  $(Z_{cm} + 3.5 \text{ \AA})$  and soon after it crosses  $(Z_{cm} + 3.5 \text{ \AA})$  (Fig. 3.3 A) thus moving out of the barrel.



### 3.4.3 Diffusion permeability

Diffusion permeability ( $p_d$ ) is defined as the exchange of individual water molecules between the extracellular and periplasmic sides at equilibrium. Diffusion permeability can be calculated using following equation:

$$p_d = \left( \frac{V_w}{N_A} \right) q_0$$

where  $V_w$  is the average volume of a single water molecule ( $18\text{cm}^3/\text{mol}$ ),  $N_A$  is the Avogadro's number and  $q_0$  is the rate of unidirectional permeation events per ns [17]. Here, diffusion permeability was calculated for last 200ns of the simulation.

### 3.4.4 Entry and exit calculation

In this calculation entry and exit from both the extracellular and periplasmic sides were calculated over last 200ns. Water molecules that fall within  $5 \text{ \AA}$  range of center of mass ( $X_{cm}$ ,  $Y_{cm}$ ,  $Z_{cm}$ ) corresponding to each entry point for every 1 ns trajectory were considered. Subsequently, it was checked whether a specific water molecule is between  $Z_{cm}$  and  $(Z_{cm} + 5 \text{ \AA})$  or between  $Z_{cm}$  and  $(Z_{cm} - 5 \text{ \AA})$ . At the extracellular entry points, if the water molecule was found to be between  $Z_{cm}$  and  $(Z_{cm} + 5 \text{ \AA})$ , it was counted as entry soon after it crosses  $(Z_{cm} - 5 \text{ \AA})$  and the reverse was considered for exit. Similarly, water molecule moving from  $(Z_{cm} - 5 \text{ \AA})$  to  $(Z_{cm} + 5 \text{ \AA})$  was accounted as entry and the reverse was counted as exit at the periplasmic side. Presence of the water molecule within the XY plane of the appropriate entry point (i.e  $X_{cm} \pm 3.5$  &  $Y_{cm} \pm 3.5$ ) during the above calculation was also considered.

### **3.4.5 Net flux calculation**

Attributing to the tortuous nature of the water conduction path inside the barrel, net flux is calculated with water entry and exit events and not with the permeation events. Once a water molecule enters the barrel, it is indistinguishable. Since the number of water molecules inside the barrel remains almost constant, a water molecule entering from one side of the barrel must be accompanied by a water molecule moving out from the other side of the barrel. Net flux is calculated by subtracting the total number of water molecules that exit from the barrel through all the entry/exit points from the total number of water molecules that enter inside the barrel through all the entry/exit points over the last 200 ns.

### **3.4.6 Normalized frequency**

Normalized frequency was calculated for entry/exit mechanisms at each entry point. Following the above entry/exit calculation method described in section 3.4.4, normalized frequency is calculated by additionally imposing a hydrogen bond distance criterion (3.5 Å) between the water oxygen atoms and the side or main chain electronegative atoms of the corresponding amino acids.

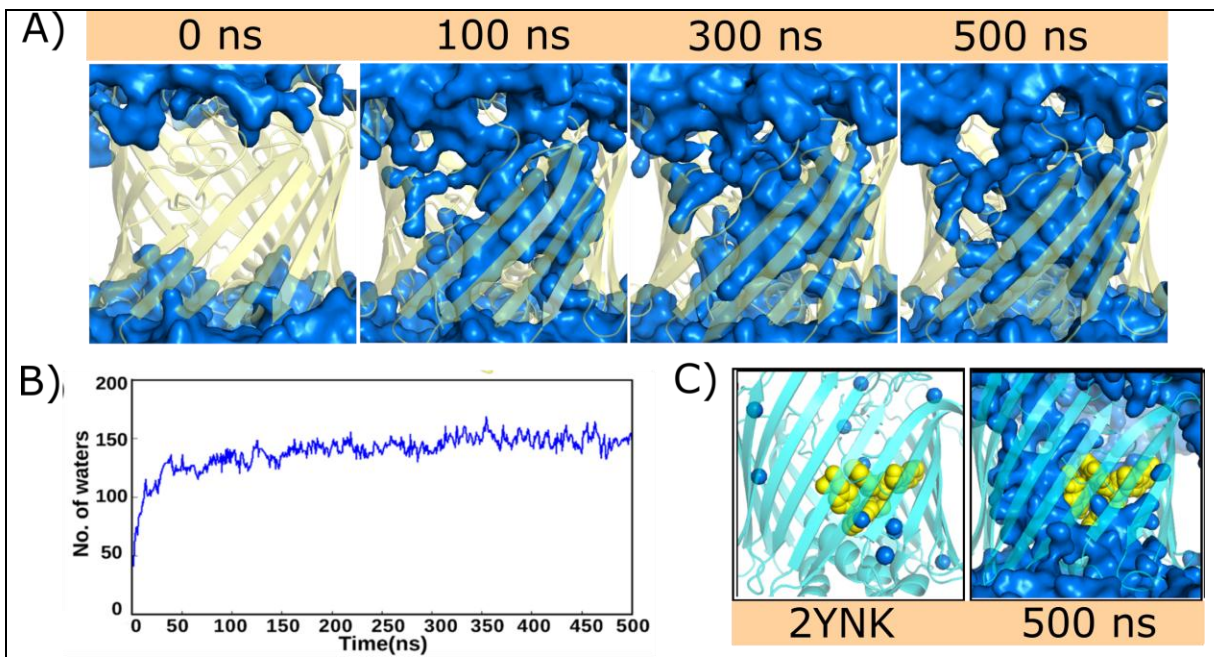
### **3.4.7 Ion coordination number**

Ion coordination number was calculated with a cut off distance for coordination bond was taken as 3.6 Å [18]. Coordination with water molecules and oxygen atoms of appropriate amino acids were calculated individually at every ns. These were summed up to get total coordination of the ion throughout the trajectory. It should be noted that the ion that is entering into the barrel was only considered for this calculation.

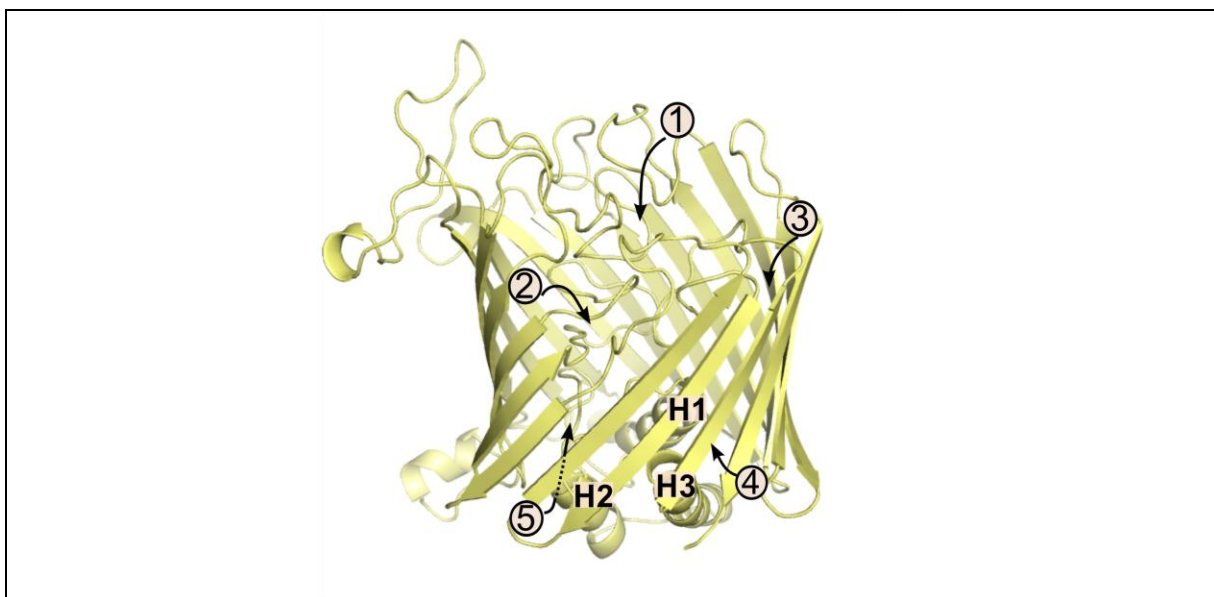
## **3.5 Results and discussion**

### **3.5.1 Dual Functional role of Wzi as a porin**

Water moves inside the barrel from both extracellular and periplasmic sides (Fig 3.4 A). The lumen of Wzi gets completely filled with water by 60 ns. During the simulation, number of water molecules inside the barrel increases with time till ~60 ns (non-equilibrated state) and then remains around 150 during the rest of the simulation (equilibrated state) (Fig 3.4 B). Crystal structure of Wzi (2YNK) has water molecules present inside the barrel, which coincides with the water conduction path observed (Fig 3.4 C). This fact strengthens the water conducting property of Wzi and shows that it acts as a porin.



**Figure 3.4. Water diffusion into the Wzi barrel.** (A) With increase in time water moves inside the barrel. Protein is shown in pale yellow cartoon representation and water is shown in blue surface representation. (B) Time vs number of water molecules inside the barrel, indicating ~150 molecules within the barrel at the end of the simulation. (C) Comparison of crystal structure (PDB ID: 2YNK) of Wzi (Left) with (Right) Wzi at 500 ns. Hydrophobic plug encompassing residues W53, F109, W174 and W175 is shown in yellow spheres. Note the water molecules (blue spheres) inside the  $\beta$ -barrel of Wzi crystal structure (cyan cartoon) coincides with the water-conducting path (blue surface) of the Wzi (cyan cartoon) during simulation.



**Figure 3.5. Diffusion points in Wzi, *E. coli*.** Three diffusion points are present at the extracellular side, entry 1, 2 and 3, and two points are present at the periplasmic side i.e. entry 4 and 5.

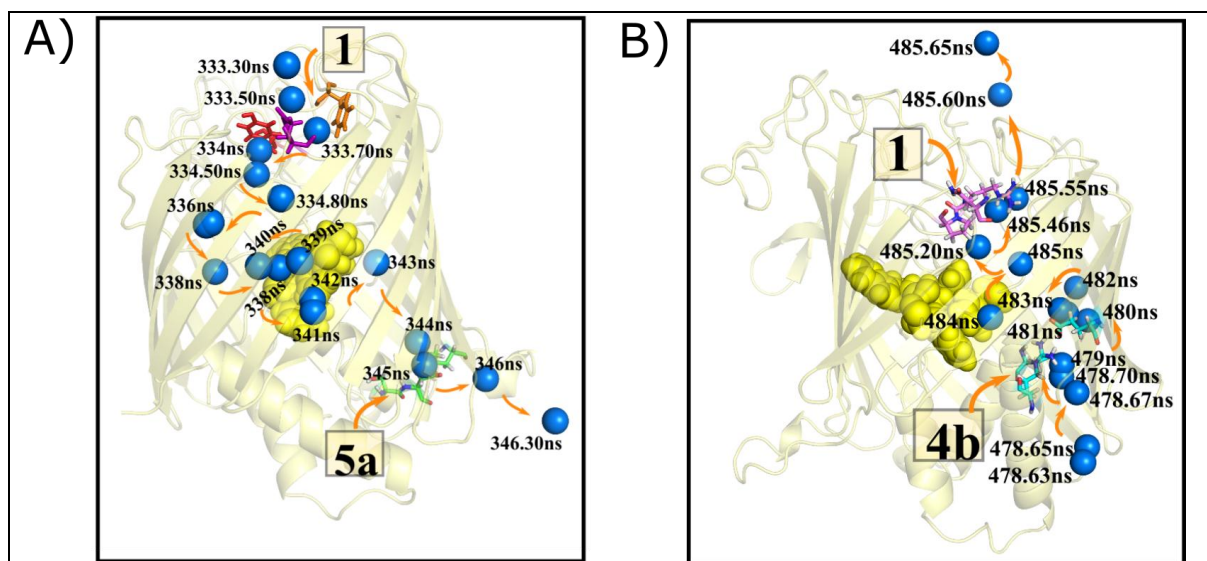


### 3.5.2 Diffusion points

Five diffusion points have been observed in Wzi of *E. coli*, three at the extracellular side and two at the periplasmic side. Three different points at the extracellular side involves conserved residues: Y380, Q102 & F427 (entry 1, Fig 3.5), N153, R190, P191, N226, Q227 & Q378 (entry 2, Fig 3.5) and T105, Q107, Q437, W455 & T457 (entry 3, Fig 3.5). At the periplasmic side, water entry at point 4 is regulated through H1, H3 &  $\beta$ -barrel encompassing residues D32, R84, R93 & E474 (entry 4, Fig 3.5), entry point 5 involves residues D42, K301, E303, H338, H47, P215, Q247 & S244 (entry 5, Fig 3.5).

### 3.5.3 Bidirectional water permeation

Complete crossing of water molecules across the barrel, from both the sides have been observed, following a nonspecific pathway (Fig 3.6). Over the last 200 ns of the simulation, 17 water molecules crosses the barrel from the periplasmic side to the extracellular side and 21 water molecules from extracellular side to the periplasmic side. Wide lumen inside the barrel on either side of the hydrophobic plug permits the water to follow nonspecific path thus contributing the small difference in the permeation events. This also reflects in the time taken by the water molecules to cross the barrel: 8–80 ns (extracellular to periplasmic side) and 6–67 ns (periplasmic to extracellular side).



**Figure 3.6. Single water molecule (blue spheres) conduction pathway.** (A) from extracellular to the periplasmic side and (B) periplasmic to extracellular side. Arrows have been marked to show water path and corresponding time is indicated. Entry points involved in conduction are showed in stick representation. Hydrophobic plug is shown as yellow spheres.

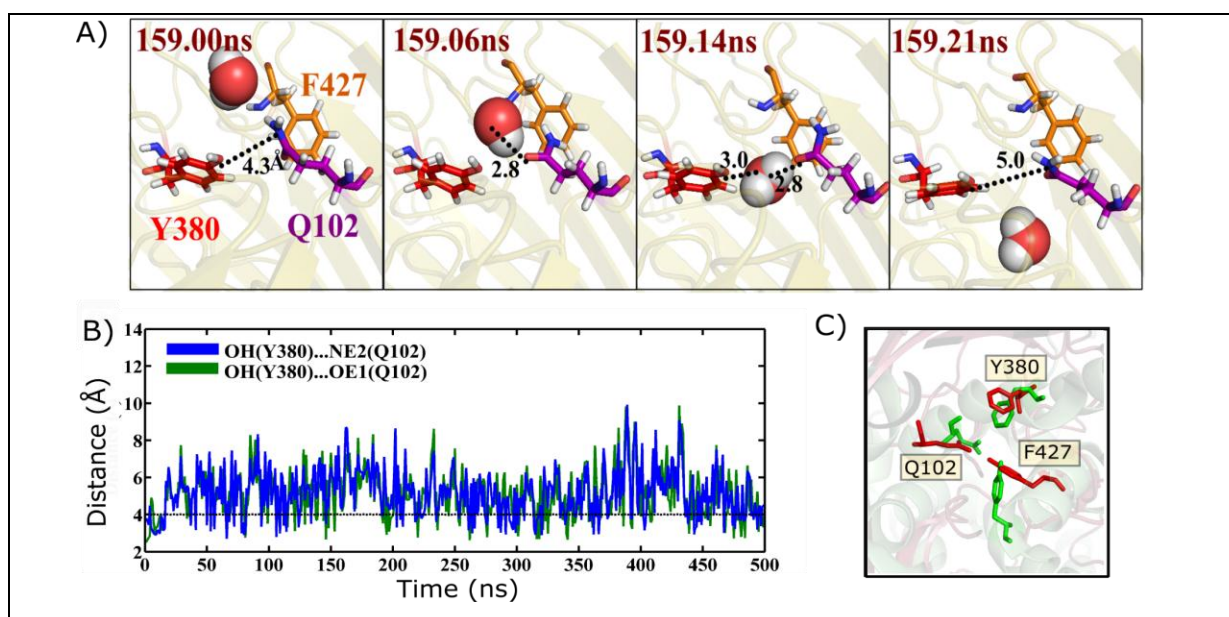
### 3.5.4 Diffusion permeability

For calculating diffusion permeability, total permeation events for last 200ns are considered i.e. 38 water molecules. So,  $q_0$  is equal to 0.095 and thus diffusion permeability is  $2.8 \times 10^{-15} \text{ cm}^3 \cdot \text{s}^{-1}$ .

### 3.5.5 Diffusion mechanisms

#### 3.5.5.1 Entry 1: YQF triad

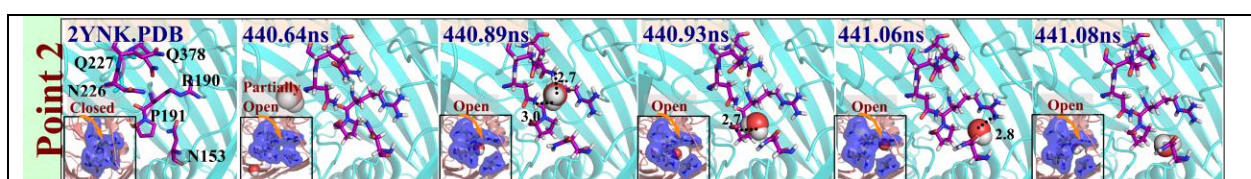
Entry point one, present on the extracellular side, involves Phe427, Tyr380 & Gln102 (Fig 3.7 A). Y380 & Q102 residues are in hydrogen bond “locked” state in crystal structure. During the “locked” state, OH group of Y380 is engaged in hydrogen bonding with NE2 and/or OE1 of Q102 resulting in blockage of water diffusion. However, transient opening and closing of the pore is involved in diffusion of water molecules inside the barrel (Fig 3.7 B). The mechanism of diffusion between Y380&Q102 is such that soon after a water molecule comes close to Q102 & Y380, it unlocks the pore by sitting in between their side chains by forming hydrogen bond(s). This is accompanied by increase in Q102&Y380 distance and consequently, the water molecule relinquishes itself from Q102&Y380 by breaking the hydrogen bond and enters into the pore (Fig. 3.7 A). This water entry mechanism is highly conserved with normalized frequency (NF) of 1. YQF triad seen here is structurally similar to the one found in the sugar binding site of sodium-galactose cotransporters (Fig 3.7 C). Thus, it is proposed that the triad may act as binding site for K30 antigen.



**Figure 3.7 Diffusion mechanism of YQF triad.** (A) Snapshots illustrating movement of water molecule into the barrel through YQ lock. (B) Distance plot to show transient opening and closing of YQ lock (C) Superposition of sugar binding site of *Vibrio parahaemolyticus* sodium-galactose cotransporter (vSGLT) (PDB ID: 3DH4, colored green) and ‘YQF’ triad of Wzi illustrating their structural similarity (PDB ID: 2YNK, colored red).

#### 3.5.5.2 Entry 2: notch region

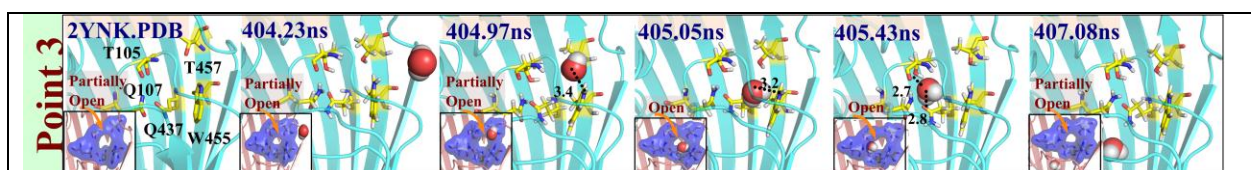
Notch region is also present on the extracellular side. Water entry via the notch is regulated by the pore formed by residues, Asn153, Arg190, Pro191, Asn226, Gln227 and Gln378 that is in closed state (Fig.3.8). Soon after a water molecule appears in the vicinity of Q227 that is protrude onto the extracellular surface, it hydrogen bonds with the amino and/or carbonyl group(s) of Q227. Subsequently, the water molecule relinquishes itself from Q227 and enters inside the barrel by breaking ND2(N226)...OE1(Q378) hydrogen bond by establishing a water mediated ND2(N226)...OE1(Q378) hydrogen bond concomitant with the complete pore opening. This event is accompanied by the movement of N153 side chain. Then, the water further diffuses into the channel, by establishing a hydrogen bond with the carbonyl group of P191 and amino group of R190. Hydrogen bonding scheme mentioned above is not conserved for entry of all water molecules.



**Figure 3.8. Water diffusion mechanism for entry 2. Snapshot from crystal structure (left) have been included. Pore opening and closing in shown in inset.**

### 3.5.5.3 Entry 3: Threonine flipping

Entry point 2, facilitates water diffusion through threonine flipping mediated by hydroxyl group of T457 (Figure 3.9). This entry point includes five residues, Thr105, Gln107, Gln437, Trp455 and Thr457, on the extracellular side. The mechanism is such that, when a water molecule enters into the channel, the hydroxyl group of T457 side chain hydrogen bonds with the water molecule and pushes it inside the channel by flipping. It then breaks the hydrogen bond with T457 and slides into the channel through aromatic W455. Hydrogen bonding with T105, Q107 and Q437 helps in diffusing the water molecule inside the barrel. Threonine flipping is conserved with NF of 0.92.



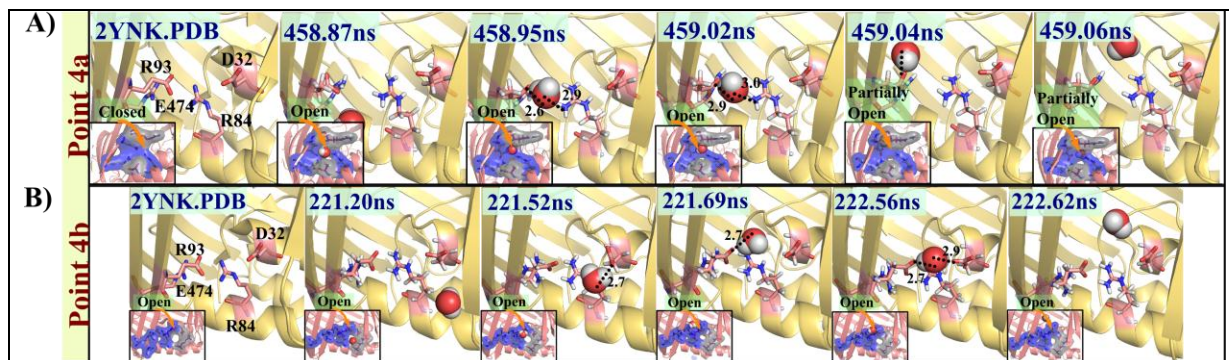
**Figure 3.9. Water diffusion mechanism for entry 3 involving residues T105, Q107, Q437, W455 & T457 (yellow sticks). Snapshot from crystal structure (left) have been included. Pore opening and closing in shown in inset.**

### 3.5.5.4 Entry 4: bifurcated salt bridge

At this periplasmic entry point, four residues, Asp32, Arg84, Arg93 and Glu474, makes two different entries viz. 4a (R84, R93 & E474) and 4b (D32, R84 & E474). When a water molecule appears in the vicinity of these residues, it hydrogen bonds initially with anyone of the amino



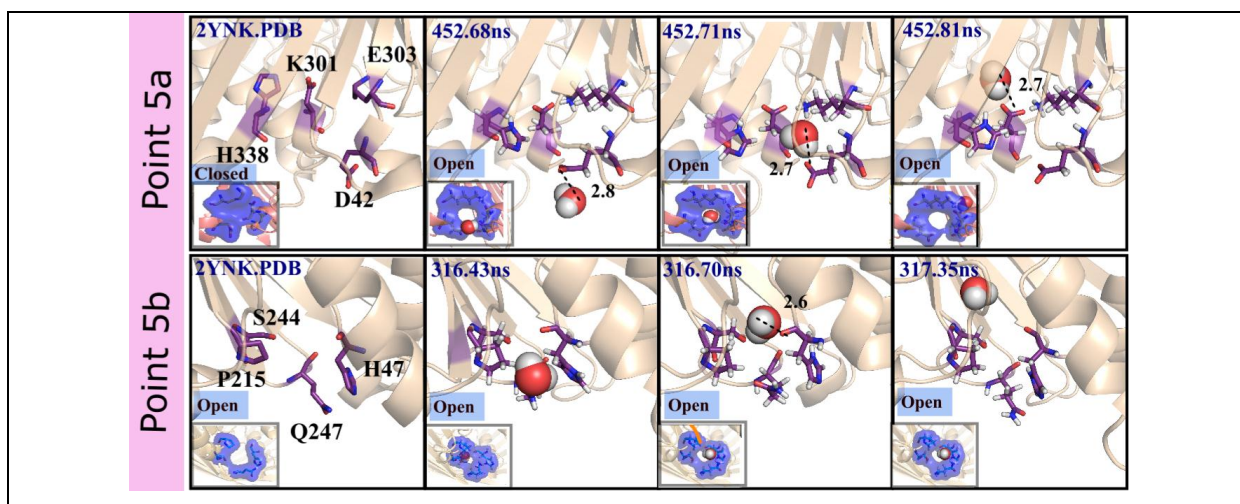
groups of R84 or R93 and subsequently side chain of E474 and diffuses slowly into the barrel without disrupting the R93 & E474 and R84 & E474 salt bridges (Fig 3.10 A). At entry 4b, water molecule first interacts with R84 or E474, and is pushed inside the barrel while interacting with D32 (Fig 3.10 B). Entry 4b contributes to the total permeation events calculated from periplasmic to extracellular side (17 water molecules). Mechanisms of both the entry points are highly conserved with a normalized frequency of 1.



**Figure 3.10. Water diffusion mechanism for entry 4a (A) and 4b (B).** Snapshot from crystal structure (left) have been included. Pore opening and closing in shown in inset. The residues directly involved in diffusion mechanisms are shown in blue surface, while grey surface depicts the residues chosen to demonstrate pore opening.

### 3.5.5.5 Entry 5: entry between helices and barrel

Point 5 involves a total of eight periplasmic residues which contributes to two adjacent entry points. Entry 5a involves Asp42, Lys301, Glu303 and His338, while entry 5b involves His47, Pro215, Ser244 and Gln247. Water entry through point 5a is initiated by the interaction of water to carboxylic group of E303 (Fig 3.11 A). Upon discharging from this hydrogen bond, water molecule moves into the D42, K301 and H338. Point 5b facilitates water entry through a pore encompassing the residues H47-P215-Q247 (Fig 3.11 B). Water entry is initiated by hydrogen bond formation between the water molecule and the side chain of Q247. Subsequently, the water moves inside the barrel by hydrogen bonding with backbone of H47 and OH of S244. These mechanisms are conserved with normalized frequency 1.



**Figure 3.11. Water diffusion mechanisms of periplasmic entry 5. Snapshots depicting water diffusion into the barrel with corresponding time indicated. Opening and closing of pore is shown in inset. Snapshot from crystal structure (left)**

It is noteworthy that all the residues mentioned above are also involved in water exit through the respective entry points.

### 3.5.6 Entry & Exit and Flux calculation

Entry/Exit from each entry point in last 200 ns is listed in table 3.1.

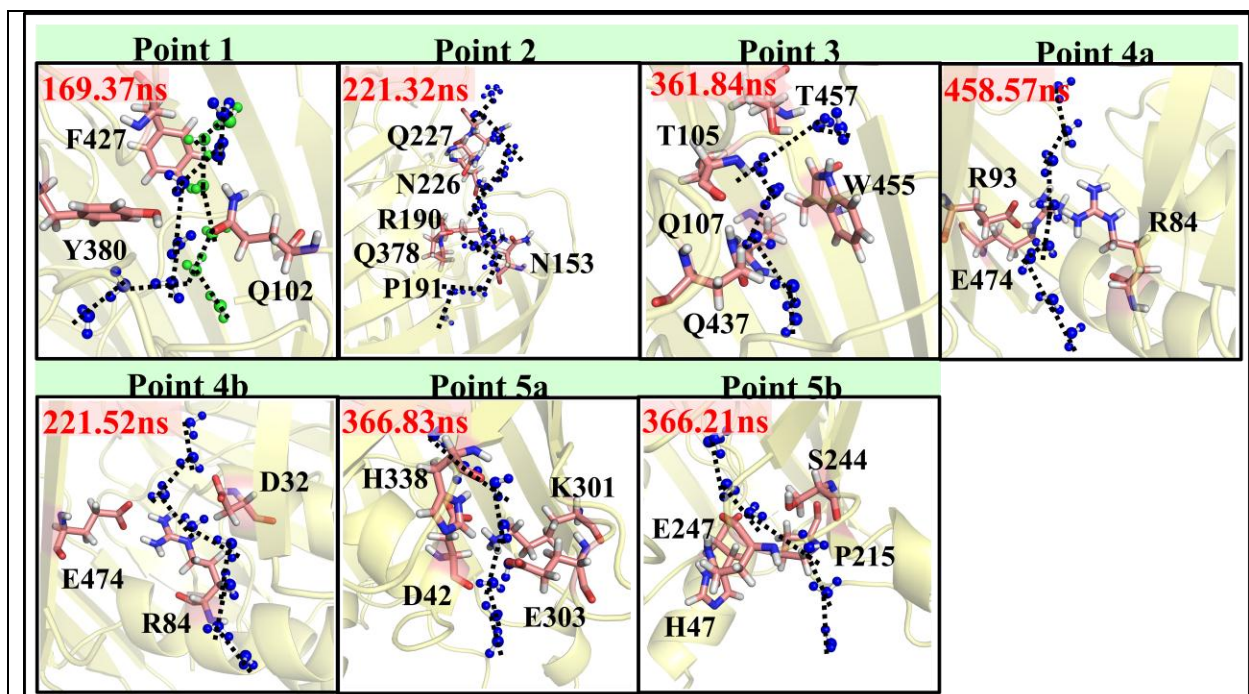
Table 3.1 Entry and exit of water molecules from each entry point in last 200 ns.		
Entry point	Entry	Exit
1	23	46
2	84	41
3	13	6
4a	4	7
4b	50	32
5a	15	31
5b	2	18
Total	191	181

So, the total number of water molecules that enter the barrel in last 200 ns is 191 while 181 water molecules exit the barrel. Thus, the net flux is  $(191-181)/200$  i.e 0.05/ns.

### 3.5.7 Transient water wires

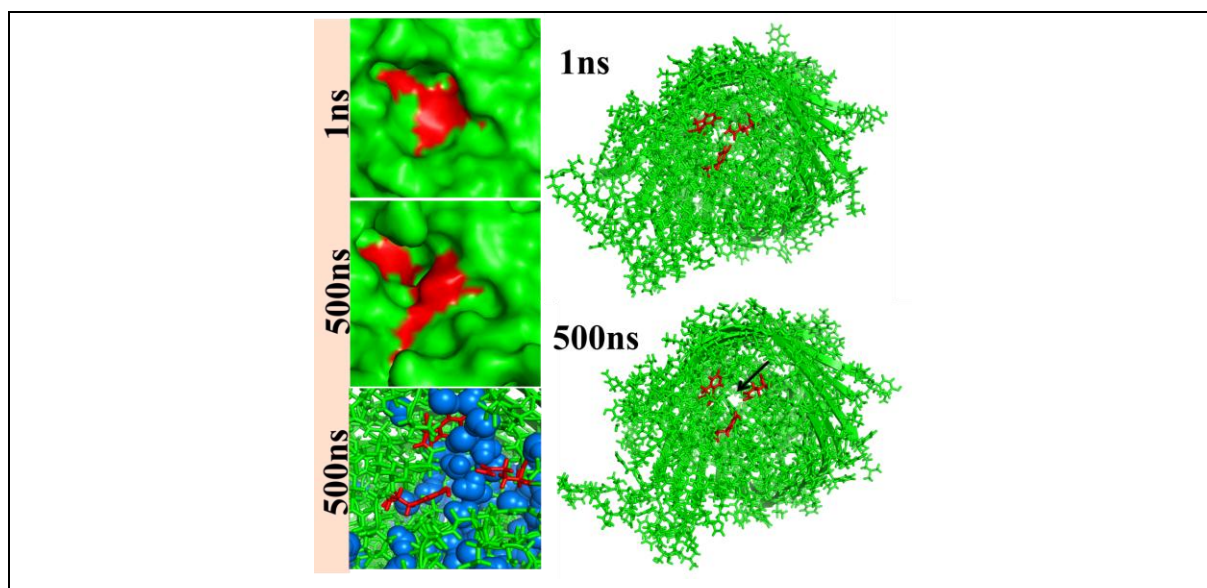
Development of transient water wires is observed at each diffusion point (Fig 3.12). As the water entry & exit happens at random time interval, they simultaneously contribute to the formation of the water wires. Breakage in these transient water wires is observed based on the entry & exit of the

water molecules. A single water wire is observed in all the diffusion points except at the extracellular point 1, where two water wires are formed one between Y380 & Q102 and other at the interspace of F427&Y380 and Q102.



**Figure 3.12. Transient water wire formation at different entry points.** Snapshots depicting transient water wires at five diffusion points with the corresponding time indicated. Water (blue colored ball and stick) oxygens within  $3.5\text{\AA}$  distances are connected by dotted lines to illustrate the formation of water wires. At point 1, water wire between Y380 & Q102 residues is shown in blue and the one at the interspace of F427&Y380 and Q102 is shown in green.





**Figure 3.13: Pore opening at extracellular YQF triad.** Snapshots of extracellular face of Wzi (green surface) at 1ns (Top-Left) and 500ns (Middle-Left & Bottom-Left) illustrating the pore opening at YQF triad (colored red). Stick representation of Figures given in Top-Left (1ns) & Middle-Left (500ns) with YQF triad colored red is shown at right hand side. Pore is indicated by black arrow.

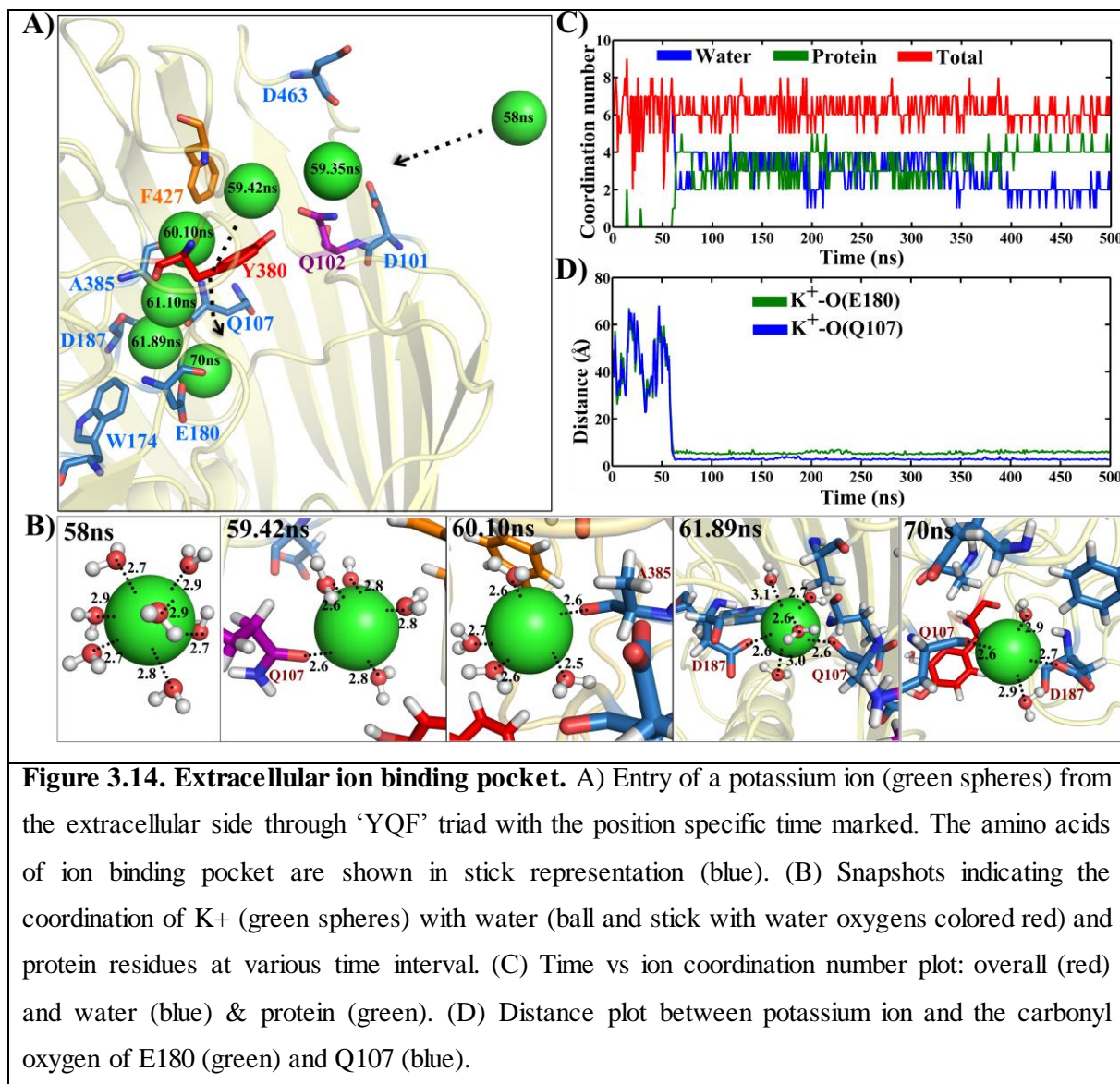
### 3.5.8 Pore opening

During the simulation, YQF triad moves in such a way that a pore is created on the extracellular side of the protein (Fig 3.13). This pore facilitates water entry/exit into the barrel as describes in section 3.5.5.1. As proposed earlier, that YQF may be the binding site of K30 antigen onto Wzi, the pore opening may facilitate this binding. This pore also coincides with ion binding pocket described in section 3.5.9.

### 3.5.9 Ion binding pocket

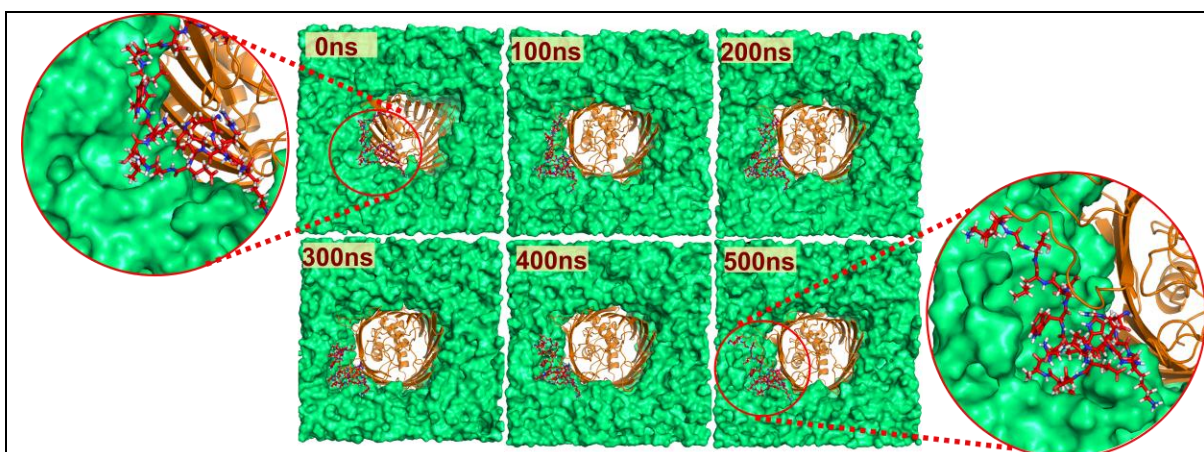
During the simulation, one potassium ion enters the barrel through YQF motif and stays strongly associated to carbonyl oxygens of E180 and Q107 throughout the simulations. Ion starts entering the barrel around 58ns (Fig 3.14 A). While entering the water shell surrounding the ion is partially stripped off and is compensated by interactions with carbonyl oxygens of the residues of ion binding pocket (Fig 3.14 B). Ion also interacts with the aromatic side chain of F427 through cation... $\pi$  interaction thus moving inside the barrel. Subsequently, it coordinating with the carbonyl oxygens of Q107, E180, D187 and A385 in an alternating fashion (Fig. 3.14 B). Coordination number of ion with water and protein was calculated, which gives a condensed coordination of  $\sim 7$  throughout the trajectory (Fig 3.14 C). A shorter distance of  $\sim 5 \text{ \AA}$  between the ion and the backbone carbonyl oxygens of E180 and Q107 beyond  $\sim 60 \text{ ns}$  indicates the strong association of the ion with the binding pocket throughout the rest of simulation (Fig 3.14 D).

The presence of cation binding site just below the YQF motif further strengthens that YQF may be the binding site for K antigen. As group I K antigens are anionic in nature so presence of a cation may facilitate this binding.



### 3.5.10 Role of extracellular loop 5 (L5)

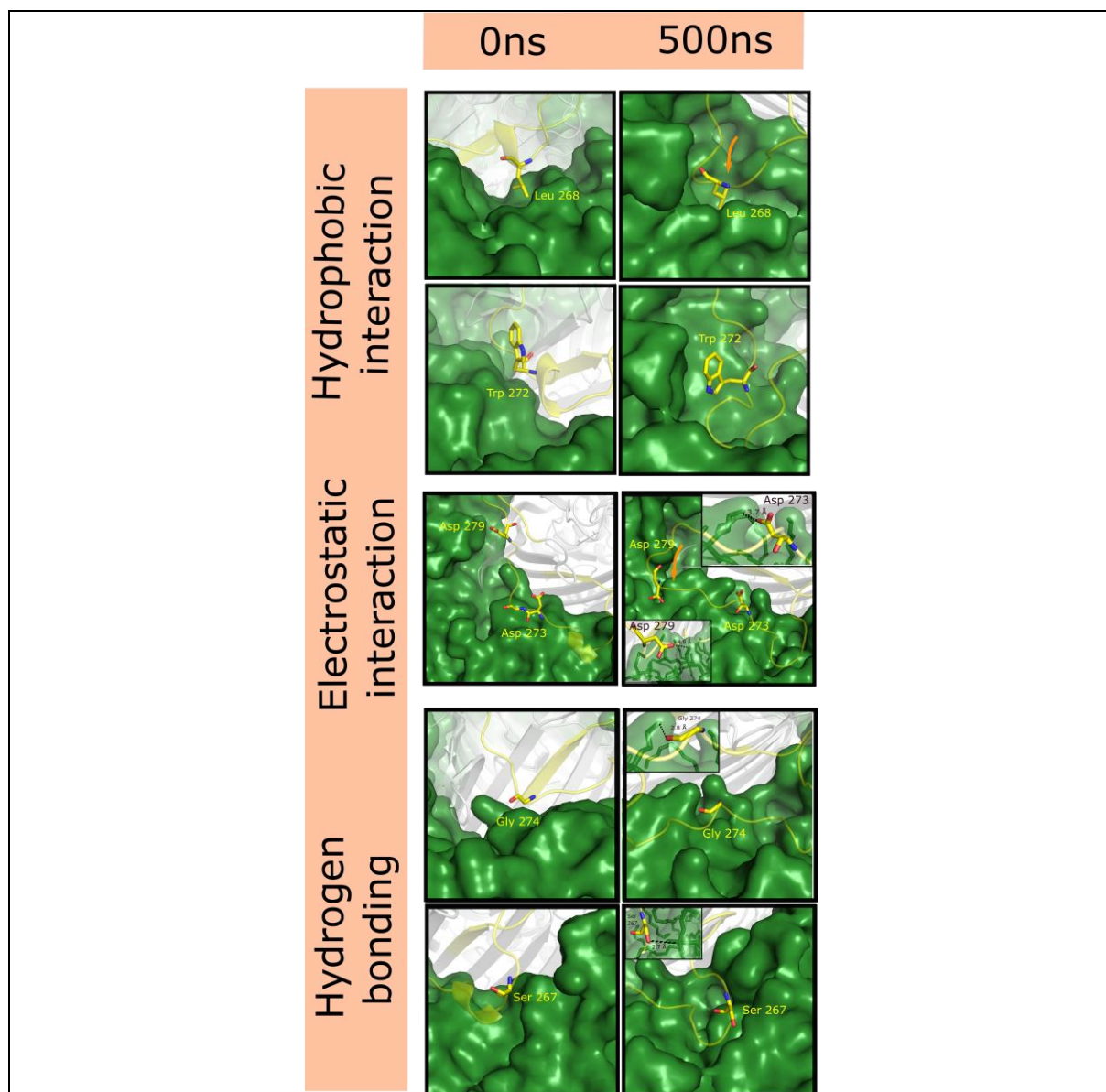
Simulation indicates that extracellular L5 plays a major role in membrane insertion. Though L5 is initially (0ns) aligned close to the barrel and is not involved in any interaction with membrane, it slowly splay out during the simulation and insert into the membrane ~50 ns (Fig 3.15). This is facilitated major conformational changes in L5 and complete burial of residues 266–276 onto the membrane surface.



**Figure 3.15. Role of extracellular Loop 5.** Snapshots depicting the splaying of L5 (red colored sticks) and its insertion onto the membrane (green surface) at different simulation time (indicated). Wzi is shown as orange colored cartoon. L5 is zoomed at 0 ns and 500 ns to illustrate the splaying of L5 onto the membrane.

Key interactions involved in the anchorage of the protein into the membrane are hydrophobic interaction, electrostatic interaction and hydrogen bonding interactions (Fig 3.16). Hydrophobic interaction is mediated by conserved W259, E266, F271 & W272 and semiconserved L268 residues. The positively charged amino head group of POPE and negatively charged carboxylic group of conserved D273 & D279 interact electrostatically. Hydrogen bonding interaction between side chain hydroxyl group of conserved S267 & phosphate oxygen of POPE and the head group of POPE & carbonyl group of G274 also participate in this anchorage.





**Figure 3.16. Loop 5 interaction with membrane.** Snapshots at 0 ns and 500 ns demonstrating interactions between L5 and membrane facilitating Wzi anchorage with a zoomed view at 500ns shown in inset: (top) hydrophobic interactions involving L268 & W272, (middle) electrostatic interactions with D273 & D279 and (bottom) hydrogen bonding interactions with G274 & S267 are shown.

### 3.6 Conclusion

Five water diffusion points are identified (two at the periplasmic and three at the extracellular sides), through which, the water enters and exits the barrel in a regulated manner. These mechanisms facilitate complete crossing of water molecules across the channel by passive diffusion. Water inside the barrel remains almost constant, but because of the nonspecific conduction pathway in lumen total permeation events remain less than entry/exit events.

Diffusion permeability of Wzi is estimated to be  $2.8 \times 10^{-15} \text{ cm}^3 \cdot \text{s}^{-1}$  (2 times lower than that of aquaporin monomers [17]).

Entry 1 ('YQF' triad motif) located at the extracellular side is identified to be one of the major diffusion points. A similar triad is found in sugar binding site of sodium-galactose cotransporter (SGLT). Thus, 'YQF' triad is proposed to be the K30CPS binding motif. Since, 'YQF' triad also participates in potassium ion conduction into the ion-binding pocket at the extracellular side. This further implicates that it may be the binding site for the negatively charged K30CPS onto Wzi.

Wzi is proposed to be an osmoregulator based on its bidirection water conduction property. Wzi may osmoregulate on basis of K antigen concentration. If K antigen is more on the bacterial surface, Wzi may transports water from periplasmic to the extracellular side to avoid the rupturing of the cell and vice-a-versa. This may further aid in keeping the K30CPS in a hydrated condition as well as prevent the accumulation of nascent K30CPS onto the bacterial surface. This study pinpoints the importance of extracellular loop 5 in the anchorage of Wzi onto the outer membrane. Thus, disrupting the interaction between K30CPS & 'YQF' triad or altering the Wzi water diffusion or disrupting the interaction between L5 & membrane may be attractive strategies to reduce the bacterial virulence.

### 3.7 References

1. Fiser, A., R.K. Do, and A. Sali, *Modeling of loops in protein structures*. Protein Sci, 2000. **9**(9): p. 1753-73.
2. Fiser, A. and A. Sali, *ModLoop: automated modeling of loops in protein structures*. Bioinformatics, 2003. **19**(18): p. 2500-1.
3. Jo, S., T. Kim, and W. Im, *Automated builder and database of protein/membrane complexes for molecular dynamics simulations*. PLoS One, 2007. **2**(9): p. e880.
4. Jo, S., et al., *CHARMM-GUI: a web-based graphical user interface for CHARMM*. J Comput Chem, 2008. **29**(11): p. 1859-65.
5. Jo, S., et al., *CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes*. Biophys J, 2009. **97**(1): p. 50-8.
6. Wu, E.L., et al., *CHARMM-GUI Membrane Builder toward realistic biological membrane simulations*. J Comput Chem, 2014. **35**(27): p. 1997-2004.
7. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water*. The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.
8. Lomize, M.A., et al., *OPM database and PPM web server: resources for positioning of proteins in membranes*. Nucleic Acids Res, 2012. **40**(Database issue): p. D370-6.

9. Chugunov, A., et al., *Lipid-II forms potential "landing terrain" for lantibiotics in simulated bacterial membrane*. Sci Rep, 2013. **3**: p. 1678.
10. Im, W., S. Seefeld, and B. Roux, *A Grand Canonical Monte Carlo-Brownian dynamics algorithm for simulating ion channels*. Biophys J, 2000. **79**(2): p. 788-801.
11. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free*. J Comput Chem, 2005. **26**(16): p. 1701-18.
12. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations*. Journal of Computational Chemistry, 1983. **4**(2): p. 187-217.
13. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program*. J Comput Chem, 2009. **30**(10): p. 1545-614.
14. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath*. The Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.
15. Evans, D.J. and B.L. Holian, *The Nose-Hoover thermostat*. The Journal of Chemical Physics, 1985. **83**(8): p. 4069-4074.
16. Parrinello, M. and A. Rahman, *Polymorphic transitions in single crystals: A new molecular dynamics method*. Journal of Applied Physics, 1981. **52**(12): p. 7182-7190.
17. Zhu, F., E. Tajkhorshid, and K. Schulten, *Theory and simulation of water permeation in aquaporin-1*. Biophys J, 2004. **86**(1 Pt 1): p. 50-7.
18. Thomas, M., D. Jayatilaka, and B. Corry, *The predominant role of coordination number in potassium channel selectivity*. Biophys J, 2007. **93**(8): p. 2635-43.



# Chapter 4

## Comparison of water diffusion mechanisms of *Wzi*, *Klebsiella pneumonia* with that of *E. coli*

### 4.1 Introduction

Gram-negative bacteria, *Klebsiella pneumonia*, is also known as a ‘superbug’ for its ability to cause range of diseases, depending upon which part of the body it affects. *K. pneumonia* causes pneumonia, blood stream infections, and meningitis. It has gained resistance to 3<sup>rd</sup> generation cephalosporins and carbapenems [1]. These resistant strains are spreading very quickly across the globe.

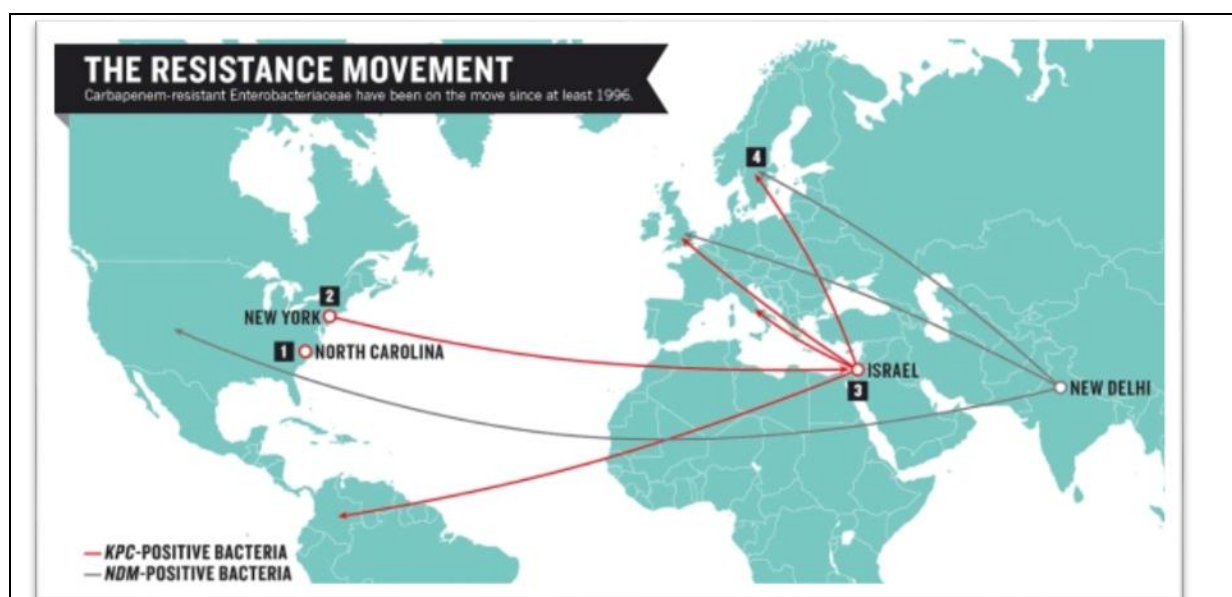


Figure 4.1. *Klebsiella* resistance movement. [Adopted from Nature, Vol 499, 394-296.]

*Klebsiella pneumoniae* carbapenemase (KPC) producing bacteria that was initially found in North Carolina (2000) soon spread to New York in 2003. By 2005, the resistant strain was found in Israel, from where, it travelled to Italy, Colombia, United Kingdom and Sweden (Fig 4.1). Similarly, New

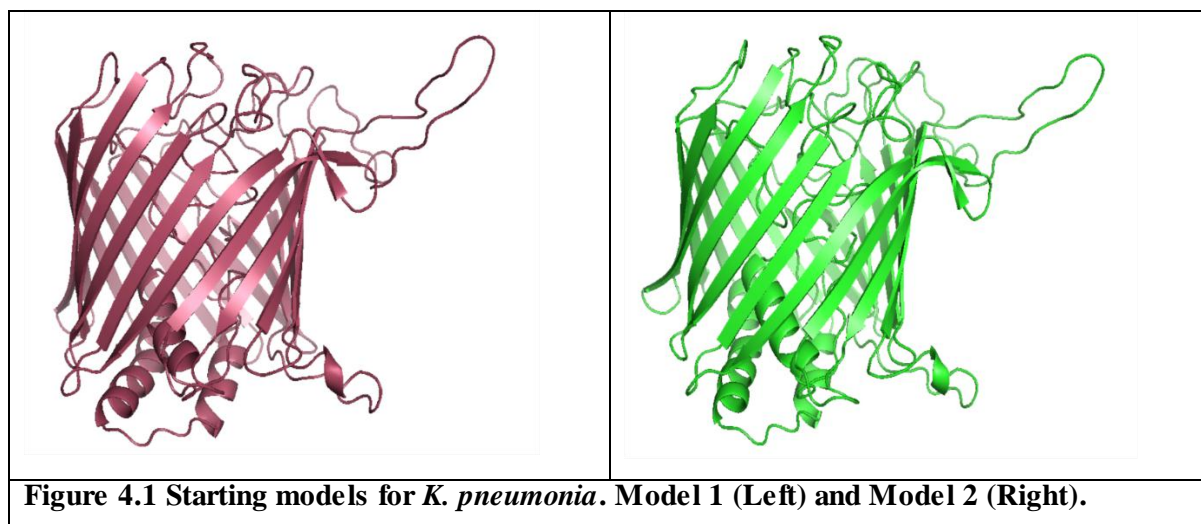
Delhi metallo- $\beta$ -lactamase (NDM) positive strain isolated in Sweden (2008) which could be traced back to India.

Wzi protein in *K. pneumonia* shares high sequence similarity with that of *E. coli*. Two such sequences sharing 98% similarity were chosen for comparison with *E. coli*. One of the models (herein model 1) has a very interesting mutation from *E. coli* sequence (Q102K). It is to be noted that Gln 102 is of the YQF motif in Wzi, *E. coli* that is one of the major diffusion points at the extracellular side and proposed site for K antigen binding onto Wzi.

In *K pneumonia*, also five entry/exit points have been observed. However, few deviations are there from that of Wzi which are explained in detailed in section 4.5.2. Ion binding pocket is conserved with same residues interacting with ion in one of the model. In another model, ion binding has not been observed, which may be evident on longer simulations. Loop 5 role in anchorage onto the membrane is observed in both the models. It is to be noted that the results shown in this chapter are from 50ns preliminary data from MD simulations using CHARMM 36[2, 3] force field.

## 4.2 System setup

*K. pneumonia* crystal structure is not available, so starting models were prepared using SWISS-MODEL [4-6] (Fig 4.1). Two sequences having 98% sequence similarity with Wzi, *E.coli* (2YNK) [7]. Thus, *E. coli* Wzi crystal structure was used as the template.



Model 1 sequence is from unknown strain. Reason behind choosing this model was Q102K mutation. It is to be noted that Gln 102 is of the YQF motif in Wzi, *E. coli* that is one of the major diffusion points at the extracellular side and proposed site for K antigen binding onto Wzi. So, to explore more about this motif and how substitution to a positively charged amino acid will affect it this sequence

was chosen. All entry point residues are conserved in model 2, which is from strain A1142 and binds to K57 CPS. Both the systems are setup using the same protocol as explained in section 3.2.

### 4.3 MD simulation

After the system is assembled, it was first subjected to 450 ps equilibration run which is divided in six steps. In the first two steps systems were subjected to 25 ps of langevin dynamics wherein the temperature is maintained at 303.15 K. This was followed by four steps of CPT dynamics (100ps each) where temperature (303.15 K) is maintained using Nose-Hoover thermostat [8]. With CPT ensemble, production run was extended upto 50 ns using CHARMM [2, 3] 36 all atoms force field.

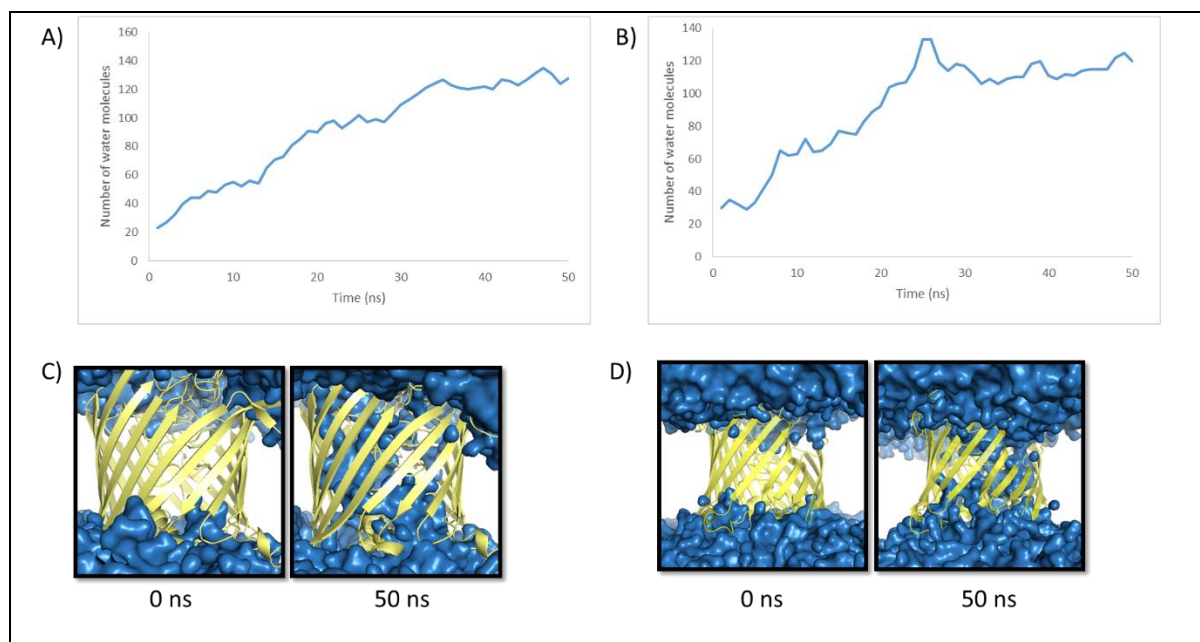
### 4.4 Analysis of trajectory

#### 4.4.1 Quantification of water inside the barrel

Wzi protein is aligned along Z-axis with the center of the barrel around 0 Å. To calculate number of water molecules inside the barrel per ns, same protocol is used as explained in section 3.4.1.

## 4.5 Results and Discussion

### 4.5.1 Wzi as a porin



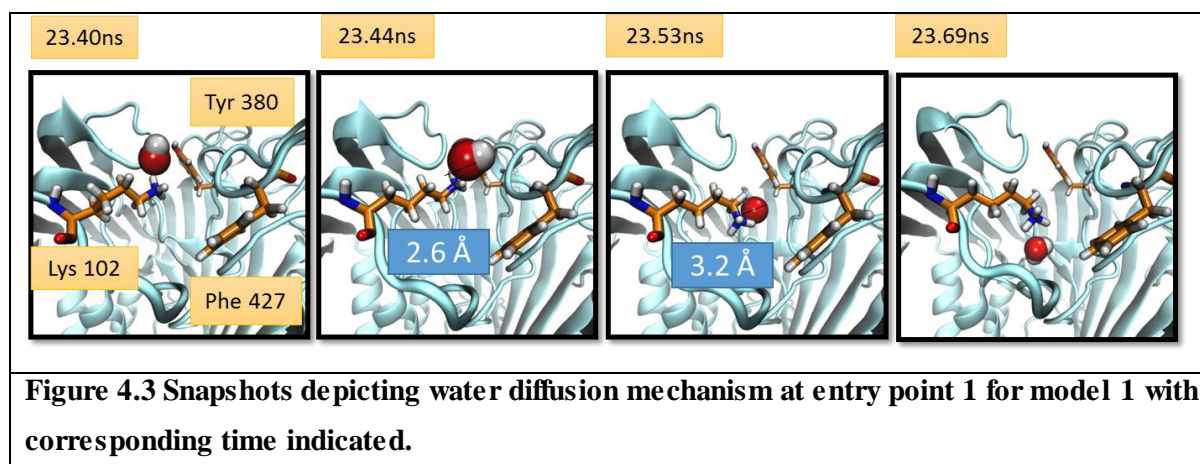
**Figure 4.2 Water moving inside the barrel.** Time vs water inside the barrel plot for model 1 (A) and model 2 (B). Snapshots for water moving inside the barrel model 1 (C) and model 2 (D).

Water moves inside the barrel from both extracellular and periplasmic sides. Water molecules inside the barrel for model 1 does not reach equilibrium during the 50 ns simulation (Fig 4.2 A). Model 2 attains equilibrium around 30 ns with an average of 120 water molecules inside the barrel in last 20ns (Fig 4.2 B). This is comparable to number of water molecules in Wzi, *E. coli* at the same time scale.

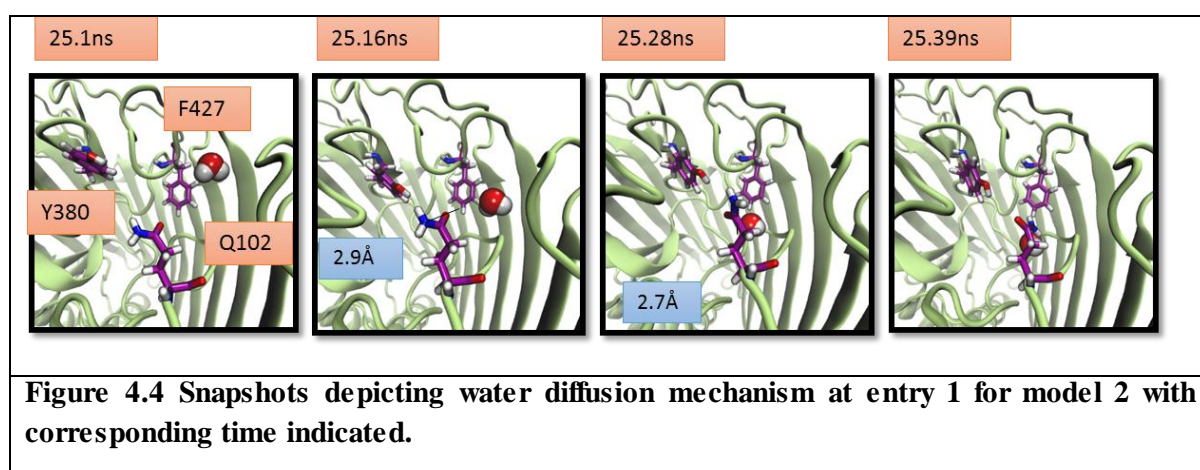
## 4.5.2 Diffusion mechanisms

### 4.5.2.1 Entry point 1

*Model 1:* Entry point 1 comprises of extracellular residues Lys102, Tyr380 and Phe427. The main mechanism by which water molecule moves into the barrel through entry 1 is by hydrogen bonding with Lys102 and moving inside from the triad (Fig 4.3). When the water molecule comes in the vicinity of YKF, it hydrogen bonds with terminal amino group of Lys 102. The amino head group moves with water molecule pushing it inside the barrel.



*Model 2:* Entry point 1 residues for model 2 are same as that of *E. coli*, i.e. Try 380, Gln 102 and Phe 427. Water molecule enters through the interface of Phe 427 and Gln 102. Oxygen atoms of water hydrogen bond with the side chain amino acids of Gln 102 while diffusing inside the barrel (Fig 4.4).

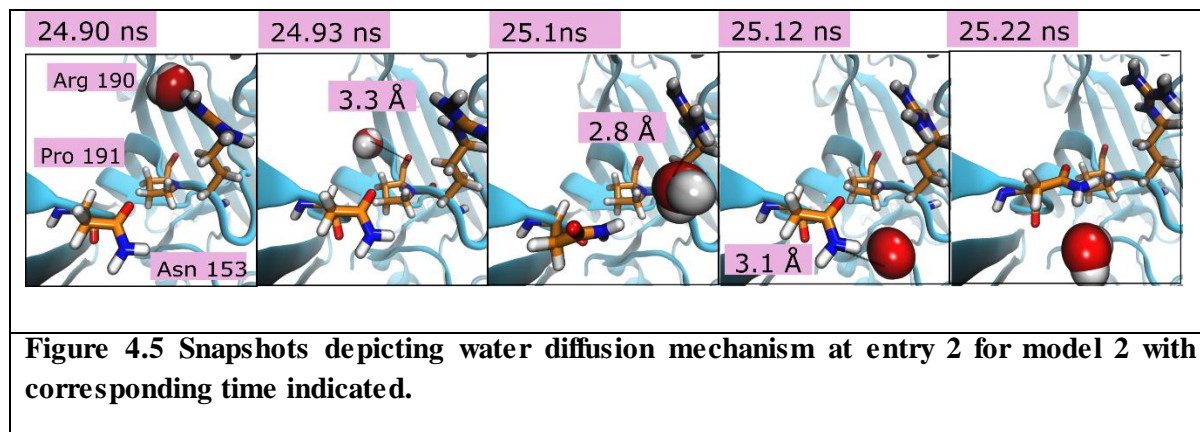


### 4.5.2.2 Entry 2: notch region

In both the models for entry 2, diffusion mechanism is similar to that of *E. coli*. However, out of six only three extracellular residues contribute to the water entry into the barrel in *K. pneumoniae*. Water

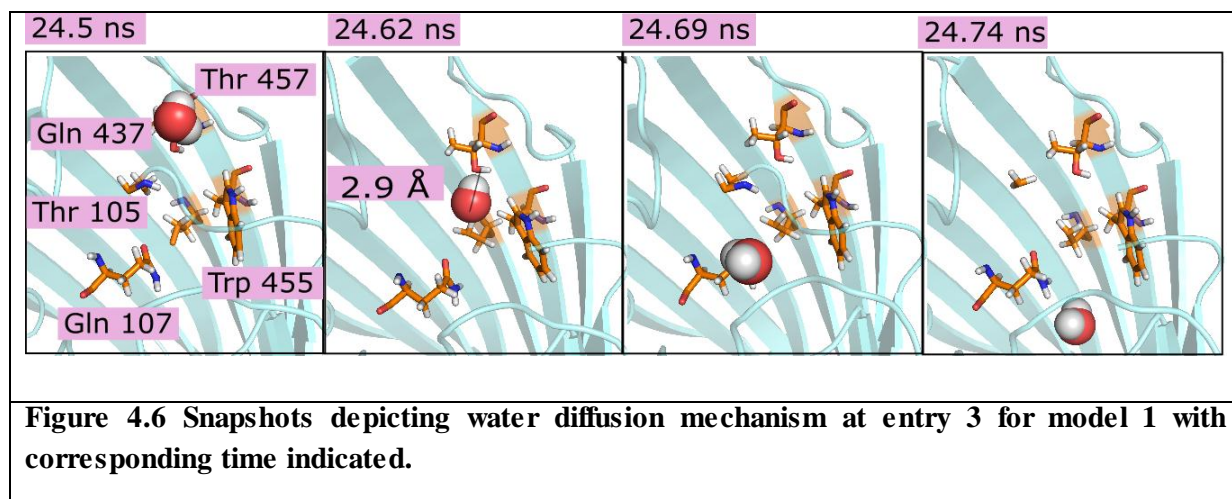


molecule moves into the barrel by making alternative hydrogen bonds with Asn 153, Pro 191 and Arg 190. Water molecule first hydrogen bonds with main chain oxygen of P191, subsequently moves and interacts with Arg 190. Finally, it hydrogen bonds with amino group of Asn153. The water molecule thus moves inside by relinquishing itself from N153.

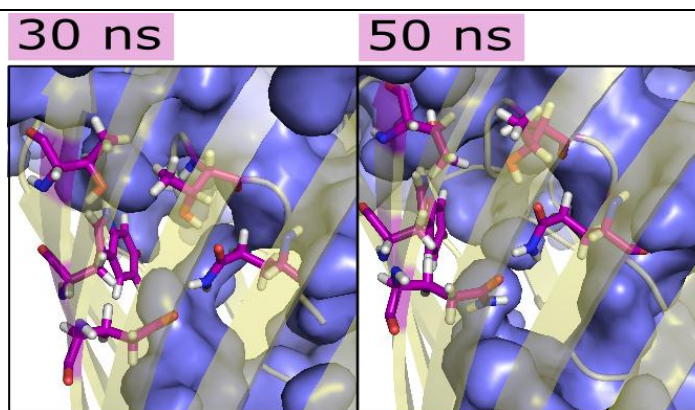


#### 4.5.2.3 Entry point 3

This entry point facilitates water entry only in model 1 but not in model 2. Extracellular residues Thr105, Gln107, Gln437, Trp455 and Thr457 are involved in water diffusion into the barrel. Water molecule enters by hydrogen bonding with hydroxyl group of threonine 457 and slides into the channel through W 455 (Fig 4.6). Threonine flipping is not observed here as in case of Wzi.



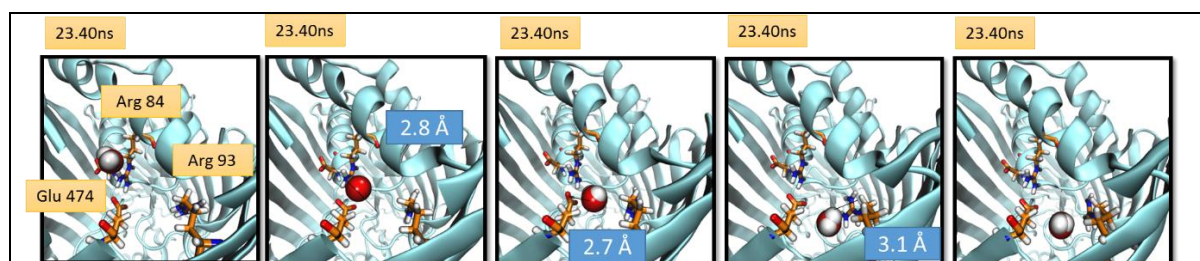
In model 2, water is not moving in or out of the barrel from entry 3. It was checked at different time intervals but no entry was observed (Fig 4.7). However, water molecules are trying to push inside but the reason why they are unable to enter is unexplored. Since, this is preliminary data, on longer simulations entry point 3 may facilitate diffusion of water molecules.



**Figure 4.7** Snapshots illustrating that water molecules are not moving in the barrel from entry 3 in model 2. Entry 2 residues are represented in purple sticks and water in blue surface.

#### 4.5.2.4 Entry point 4

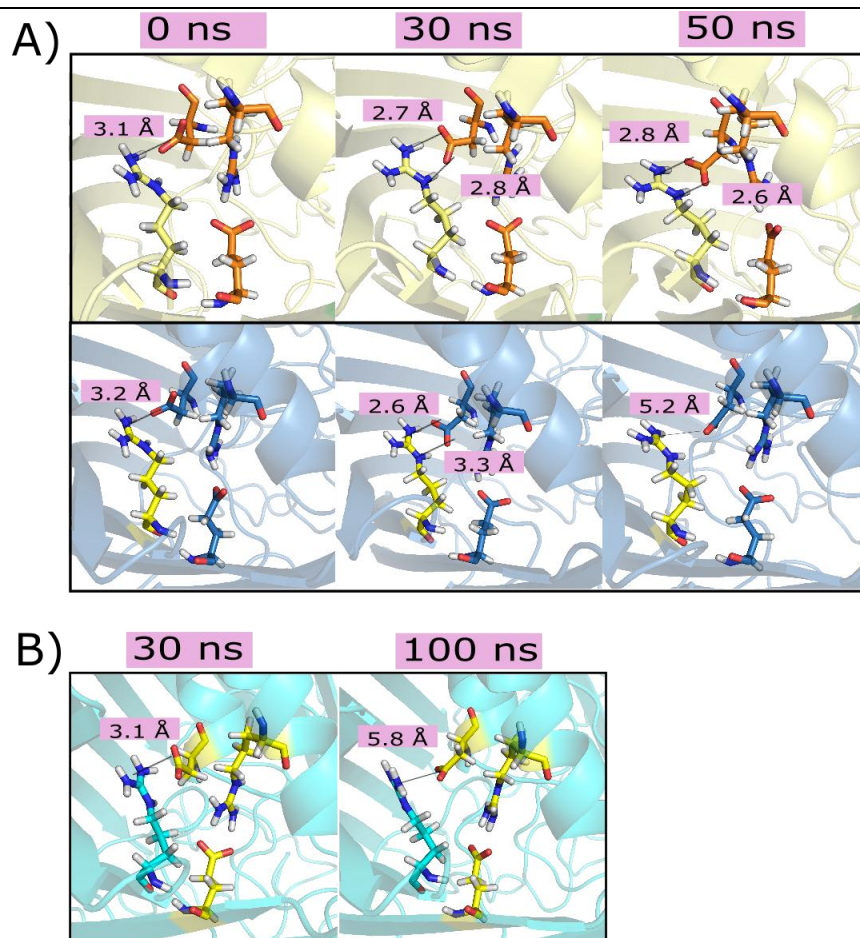
Entry 4 at periplasmic side includes residues Asp32, Arg84, Arg93 and Glu474, makes two different entries viz. 4a (R84, R93 & E474) and 4b (D32, R84 & E474). Water entry into the barrel is observed from 4a in both the models. While entering into the barrel, the water molecule initially hydrogen bonds with amino group of Arg 84. Subsequently, water molecule relinquish itself and interacts with side chain of Glu 474. The water molecule finally moves into the barrel by hydrogen bonding with amino group of Arg 93 (Fig 4.8).



**Figure 4.8** Snapshots depicting water diffusion mechanism at entry 4a for model 1 with corresponding time indicated.

Entry 4b is the major periplasmic entry/exit point in *E. coli*. However, in *K. pneumonia* water entry through 4b is a rare event. The reason behind this is strong hydrogen bonding between amino group of Arg 449 and side chain of Asp 32 (Fig. 4.9). However, in model two this hydrogen bonding breaks at the end of simulation, entry of water molecules was not observed (Fig 4.9 A). This hydrogen bonding is prevalent in *E. coli* as well. It only breaks around 100 ns facilitating entry/ exit of water molecules (Fig 4.9 B).

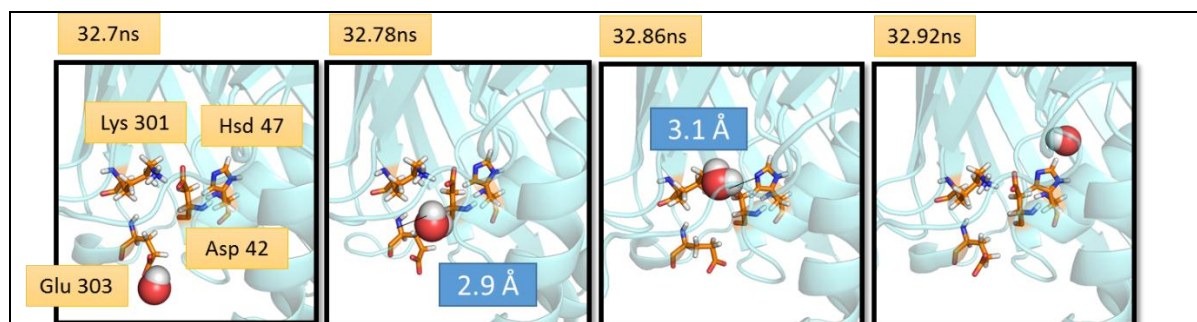




**Figure 4.9. Hydrogen bonding between side chain of Arg 449 and Asp 42.** (A) Snapshots of model 1 (top) and model 2 (bottom) showing strong association between Arg 449 and Asp 42. (B) Snapshots of *E. coli* showing hydrogen bonding till 100 ns. Corresponding time have been indicated.

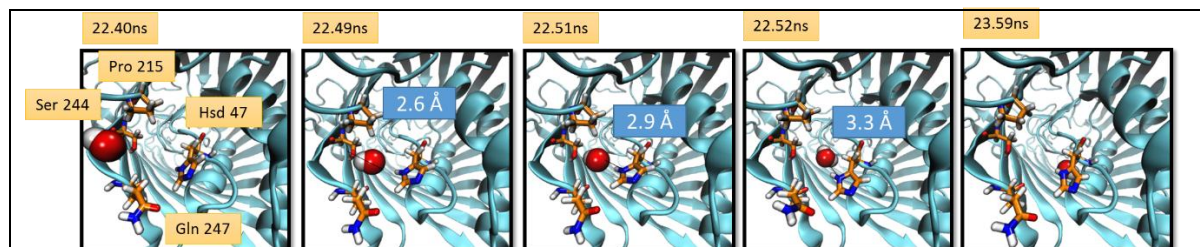
#### 4.5.2.5 Entry point 5

This periplasmic entry point facilitates entry through 5a involving residues Asp42, Lys301, Glu303 and His338, and entry 5b comprising of residues His47, Pro215, Ser244 and Gln247.



**Figure 4.10 Snapshots depicting water diffusion mechanism at entry 5a for model 1 with corresponding time indicated.**

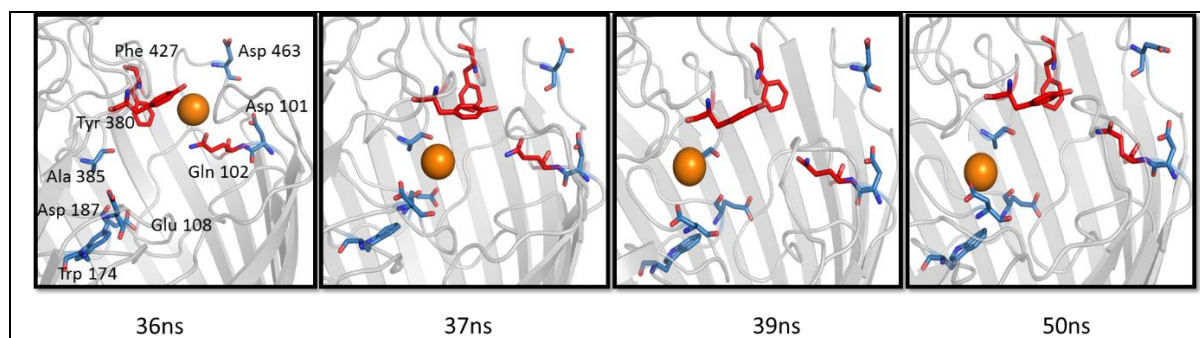
Entry mechanisms for both entry 5a & 5b are conserved. Water entry through point 5a is initiated by the interaction of water to carboxylic group of E303 (Fig 4.10). Upon discharging from this hydrogen bond, water molecule moves into the H338 and thus moves inside the barrel. Point 5b hydrogen bond formation between the water molecule and the side chain of Gln 247, Ser 244 and Hsd 47 (Fig 4.11).



**Figure 4.11 Snapshots depicting water diffusion mechanism at entry 5b for model 1 with corresponding time indicated.**

#### 4.5.3 Ion binding on the extracellular side

Wzi, *K. pneumonia* model 9 also harbors one potassium ion in the ion binding site, same as in *E. coli*. Ion starts to enter the barrel around 36 ns and moves inside through YQF motif (Fig 4.12). It stays in the ion binding pocket for the rest of the simulation. Ion coordinates with the carbonyl oxygens of Q108, E180, D187 and A385 in an alternating fashion. Ion binding pocket just below YQF motif strengthens the YQF as proposed binding site.

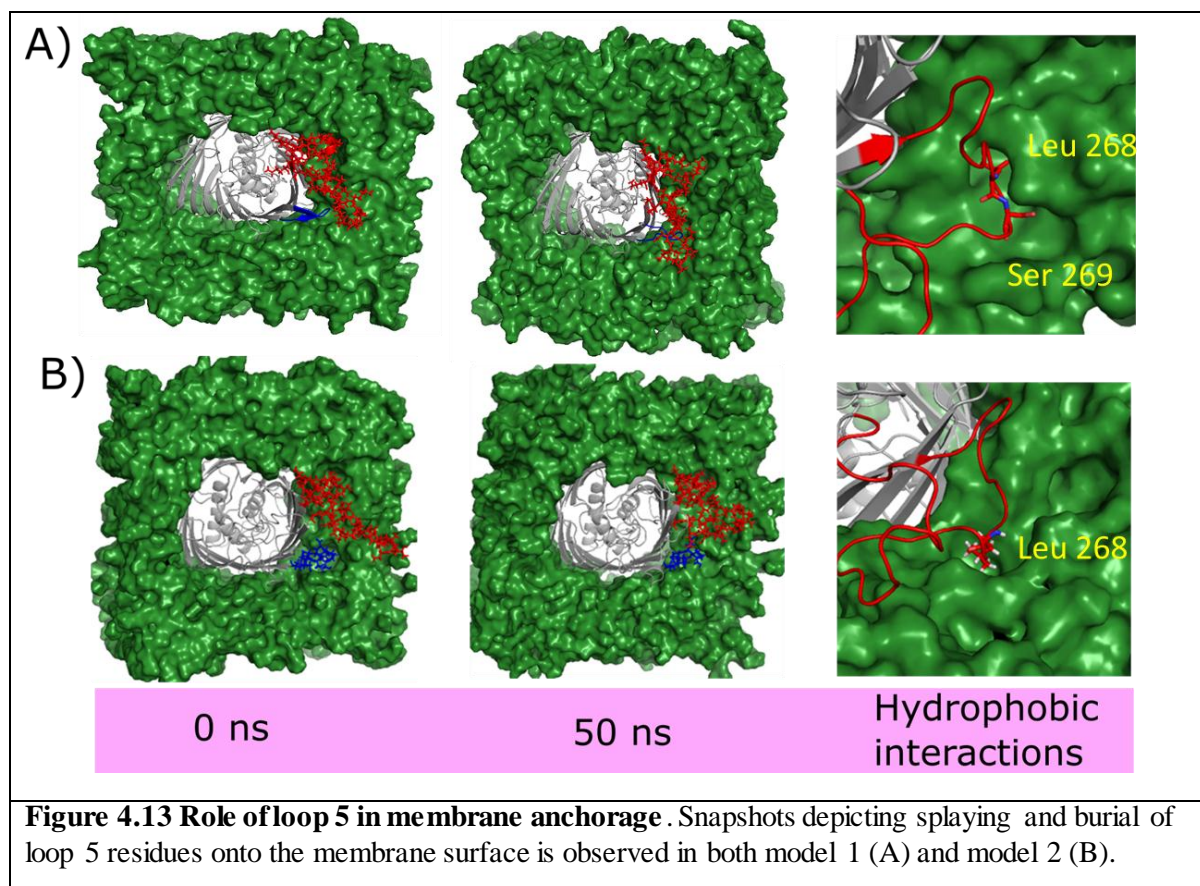


**Fig 4.12 Snapshots depicting ion binding pocket in model 2. Ion is shown in orange sphere, YQF motif is shown in red sticks and other residues involved in coordination are shown in blue colored sticks.**

In model 1, no ion binding site was observed. This may be due to Q102K mutation. The presence of positively charged amino acid may interfere with ion entry. Also, lysine 102 may help in binding to positively charged K antigen.

#### 4.5.4 Role of extracellular loop 5

Extracellular loop 5 splays out and gets embedded in the membrane in both the models within 25 ns (Fig 4.13). However, complete burial is not observed as in case of *E.coli*. This may be due to the small simulation time, on higher time scale loop 5 may show complete burial. Key interaction in both *Klebsiella* models is hydrophobic interactions. Conserved residues Leu 268 and Ser 269 participates in hydrophobic interactions (Fig 4.13). Loop 5 interaction with membrane strengthens the fact that loop 5 plays key role in anchorage of Wzi onto the membrane.



#### 4.6 Conclusion

Diffusion mechanisms of Wzi in *Klebsiella pneumonia* varies from that of *E. coli*, however that can be attributed to the small simulation time and difference in simulation engines. But, major inference remains the same as in case of *E. coli*. Both the systems of *Klebsiella pneumonia* does conduct water. Water enters into the barrel from both the sides of the barrel. Five water diffusion points have been observed (two at the periplasmic and three at the extracellular sides), similar to *E. coli*.

Ion binding pocket on the extracellular side (below YQF) was only found in model 2. In *E. coli* and *K. pneumonia* model 1, ion binding may be essential for K antigen binding onto Wzi. However, *K. pneumonia* model 1 may not need potassium ion because of presence of positively charged amino acid present in the proposed binding site. This further implicates that it may be the binding site for the negatively charged K antigen onto Wzi.



Loop 5 interaction with membrane was also found in *K. pneumonia* models. However, complete burial was not observed due to small time scale. Extending the simulations to  $\mu$ s scale is essential for complete analysis of both the models.

#### 4.7 References

1. WHO, *Antimicrobial resistance: Global report on surveillance*. 2014.
2. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations*. Journal of Computational Chemistry, 1983. **4**(2): p. 187-217.
3. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program*. J Comput Chem, 2009. **30**(10): p. 1545-614.
4. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. Bioinformatics, 2006. **22**(2): p. 195-201.
5. Guex, N., M.C. Peitsch, and T. Schwede, *Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective*. Electrophoresis, 2009. **30 Suppl 1**: p. S162-73.
6. Biasini, M., et al., *SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information*. Nucleic Acids Res, 2014. **42**(Web Server issue): p. W252-8.
7. Bushell, S.R., et al., *Wzi is an outer membrane lectin that underpins group 1 capsule assembly in Escherichia coli*. Structure, 2013. **21**(5): p. 844-53.
8. Evans, D.J. and B.L. Holian, *The Nose-Hoover thermostat*. The Journal of Chemical Physics, 1985. **83**(8): p. 4069-4074.

## **Publications**

1. Sachdeva, S. *et al.* Key diffusion mechanisms involved in regulating bidirectional water permeation across *E. coli* outer membrane lectin. *Sci. Rep.* **6**, 28157; doi: 10.1038/srep28157 (2016).