

Syracuse University

**SURFACE**

---

Dissertations - ALL

SURFACE

---

May 2019

# COMPUTATIONAL MODELING OF BACTERIAL OUTER MEMBRANES AND DEVELOPMENT OF HIGH-THROUGHPUT SCREENING PLATFORM FOR ANTIBIOTICS

Huilin Ma  
*Syracuse University*

Follow this and additional works at: <https://surface.syr.edu/etd>



Part of the [Engineering Commons](#)

---

## Recommended Citation

Ma, Huilin, "COMPUTATIONAL MODELING OF BACTERIAL OUTER MEMBRANES AND DEVELOPMENT OF HIGH-THROUGHPUT SCREENING PLATFORM FOR ANTIBIOTICS" (2019). *Dissertations - ALL*. 1005.

<https://surface.syr.edu/etd/1005>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact [surface@syr.edu](mailto:surface@syr.edu).

## Abstract

Antibiotic resistance is a major health challenge because it limits the treatment options for common infectious diseases and will cause 10 million deaths each year after 2050. There is an urgent need to reduce the misuse of antibiotics and seek new classes of antibiotics that induce less or no resistance. Despite the push for new therapeutics, there has been a precipitous decline in the number of newly approved antibacterial drugs due to a limited understanding of how bacteria adapt to the chemical stress stimuli. The development of antimicrobial resistance is especially true for Gram-negative bacteria that develop resistance to antibiotics readily due to their unique highly charged outer membrane. Structurally, the Gram-negative bacteria is highly asymmetric bilayer that comprises of an inner leaflet of phospholipids and an outer leaflet of lipopolysaccharides. Embedded in the bilayer are outer membrane proteins (OMPs) that form pores to allow passage of nutrients and other small molecules through the cell wall. In addition to the outer membrane, the Gram-negative bacteria have a thin peptidoglycan layer and an inner phospholipid membrane that surrounds the cytosol. All potential small molecules have to navigate through all three layers of the Gram-negative bacterial cell wall before targeting the cellular functions. There is, however, limited understanding of the chemical specificity, structure, and functional aspects of each layer in the cell wall. To enhance our understanding of the bacterial cell wall, we first developed molecular models of ten commensal or human pathogenic bacterial species: *Pseudomonas aeruginosa*, *Escherichia coli*, *Helicobacter pylori*, *Porphyromonas gingivalis*, *Bacteroides fragilis*, *Bordetella pertussis*, *Chlamydia trachomatis*, *Campylobacter jejuni*, *Neisseria meningitidis*, and *Salmonella minnesota*. Second, we studied the self-assembly of OMPs that in some cases form trimers in the outer membranes to perform their function. In the third step, we combined the outer membrane models and the OMPs to build a computational screening platform

to quantify the transport properties of molecules across a bacterial outer membrane. The goal of the computational platform is to provide high-throughput screening of vast libraries of small molecules that have the potential of being active antibacterial agents against Gram-negative bacteria. A computational platform has merit to producing reliable first-round screening of molecules at a fraction of the cost in the otherwise expensive drug-discovery pipeline.

**COMPUTATIONAL MODELING OF BACTERIAL OUTER MEMBRANES AND  
DEVELOPMENT OF HIGH-THROUGHPUT SCREENING PLATFORM FOR  
ANTIBIOTICS**

By

Huilin Ma

B.S. Bioengineering, China Pharmaceutical University, 2013

M.S. Biomedical Engineering, Syracuse University, 2015

Dissertation

Submitted in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy in Bioengineering

Syracuse University

May 2019



Copyright© Huilin Ma 2019

All Rights Reserved

## Acknowledgement

First, I will like to thank my advisor, Professor Shikha Nangia, for support, enthusiasm, suggestions, and patient guidance. She was the most important person during my Ph.D. —this thesis was not possible without her. Professor Nangia is very passionate about her work and has lot of great ideas, which were very helpful and instructive for my work. Her passion for research inspires everyone in our lab. I was formerly an experimentalist and knew nothing about molecular dynamics, Professor Nangia gave me the opportunity to grow this field and appreciate the combined power of experiments and simulations, which is a powerful approach to conduct meaningful research.

I will like to thank my lovely family for their support and understanding in these six years. My father, Jiankang Ma, my mother, Shufang Wu and my dear wife, Huan Zhang for their unconditional support, encouragement and endless love.

I will also like to thank all the committee members—Professors Shalabh Maroo, Dacheng Ren, Julie Hasenwinkel, Ian Hosein and Pranav Soman—for their time and support.

Many thanks to our collaborator Dr. Mitragotri's and his group at Harvard University. I appreciate their help in providing experimental data for corroborating my computational work. I feel fortunate to have worked with them and have them as my coauthors.

In the five years of a graduate student in our lab, I have several amazing colleagues. I would like thank, Wenjuan Jiang, Flaviyan Jerome Irudayanathan, Xiaoyi Wang, Nan Wang, Nandhini Rajagopal, Amogh Srihari, Katie Piston, Yinghui Dai. A special thanks to the 2017 cohort of undergraduate students, Dan Cummins, Natalie Edelstein, Jerry Gomez, Aliza Khan, Masud Llewellyn, Tara Picudella, Sarah Willsey, who helped me in my research projects, provides strong support, and showed their enormous talents to finish a paper within 10 weeks! I am also thankful

to another cohort of undergraduate students, Meishan Wu, Tory Welsch, Colleen Cassidy, and Soor Vora for working with me in the 2017 summer research program.

Finally, I would like to thank the Biomedical and Chemical Engineering department and Syracuse University. I am proud to be part of them. Many thanks to the financial support provided by the Department of Biomedical and Chemical Engineering at Syracuse University. My research was also conducted with funds from the National Science Foundation. I also want to thank Syracuse University Research Computing and XSEDE supercomputing facility for providing computational resources for my work.

## Table of contents

Acknowledgement .....	v
List of Figures .....	xi
List of Tables .....	xvii
Chapter 1 Background .....	1
1.1 Antibiotic Resistance .....	2
1.2 Mechanisms of Antibiotic Resistance .....	3
1.2.1. Modification of Antibiotic molecule structures .....	4
1.2.2. Modification of Antibiotics' Target Sites structures .....	5
1.2.3. Decreased Antibiotic Penetration .....	5
1.2.4. Efflux pumps .....	6
1.3 Current Study of Antibiotic Resistance .....	7
1.4 Overcome Antibiotic Resistance using Computational approaches .....	8
Reference .....	11
 Chapter 2 Basic Principles of Molecular Dynamics .....	 23
2.1 Equation of Motion .....	24
2.2 Force Fields .....	27
2.2.1. All-atom force field .....	27
2.2.2. Coarse grain force field .....	28
2.3 Ensembles .....	29
2.3.1 Berendsen thermostat .....	31
2.3.2 Berendsen barostat .....	31
2.4 Energy minimization .....	32
2.5 Software to perform MD simulations: GROMACS .....	33
Reference .....	34
 Chapter 3 Modeling diversity in structures of bacterial outer membrane lipids .....	 40
3.1 Abstract .....	41
3.2 Introduction .....	42

3.3 Background .....	46
3.4 Methods.....	54
3.4.1 Parameterization.....	54
3.4.2 Simulation and analysis details .....	57
3.5 Results.....	60
3.5.1 Bonded parameters.....	60
3.5.2 Area per lipid (AL) and phase-transition temperature ( $T_m$ ) .....	65
3.5.3 Membrane thickness .....	68
3.5.4 Density profile .....	69
3.5.5 Radial distribution function .....	72
3.5.6 Diffusion coefficient .....	72
3.6 Discussion.....	74
3.6.1 Effect of number of acyl chains .....	74
3.6.2 Effect of acyl chain length .....	75
3.6.3 Effect of phosphorylation and counterions .....	76
3.7 Conclusions.....	79
Reference .....	80

Chapter 4 Dynamics of OmpF trimer formation in the bacterial outer membrane of <i>Escherichia coli</i> .....	84
4.1 Abstract.....	85
4.2 Introduction.....	86
4.3 Methods.....	90
4.3.1 Parameterization.....	90
4.3.2 Simulation details.....	95
4.4 Results and Discussion .....	99
4.4.1 Force field development and validation of membrane properties.....	99
4.4.1.1 Area per lipid .....	100
4.4.1.2 Phase transition temperature.....	102
4.4.1.3 Other structural properties .....	103
4.4.2 OmpF self-assembly simulations .....	105
4.4.2.1 Dimer and trimer formation.....	105

4.4.2.2 Comparison of self-assembled OmpF trimer and native structure .....	111
4.4.2.3 Mechanism and thermodynamics of OmpF oligomerization.....	112
4.4.2.4 Interaction of OmpF timer with membrane lipids .....	114
4.5 Conclusions.....	117
Reference .....	119
 Chapter 5 The mechanism of antimicrobial activity of choline-based ionic liquids (CAGE) .....	126
5.1 Abstract .....	127
5.2 Introduction.....	128
5.3 Methods.....	130
5.4 Results .....	131
5.5 Discussion .....	135
5.5 Conclusions.....	137
Reference .....	138
 Chapter 6 A computational platform for accelerating antibiotics discovery .....	143
6.1 Abstract .....	144
6.2 Introduction.....	145
6.3 Methods.....	150
6.3.1 Simulation details.....	150
6.3.2 Method development.....	151
6.3.3 CLASP performance .....	153
6.3.4 Carbapenem coarse graining .....	153
6.3.5 Analysis.....	155
6.4 Results .....	155
Case study 1 Doripenem and Ertapenem .....	156
Case study 2 Panipenem, Biapenem and Meropenem .....	163
Case study 3 Imipenem .....	169
6.5 Discussion .....	172
6.6 Conclusions.....	174

Reference.....	175
Chapter 7 Conclusions and Future work.....	181
7.1 Conclusions.....	182
7.2 Future work.....	183
Appendix: Tables of parameters for coarse grained <i>E.coli</i> lipid A and core oligosaccharides .....	184
VITA.....	189

## LIST OF FIGURES

Figure 3- 1 Lipid A template structure .....	43
Figure 3- 2 Atomistic structure of <i>Helicobacter pylori</i> Lipid A.....	47
Figure 3- 3 Atomistic structure of <i>Porphyromonas gingivalis</i> Lipid A.....	48
Figure 3- 4 Atomistic structure of <i>Bacteroides fragilis</i> Lipid A. ....	49
Figure 3- 5 Atomistic structure of <i>Bordetella pertussis</i> Lipid A.....	50
Figure 3- 6 Atomistic structure of <i>Chlamydia trachomatis</i> Lipid A .....	51
Figure 3- 7 Atomistic structure of <i>Campylobacter jejuni</i> Lipid A. ....	52
Figure 3- 8 Atomistic structure of <i>Neisseria meningitides</i> Lipid A.....	53
Figure 3- 9 Atomistic structure of <i>Salmonella Minnesota</i> Lipid A.....	54
Figure 3- 10 Coarse-grained mapping scheme for lipid A tails of (A) <i>N. meningitidis</i> , (B) <i>H. pylori</i> , (C) <i>P. gingivalis</i> , (D) <i>B. fragilis</i> , (E) <i>B. pertussis</i> , (F) <i>C. trachomatis</i> , (G) <i>C. jejuni</i> , and (H) <i>S.</i> <i>minnesota</i> . ....	56
Figure 3- 11 Average disaccharide head group bond distance frequency distribution for (A) <i>H.</i> <i>pylori</i> (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N.</i> <i>meningitides</i> , and (H) <i>S. Minnesota</i> .....	61
Figure 3- 12 Average disaccharide head group angle frequency distribution for (A) <i>H. pylori</i> (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N.</i> <i>meningitides</i> , and (H) <i>S. Minnesota</i> .....	62
Figure 3- 13 Average acyl chain bond distance frequency distributions for (A) <i>H. pylori</i> , (B) <i>P.</i> <i>gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N. meningitidis</i> , and (H) <i>S. Minnesota</i> .....	63



Figure 3- 14 Average acyl angle frequency distribution for (A) <i>H. pylori</i> (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N. meningitides</i> , and (H) <i>S. Minnesota</i> .....	64
Figure 3- 15 Dihedral angle fluctuations (degrees) as a function of simulation time for (A) <i>H. pylori</i> (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N. meningitides</i> , and (H) <i>S. Minnesota</i> .....	64
Figure 3- 16 Area per lipid ( $A_L$ ) of Lipid A as a function of temperature for (A) Set II and (B) Set VII bacterial outer membrane. Color scheme: <i>H. pylori</i> (brown), <i>P. gingivalis</i> (red), <i>B. fragilis</i> (purple), <i>B. pertussis</i> (gray), <i>C. trachomatis</i> (black), <i>C. jejuni</i> (orange), <i>N. meningitides</i> (yellow), and <i>S. Minnesota</i> (green). .....	65
Figure 3- 17 Phase transition temperature ( $T_m$ ) for Set II membranes determined by the change in the $A_L$ ( $\text{nm}^2$ ) versus $T$ (K).....	66
Figure 3- 18 Snapshots of thermal phase transition of (A) <i>H. pylori</i> , (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N. meningitides</i> , and (H) <i>S. Minnesota</i> .....	67
Figure 3- 19 Membrane thickness (nm) as a function of $T$ (K) for (A) <i>H. pylori</i> , (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N. meningitides</i> , and (H) <i>S. Minnesota</i> in Set II.....	68
Figure 3- 20 Density profile of key components of (A) <i>H. pylori</i> , (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> (E) <i>C. trachomatis</i> (F) <i>C. jejuni</i> , (G) <i>N. meningitides</i> , and (H) <i>S. minnesota</i> membranes (Set VI).. .....	70

Figure 3- 21 RDFs for set II ( $\text{Na}^+$ , dashed) and set VI ( $\text{Ca}^{2+}$ , solid) for phosphate (black) and carboxylate (red) for (A) <i>H. pylori</i> , (B) <i>P. gingivalis</i> , (C) <i>B. fragilis</i> , (D) <i>B. pertussis</i> , (E) <i>C. trachomatis</i> , (F) <i>C. jejuni</i> , (G) <i>N. meningitidis</i> , and (H) <i>S. minnesota</i> .	72
Figure 3- 22 Diffusion coefficient ( $\text{cm}^2\text{s}^{-1}$ ) of for lipid A molecules (Set I) at 275 K.	73
Figure 3- 23 Density profiles of key components of (A) <i>B. pertussis</i> and (B) <i>C. trachomatis</i> membranes (set VI).	76
Figure 3- 24 Lipid A headgroup phosphate –phosphate ( $g_{\text{P-P}}(r)$ ; solid lines) and phosphate – $\text{Ca}^{2+}$ counterion ( $g_{\text{P-Ca}^{2+}}(r)$ ; dashed lines) radial distribution functions for (A) <i>H. pylori</i> , (B) <i>C. jejuni</i> , and (C) <i>N. meningitidis</i> .	78
Figure 4- 1 The chemical structure of <i>E. coli</i> lipid A domain.	91
Figure 4- 2 Representative chemical structure of <i>E.coli</i> (a) LPA, (b) LPC and (c) coarse grain mapping of LPC.	92
Figure 4- 3 (A) $\text{PO}_4$ , (B) XYA, (C) SYB, (D) LKO, (E) 0KO, (F) Gal 1, (G) Gal 2, (H) Glc1, (I) Glc 2, (J) Hep 1,(K) Hep 2,(L) Hep 3,(M) LP1,(N) LP2, (O) LP3. The coarse grained mapping scheme showing Martini beads of types Qa (black), P1 (yellow), P2 (blue), P4 (orange), P5 (green), N0 (purple), Na (light blue), and C1(grey) and overlaid on the atomistic structure in ball and stick representation with carbon, oxygen, and linking bonds.	93
Figure 4- 4 (A) Average acyl bond and B) angle frequency distribution. Average disaccharide head group bond (C) and angle frequency distribution (D).	94
Figure 4- 5 Dihedral angle fluctuations (degrees) as a function of simulation time.	95
Figure 4- 6 Simulation setup and self-assembly of OmpF trimer.	97

Figure 4- 7 (A) $A_L$ of outer leaflet lipids in MT1 (orange), MT2 (cyan), and MT3 (green) membranes in (B) Effect of temperature on $A_L$ of outer membrane (same color scheme). (C) Hydrophobic thickness of membrane in Systems 1 and 2. ....	101
Figure 4- 8 Density profile of (a) MT1, (b) MT2 and (c) MT3. ....	104
Figure 4- 9 Distances between interfacial G19 and G135 residues in Fig.4-6b (purple), Fig.4-6c (green) and Fig. 4-6d (blue) .....	106
Figure 4- 10 Average partial number density of membrane components (a) LPA (outer leaflet), (b) DPPE (inner leaflet), (c) POPG (inner leaflet), (d) CDL2 (inner leaflet), and (e) ions (membrane surface) during two microsecond intervals of trimer formation (System 3).....	107
Figure 4- 11 Reversed mapped snapshot of the self-assembled OmpF-OmpF' dimer (cyan; surface representation) in System 3. Difference in (a) pore lumens (arrows) and (b) the overall root-mean-square deviation of individual OmpF (orange) and OmpF' (black) units in the dimer as a function of time. The G19 (red) and G135 (yellow) residues are shown as beads. ....	108
Figure 4- 12 Average partial number density of membrane components (a) LPA (outer leaflet), (b) DPPE (inner leaflet), (c) POPG (inner leaflet), (d) CDL2 (inner leaflet), (e) water (membrane midplane), and (f) ions (membrane surface) during two microsecond intervals of trimer formation (System 6). ....	109
Figure 4- 13 Dynamical properties of OmpF monomers <sup>a</sup> (orange, System 3), dimer <sup>b</sup> (green, System 3) and trimer <sup>c</sup> (blue, System 6) during self-assembly. (a) Root-mean-square fluctuations of backbone beads, (b) G19–G135 pair separation distance, (c) overall root-mean-square deviation, and (d) $R_g$ of the protein as a function of time. In panel (a), the shaded (light blue) regions show the OmpF loop domains (L1–L8). ....	111

Figure 4- 14 Potential of mean force profiles for unbinding of the dimer (System 8) and trimer (System 9) interfaces along the interseparation distance coordinate of the interacting pair. Error bars calculated by Bayesian bootstrapping method are shown in the same color. ....	113
Figure 4- 15 A schematic showing two possible pathways that lead to OmpF trimer formation. ....	114
Figure 4- 16 Dimer formation in different membranes. ....	116
Figure 5- 1 Comparison of CAGE variants (a) Choline bicarbonate (b) 2:1 and (c) 1:1. ....	133
Figure 5- 2 Comparison of CAGE variants (a) 1:2 (b) 1:4, and (c) pure geranic acid. ....	134
Figure 5- 3 Comparison of Na <sup>+</sup> 1: 4 Sodium Geranate.....	135
Figure 6- 1 CLASP workflow.....	152
Figure 6- 2 The coarse grained mapping scheme showing Martini beads of types Q/SQ (cyan), P/SP (orange), C/SC (grey), N/SN (green), A) Doripenem B) Ertapenem C) Biapenem D) Panipenem E) Meropenem F) Imipenem.....	154
Figure 6- 3 PMF of Doripenem (Navy blue) and Ertapenem (red). ....	156
Figure 6- 4 Amino acid – Carbapenem contacts A) Doripenem B) Ertapenem. Tyrosine (blue), Arginine (red), Phenylalanine (yellow), Threonine (green), Histidine (light green), Lysine (orange). ....	158
Figure 6- 5 Tyrosine - Carbapenem contact A) Tyrosine – Doripenem contact B) Tyrosine – Ertapenem contact C) Positions of tyrosine with highest contact.....	161
Figure 6- 6 Arginine - Carbapenem contact A) Arginine – Doripenem contact B) Arginine – Ertapenem contact.....	162
Figure 6- 7 D <sub>ij</sub> of A) Doripenem and B) Ertapenem.....	162
Figure 6- 8 PMF of Panipenem (orange), Biapenem (purple) and Meropenem (green). ....	163

Figure 6- 9 Amino acid – Carbapenem contacts A) Panipenem B) Biapenem C) Meropenem Tyrosine (blue), Arginine (red), Phenylalanine (yellow), Threonine (green), Histidine (light green), Lysine (orange).....	164
Figure 6- 10 Tyrosine - Carbapenem contact A) Tyrosine – Panipenem contact B) Tyrosine – Biapenem contact C) Tyrosine – Meropenem contact.....	165
Figure 6- 11 Arginine - Carbapenem contact A) Arginine –Panipenem contact B) Arginine – Biapenem contact C) Arginine – Meropenem contact.....	166
Figure 6- 12 $D_{ij}$ of A) Panipenem and B) Biapenem C) Meropenem.....	167
Figure 6- 13 Panipenem penetrating OccD1 constriction zone process. Major arginine involved in this process are colored red.....	169
Figure 6- 14 PMF of Imipenem. ....	170
Figure 6- 15 Amino acid – Imipenem contact. ....	170
Figure 6- 16 A) Tyrosine - Imipenem contact B) Arginine – Imipenem contact. ....	171
Figure 6- 17 $D_{ij}$ of Imipenem.....	172

## LIST OF TABLES

Table 3- 1 Summary of chemical structure of lipid A in various species of Gram-negative bacteria. Labels A, A', B, B', C, C', D, and D' correspond to acyl chains shown in Figure 3-1. Label P denotes the total number of phosphates.....	45
Table 3- 2 System Details of the Membrane Simulations Involving DPPE in the Inner Leaflet <sup>a</sup> .....	57
Table 3- 3 System Details of the Membrane Simulations Involving a Complex Composition of the Inner Leaflet.....	58
Table 3- 4 Key Properties of the Membranes and Comparison of the Phase Transition Temperatures from Sets II and VII with Available Experimental Data .....	69
Table 4- 1 Details of MT1–MT3 asymmetric membrane systems (without OmpF proteins).....	96
Table 4- 2 System details of OmpF-membrane simulations.....	97
Table 4- 3 Comparison of area per lipid of lipopolysachharides, phase transition temperature and membrane thickness, and hydrophobic thickness data of MT1, MT2, and MT3 models with available data in the literature. ....	102
Table 4- 4 RMSD (Å) of self-assembled trimer (OmpF-OmpF'-OmpF'') relative to the X-ray crystal structure (pdb:4LSF).....	112
Table 6- 1 System details of Carbapenem-OccD1 simulations .....	151
Table 6- 2 Simulation Performance .....	153

## **CHAPTER 1**

# **BACKGROUND**

## 1.1 Antibiotics Resistance

The discovery of antibiotics in the last century was one of the greatest achievements in the medical world. Before that, a simple scratch may lead to death and infections were the most difficult problem to overcome during surgery.<sup>1</sup> From the 1920s to the 1980s, many different categories of antibiotics were introduced to the market to make some common infections easy to treat.<sup>2</sup> However, we are now at the same point as 70 years ago that infections are becoming harder to treat when bacteria can protect themselves from antibiotics, and our most powerful antibiotics are becoming ineffective, which is called antibiotic resistance.<sup>3-6</sup> Antibiotic resistance occurs when microorganisms gain the ability to stop the antimicrobial from working against it. Standard treatments are becoming less effective, and infections are becoming harder to treat. Even though the occurrence of antibiotic resistance is a natural process, the overuse and misuse of antibiotics are becoming the primary reasons for accelerating this process to causes higher patient mortality and treatment expenses.<sup>7</sup> Consequently, many infectious diseases could no longer being effectively treated by available antibiotics. Now, about 2 million people in United States have hospital-acquired infections, resulting in 99000 deaths per year.<sup>8</sup> It has been estimated that by 2050, 10 million deaths will be caused annually by antibiotic resistance and US\$ 100 trillion in losses if no action is taken.<sup>9</sup> So antibiotic resistance has become one of the biggest threats to human health.<sup>10-</sup>

12

Even though the situation is very serious, the development of new antibiotics has slowed to a standstill and cannot catch up with the emergence of resistant bacteria.<sup>13-17</sup> The last discovery of a new class of antibiotics was in 1987, since then there is a huge void in the history of antibiotics development.<sup>18</sup> Unfortunately, many major pharmaceutical companies are dropping antibiotics development programs now, the development of one FDA-approved antibiotic needs at least 10



years and over 1 billion dollars.<sup>19</sup> Potential antibiotics need to be screened out from more than thousands of small molecules, there is a lack of efficient screening technique, which can help lower the investment and find out the most promising candidates.<sup>20</sup> The low success rate makes the cost higher than expectation, antibiotic resistance develops fast after new antibiotics being introduced, which makes antibiotics a short-term drug and profits are also reduced.<sup>13, 15, 16</sup>. Workable guidance of describing how to design antibiotic clinical trials from US Food and Drug Administration has been long delayed.<sup>21</sup> Even though there are some new antibiotics being introduced to the market, physicians would prefer to use them when the worse situation happens because they are always worry about the development of new resistance. These factors reduce the enthusiasm and motivation of pharmaceutical companies to develop new antibiotics. So now, if we do not take actions, we would have to face the same situation that there were no appropriate treatments for infections.<sup>22-25</sup>

## **1.2 Mechanisms of Antibiotic Resistance**

To better understand how to overcome antibiotic resistance, it is necessary to know how bacteria develop antibiotic resistance. Bacteria, large group of unicellular microorganisms, which are thought to be the earliest life forms on earth.<sup>26-28</sup> After a long period of evolution, bacteria have gained the ability to adapt to hostile environment and sophisticated mechanisms of resistance to some harmful naturally-produced molecules, which are the main source of antibiotics.<sup>2</sup> Resistance to one molecule can be accomplished by more than one biochemical pathway and bacteria can take a variety of mechanisms to defend themselves from harmful molecules.<sup>7</sup> To provide a comprehensive way of explaining the mechanisms, people have categorized them as 1) Modification of Antibiotic Molecule structures, 2) Modification of Antibiotics' Target Sites structures 3) Decreased Antibiotic Penetration 4) Efflux pump to extrude the antimicrobial

compound.<sup>29</sup> For example, resistance to  $\beta$ -lactams in Gram-negative bacteria is mainly because of the effect of  $\beta$ -lactamases, which can break the structure of  $\beta$ -lactams.<sup>30-33</sup> In Gram-positive bacteria, the resistance is due to the modification of penicillin-binding proteins, which are the main target of  $\beta$ -lactams.<sup>34-36</sup> Besides these two mechanisms, Gram-negative bacteria have gained advanced antibiotic resistance mechanisms due to their unique cell envelop structure, the special cell wall prevent the entry of most harmful molecules and this cell wall has a group of proteins called efflux pumps, which can pump out some toxic molecules that are in the periplasmic space of Gram-negative bacteria.<sup>37-42</sup> These advanced mechanisms make infections caused by Gram-negative bacteria even harder to treat. But this outer membrane and proteins are absent in Gram-positive bacteria, which explains why most Gram-negative bacteria are more pathogenic.<sup>43-45</sup>

### **1.2.1 Modification of Antibiotic Molecule structures**

Both Gram-negative and Gram-positive bacteria can produce some specific enzymes that can add some chemical moieties to the harmful compound or directly degrade these molecules.<sup>29, 46, 47</sup> A classic example of adding chemical moieties is aminoglycoside modifying enzymes (AMEs), aminoglycoside is a traditional category of antibiotics that can inhibit protein synthesis in bacteria.<sup>48, 49</sup> AMEs can covalently modify the amino or hydroxyl groups. These modifications lead to weaker affinity between the drug and its target due to the steric hindrance, resulting in higher MIC. AMEs have become the leading cause of aminoglycoside resistance.<sup>50</sup>

In addition to modification, bacteria can also destroy the structure of some harmful compounds. This is achieved by  $\beta$ -lactamases, which are able to destroy the amide bond of the  $\beta$ -lactam ring.<sup>30, 31, 33</sup> In order to prevent the effect of  $\beta$ -lactamases, new  $\beta$ -lactams were introduced to the market with less susceptibility to  $\beta$ -lactamases. However, new  $\beta$ -lactamases were also appeared to be able to hydrolyze the new  $\beta$ -lactams.<sup>32</sup> Now even though people have developed more generations

of  $\beta$ -lactams,  $\beta$ -lactamases that can destroy any compounds were also found in bacteria. This is a typical example of antibiotic driven adaptive bacterial evolution.<sup>51</sup>

### **1.2.2 Modification of Antibiotics' Target Sites structures**

Most antibiotics have their specific targets inside the bacteria to take effect by binding to them.<sup>2, 52, 53</sup> To interfere the binding of antibiotics and their targets, bacteria can protect the targets or modify the structure of these targets to decrease the binding affinity.<sup>54-56</sup> For example, a best-studied example of target protection mechanism involves tetracycline resistance determinants Tet(M) and Tet(O).<sup>57-60</sup> They are widely distributed among many different bacteria species. Tet(M) and Tet(O) show different ways to protect the tetracycline target in ribosome.<sup>61</sup> Tet(M) can directly dislodge tetracycline from its binding site in ribosome by forming interaction between the domain IV of the 16S rRNA and the tetracycline binding site, the formation of this interaction can also change the conformation of ribosome to reduce the probability of rebinding. TetO has the same binding site as tetracycline and thus is able to compete with tetracycline for the site to reduce the binding of tetracycline.<sup>62-64</sup>

Moreover, the structure of target sites can also be altered by bacteria. This process is achieved by enzymes. For example, the effect of erythromycin ribosomal methylation genes has been well studied. An enzyme encoded by these genes can mono- or dimethylate an adenine residue in position A 2058 of the domain V of the 23rRNA of the 50S ribosomal subunit, which is the main binding site for macrolides, lincosamides, and streptogramin B antibiotics.<sup>65-68</sup>

### **1.2.3 Decreased Antibiotic Penetration**

As mentioned above, most antibiotics have their intracellular bacterial hit targets. To get into the bacteria, antibiotics have to penetrate the bacterial cell wall. The cell wall of both Gram-negative

and Gram-positive bacteria can prevent the entry of antibiotics to reach their intracellular targets.<sup>69</sup> It is worth noting that Gram-negative bacteria are becoming resistant to most available antibiotic drugs.<sup>42, 70</sup> The complicated outer membrane of Gram-negative bacteria was found to be able to help them develop antibiotic resistance easily and quickly.<sup>71-74</sup> The outer membrane of Gram-negative bacteria comprises lipopolysaccharides (LPS) as outer leaflet, mixture of phospholipids as inner leaflet and outer membrane proteins (OMPs) as pores to allow transport of specific small molecules.<sup>75, 76</sup> The rigid outer membrane forms the first line of defense, which makes the Gram-negative bacteria much harder to treat than Gram-positive bacteria. Some hydrophilic antibiotics such as  $\beta$ -lactams, tetracyclines and fluoroquinolones are found to be hard to penetrate the hydrophobic outer membrane of Gram-negative bacteria.<sup>71</sup> Vancomycin, which is another example, cannot penetrate the outer membrane of Gram-negative bacteria, so it is ineffective to treat infections caused by Gram-negative bacteria.<sup>77, 78</sup>

#### **1.2.4 Efflux pumps**

Efflux pumps are proteinaceous transporters found in bacterial cell envelop.<sup>79</sup> There are 5 major families of bacterial efflux pumps. 1) The major facilitator superfamily (MFS), 2) The resistance-nodulation-cell division superfamily (RND), 3) The small multidrug resistance family (SMR), 4) The ATP-binding cassette superfamily (ABC), 5) The Multi antimicrobial extrusion protein family (MATE).<sup>80</sup> MFS is mainly found in Gram-positive bacteria and RND is unique to Gram-negative bacteria.<sup>81</sup> In 1980, scientists found that in *E.coli*, its efflux pump was able to extrude tetracycline.<sup>82</sup> Since then, many different kinds of efflux pumps have been identified. Their ability to pump toxic compounds out of bacterial cells have become another leading cause of the development of antibiotic resistance.<sup>37, 80, 83, 84</sup>

### 1.3 Current Study of Antibiotics Resistance

Bacteria mostly develop antibiotic resistance through these 4 mechanisms above. The decreased antibiotic penetration was found to be the major difficulty for most antibiotics, which results in poor effectiveness of most antibiotics. However, current drug design is largely based on the interactions of the molecules with their target sites inside the cells, but the efficacy of antibiotics also depends on the influx mediated by membrane lipids and porins<sup>85</sup>. Since most antibiotics need to enter bacterial cells and bind to their target sites, they must be able to penetrate the bacterial cell envelope. Hydrophobic antibiotics can take a lipid-mediated pathway and porins-mediate pathway is major pathway for hydrophilic antibiotics<sup>74</sup>.

There have been number of experimental and computational studies conducted in understanding the permeability of bacterial lipid membranes and the selectivity of porins. These high diversities of lipid composition and porins types of bacterial outer membranes make huge differences of their permeability. Understanding these differences would lead to more targeted antibiotics structure design.

Nikado<sup>44</sup> et al. found that some hydrophobic antibiotic such as aminoglycosides, macrolides, rifamycins, novobiocin, fusidic acid, cationic peptides are able to penetrate the cell envelop by diffuse through the lipids. However, the core oligosaccharide region of bacterial outer membrane provides a barrier to hydrophobic antibiotics since it contains 6 to 10 sugars. Bacteria mutants with truncated core show high sensitivity to lipophylic agents<sup>86</sup>.

Eren<sup>85</sup> et al. summarized the substrate specificity of Outer Membrane Carboxylate Channels (Occ) of *Pseudomonas aeruginosa* to help understand how this certain bacterium takes majority of small molecules using these channels. They successfully proved that a carboxyl group in the substrates is necessary for them to effectively transport Occ. They also identified the substrate specificities

of the two subfamilies of Occ, which includes 9 different porins. These results here revealed the complexity of the selectivity of porins and the necessity of understanding it, which also lead to rational design of novel antibiotics to fight against *Pseudomonas aeruginosa*.

Modi<sup>87</sup> et al conducted x-ray crystallography, electrophysiology and molecular dynamics simulations to study the outer membrane channels OprP and OprO with high similarity from *Pseudomonas aeruginosa*. They identified the amino acids differences between these two channels at the constriction region, which result in very different selectivity of these two channels. Double mutants of these two porins were generated to understand the functions of these amino acids in determining the channel specificity. These results they provided proposed another promising strategy of modifying specific amino acids to obtain desired channel specificity.

#### **1.4 Overcome Antibiotic Resistance using Computational approaches**

Since the outer membrane of Gram-negative bacteria is the leading cause of antibiotic resistance, and it has attracted a lot of attentions of scientists. Understanding the interactions between this unique membrane and small molecules is essential to guide the development process of new antibiotics. To achieve this goal, many potential candidates need to be tested from some small molecules libraries, e.g. Microsource SPECTRUM Collection<sup>88</sup>, ChemBridge Diversity Set Library<sup>89</sup>, etc. However, the traditional methods have many limitations, and they take longer time, which indirectly leads to an increase in cost. To facilitate this process, computational approaches should be employed to comprehensively understand the details at the molecular level and screen out the most potential antibiotics candidates from these libraries to boost the drug discovery pipeline. Even though there are a lot of experiments going on to study antibiotic resistance, there are still many misunderstandings and unclear areas about the specific process of antibiotic

resistance.<sup>5, 9, 11, 29, 46</sup> Most experimental data can only give us macro conclusions, it is hard for experiments to explore the detailed information in molecular level.

Recently, with the development of Computer-aided drug design (CADD), computational approaches have been widely used to guide and accelerate the early-stage development of new compounds more efficient and cheaply.<sup>90-92</sup> Molecular dynamics (MD), which is a computer-based simulation method to study chemical systems and provide physical behavior of each atom or molecule in the system for nanoseconds to microseconds, can be employed to gain insights into the actions of small molecules on bacterial outer membrane or membrane proteins in molecular level to help develop robust antibiotics.<sup>93-96</sup>

All-atom MD and Coarse-grained MD are being widely used to perform simulations. All-atom MD provides us detailed interactions between each atom but needs more computing resources and time, so it is too expensive for all-atom MD to achieve long simulation time and simulate complicated systems.<sup>97</sup> MD in coarse-grained (CG) level could help achieve longer time scale and larger system sizes by merging several atoms into one bead to reduce the degree of freedom so that people can explore more complex systems.<sup>98</sup> Martini force field is a popular used CG force field that provides us a variety of parameters of LPS, membrane lipids, amino acids, proteins, solvents and ions.<sup>99-104</sup> The details about how MD works will be discussed immediately in the next chapter.

However, when we started, Martini force field did not provide parameters for bacterial LPS. To fill this gap, we built force field parameters for 10 different common pathogenic or non-pathogenic bacteria, available experimental data was used to validate our CG force field parameters.<sup>105, 106</sup> Proteins were also studied to show the ability Martini force field to reproduce the properties and behavior of common membrane proteins.<sup>107</sup> After that, we built our CG representation of

simulation systems comprising Gram-negative bacterial outer membrane with or without OMPs embedded depending on the specific bacteria, target molecule and solvents to explore the exhaustive transport behavior of target molecule through the membrane or OMPs into the periplasmic space. The thermodynamic and kinetic data of the transport can be obtained by our CG molecular simulations as well as the molecule-membrane/molecule-protein interactions. To alleviate the burden of building new systems with different bacteria membranes, OMPs and small molecules, we made an automatic simulation control algorithm to easily achieve our goals, which is a computational automated screening platform that can quickly generate the data for further analysis. These parameters of Gram-negative bacteria as well as the computational platform can better guide the modification of existing antibiotics, design of new antibiotics and faster the process of finding new antibiotics.



## REFERENCE

1. Demain, A. L., Antibiotics: Natural Products Essential to Human Health. *Medicinal Research Reviews* **2009**, 29 (6), 821-842.
2. Walsh, C., *Antibiotics: actions, origins, resistance*. American Society for Microbiology (ASM): 2003.
3. Stewart, P. S.; Costerton, J. W., Antibiotic resistance of bacteria in biofilms. *Lancet* **2001**, 358 (9276), 135-138.
4. Neu, H. C., THE CRISIS IN ANTIBIOTIC-RESISTANCE. *Science* **1992**, 257 (5073), 1064-1073.
5. Levy, S. B.; Marshall, B., Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine* **2004**, 10 (12), S122-S129.
6. Cohen, S. N.; Chang, A. C. Y.; Hsu, L., NONCHROMOSOMAL ANTIBIOTIC RESISTANCE IN BACTERIA - GENETIC TRANSFORMATION OF ESCHERICHIA-COLI BY R-FACTOR DNA. *Proceedings of the National Academy of Sciences of the United States of America* **1972**, 69 (8), 2110-&.
7. Davies, J.; Davies, D., Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews* **2010**, 74 (3), 417-+.
8. Revelas, A., Healthcare-associated infections: A public health problem. *Nigerian medical journal: journal of the Nigeria Medical Association* **2012**, 53 (2), 59.
9. O'Neill, J., Antimicrobial resistance: tackling a crisis for the health and wealth of nations. *Rev. Antimicrob. Resist* **2014**, 20, 1-16.
10. Byarugaba, D., Antimicrobial resistance in developing countries and responsible risk factors. *International journal of antimicrobial agents* **2004**, 24 (2), 105-110.

11. Li, B.; Webster, T. J., Bacteria antibiotic resistance: New challenges and opportunities for implant - associated orthopedic infections. *Journal of Orthopaedic Research®* **2018**, 36 (1), 22-32.
12. Prestinaci, F.; Pezzotti, P.; Pantosti, A., Antimicrobial resistance: a global multifaceted phenomenon. *Pathogens and global health* **2015**, 109 (7), 309-318.
13. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L., Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature reviews Drug discovery* **2007**, 6 (1), 29.
14. Silver, L. L., Challenges of antibacterial discovery. *Clinical microbiology reviews* **2011**, 24 (1), 71-109.
15. Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K., The challenge of new drug discovery for tuberculosis. *Nature* **2011**, 469 (7331), 483.
16. Livermore, D. M.; Discovery, B. S. f. A. C. W. P. o. T. U. N. R. A. D.; Development; Blaser, M.; Carrs, O.; Cassell, G.; Fishman, N.; Guidos, R.; Levy, S.; Powers, J.; Norrby, R., Discovery research: the scientific challenge of finding new antibiotics. *Journal of antimicrobial chemotherapy* **2011**, 66 (9), 1941-1944.
17. Brown, E. D.; Wright, G. D., Antibacterial drug discovery in the resistance era. *Nature* **2016**, 529 (7586), 336.
18. Roundtable, C. S.; on Earth, D.; Council, N. R., Challenges In Overcoming Antibiotic Resistance. **2014**.
19. Sukkar, E., Why are there so few antibiotics in the research and development pipeline. *Pharm. J* **2013**, 291, 520.

20. Lewis, K., Platforms for antibiotic discovery. *Nature reviews Drug discovery* **2013**, *12* (5), 371.
21. Golkar, Z.; Bagasra, O.; Pace, D. G., Bacteriophage therapy: a potential solution for the antibiotic resistance crisis. *The Journal of Infection in Developing Countries* **2014**, *8* (02), 129-136.
22. Tacconelli, E.; Carrara, E.; Savoldi, A.; Harbarth, S.; Mendelson, M.; Monnet, D. L.; Pulcini, C.; Kahlmeter, G.; Kluytmans, J.; Carmeli, Y., Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet Infectious Diseases* **2018**, *18* (3), 318-327.
23. Smith, P. A.; Koehler, M. F.; Girgis, H. S.; Yan, D.; Chen, Y.; Chen, Y.; Crawford, J. J.; Durk, M. R.; Higuchi, R. I.; Kang, J., Optimized arylomycins are a new class of Gram-negative antibiotics. *Nature* **2018**, *561* (7722), 189.
24. Drown, B. S.; Hergenrother, P. J., Going on offense against the gram-negative defense. *Proceedings of the National Academy of Sciences* **2018**, 201807278.
25. Liu, R.; Miller, P. A.; Vakulenko, S. B.; Stewart, N. K.; Boggess, W. C.; Miller, M. J., A Synthetic Dual Drug Sideromycin Induces Gram-Negative Bacteria To Commit Suicide with a Gram-Positive Antibiotic. *Journal of medicinal chemistry* **2018**, *61* (9), 3845-3854.
26. Doetsch, R. N.; Cook, T. M., *Introduction to bacteria and their ecobiology*. Springer Science & Business Media: 2012.
27. Russell, A., Introduction of biocides into clinical practice and the impact on antibiotic - resistant bacteria. *Journal of Applied Microbiology* **2002**, *92*, 121S-135S.
28. Donohue, D.; Salminen, S., Safety of probiotic bacteria. *Asia pacific journal of clinical nutrition* **1996**, *5*, 25-28.

29. Munita, J. M.; Arias, C. A., Mechanisms of antibiotic resistance. *Microbiology spectrum* **2016**, *4* (2).
30. Bush, K.; Jacoby, G. A., Updated functional classification of  $\beta$ -lactamases. *Antimicrobial agents and chemotherapy* **2010**, *54* (3), 969-976.
31. Bradford, P. A., Extended-spectrum  $\beta$ -lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. *Clinical microbiology reviews* **2001**, *14* (4), 933-951.
32. Jacoby, G. A.; Munoz-Price, L. S., The new  $\beta$ -lactamases. *New England Journal of Medicine* **2005**, *352* (4), 380-391.
33. Sykes, R.; Matthew, M., The  $\beta$ -lactamases of gram-negative bacteria and their role in resistance to  $\beta$ -lactam antibiotics. *Journal of antimicrobial chemotherapy* **1976**, *2* (2), 115-157.
34. Tomasz, A.; Drugeon, H.; De Lencastre, H.; Jabes, D.; McDougall, L.; Bille, J., New mechanism for methicillin resistance in *Staphylococcus aureus*: clinical isolates that lack the PBP 2a gene and contain normal penicillin-binding proteins with modified penicillin-binding capacity. *Antimicrobial agents and chemotherapy* **1989**, *33* (11), 1869-1874.
35. Malouin, F.; Bryan, L., Modification of penicillin-binding proteins as mechanisms of beta-lactam resistance. *Antimicrobial agents and chemotherapy* **1986**, *30* (1), 1.
36. Zapun, A.; Contreras-Martel, C.; Vernet, T., Penicillin-binding proteins and  $\beta$ -lactam resistance. *FEMS microbiology reviews* **2008**, *32* (2), 361-385.
37. Nikaido, H., Multidrug efflux pumps of gram-negative bacteria. *Journal of bacteriology* **1996**, *178* (20), 5853.

38. Hancock, R. E., Resistance mechanisms in *Pseudomonas aeruginosa* and other nonfermentative gram-negative bacteria. *Clinical Infectious Diseases* **1998**, 27 (Supplement\_1), S93-S99.
39. Hancock, R. E., Peptide antibiotics. *The Lancet* **1997**, 349 (9049), 418-422.
40. Poole, K., Multidrug resistance in Gram-negative bacteria. *Current opinion in microbiology* **2001**, 4 (5), 500-508.
41. Poole, K., Efflux-mediated resistance to fluoroquinolones in gram-negative bacteria. *Antimicrobial agents and chemotherapy* **2000**, 44 (9), 2233-2241.
42. Beveridge, T. J., Structures of gram-negative cell walls and their derived membrane vesicles. *Journal of bacteriology* **1999**, 181 (16), 4725-4733.
43. Pagès, J.-M.; James, C. E.; Winterhalter, M., The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nature Reviews Microbiology* **2008**, 6 (12), 893.
44. Nikaido, H.; Vaara, M., Molecular basis of bacterial outer membrane permeability. *Microbiological reviews* **1985**, 49 (1), 1.
45. Tagg, J. R.; Dajani, A. S.; Wannamaker, L. W., Bacteriocins of gram-positive bacteria. *Bacteriological reviews* **1976**, 40 (3), 722.
46. Wright, G. D., Molecular mechanisms of antibiotic resistance. *Chemical communications* **2011**, 47 (14), 4055-4061.
47. Reynolds, P. E., Structure, biochemistry and mechanism of action of glycopeptide antibiotics. *European Journal of Clinical Microbiology and Infectious Diseases* **1989**, 8 (11), 943-950.

48. Ramirez, M. S.; Tolmasky, M. E., Aminoglycoside modifying enzymes. *Drug Resistance Updates* **2010**, *13* (6), 151-171.
49. Wright, G. D., Aminoglycoside-modifying enzymes. *Current opinion in microbiology* **1999**, *2* (5), 499-503.
50. Robicsek, A.; Strahilevitz, J.; Jacoby, G. A.; Macielag, M.; Abbanat, D.; Park, C. H.; Bush, K.; Hooper, D. C., Fluoroquinolone-modifying enzyme: a new adaptation of a common aminoglycoside acetyltransferase. *Nature Medicine* **2006**, *12* (1), 83.
51. Andersson, D. I., Persistence of antibiotic resistant bacteria. *Current opinion in microbiology* **2003**, *6* (5), 452-456.
52. Kohanski, M. A.; Dwyer, D. J.; Collins, J. J., How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology* **2010**, *8* (6), 423.
53. Poehlsgaard, J.; Douthwaite, S., The bacterial ribosome as a target for antibiotics. *Nature Reviews Microbiology* **2005**, *3* (11), 870.
54. Leclercq, R.; Courvalin, P., Bacterial resistance to macrolide, lincosamide, and streptogramin antibiotics by target modification. *Antimicrobial agents and chemotherapy* **1991**, *35* (7), 1267.
55. Wilson, D. N., Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Reviews Microbiology* **2014**, *12* (1), 35.
56. Džidić, S.; Šušković, J.; Kos, B., Antibiotic resistance mechanisms in bacteria: biochemical and genetic aspects. *Food Technology & Biotechnology* **2008**, *46* (1).
57. Roberts, M., Tetracycline resistance determinants: mechanisms of action, regulation of expression, genetic mobility, and distribution. *FEMS microbiology reviews* **1996**, *19* (1), 1-24.

58. Levy, S. B.; McMurry, L. M.; Barbosa, T. M.; Burdett, V.; Courvalin, P.; Hillen, W.; Roberts, M. C.; Rood, J. I.; Taylor, D. E., Nomenclature for new tetracycline resistance determinants. *Antimicrobial agents and chemotherapy* **1999**, *43* (6), 1523-1524.
59. Connell, S. R.; Tracz, D. M.; Nierhaus, K. H.; Taylor, D. E., Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrobial agents and chemotherapy* **2003**, *47* (12), 3675-3681.
60. Taylor, D. E.; Chau, A., Tetracycline resistance mediated by ribosomal protection. *Antimicrobial agents and chemotherapy* **1996**, *40* (1), 1.
61. Bergeron, J.; Ammirati, M.; Danley, D.; James, L.; Norcia, M.; Retsema, J.; Strick, C. A.; Su, W.-G.; Sutcliffe, J.; Wondrack, L., Glycylcyclines bind to the high-affinity tetracycline ribosomal binding site and evade Tet (M)-and Tet (O)-mediated ribosomal protection. *Antimicrobial agents and chemotherapy* **1996**, *40* (9), 2226-2228.
62. Aires, J.; Doucet-Populaire, F.; Butel, M., Tetracycline resistance mediated by tet (W), tet (M), and tet (O) genes of Bifidobacterium isolates from humans. *Applied and environmental microbiology* **2007**, *73* (8), 2751-2754.
63. Connell, S. R.; Trieber, C. A.; Dinos, G. P.; Einfeldt, E.; Taylor, D. E.; Nierhaus, K. H., Mechanism of Tet (O) - mediated tetracycline resistance. *The EMBO journal* **2003**, *22* (4), 945-953.
64. Giovanetti, E.; Brenciani, A.; Lupidi, R.; Roberts, M. C.; Varaldo, P. E., Presence of the tet (O) gene in erythromycin-and tetracycline-resistant strains of Streptococcus pyogenes and linkage with either the mef (A) or the erm (A) gene. *Antimicrobial agents and chemotherapy* **2003**, *47* (9), 2844-2849.

65. Weisblum, B., Erythromycin resistance by ribosome modification. *Antimicrobial agents and chemotherapy* **1995**, 39 (3), 577.
66. Lai, C.; Weisblum, B., Altered methylation of ribosomal RNA in an erythromycin-resistant strain of *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences* **1971**, 68 (4), 856-860.
67. Skinner, R.; Cundliffe, E.; Schmidt, F., Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *Journal of Biological Chemistry* **1983**, 258 (20), 12702-12706.
68. Lai, C.-J.; Dahlberg, J. E.; Weisblum, B., Structure of an inducibly methylatable nucleotide sequence in 23S ribosomal ribonucleic acid from erythromycin-resistant *Staphylococcus aureus*. *Biochemistry* **1973**, 12 (3), 457-460.
69. Salton, M., Studies of the bacterial cell wall: IV. The composition of the cell walls of some gram-positive and gram-negative bacteria. *Biochimica et biophysica acta* **1953**, 10, 512-523.
70. Hancock, R. E., The bacterial outer membrane as a drug barrier. *Trends in microbiology* **1997**, 5 (1), 37-42.
71. Hancock, R. E.; Bell, A., Antibiotic uptake into gram-negative bacteria. In *Perspectives in Antiinfective Therapy*, Springer: 1989; pp 42-53.
72. Zimmermann, W.; Rosselet, A., Function of the outer membrane of *Escherichia coli* as a permeability barrier to beta-lactam antibiotics. *Antimicrobial agents and chemotherapy* **1977**, 12 (3), 368-372.
73. Vaara, M., Outer membrane permeability barrier to azithromycin, clarithromycin, and roxithromycin in gram-negative enteric bacteria. *Antimicrobial agents and chemotherapy* **1993**, 37 (2), 354-356.



74. Delcour, A. H., Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2009**, *1794* (5), 808-816.
75. Nikaido, H.; Nakae, T., The outer membrane of Gram-negative bacteria. In *Advances in microbial physiology*, Elsevier: 1980; Vol. 20, pp 163-250.
76. Costerton, J.; Ingram, J.; Cheng, K., Structure and function of the cell envelope of gram-negative bacteria. *Bacteriological reviews* **1974**, *38* (1), 87.
77. Zhou, A.; Kang, T. M.; Yuan, J.; Beppler, C.; Nguyen, C.; Mao, Z.; Nguyen, M. Q.; Yeh, P.; Miller, J. H., Synergistic interactions of vancomycin with different antibiotics against *Escherichia coli*: trimethoprim and nitrofurantoin display strong synergies with vancomycin against wild-type *E. coli*. *Antimicrobial agents and chemotherapy* **2015**, *59* (1), 276-281.
78. Vaara, M.; Nurminen, M., Outer membrane permeability barrier in *Escherichia coli* mutants that are defective in the late acyltransferases of lipid A biosynthesis. *Antimicrobial agents and chemotherapy* **1999**, *43* (6), 1459-1462.
79. Malléa, M.; Mahamoud, A.; Chevalier, J.; Alibert-Franco, S.; Brouant, P.; Barbe, J.; Pagès, J.-M., Alkylaminoquinolines inhibit the bacterial antibiotic efflux pump in multidrug-resistant clinical isolates. *Biochemical Journal* **2003**, *376* (3), 801-805.
80. Webber, M.; Piddock, L., The importance of efflux pumps in bacterial antibiotic resistance. *Journal of antimicrobial chemotherapy* **2003**, *51* (1), 9-11.
81. Poole, K., Efflux-mediated multiresistance in Gram-negative bacteria. *Clinical Microbiology and infection* **2004**, *10* (1), 12-26.
82. McMurtry, L.; Petrucci, R. E.; Levy, S. B., Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **1980**, *77* (7), 3974-3977.

83. Stavri, M.; Piddock, L. J.; Gibbons, S., Bacterial efflux pump inhibitors from natural sources. *Journal of antimicrobial chemotherapy* **2006**, *59* (6), 1247-1260.
84. Murakami, S.; Nakashima, R.; Yamashita, E.; Yamaguchi, A., Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* **2002**, *419* (6907), 587.
85. Eren, E.; Vijayaraghavan, J.; Liu, J.; Cheneke, B. R.; Touw, D. S.; Lepore, B. W.; Indic, M.; Movileanu, L.; Van den Berg, B., Substrate specificity within a family of outer membrane carboxylate channels. *PLoS biology* **2012**, *10* (1), e1001242.
86. Vaara, M., Agents that increase the permeability of the outer membrane. *Microbiology and Molecular Biology Reviews* **1992**, *56* (3), 395-411.
87. Modi, N.; Ganguly, S.; Bárcena-Uribarri, I.; Benz, R.; van den Berg, B.; Kleinekathöfer, U., Structure, dynamics, and substrate specificity of the OprO porin from *Pseudomonas aeruginosa*. *Biophysical journal* **2015**, *109* (7), 1429-1438.
88. Weisman, J. L.; Liou, A. P.; Shelat, A. A.; Cohen, F. E.; Kiplin Guy, R.; DeRisi, J. L., Searching for new antimalarial therapeutics amongst known drugs. *Chemical biology & drug design* **2006**, *67* (6), 409-416.
89. Reynès, C.; Host, H.; Camproux, A.-C.; Laconde, G.; Leroux, F.; Mazars, A.; Deprez, B.; Fahraeus, R.; Villoutreix, B. O.; Sperandio, O., Designing focused chemical libraries enriched in protein-protein interaction inhibitors using machine-learning methods. *PLoS computational biology* **2010**, *6* (3), e1000695.
90. Åqvist, J.; Medina, C.; Samuelsson, J.-E., A new method for predicting binding affinity in computer-aided drug design. *Protein Engineering, Design and Selection* **1994**, *7* (3), 385-391.
91. Lill, M. A.; Danielson, M. L., Computer-aided drug design platform using PyMOL. *Journal of computer-aided molecular design* **2011**, *25* (1), 13-19.

92. Haney, E. F.; Brito-Sánchez, Y.; Trimble, M. J.; Mansour, S. C.; Cherkasov, A.; Hancock, R. E., Computer-aided discovery of peptides that specifically attack bacterial biofilms. *Scientific reports* **2018**, *8* (1), 1871.
93. Haloi, N.; Shekhar, M.; Drown, B. S.; Hergenrother, P. J.; Tajkhorshid, E., Antibiotic Permeation across the Bacterial Outer Membrane Porin. *Biophysical journal* **2018**, *114* (3), 226a.
94. Vaiana, A.; Westhof, E.; Auffinger, P., A molecular dynamics simulation study of an aminoglycoside/A-site RNA complex: conformational and hydration patterns. *Biochimie* **2006**, *88* (8), 1061-1073.
95. Hajjar, E.; Mahendran, K. R.; Kumar, A.; Bessonov, A.; Petrescu, M.; Weingart, H.; Ruggerone, P.; Winterhalter, M.; Ceccarelli, M., Bridging timescales and length scales: from macroscopic flux to the molecular mechanism of antibiotic diffusion through porins. *Biophysical journal* **2010**, *98* (4), 569-575.
96. Kulik, M.; Mori, T.; Sugita, Y.; Trylska, J., Molecular Dynamics Simulations of a Riboswitch Binding Aminoglycoside Antibiotics. *Biophysical journal* **2018**, *114* (3), 433a.
97. Huang, J.; MacKerell Jr, A. D., CHARMM36 all - atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry* **2013**, *34* (25), 2135-2145.
98. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H., The MARTINI force field: coarse grained model for biomolecular simulations. *The journal of physical chemistry B* **2007**, *111* (27), 7812-7824.
99. Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J., The MARTINI coarse-grained force field: extension to proteins. *Journal of chemical theory and computation* **2008**, *4* (5), 819-834.

100. Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J., Polarizable water model for the coarse-grained MARTINI force field. *PLoS computational biology* **2010**, *6* (6), e1000810.
101. de Jong, D. H.; Singh, G.; Bennett, W. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J., Improved parameters for the martini coarse-grained protein force field. *Journal of chemical theory and computation* **2012**, *9* (1), 687-697.
102. López, C. A.; Rzepiela, A. J.; De Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J., Martini coarse-grained force field: extension to carbohydrates. *Journal of chemical theory and computation* **2009**, *5* (12), 3195-3210.
103. López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J., Martini force field parameters for glycolipids. *Journal of chemical theory and computation* **2013**, *9* (3), 1694-1708.
104. Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J., Martini coarse-grained force field: extension to DNA. *Journal of chemical theory and computation* **2015**, *11* (8), 3932-3945.
105. Ma, H.; Irudayanathan, F. J.; Jiang, W.; Nangia, S., Simulating Gram-negative bacterial outer membrane: a coarse grain model. *The journal of physical chemistry B* **2015**, *119* (46), 14668-14682.
106. Ma, H.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan, A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S., Modeling diversity in structures of bacterial outer membrane lipids. *Journal of chemical theory and computation* **2017**, *13* (2), 811-824.
107. Ma, H.; Khan, A.; Nangia, S., Dynamics of OmpF trimer formation in the bacterial outer membrane of Escherichia coli. *Langmuir* **2017**, *34* (19), 5623-5634.

## **CHAPTER 2**

### **Basic Principles of Molecular Dynamics**

Understanding the microscopic interactions of atoms and molecules have always been a problem in the scientific community due to their tiny size. The basic idea of MD simulation is to calculate the interactions between the particles and integrate the equations of motion to explore the microscopic properties and behavior of some systems that scientists have interested in.<sup>1</sup> Classical mechanics, statistical mechanics, thermal dynamics and kinetic theory are being used to obtain the microscopic properties and corresponding macroscopic properties can be calculated by sampling and averaging the microscopic behavior.<sup>2</sup>

When performing MD simulations, all particles in a specific system can move and interact for a fixed period of time, generating the trajectories by solving Newton's equations of motion.<sup>3</sup> During each step, the position, velocity, acceleration of every particle will be calculated and used to predict the next position, velocity and acceleration. By repeating the process for all particles, the dynamic evolution of the system can be revealed.<sup>4</sup> The potential function, which is also known as force field, is used to determine the forces acting on each particle. Potentials are defined to be able to reproduce or mimic the structural or conformational changes to study a complex biological system.<sup>1, 5, 6</sup>

In this chapter, a summary of key concepts and theoretical basis is introduced to clarify how MD simulation works. Some commonly used algorithms, sample applications, limitations of MD, software and force fields are also discussed briefly.

## **2.1 Equations of Motion**

The basic idea of MD simulation is to solve the Newtown's equations of motion.

$$F_i = m_i a_i \quad (1.1)$$

Where  $F_i$  is the net force acting on an atom or molecule,  $m_i$  is the mass of the atom and  $a$  is the acceleration.

Alternatively,

$$m_i \frac{d^2 r_i}{dt^2} = F_i \quad (1.2)$$

The forces acting on the atoms are usually derived from a potential energy. And the potential energy function  $U$  includes bonded and non-bonded elements as described below<sup>7</sup>.

$$m_i \frac{d^2 r_i}{dt^2} = - \frac{dU}{dr_i} \quad (1.3)$$

$$U = U_{bonded} + U_{non-bonded}$$

Firstly, we will consider the intramolecular bonding interactions including interatomic distance, bend angles and torsion angles.

$$U_{bonded} = \frac{1}{2} \sum_{bonds} k_{ij}^r (r_{ij} - r_{eq})^2 + \frac{1}{2} \sum_{angles} k_{ijk}^\theta (\theta_{ijk} - \theta_{eq})^2 + \frac{1}{2} \sum_{torsion} \sum_m k_{ijkl}^{\phi,m} (1 + \cos(m\Phi_{ijkl} - Ym)) \quad (1.4)$$

Bonds typically involve the separation  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  between adjacent pairs of atoms and a harmonic form with specified equilibrium separation, the bend angle  $\theta_{ijk}$  are between continuous bond vectors such as  $\mathbf{r}_i - \mathbf{r}_j$  and  $\mathbf{r}_j - \mathbf{r}_k$ , the torsion angles  $\Phi_{ijkl}$  are defined in terms of three connected bonds. The torsional potential typically involves an expansion in periodic functions of order  $m = 1, 2, \dots$

Non-bonded interactions include Van der Waals Potential and electric potential energy.

$$U_{non-bonded} = \sum_{atom\ pairs} 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{atom\ pairs} k \frac{q_i q_j}{r_{ij}} \quad (1.5)$$

The first term describes the Van der Waals repulsive and attractive interatomic forces in the form of the Lennard – Jones 12-6 potential, and the send term is the Coulomb electrostatic potential.

Now the force can be calculated using equation 1.3 when total potential energy is defined and known. Together with the Taylor series expansions<sup>8</sup>:

$$r_i(t_0 + \Delta t) = r_i(t_0) + v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)(\Delta t)^2 \quad 1.6$$

$$r_i(t_0 - \Delta t) = r_i(t_0) - v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)(\Delta t)^2 \quad 1.7$$

1.6 + 1.7:

$$\begin{aligned} r_i(t_0 + \Delta t) &= r_i(t_0) + v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)(\Delta t)^2 \\ + \quad r_i(t_0 - \Delta t) &= r_i(t_0) - v_i(t_0)\Delta t + \frac{1}{2}a_i(t_0)(\Delta t)^2 \end{aligned}$$

we can get

$$r_i(t_0 + \Delta t) = -r_i(t_0 - \Delta t) + 2r_i(t_0) + \frac{1}{2}a_i(t_0)(\Delta t)^2$$

where  $r$  is the position,  $v$  is the velocity (the first derivative with respect to time),  $a$  is the acceleration.  $-r_i(t_0 - \Delta t)$  is the position at  $(t_0 - \Delta t)$ ,  $r_i(t_0)$  is the positions at  $t_0$ ,  $a_i$  is the acceleration at  $t_0$ .

$\Delta t$  is an important parameter that we define in MD simulations that is known. Mass of atom or particle is known from periodic table of elements. Acceleration can be obtained from equation 1.1 because force is known when potential energy is defined. When starting a MD simulation, the initial configuration tells us the initial  $r_i(0)$ , the distribution of velocities for all atoms at a certain temperature we pre-set follow the Maxwell–Boltzmann distribution, which predicts the most probable and average velocity.<sup>9</sup> So after  $\Delta t$ , all atoms' new position, net force, velocity and acceleration can be determined. By defining the total simulation time, this step will be repeated many time and finally the movement of atoms can be simulated.



## **2.2 Force fields**

For MD simulations, force fields provide potential functions mentioned in equation 1.3 and parameter sets used in the potential functions to obtain the interactions between the atoms or coarse-grained molecules in the systems.<sup>10</sup> For most force fields, the parameter sets of potential energy can be derived from experimental data and quantum mechanics. There are many different force fields available now, such as OPLS (the Optimized potential for liquid simulation)<sup>11</sup>, ECEPP (the empirical conformational energy program for peptides)<sup>12</sup>, AMBER (assisted model building with energy refinement)<sup>13, 14</sup>, CHARMM (chemistry at Harvard molecular mechanics)<sup>15</sup>, GROMOS (GRoningen MOlecular Simulation force field)<sup>16</sup>, Martini coarse grain force field<sup>17</sup>, etc. Different force fields may differ in many aspects and also at different levels. These force fields were developed by different research groups and each force field has their own emphasis on specific biomolecules or systems. They all show different performance of simulating proteins, organic solvents, solvents, nucleic acids, etc. Because the research groups who developed force fields have their own research interests, the force fields were tuned for specific types of problems. They are also compatible with different software. So the choice of force field depends on the actual simulated system.<sup>18-20</sup> Most force fields can be classified into two types: All-atom force field and Coarse Grain force field.

### **2.2.1 All-atom force field**

All-atom force field, as the name suggests, provides parameter sets for every single atom in a system. GROMOS force field, which is an All-atom force field, was developed at University of Groningen and Swiss Federal Institute of Technology. GROMOS force field is suitable to simulate most small molecules, solvent and proteins. It can also be easily converted to coarse grain force field. The potential functions of GROMOS force field is described as equation 1.4 and 1.5, which apply

to most All-atom force fields even though they have different parameter sets from different sources used in these potential functions.<sup>16</sup> Some other all-atom force field hydrogen bonding potential like AMBER force field.<sup>13</sup> All-atom force field describes the detailed intermolecular and intramolecular potential, so it provides accurate behavior of all atoms and molecules in a system. However, due to extensive calculations of all-atom force fields, they cannot be applied to large biological systems (> 100 nm) and some biological phenomenon that cost more than 100 ns.<sup>21</sup>

### **2.2.2 Coarse grain force field**

Coarse-grained force field, on the contrary, is widely used in long-time simulations of macromolecules such as lipids, sugars, sterols, polymers, proteins, nucleic acids, and multi-component complexes. The MARTINI force field, which was developed by Marrink and coworkers at the University of Groningen.<sup>17</sup> After 2007, MARTINI force field was extended to a variety of biomolecules that we needed to use in my work.<sup>22-27</sup>

MARTINI force field is one of the most popular coarse grain force field now. Based on all-atom force field, MARTINI force field employs certain rules to coarse grain the atoms, on average 3 to 4 heavy atoms are mapped into one MARTINI coarse grained (CG) bead. This mapping reduces the number of particles being simulated in a system and also the degree of freedom to reduce the amount of calculations.

To accurately reproduce the interatomic and intraatomic interactions from all-atom force field, CG beads in Martini force field are classified into 4 major types and 18 subtypes: polar (P), nonpolar (N), apolar (C) and charged (Q). Each major type has its own subtypes to represent specific chemical units or functional groups based on their polarity. The combination of these 18 subtypes can represent a lot of biomolecules. Because hydrogens are ignored in coarse grain force fields,

based on the ability of forming hydrogen bonding, some functional groups can be divided into 4 kinds: d (donor), A (acceptor), da (both) and 0 (none).

MARTINI force field still uses similar bonded and non-bonded potential energy described in equation 1.3, 1.4 and 1.5. Comparisons of MARTINI force field and some other all-atom force fields have been done to verify the ability of MARTINI force field to reproduce the accurate behavior of lipid bilayers, formation of vesicles, bacterial lipopolysaccharides membrane, etc. The simplified CG model helps MARTINI force field to perform simulations of large systems ( $> 100$  nm) and longer time scale ( $> 100$  ns), which is the best choice for my work to study the behavior of complex membrane and proteins.<sup>28-31</sup>

## 2.3 Ensembles

In MD simulations, sometimes we have to keep some variables of the system constant to mimic the real experimental conditions. Different statistical ensembles can be generated by controlling the energy  $E$ , volume  $V$ , temperature  $T$ , pressure  $P$ , and number of particles  $N$ . After fixing one or more variables mentioned above, structural, energetic, and dynamic properties can be obtained based on the averages or the fluctuations of these fixed and unfixed variables. The most common used ensembles in my study are NVT (constant number of particles, constant volume, constant temperature)<sup>32</sup> and NPT (constant number of particles, constant pressure, constant temperature)<sup>33</sup>. NVT ensemble, also known as the canonical ensemble, the volume and temperature are fixed during the simulation process. NVT ensemble can be obtained using direct temperature scaling at the initialization stage and temperature-bath coupling during the simulation to control the system temperature. Volume is always fixed during the whole process. NVT ensembles are mostly used when performing conformational searches of molecules in vacuum regardless of defining volume,

pressure, and density. In this condition, pressure is not a significant factor and NVT could provide the advantage of less perturbation of the trajectory without coupling to a pressure bath.

NPT ensemble, known as isothermal-isobaric ensemble, the temperature and pressure are fixed. Isothermal-isobaric ensemble controls both temperature and pressure, which applies to most cases of my study when pressure is an important factor and one atmosphere is always used to mimic the realistic condition. In NPT ensemble, the volume of simulation systems is allowed to change to adjust the pressure to preset value. This ensemble considers correct pressure, volume, and densities are important factors. NPT ensemble can also be used to obtain another NVT ensemble when desired temperature and pressure are needed to be achieved.

To achieve desired temperature and pressure, external heat and pressure bath controls are supplied, they are also known as “thermostat” and “barostat”.<sup>34, 35</sup> There are many frequently-used thermostat and barostat, such as Nosé-Hoover Thermostat<sup>36</sup>, Parrinello-Rahman barostat<sup>37</sup>, Anderson Thermostat<sup>38</sup> and Berendsen thermostat/barostat<sup>39</sup>, etc. They have their own advantages and disadvantages in different situations.

The choice of thermostat and barostat depends on the actual situation and the original design purpose of these algorithms. These methods were designed based on specific systems. Based on other people’s experience and the systems I was studying, Berendsen thermostat and Berendsen barostat were most frequently used to control the temperature and pressure as recommended by the force field I was using and can be widely used for different systems. Another important reason is the potential energy of the coarse-grained force field I was using was developed to take into account the suitability of these types of thermostat and barostat. Some tests have been done to prove the good performance of these two methods to achieve desired temperature and pressure with less fluctuations.

### 2.3.1 Berendsen thermostat

Berendsen thermostat is a coupling to an external ‘heat bath’ with given temperature  $T_D$ . The function of the heat bath is to compensate for missing or removed excess energy for the system.

The rate of change of the actual temperature  $T_A$  is related to the preset temperature.

$$\frac{dT}{dt} = \frac{T_D - T_A}{\tau} \quad 1.8$$

Based on the equation above, the effect of Berendsen thermostat is that a deviation of the system temperature from initial  $T_A$  is slowly corrected to the desired temperature  $T_D$ .  $\tau$  is a time constant meaning the strength of the coupling between the system and the heat bath and should be preset before starting a simulation, also known as relaxation time.

The temperature change between successive steps based on equation 1.8 is

$$\Delta T = \frac{\Delta t}{\tau} (T_D - T_A) \quad 1.9$$

A factor  $\lambda$  is defined as:

$$\lambda = [1 + \frac{\Delta t}{\tau} (\frac{T_D}{T_A} - 1)] \quad 1.10$$

Because the temperature of a system is corresponding to its kinetic energy, or velocity. So the temperature of the system can be adjusted by scaling the velocity by the factor  $\lambda$  defined in equation 1.10.

$$v_{new} = v\sqrt{\lambda}$$

Based on experience before, when  $\frac{\Delta t}{\tau} = 0.0025$ , a good performance of Berendsen thermostat can be achieved.

### 2.3.2 Berendsen barostat

Similar as thermostat, a system can be coupled to an external ‘pressure bath’ with given pressure  $P_D$  when using NPT ensemble. Now, the atomic coordinates and the box vectors are rescaled at

each step instead of velocity in temperature coupling. It has the effect of a first-order kinetic relaxation of the actual pressure  $P_A$  towards a given pressure  $P_D$  by

$$\frac{dP}{dt} = \frac{P_D - P_A}{\tau} \quad 1.11$$

The rescaling of atomic coordinates and the box vectors are achieved by a factor  $\mu$ , which is defined as

$$\mu_{ij} = \delta_{ij} - \frac{\Delta t}{3\tau} \beta_{ij} (P_D - P_A) \quad 1.12$$

Here  $\delta_{ij}$  is the Kronecker delta,  $\tau$  is the pressure coupling time constant and  $\beta$  is the isothermal compressibility of the system, which can be obtained from experiments. For example, when water at 1 atm and 300 K,  $\beta = 4.6 \times 10^{-10} \text{ Pa}^{-1} = 4.6 \times 10^{-5} \text{ bar}^{-1}$ . These two parameters must be set before starting the simulation.

## ***2.4 Energy minimization***

In MD simulations, the stable conformers of a system is necessary to perform the simulations and understand the microscopic and macroscopic properties. When building simulation systems using some computational chemistry software packages, the initial geometries are always not at the stable state, some high energy conformers exist due to the molecular overlap, unreasonable structure, etc. These high energy conformers can cause the simulations to terminate. In order to remove these bad structures to get the stable conformers and normal simulations run, energy minimization is performed before normal simulations.<sup>40-42</sup>

Steepest descent algorithm is being used widely to energy-minimize the initial configuration. It can remove some molecular overlap, bad contacts and adjust bond lengths and angle to suitable values to create a relatively low energy initial conformation. The basic idea of steepest descent is to find the net force in the systems based on the potential energy, then moving in the direction of the force, just like walking straight down the hill in a geographical contour to find the conformer

with lowest energy. The direction is determined by the maximum force between atoms. So from the initial configuration, the system can quickly get rid out of the conformer with highest energy. Each step, the new maximum force will be determined and the searching direction is changing with the maximum force, by repeating the process, the maximum force is becoming smaller and smaller until a preset threshold is achieved or close to zero, then the configuration at this moment can be used as the starting configuration for next normal MD simulations.<sup>43, 44</sup>

## **2.5 Software to perform MD simulations: GROMACS**

GRONingen MACHine for Chemical Simulations (GROMACS) is a computational molecular dynamics software package that is being used widely in the world.<sup>45</sup> It was developed by Biophysical Chemistry department of University of Groningen in 1991, and now it is maintained by global universities and research centers. GROMACS is free and open-source software and one of the most popular package for performing MD simulations for biological systems, such as proteins, lipids, solvents, DNA/RNA and polymers.<sup>46</sup> After years of development and update, a rich set of calculation types, preparation and analysis tools are provided by GROMACS.<sup>44, 47</sup> After version 5, it is updated with several new and enhanced parallelization algorithms to significantly improved computational efficiency.<sup>48</sup> Moreover, Martini force field, which was developed in the same university, uses the GROMACS infrastructure to implement coarse-grained physics models that can reach much longer time scale and larger system size.

## REFERENCE

1. Allen, M. P., Introduction to molecular dynamics simulation. *Computational soft matter: from synthetic polymers to proteins* **2004**, 23, 1-28.
2. Balescu, R., Equilibrium and nonequilibrium statistical mechanics. *NASA STI/Recon Technical Report A* **1975**, 76.
3. Karplus, M.; Petsko, G. A., Molecular dynamics simulations in biology. *Nature* **1990**, 347 (6294), 631.
4. Haile, J. M., *Molecular dynamics simulation: elementary methods*. Wiley New York: 1992; Vol. 1.
5. Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F., A generalized reaction field method for molecular dynamics simulations. *The Journal of chemical physics* **1995**, 102 (13), 5451-5459.
6. Maple, J. R.; Dinur, U.; Hagler, A. T., Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proceedings of the National Academy of Sciences* **1988**, 85 (15), 5350-5354.
7. Berendsen, H. J.; van der Spoel, D.; van Drunen, R., GROMACS: a message-passing parallel molecular dynamics implementation. *Computer physics communications* **1995**, 91 (1-3), 43-56.
8. Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K., Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molecular Simulation* **1991**, 6 (1-3), 121-142.
9. Silva Jr, R.; Plastino, A.; Lima, J., A Maxwellian path to the q-nonextensive velocity distribution function. *Physics Letters A* **1998**, 249 (5-6), 401-408.



10. Leach, A. R.; Leach, A., *Molecular modelling: principles and applications*. Pearson education: 2001.
11. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **1996**, *118* (45), 11225-11236.
12. Arnautova, Y. A.; Jagielska, A.; Scheraga, H. A., A new force field (ECEPP-05) for peptides, proteins, and organic molecules. *J. Chem. Phys* **2006**, *110* (10), 5025-5044.
13. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of computational chemistry* **2004**, *25* (9), 1157-1174.
14. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006**, *65* (3), 712-725.
15. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D., CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *Journal of Computational Chemistry* **2010**, *31* (4), 671-690.
16. Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F., A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry* **2004**, *25* (13), 1656-1676.
17. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H., The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* **2007**, *111* (27), 7812-7824.

18. Ponder, J. W.; Case, D. A., Force fields for protein simulations. *Protein Simulations* **2003**, 66, 27-+.
19. Gerben, S. R.; Lemkul, J. A.; Brown, A. M.; Bevan, D. R., Comparing atomistic molecular mechanics force fields for a difficult target: a case study on the Alzheimer's amyloid beta-peptide. *Journal of Biomolecular Structure & Dynamics* **2014**, 32 (11), 1817-1832.
20. Smith, M. D.; Rao, J. S.; Segelken, E.; Cruz, L., Force-Field Induced Bias in the Structure of A beta(21-30): A Comparison of OPLS, AMBER, CHARMM, and GROMOS Force Fields. *Journal of Chemical Information and Modeling* **2015**, 55 (12), 2587-2595.
21. Secrier, M.; Schneider, R., Visualizing time-related data in biology, a review. *Briefings in Bioinformatics* **2014**, 15 (5), 771-782.
22. Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J., The MARTINI coarse-grained force field: extension to proteins. *Journal of chemical theory and computation* **2008**, 4 (5), 819-834.
23. Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J., Polarizable water model for the coarse-grained MARTINI force field. *PLoS computational biology* **2010**, 6 (6), e1000810.
24. de Jong, D. H.; Singh, G.; Bennett, W. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J., Improved parameters for the martini coarse-grained protein force field. *Journal of Chemical Theory and Computation* **2012**, 9 (1), 687-697.
25. López, C. A.; Rzepiela, A. J.; De Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J., Martini coarse-grained force field: extension to carbohydrates. *Journal of Chemical Theory and Computation* **2009**, 5 (12), 3195-3210.

26. López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J., Martini force field parameters for glycolipids. *Journal of chemical theory and computation* **2013**, 9 (3), 1694-1708.
27. Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J., Martini coarse-grained force field: extension to DNA. *Journal of chemical theory and computation* **2015**, 11 (8), 3932-3945.
28. Ma, H.; Irudayanathan, F. J.; Jiang, W.; Nangia, S., Simulating Gram-negative bacterial outer membrane: a coarse grain model. *J. Chem. Phys* **2015**, 119 (46), 14668-14682.
29. Ma, H.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan, A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S., Modeling diversity in structures of bacterial outer membrane lipids. *Journal of chemical theory and computation* **2017**, 13 (2), 811-824.
30. Ma, H.; Khan, A.; Nangia, S., Dynamics of OmpF trimer formation in the bacterial outer membrane of Escherichia coli. *Langmuir* **2017**, 34 (19), 5623-5634.
31. Ibsen, K. N.; Ma, H. L.; Banerjee, A.; Tanner, E. E. L.; Nangia, S.; Mitragotri, S., Mechanism of Antibacterial Activity of Choline-Based Ionic Liquids (CAGE). *Acs Biomaterials Science & Engineering* **2018**, 4 (7), 2370-2379.
32. Nose, S., A MOLECULAR-DYNAMICS METHOD FOR SIMULATIONS IN THE CANONICAL ENSEMBLE. *Molecular Physics* **1984**, 52 (2), 255-268.
33. Evans, D. J.; Morriss, G. P., THE ISOTHERMAL ISOBARIC MOLECULAR-DYNAMICS ENSEMBLE. *Physics Letters A* **1983**, 98 (8-9), 433-436.
34. Hunenberger, P., Thermostat algorithms for molecular dynamics simulations. In *Advanced Computer Simulation Approaches for Soft Matter Sciences I*, Holm, C.; Kremer, K., Eds. 2005; Vol. 173, pp 105-147.

35. Martyna, G. J.; Tobias, D. J.; Klein, M. L., CONSTANT-PRESSURE MOLECULAR-DYNAMICS ALGORITHMS. *Journal of Chemical Physics* **1994**, *101* (5), 4177-4189.
36. Martyna, G. J.; Klein, M. L.; Tuckerman, M., NOSE-HOOVER CHAINS - THE CANONICAL ENSEMBLE VIA CONTINUOUS DYNAMICS. *Journal of Chemical Physics* **1992**, *97* (4), 2635-2643.
37. Melchionna, S.; Ciccotti, G.; Holian, B. L., HOOVER NPT DYNAMICS FOR SYSTEMS VARYING IN SHAPE AND SIZE. *Molecular Physics* **1993**, *78* (3), 533-544.
38. Andersen, H. C., MOLECULAR-DYNAMICS SIMULATIONS AT CONSTANT PRESSURE AND-OR TEMPERATURE. *Journal of Chemical Physics* **1980**, *72* (4), 2384-2393.
39. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., MOLECULAR-DYNAMICS WITH COUPLING TO AN EXTERNAL BATH. *Journal of Chemical Physics* **1984**, *81* (8), 3684-3690.
40. Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D., ITERATIVE MINIMIZATION TECHNIQUES FOR ABINITIO TOTAL-ENERGY CALCULATIONS - MOLECULAR-DYNAMICS AND CONJUGATE GRADIENTS. *Reviews of Modern Physics* **1992**, *64* (4), 1045-1097.
41. Jorgensen, W. L.; Tiradorives, J., THE OPLS POTENTIAL FUNCTIONS FOR PROTEINS - ENERGY MINIMIZATIONS FOR CRYSTALS OF CYCLIC-PEPTIDES AND CRAMBIN. *Journal of the American Chemical Society* **1988**, *110* (6), 1657-1666.
42. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., CHARMM - A PROGRAM FOR MACROMOLECULAR ENERGY, MINIMIZATION, AND DYNAMICS CALCULATIONS. *Journal of Computational Chemistry* **1983**, *4* (2), 187-217.

43. Armijo, L., MINIMIZATION OF FUNCTIONS HAVING LIPSCHITZ CONTINUOUS FIRST PARTIAL DERIVATIVES. *Pacific Journal of Mathematics* **1966**, 16 (1), 1-&.
44. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation* **2008**, 4 (3), 435-447.
45. Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R., GROMACS - A MESSAGE-PASSING PARALLEL MOLECULAR-DYNAMICS IMPLEMENTATION. *Computer Physics Communications* **1995**, 91 (1-3), 43-56.
46. Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* **2005**, 26 (16), 1701-1718.
47. Lindahl, E.; Hess, B.; van der Spoel, D., GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling* **2001**, 7 (8), 306-317.
48. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, 1-2, 19-25.

## CHAPTER 3

### Modeling diversity in structures of bacterial outer membrane lipids

---

Ma, H.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan, A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S., Modeling diversity in structures of bacterial outer membrane lipids. *Journal of chemical theory and computation* **2017**, *13* (2), 811-824.

### 3.1 ABSTRACT

Lipopolysaccharides (LPS) are vital components of the outer membrane of Gram-negative bacteria, and they act as extremely strong stimulators of innate immunity in diverse eukaryotic species. The primary immunostimulatory center of the LPS molecule is lipid A—a disaccharide-bound lipophilic domain. Considering the broad diversity in bacterial species, there are variations in the lipid A structure and their immunogenic potency. In this work, we model the lipid A structures of eight commensal or human pathogenic bacterial species: *Salmonella minnesota*, *Neisseria meningitidis*, *Porphyromonas gingivalis*, *Campylobacter jejuni*, *Bordetella pertussis*, *Helicobacter pylori*, *Bacteroides fragilis*, and *Chlamydia trachomatis*. The membrane properties of lipid A from these bacterial species were characterized and compared using molecular simulations. Molecular and structural insights provided reveal the diversity of in bacterial outer membrane lipids and their contribution to human disease and immunity.

### 3.2 INTRODUCTION

Gram-negative bacteria have evolved to protect themselves from hostile environments by developing a protective outer membrane. Primary component of the outer membrane are the lipopolysaccharide (LPS) macromolecules that due to their unique chemical structure provide a negatively charged envelope around the bacterial cell.<sup>1</sup> The LPS has three distinct domains—lipid A, core oligosaccharide, and O-antigen polysaccharide. Although all three domains have integral roles in the outer membrane, the amphiphilic lipid A domain plays a key role in anchoring the LPS to the membrane via its hydrophobic interactions.<sup>2</sup>

Additionally, lipid A is a well-established endotoxin that stimulates innate immunity in diverse eukaryotic species.<sup>3</sup> Lipid A is highly conserved among bacterial species, and due to its distinctive molecular structure it is recognized as a pathogen-associated molecule by Toll-like receptor 4/myeloid differentiation factor 2 (TLR4/MD2) present on host immune cells.<sup>4</sup> As a response to lipid A, the host cells secrete pro-inflammatory cytokines to neutralize the bacteria and their endotoxic effects.

Structurally, lipid A molecule consists of a hydrophilic 1,4'-bisphosphorylated disaccharide head and variable numbers of saturated fatty acid tails (Figure 3-1). The head unit consists of hydrophilic  $\beta$  (1 $\rightarrow$ 6)-linked D-glucosamine disaccharide head group with  $\alpha$  phosphate group at position 1 of the proximal, reducing glucosamine residue (GlcN I) and an ester bound phosphate group at position 4' of the distal, non-reducing glucosamine residue (GlcN II). The hydrophobic tails comprise of four primary (R)-3-hydroxyacyl residues (labeled A, B, C and D) at the positions 2 and 3 as well as 2' and 3' via amide or ester linkages, along with four secondary fatty acid



chains (A', B', C' and D'). The -R functional group (at position 6') is the binding site of core oligosaccharide domain.

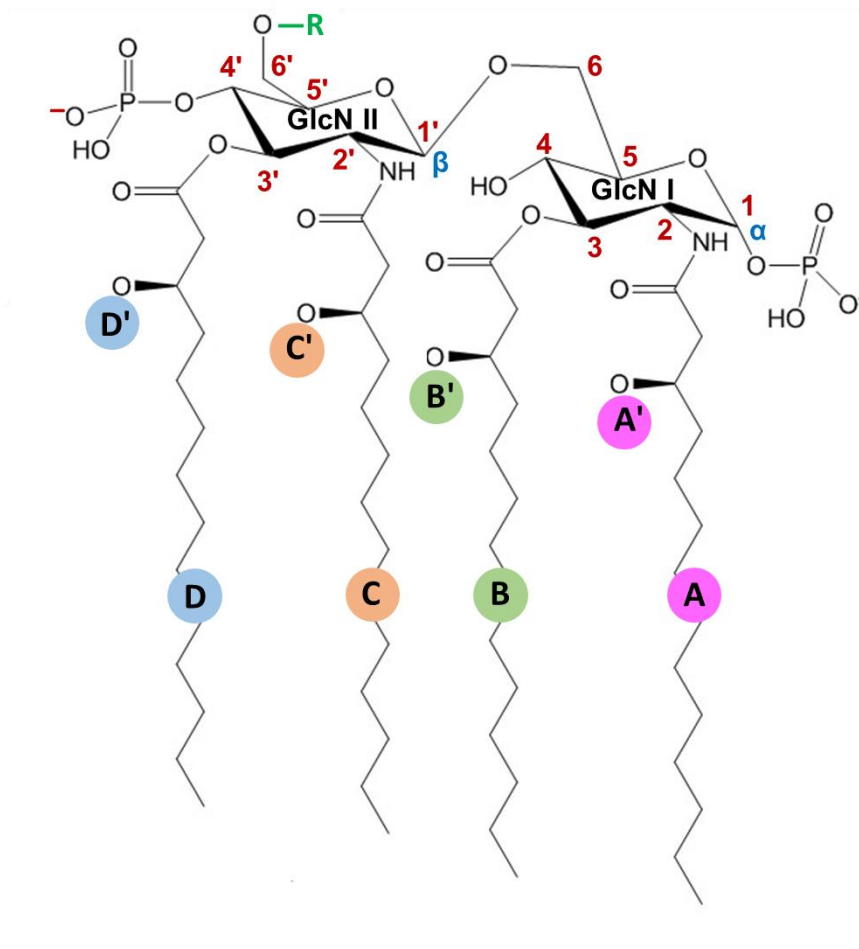


Figure 3-1. Lipid A template structure.

Despite the well-defined structural template, lipid A structures vary among bacterial species. Furthermore, to evade detection by the host immune system, bacteria undergo subtle modifications to alter their quintessential primary lipid A template—in terms of glucosamine head group, degree of phosphorylation, presence of phosphate substituents, as well as the nature, number, location, and length of acyl chains.<sup>5,6</sup> Often these structural modifications are employed as an active response to changing environmental chemical stresses.<sup>7, 8</sup> The structural modifications directly

affect pathogenesis by changing outer membrane permeability and promoting resistance to antimicrobial peptides. There is therefore, a need to understand structure property relationships between the lipid A structures and the properties they confer to the outer membrane of a bacterial species.

Experimental characterization of LPS remains challenging due to the complexity and heterogeneity of the bacterial membrane. The isolation of a LPS macromolecule is non-trivial considering the amphiphilic nature of lipid A that causes the formation of micelles. Determination of high-resolution LPS structure requires iterative extractions followed by refinement and fragmentation. Such advances in extraction methods coupled with improved characterization techniques such as mass spectrometry (MS), matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) have been invaluable. In order to expand beyond experimentally determined static structural properties of these lipids, complementary computational approaches are now being employed to assess the dynamical and thermodynamical properties of the bacterial membranes.

Molecular simulations have become an indispensable tool to understand both the dynamic and nanoscale organization of bacterial membrane structures. Although multiple computational techniques have been employed to investigate these membranes, coarse-grain representation provides an equitable balance between (1) the complexity and chemical specificity of membrane lipids and (2) the length and timescales required to characterize these systems.<sup>9-11</sup> In our previous work, we adopted a multiscale approach to bridge atomistic and coarse-grained representations by developing force field parameters for *Pseudomonas aeruginosa* LPS macromolecule.<sup>9</sup>

In this work, we extend the MARTINI force field parameters<sup>10</sup> for a library of eight commensal or human pathogenic Gram-negative bacteria species: *Helicobacter pylori*, *Porphyromonas gingivalis*, *Bacteroides fragilis*, *Bordetella pertussis*, *Chlamydia trachomatis*, *Campylobacter jejuni*, *Neisseria meningitidis*, and *Salmonella minnesota*. After 80 independent simulations and close to 150  $\mu$ s of total simulation time, this library of representative bacterial lipids, provides an excellent example of the structure-property relationship of lipid A structural modifications and the impact they have on the bacterial outer membrane properties. The results highlight the role of acyl chain lengths, number of chains, and phosphorylation state in regulating the phase transition temperature of the membrane, along with the role of membrane composition and charge of the counterions on membrane permeability. Prior to discussing the results, the background information on the eight bacterial species, their differences in preferred habitat, and lipid A structure (Table 3-1) are briefly discussed.

Table 3-1 Summary of chemical structure of lipid A in various species of Gram-negative bacteria. Labels A, A', B, B', C, C', D, and D' correspond to acyl chains shown in Figure 3-1.

Label P denotes the total number of phosphates.

Species	Saturated acyl chain lengths								Total chains	P
	A	A'	B	B'	C	C'	D	D'		
<i>H. pylori</i>	18	–	16	–	18	18	–	–	4	1
<i>P. gingivalis</i>	17	–	16	–	17	16	15	–	5	1
<i>B. fragilis</i>	16	–	15	–	17	15	16	–	5	1
<i>B. pertussis</i>	14	–	10	–	14	14	14	–	5	2
<i>C. trachomatis</i>	20	–	14-15	–	20	18-21	14-16	–	5	2
<i>C. jejuni</i>	14	–	14	–	14	16	14	16	6	2
<i>N. meningitidis</i>	14	12	12	–	14	12	12	–	6	4
<i>S. minnesota</i>	14	16	14	–	14	12	14	14	7	2

### 3.3 BACKGROUND

#### *Helicobacter pylori*

*H. pylori* are spiral, rod-shaped bacteria that live in the upper gastrointestinal tract. It is associated with a variety of gastrointestinal diseases such as peptic ulcers, gastric adenocarcinoma and can lead to stomach cancer.<sup>12</sup> *H. pylori* infection is found in over 50 percent of the world's population, especially among the young, and is transmitted through direct human contact. *H. pylori* can often be a lifelong infection in many of its hosts. The outer cell membrane of *H. Pylori* is similar to that of other Gram-negative bacteria. The temperature range supporting *H. Pylori*'s growth is 307 K to 313 K, with the optimum temperature being 310 K, which is the average temperature of the human body. The chemical structure of its lipid A has glucosamine  $\beta$ -(1-6) disaccharide with phosphate at position 1 and four acyl chains (Figure 3-2). The acyl groups are (*R*)-3-hydroxyoctadecanoic acid, (*R*)-3-hydroxyhexadecanoic acid, and (*R*)-3-(octadecanoyloxy)octadecanoic acid at the 2-, 3- and 2'-positions, respectively.<sup>13</sup>

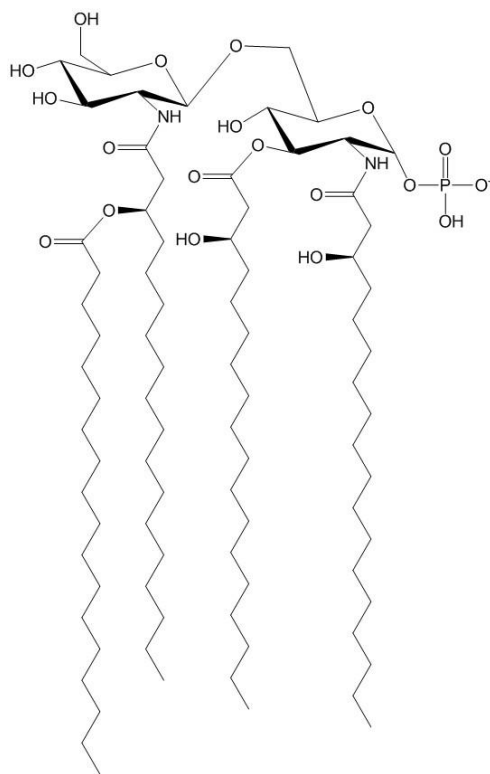


Figure 3-2. Atomistic structure of *Helicobacter pylori* Lipid A.

### *Porphyromonas gingivalis*

*P. gingivalis* is a non-motile, Gram-negative, endotoxic, anaerobic bacillus of the phylum *Bacteroidetes* found in gingival tissue and in atheromatous plaque and thrives best at 310 K. It is a suspected periodontal pathogen because it produces collagenase; however, about 25% of people without periodontitis test positive for *P. gingivalis*, while 21% of patients with periodontitis test negative for *P. gingivalis*.<sup>61</sup> The chemical structure of its lipid A comprises a hydrophilic  $\beta$ -(1,6)-linked D-glucosamine disaccharide head that is monophosphorylated at position 1, and the hydrophobic N- and/or O-acylation at positions 2, 3, 2', and 3' (Figure 3-3).<sup>15</sup>

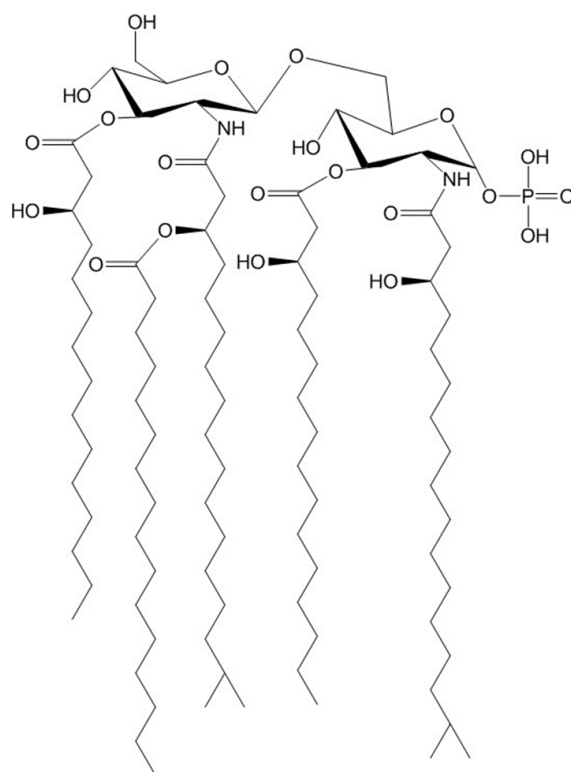


Figure 3-3. Atomistic structure of *Porphyromonas gingivalis* Lipid A.

### *Bacteroides fragilis*

The anaerobic bacteria *B. fragilis* is part of the normal microflora of the human large intestine. It is the most frequent cause of abdominal and wound infection post-surgical procedures of the gastrointestinal or urogenital tract. *B. fragilis* is a unique enterobacteria with low endotoxicity, primarily attributed due its monophosphorylated lipid A that has five acyl residues, which are relatively long chains, each with 15-17 carbon atoms (Figure 3-4). The (R)-3-hydroxyhexadecanoic acid and (R)-3-hydroxypentadecanoic acid residues are present at the positions 3' and 3 of the distal GlcN and reducing GlcN groups, respectively. The amino group at 3' position carries (R)-3-(13-methyltetradecanoyloxy)-15-methylhexadecanoic acid and the pne at position 3 carries (R)-3-hydroxyhexadecanoic acid.<sup>16</sup>

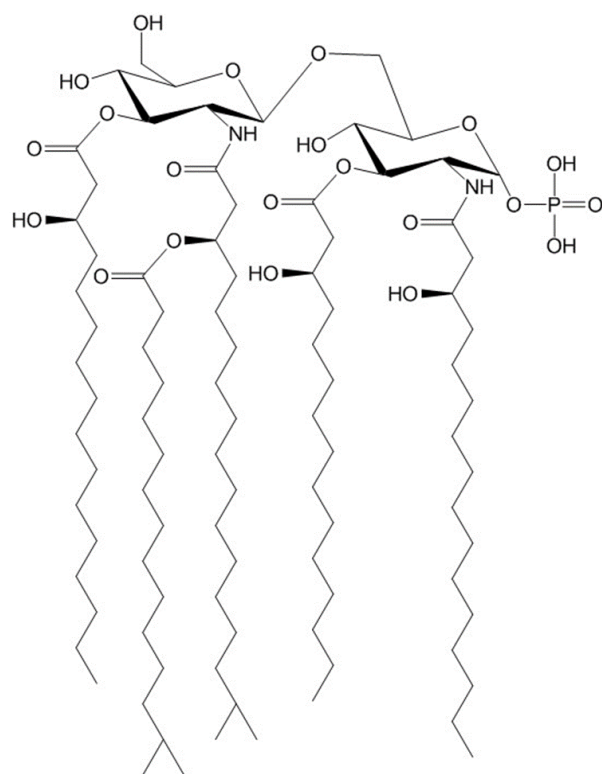


Figure 3-4. Atomistic structure of *Bacteroides fragilis* Lipid A.

### *Bordetella pertussis*

*B. pertussis* causes pertussis, a highly contagious respiratory infection commonly known as whooping cough because of the characteristic sound patients make when they inhale. Transmission between people most commonly occurs by coughing or sneezing. Its lipid A structure contains a common bisphosphorylated disaccharide head group with hydroxytetradecanoic acid in the amide as well at the 3' position (Figure 3-5).<sup>17</sup> The shorter acyl chains enable bacteria to escape the receptor signaling system.

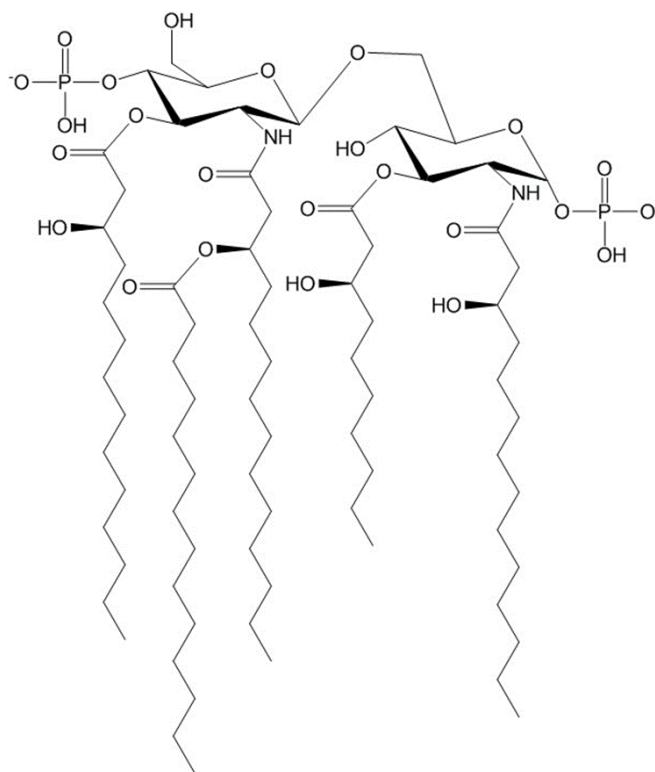


Figure 3-5. Atomistic structure of *Bordetella pertussis* Lipid A.

### *Chlamydia trachomatis*

*C. trachomatis* is the most common cause of sexually transmitted bacterial infection, with more than 90 million new cases annually worldwide.<sup>18</sup> *C. trachomatis* also is a cause of pelvic inflammatory disease in women, increases transmission of HIV, and is a significant cause of blindness in the developing world, where treatment is largely absent. Members of the *Chlamydiae* genus are obligate intracellular parasites, and *C. trachomatis* is specifically reliant on human cells to carry out its life cycle. The physiological effects of *C. trachomatis*, like all Gram-negative bacteria, are invoked by its lipid A component. Mass spectrometry shows that the LPS of *C. trachomatis* is composed mainly of a glucosamine disaccharide with five-fatty acid chains and one phosphate head (Figure 3-6). The long fatty acid chains of *C. trachomatis* (up to 21 carbons) are



anomalous to most Gram-negative bacteria lipid A components, and is thought to bring about its relatively low toxicity.<sup>19,20</sup>

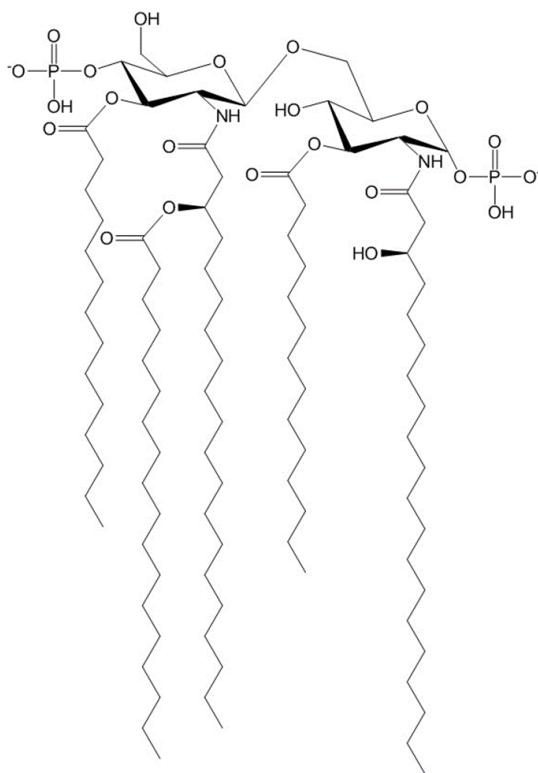


Figure 3-6. Atomistic structure of *Chlamydia trachomatis* Lipid A.

### *Campylobacter jejuni*

*C. jejuni* is a microaerobic strain of proteobacteria with a helical shape. It is primarily responsible for food borne bacterial gastroenteritis.<sup>21</sup> *C. jejuni* is often found in animal feces and is transmitted easily between animals and humans. Its capacity to form a biofilm increases the survival of *C. jejuni* under detrimental conditions; when in a biofilm, the bacteria is one-thousand times more resistant to disinfectants.<sup>22</sup> The structure of *C. jejuni* lipid A is similar to others studies in this work, except that one of the glucosamine residues of the lipid A backbone is replaced by of a GlcN3N monosaccharide, a phosphorylated 2,3 diamino-2,3-dideoxy-D-glucose (GlcN3N) disaccharide (Figure 3-7).

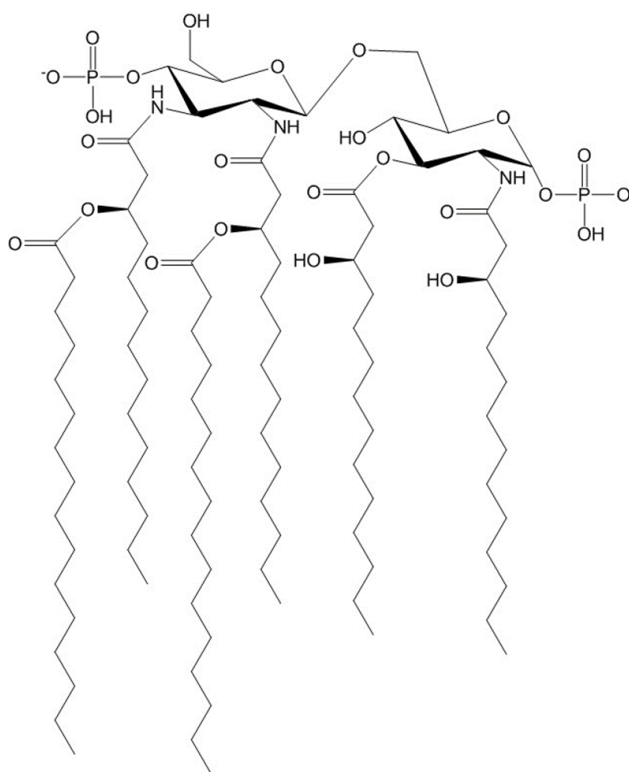


Figure 3-7. Atomistic structure of *Campylobacter jejuni* Lipid A.

### *Neisseria meningitidis*

*N. meningitidis* is a leading cause of bacterial meningitis and sepsis worldwide.<sup>23</sup> Meningococcal LPS has a bisphosphorylated disaccharide head group with 12:0(3-OH) acyl chains bound to each of the two hydroxyl groups at positions 3 and 3', and 14:0(3-OH) acyl chains linked to the amino groups at positions 2 and 2', and the hydroxyl groups of the amide-linked chains acylated by 12:0 carbon tails.<sup>24</sup> Additionally, O-phosphorylethanolamine residues cap the phosphates at positions 1 and 4' (Figure 3-8).

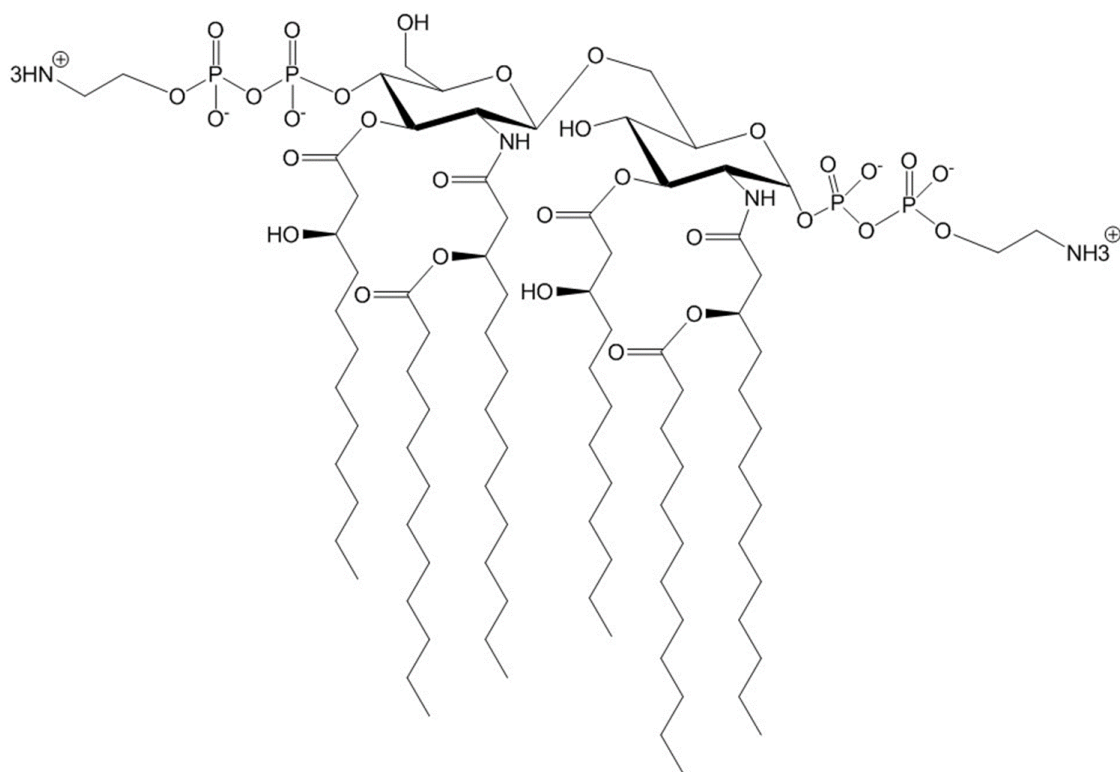


Figure 3-8. Atomistic structure of *Neisseria meningitidis* Lipid A.

### *Salmonella minnesota*

Typically *S. minnesota*, the second leading cause of intestinal infections, is transmitted through ingestion of contaminated food. *S. minnesota* infection commonly occurs in the intestinal tract and is associated with bloody diarrhea, abdominal cramps, and other related symptoms. Most *Salmonella* serotypes are able to grow and thrive in environments whose temperature falls between 280 K to 321 K. *S. minnesota* lipid A has a typical 1,4'-bisphosphorylated disaccharide head group with seven acyl chains that are 12–14 carbons in length. Position 2 and 3 have (R)-3-hydroxy fatty acids and 2' and 3' have (R)-3-acyloxyacyl residues. Additionally hexadecanoic acid and dodecanoic acid residues are on the (R)-3-hydroxytetradecanoic acid at positions 2 and 2', respectively (Figure 3-9).<sup>25</sup>

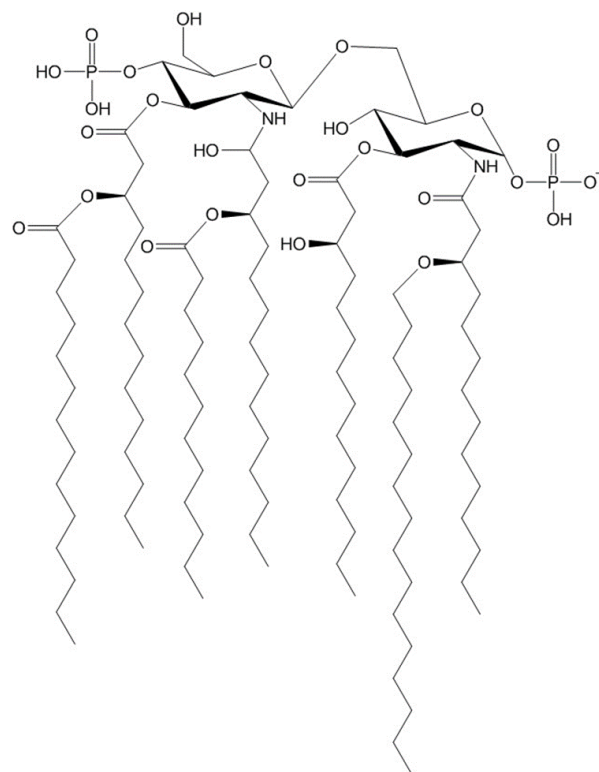


Figure 3-9. Atomistic structure of *Salmonella Minnesota* Lipid A.

### 3.4 METHODS

#### 3.4.1 Parameterization

The coarse-grain parameterization of the library of eight lipid A molecules is developed on the Martini many-to-one mapping approach,<sup>10</sup> and in most cases four heavy atoms are mapped into one bead. The structural similarities in the disaccharide head groups and the dissimilarities in the phosphorylation state and acyl chain patterns have been incorporated in the parameterization (see Table 3-1). The proximal reducing (GlcN I) and non-reducing (GlcN II) glucosamine residues were mapped individually to four beads with bead types ranging from P1–P4, based on the number of hydroxyl groups. The phosphates at positions 1 and 4' were assigned a Qa bead type with a unit negative charge. The beads linking the acyl carbon chains via amide or ester linkages at positions 2, 3, 2' and 3' were assigned Na bead type. For *N. meningitidis* lipid A, the additional  $\text{NH}_3^+$  groups

linked to the phosphates were assigned Qd bead type. The acyl chain beads were assigned C1 bead type. Figure 3-10 shows the CG mapping of *N. meningitidis* with bead assignments for the disaccharide head group and six acyl chains. For the remaining lipids in the library, the CG mapping of the acyl chains is also depicted in Figure 3-10.

Monovalent ( $\text{Na}^+$ ) or divalent ( $\text{Ca}^{2+}$ ) counterions were used to make the systems electrically neutral. This ion parameterization accounts for the first hydration shell around the ion, and both ions were assigned the Qd bead type. As in our previous work, no additional parameterization of  $\text{Ca}^{2+}$  ions was performed, and the only difference in  $\text{Na}^+$  and  $\text{Ca}^{2+}$  was their net integral charge.

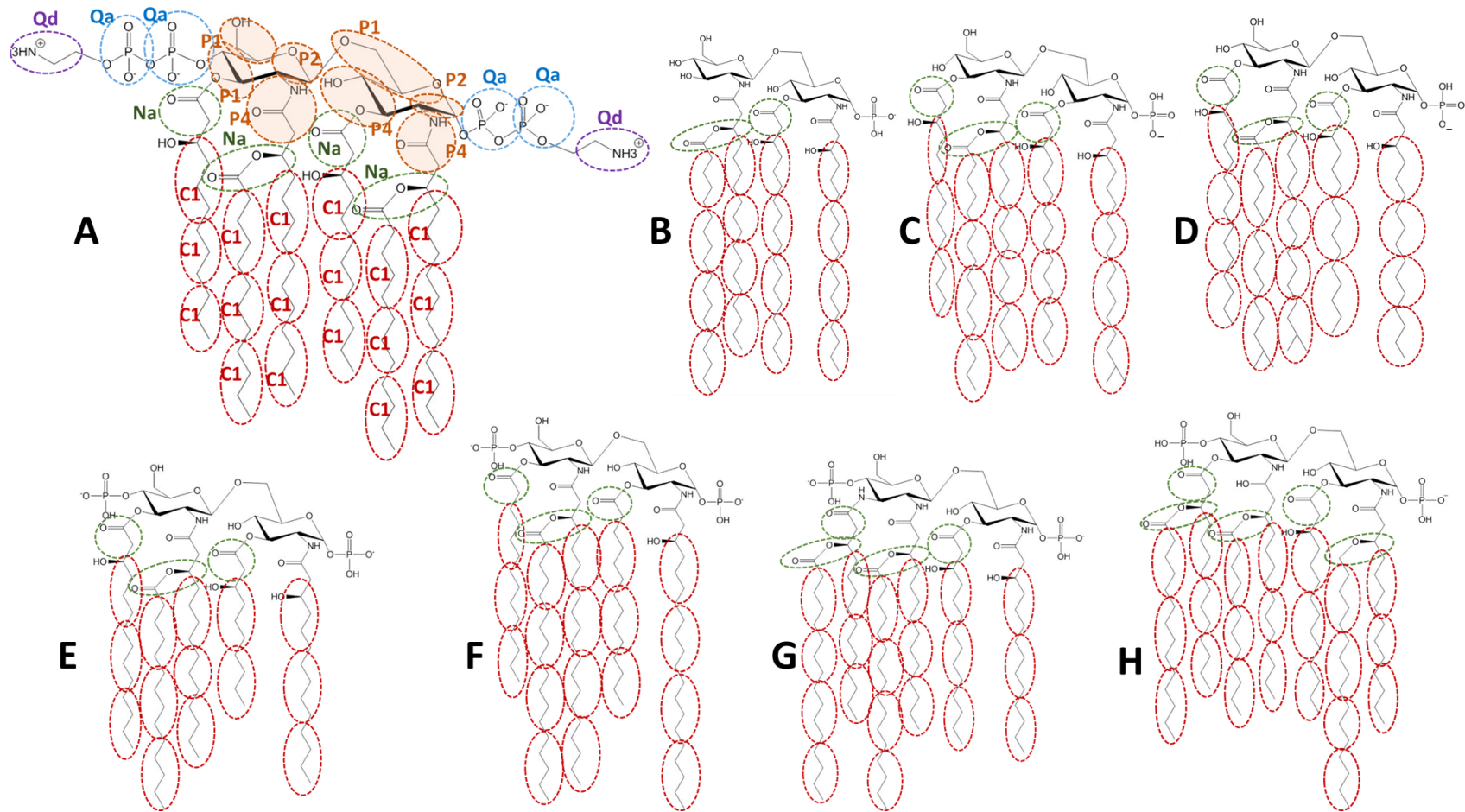


Figure 3-10. Coarse-grained mapping scheme for lipid A tails of (A) *N. meningitidis*, (B) *H. pylori*, (C) *P. gingivalis*, (D) *B. fragilis*, (E) *B. pertussis*, (F) *C. trachomatis*, (G) *C. jejuni*, and (H) *S. minnesota*. The bead types are shown in bold.

### 3.4.2 Simulation and analysis details

Eight sets of simulations, which include variation of membrane composition, membrane size, solvent, counterions, and temperature were performed for each of the eight membrane systems (Tables 3-2 and 3-3). The simulations were performed using the molecular dynamics engine GROMACS, version 5.1.2. The workflow of the simulations involved the initial construction of membrane, energy minimization, short isothermal-isochoric (*NVT*) and isothermal-isobaric (*NPT*) equilibrations runs, and long-production *NPT* runs.

Table 3-2 System Details of the Membrane Simulations Involving DPPE in the Inner Leaflet<sup>a</sup>

set	species	inner	outer		water	ion	no. of waters	no. of ions	T (K)
		DPPE	lipid A	DPPE					
I <sup>b</sup>	<i>H. pylori</i>	224	72	8	W	Na <sup>+</sup>	4420	72	310
	<i>P. gingivalis</i> <sup>c</sup>	188	72	8	W	Na <sup>+</sup>	4419	72	310
	<i>B. fragilis</i>	187	72	7	W	Na <sup>+</sup>	4535	72	310
	<i>B. pertussis</i>	176	64	16	W	Na <sup>+</sup>	4484	128	310
	<i>C. trachomatis</i>	188	72	8	W	Na <sup>+</sup>	4341	144	310
	<i>C. jejuni</i> <sup>d</sup>	224	72	8	W	Na <sup>+</sup>	4341	144	310
	<i>N. meningitidis</i>	178	57	6	W	Na <sup>+</sup>	4546	114	310
	<i>S. minnesota</i>	260	72	8	W	Na <sup>+</sup>	4344	144	310
II	<i>H. pylori</i>	224	72	8	W	Na <sup>+</sup>	4420	72	315 → 360
	<i>P. gingivalis</i>	188	72	8	W	Na <sup>+</sup>	4419	72	275 → 360
	<i>B. fragilis</i>	187	72	7	W	Na <sup>+</sup>	4535	72	275 → 360
	<i>B. pertussis</i>	176	64	16	W	Na <sup>+</sup>	4484	128	275 → 360
	<i>C. trachomatis</i>	188	72	8	W	Na <sup>+</sup>	4341	144	275 → 360
	<i>C. jejuni</i>	224	72	8	W	Na <sup>+</sup>	4341	144	275 → 360
	<i>N. meningitidis</i>	178	57	6	W	Na <sup>+</sup>	4546	114	275 → 360
	<i>S. minnesota</i>	260	72	8	W	Na <sup>+</sup>	4344	144	275 → 360
III	<i>H. pylori</i>	224	72	8	W	Ca <sup>2+</sup>	4420	36	275 → 360
	<i>P. gingivalis</i>	188	72	8	W	Ca <sup>2+</sup>	4419	36	275 → 360
	<i>B. fragilis</i>	187	72	7	W	Ca <sup>2+</sup>	4535	36	275 → 360
	<i>B. pertussis</i>	176	64	16	W	Ca <sup>2+</sup>	4484	64	275 → 360
	<i>C. trachomatis</i>	188	72	8	W	Ca <sup>2+</sup>	4341	72	275 → 360
	<i>C. jejuni</i>	224	72	8	W	Ca <sup>2+</sup>	4341	72	275 → 360
	<i>N. meningitidis</i>	178	57	6	W	Ca <sup>2+</sup>	4546	57	275 → 360
	<i>S. minnesota</i>	260	72	8	W	Ca <sup>2+</sup>	4344	72	275 → 360
IV	<i>H. pylori</i>	152	72	8	PW	Na <sup>+</sup>	4429	72	310
	<i>P. gingivalis</i>	188	72	8	PW	Na <sup>+</sup>	4407	72	310
	<i>B. fragilis</i>	264	96	24	PW	Na <sup>+</sup>	8937	113	310
	<i>B. pertussis</i>	188	72	8	PW	Na <sup>+</sup>	4450	144	310
	<i>C. trachomatis</i>	188	72	8	PW	Na <sup>+</sup>	4339	144	310
	<i>C. jejuni</i>	224	72	8	PW	Na <sup>+</sup>	4340	144	310
	<i>N. meningitidis</i>	177	57	6	PW	Na <sup>+</sup>	4544	114	310
	<i>S. minnesota</i>	260	72	8	PW	Na <sup>+</sup>	4343	144	310
V	<i>H. pylori</i>	152	72	8	PW	Ca <sup>2+</sup>	4429	36	310
	<i>P. gingivalis</i>	188	72	8	PW	Ca <sup>2+</sup>	4407	36	310
	<i>B. fragilis</i>	264	96	24	PW	Ca <sup>2+</sup>	8937	56	310
	<i>B. pertussis</i>	188	72	8	PW	Ca <sup>2+</sup>	4450	72	310
	<i>C. trachomatis</i>	188	72	8	PW	Ca <sup>2+</sup>	4339	72	310
	<i>C. jejuni</i>	224	72	8	PW	Ca <sup>2+</sup>	4340	72	310
	<i>N. meningitidis</i>	177	57	6	PW	Ca <sup>2+</sup>	4544	57	310
	<i>S. minnesota</i>	177	57	6	PW	Ca <sup>2+</sup>	4544	57	310

<sup>a</sup>The simulations were performed in duplicate. <sup>b</sup>Set I was simulated for 10  $\mu$ s. <sup>c</sup>The simulation was repeated after addition of the 4'-phosphate and adjustment of the counterions. <sup>d</sup>The simulation was repeated after deletion of the 4'-phosphate and adjustment of the counterions.

Table 3-3 System Details of the Membrane Simulations Involving a Complex Composition of the Inner Leaflet<sup>a</sup>

set	species	inner			outer		water	ion	no. of waters	no. of ions	T (K)
		DPPE	POPG	CDL2	lipid A	DPPE					
VI <sup>b</sup>	<i>H. pylori</i>	118	33	16	87	9	PW	Ca <sup>2+</sup>	4318	76	310
	<i>P. gingivalis</i>	118	33	16	70	8	PW	Ca <sup>2+</sup>	4378	67	310
	<i>B. fragilis</i>	117	33	16	68	12	PW	Ca <sup>2+</sup>	4704	68	310
	<i>B. pertussis</i>	118	33	16	70	8	PW	Ca <sup>2+</sup>	4627	102	310
	<i>C. trachomatis</i>	118	33	16	70	8	PW	Ca <sup>2+</sup>	4283	103	310
	<i>C. jejuni</i>	137	39	19	69	7	PW	Ca <sup>2+</sup>	4261	110	310
	<i>N. meningitidis</i>	137	39	19	68	10	PW	Ca <sup>2+</sup>	4619	106	310
	<i>S. minnesota</i>	137	39	19	60	4	PW	Ca <sup>2+</sup>	4276	110	310
VII	<i>H. pylori</i>	118	33	16	87	9	PW	Ca <sup>2+</sup>	4318	76	275 → 330
	<i>P. gingivalis</i>	118	33	16	70	8	PW	Ca <sup>2+</sup>	4378	67	275 → 330
	<i>B. fragilis</i>	117	33	16	68	12	PW	Ca <sup>2+</sup>	4704	68	275 → 330
	<i>B. pertussis</i>	118	33	16	70	8	PW	Ca <sup>2+</sup>	4627	102	250 → 330
	<i>C. trachomatis</i>	118	33	16	70	8	PW	Ca <sup>2+</sup>	4283	103	250 → 330
	<i>C. jejuni</i>	137	39	19	69	7	PW	Ca <sup>2+</sup>	4261	110	275 → 330
	<i>N. meningitidis</i>	137	39	19	68	10	PW	Ca <sup>2+</sup>	4619	106	250 → 330
	<i>S. minnesota</i>	137	39	19	60	4	PW	Ca <sup>2+</sup>	4276	110	275 → 330

<sup>a</sup>The simulations were performed in duplicate. <sup>b</sup>Set VI was simulated for 10  $\mu$ s.

For each simulation run, the membrane was built using a python script, which is a locally modified enhanced version *insane*, a versatile membrane-building tool routinely used in constructing coarse-grained membranes. The library of eight bacterial lipids have been coded in the freely distributed *insane* script programmed in python. The workflow of the *insane* script was not changed from the published version. The command line syntax for building the membranes, and the associated topology files are provided in the Supporting Information.

The outer leaflet of the membrane is a mixture of lipid A and 1,2-dihexadecanoyl-sn-glycero-3-phosphoethanolamine (DPPE) in 9:1 ratio for all seven sets. For the inner leaflet either a pure DPPE (Sets I-V) or a mixture of DPPE, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), and Cardiolipin (CDL2) in the ratio 7:2:1 (Set VI-VII) was used. POPE and CDL2 lipids have  $-1$  and  $-2$  charge, respectively. In generating the membranes, the total number of acyl chains in the inner and outer leaflet were kept the same to avoid unphysical bending of the membrane. The membranes were solvated with either standard water (W) or polarizable (PW) Martini water



as specified (Table 3-2 and 3-3). All systems were made charge neutral by adding  $\text{Na}^+$  or  $\text{Ca}^{2+}$  counterions. Details of the solvent, number of ions, and membrane composition are provided. Periodic boundary conditions were applied in all three dimensions.

Energy minimization was performed using the steepest-decent algorithm with a 20 fs time-step until the maximum force on any bead was below the tolerance parameter of  $10 \text{ kJmol}^{-1}\text{nm}^{-1}$ . The *NVT* and *NPT* simulation runs were performed for 0.2  $\mu\text{s}$ . The production simulations were run for at least 2  $\mu\text{s}$  and up to 10  $\mu\text{s}$  in some cases (Table 3-2 and 3-3) with a 20 fs time-step. Semi-isotropic pressure coupling was used, and systems were maintained at 1 bar using the Berendsen barostat with time constant,  $\tau_p = 4.0 \text{ ps}$ . Temperature was maintained at 310 K by independently coupling the lipids and solvent to an external velocity rescaling thermostat with  $\tau_T = 1.0 \text{ ps}$ . The heating scans were performed for a wider temperature range, varying from 275–360 K (Table 2 and 3). The neighbor list was updated every 25 steps using a cutoff equal to 1.4 and 1.2 nm for short-range van der Waals and electrostatic cutoff, respectively. For simulations with polarizable water, PME was used for the long-range electrostatics, with an electrostatic screening constant  $\epsilon_r = 2.5$ .

The structural and dynamic properties of the membranes were compared by computing area per lipid, membrane thickness, density profiles, order parameters, and phase transition temperatures, and diffusion coefficients. The membrane microstructure was quantified by the average area per lipid ( $A_L$ ) and membrane thickness ( $D_M$ ), and hydrophobic thickness ( $D_H$ ). For bacterial membranes, the  $A_L$  values was computed by dividing the cross-sectional area of the membrane by the number of lipid A molecules in the leaflet. Standard utilities available in the GROMACS software suite were employed for all the quantities described above. To determine  $T_m$ , the characteristic phase-transition temperature values for the model systems, we performed annealing

simulations starting from well-equilibrated configurations to mimic phase transition conditions. The heating scans were performed over the 275–360 K temperature range with intermediate temperatures of 292, 309, 326, and 343 K and 2  $\mu$ s of simulation time.

## 3.5 RESULTS

### 3.5.1 Bonded parameters

The bond distances and bond angles analysis for all eight bacterial lipid A membranes was performed using an identical protocol. For the ease of comparison, the lipid A analysis was divided into two parts-the head group and acyl chains. Given that the proximal reducing (GlcN I) and non-reducing (GlcN II) glucosamine residues in the lipid A head group are the same or slightly different (in *C. jejuni*) in the eight bacterial lipids, the average bond distance frequency distribution is very similar (Figure 3-11). A unimodal frequency distribution of the bonded pairs centered at  $0.30 \pm 0.01$  nm shows that the bonded pairs in a saccharide head group range between 0.29-0.39 nm both in atomistic and CG simulations. A similar frequency distribution of the average internal angles also shows a unimodal distribution centered at  $75.2^\circ \pm 1.2^\circ$  for all eight lipid A membranes (Figure 3-12). These results are consistent with analysis reported earlier for *P. aeruginosa* CG parameterization.<sup>56</sup> Although, the similarity in the bond and angle distribution is expected, the results demonstrate the variations in the structures and the influence that phosphorylation state and acylation pattern on the overall properties of the membrane.

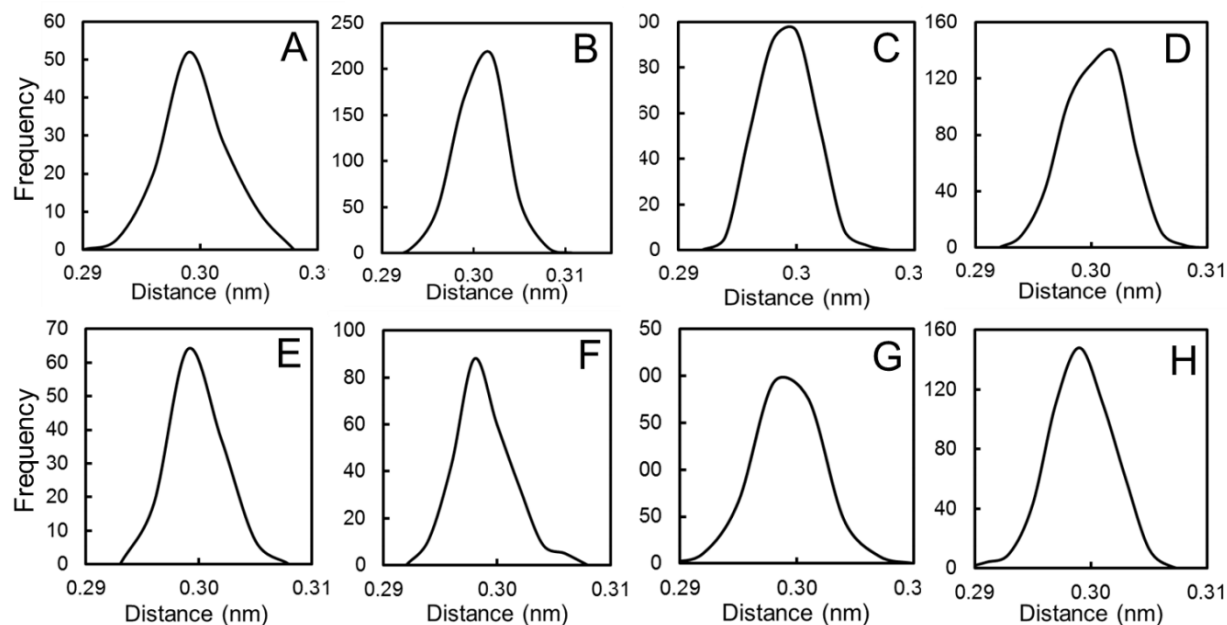


Figure 3-11. Average disaccharide head group bond distance frequency distribution for (A) *H. pylori* (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitidis*, and (H) *S. Minnesota*.

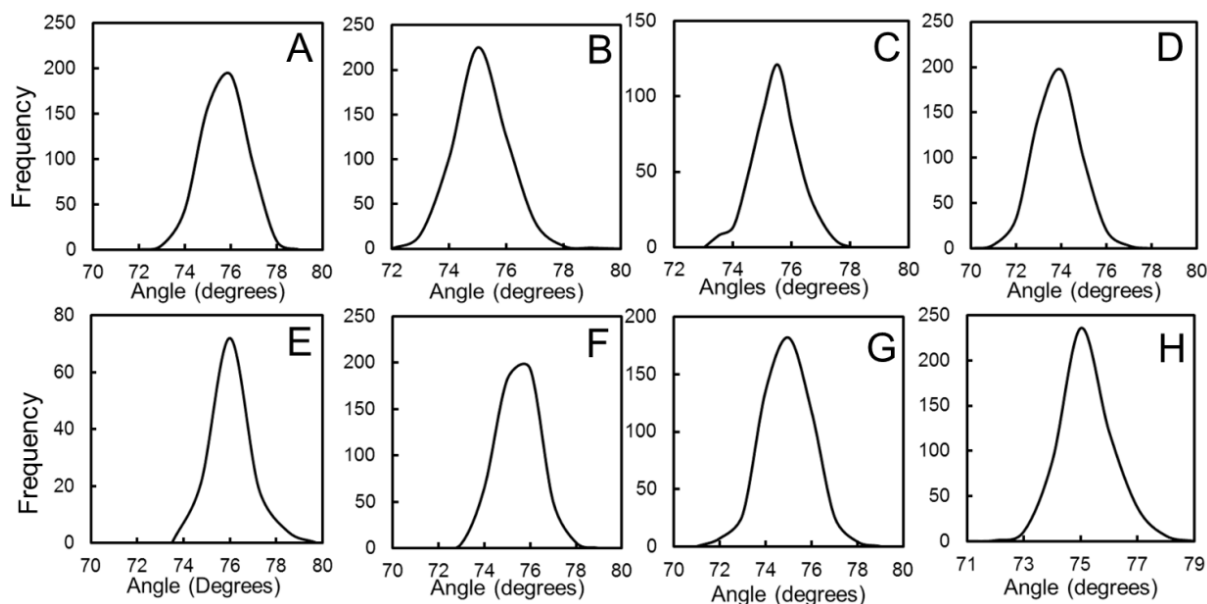


Figure 3-12. Average disaccharide head group angle frequency distribution for (A) *H. pylori* (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitidis*, and (H) *S. Minnesota*.

For the acyl chains in all the systems, the average bond distance distribution shows unimodal curve centered at  $0.442 \pm 0.012$  nm (Figure 3-13). This average bond distance is  $+0.142$  nm larger than the head group bond distance because unlike the tails, head group beads are smaller and do not always follow the 4-to-1 mapping prescription. The acyl chain bond angle distribution is unimodal for all the lipid A membranes, but the location of the peak depends on the specific bacterial lipid A structure (Figure 3-14). For example, lipid A structures that have 17-21 carbon acyls chains (*H. pylori*, *P. gingivalis*, *B. fragilis*, and *C. trachomatis*) have peaks centered at  $158 \pm 2^\circ$ , while structures with shorter 14-16 carbon acyl chains (*C. jejuni*, *N. meningitidis*, and *S. minnesota*) have peaks at  $151 \pm 1^\circ$ , and *B. pertussis* with shortest 10-14 carbon acyl chains has peak at  $145^\circ$ . Despite having exactly same bond angle parameters for the acyl beads, the variation in the average angle with the acyl chain length as significant implication on the membrane properties. It not only demonstrates

that the CG parameterization is able to capture the molecular differences in these lipid A structures, but it also validates that the membrane properties predicted by the force field are reliable.

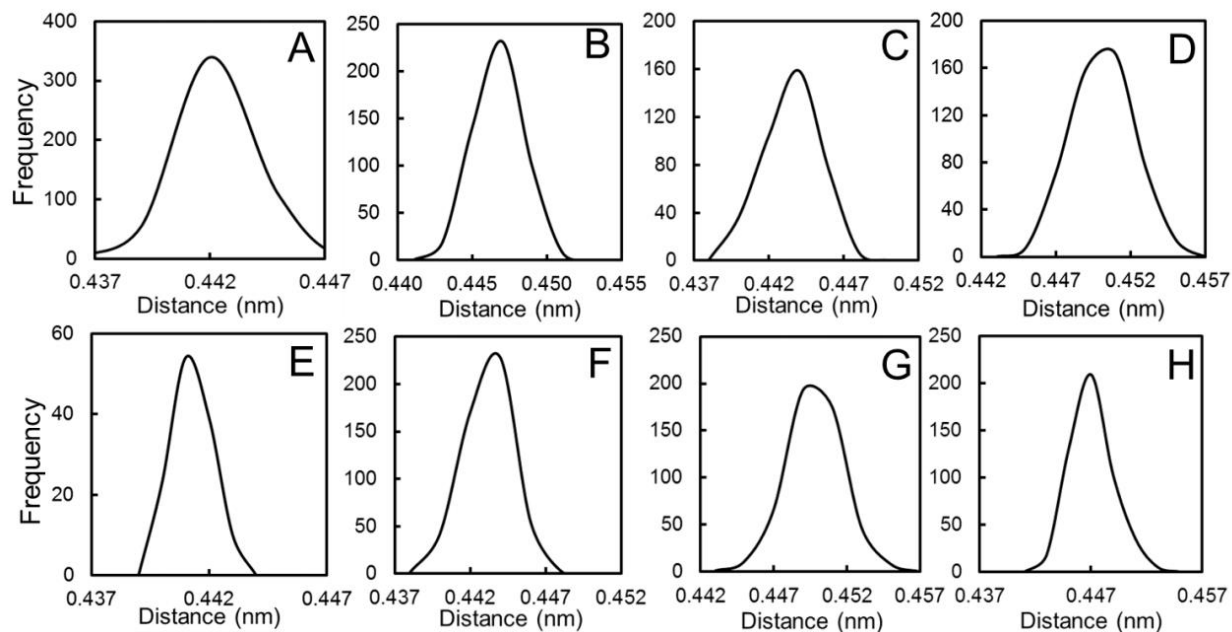


Figure 3-13. Average acyl chain bond distance frequency distributions for (A) *H. pylori*, (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitidis*, and (H) *S. minnesota*.

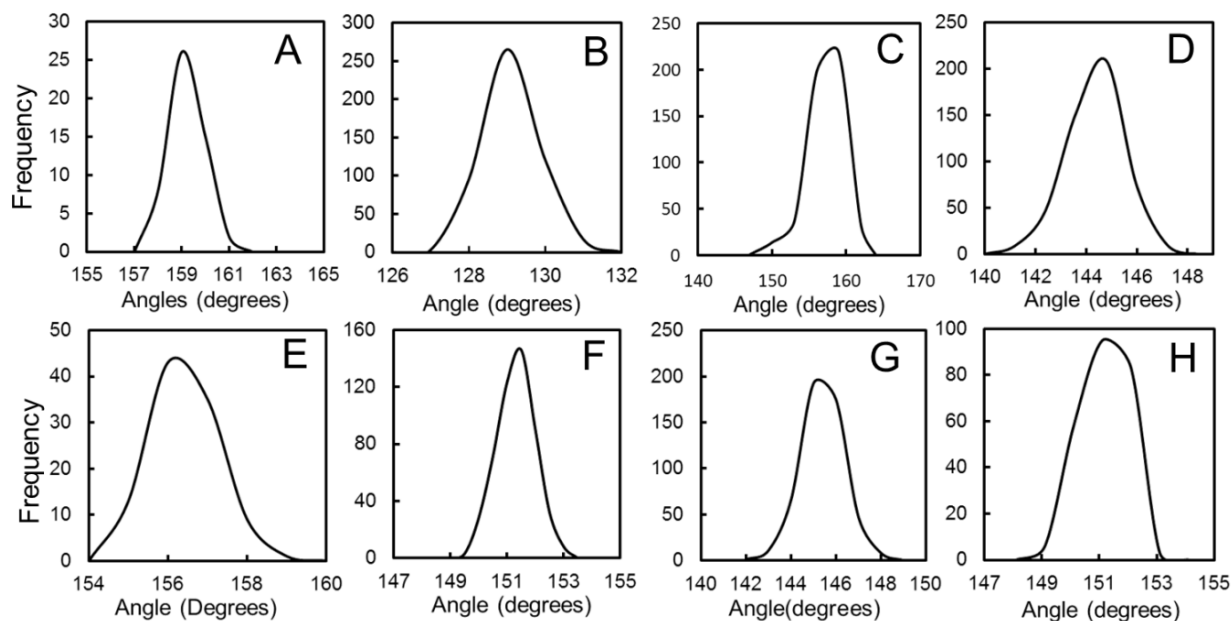


Figure 3-14. Average acyl angle frequency distribution for (A) *H. pylori* (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitides*, and (H) *S. Minnesota*.

The dihedral angles were not included for the disaccharide head group or the acyl chain beads. This choice was based on the earlier reports in the literature, where including of dihedral angle parameters required the use of an order of magnitude smaller time steps, not optimal for CG simulations. Despite the absence of explicit dihedral angle parameters, the average dihedral angle was computed for 2  $\mu$ s trajectory. In all eight membranes (Set II) the acyl chains are linear with average dihedral angle of  $180^\circ \pm 11^\circ$  or  $(0^\circ \pm 11)$  through the trajectory (Figure 3-15).

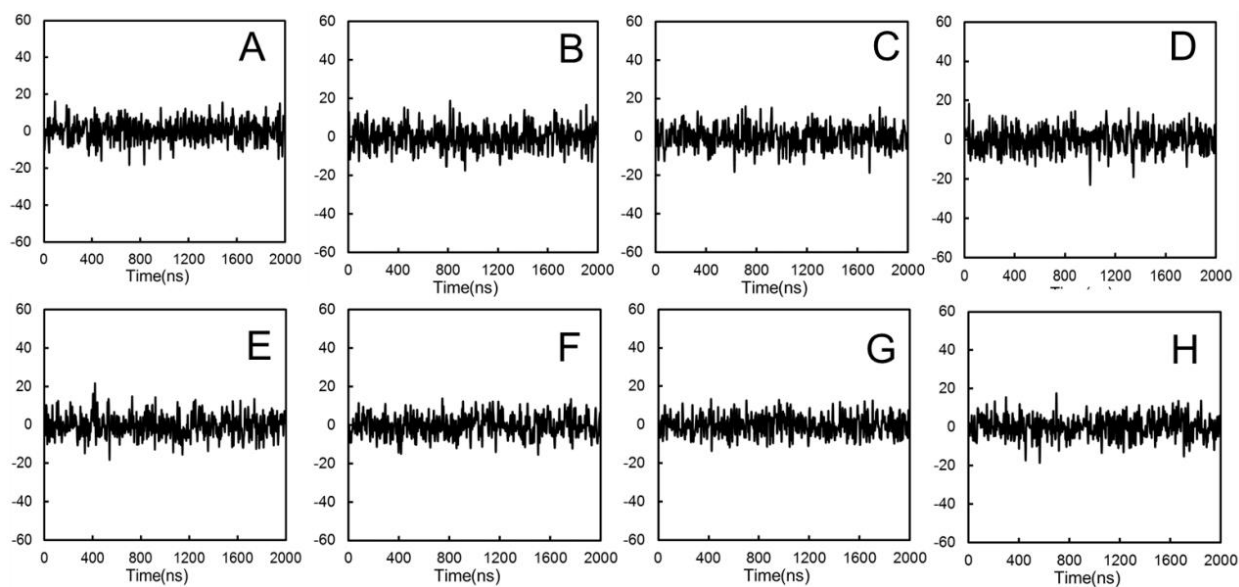


Figure 3-15. Dihedral angle fluctuations (degrees) as a function of simulation time for (A) *H. pylori* (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitides*, and (H) *S. Minnesota*.

### 3.5.2 Area per lipid ( $A_L$ ) and phase-transition temperature ( $T_m$ )

In general, the  $A_L$  for a lipid increases with increases in temperature as it acquires higher thermal energy, and if varied over a long-enough temperature range, the lipids undergo phase transition marked by a sharp increase in the  $A_L$  versus  $T$  plot. In this work, the variation in  $A_L$  for all eight membranes was computed over 275-360 K temperature range (Figure 3-16) to determine the phase transition melting temperature ( $T_m$ ). To determine the  $T_m$  more precisely, the change in  $A_L$  ( $\Delta A_L$ ) as a function of temperature was computed as a function of  $T$ , where the peak in the curve reflects a sharp change in the area per lipid over a small change in temperature for an individual membrane (Figure 3-17).

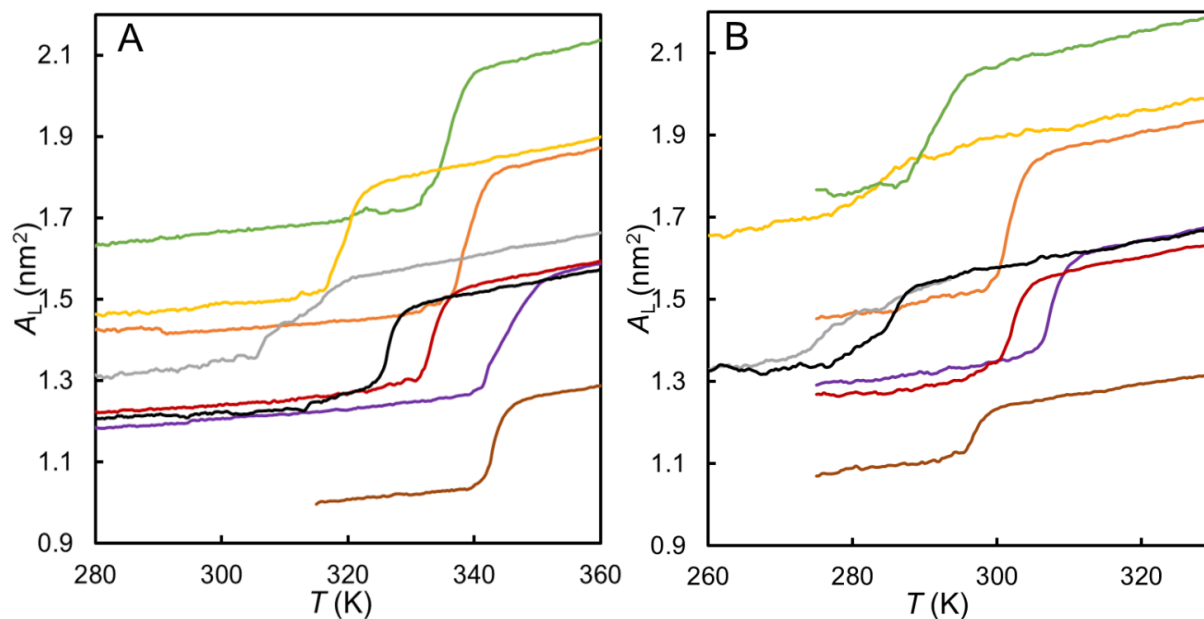


Figure 3-16. Area per lipid ( $A_L$ ) of Lipid A as a function of temperature for (A) Set II and (B) Set VII bacterial outer membrane. Color scheme: *H. pylori* (brown), *P. gingivalis* (red), *B. fragilis* (purple), *B. pertussis* (gray), *C. trachomatis* (black), *C. jejuni* (orange), *N. meningitides* (yellow), and *S. Minnesota* (green).

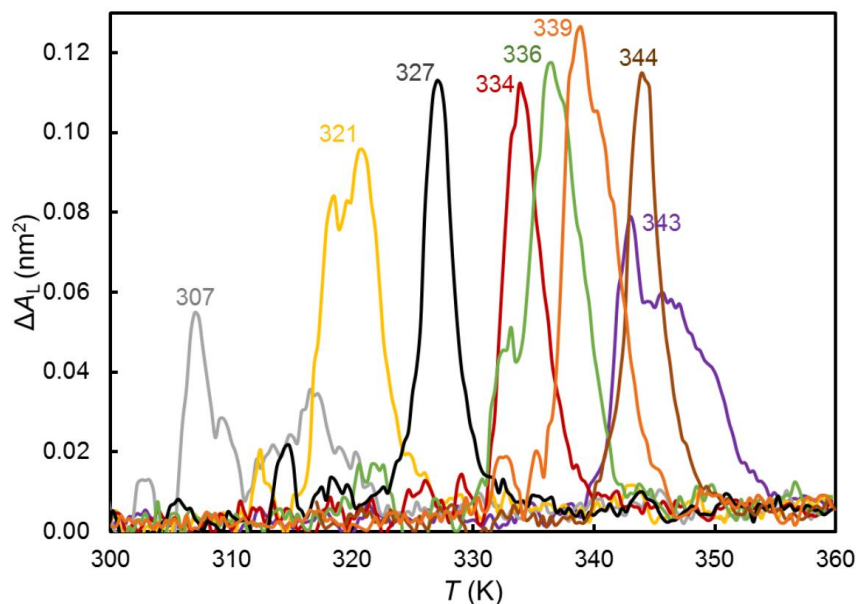


Figure 3-17. Phase transition temperature ( $T_m$ ) for Set II membranes determined by the change in the  $A_L$  ( $\text{nm}^2$ ) versus  $T$  (K). The  $T_m$  values (K) are labeled for each curve. Color scheme for the lines and labels: *H. pylori* (brown), *P. gingivalis* (red), *B. fragilis* (purple), *B. pertussis* (gray), *C. trachomatis* (black), *C. jejuni* (orange), *N. meningitidis* (yellow), and *S. Minnesota* (green).

As with the  $A_L$  values of the lipids, characteristic changes in the lipid tails were observed for the membranes below and above their  $T_m$  values. The tetra-acylated *H. pylori* lipid A has the smallest  $A_L$  value compared to the penta-, hexa- and hepta-acylated lipid A. The  $A_L$  values of the penta-acylated lipid A (*P. gingivalis*, *B. fragilis*, *C. trachomatis*, and *B. pertussis*) are 1.2–1.3  $\text{nm}^2$  in the ordered phase, below their phase transition temperature. In the disordered phase, about 10 K above the  $T_m$ , the  $A_L$  values increase to 1.5–1.6  $\text{nm}^2$ . Upon increasing the number of tails to six the  $A_L$  values increase for both *C. jejuni* and *N. meningitidis* membranes. The computed  $A_L$  values are in the range 1.45–1.48  $\text{nm}^2$  and 1.85–1.9  $\text{nm}^2$ , 10 K below and above their  $T_m$ , respectively. The hepta-acylated *S. minnesota*, has the highest  $A_L$  of 1.65  $\text{nm}^2$  and 2.0–2.1  $\text{nm}^2$  10 K below and above



the  $T_m$ . Snapshots of the membranes below the  $T_m$  show ordered and fully extended lipid tails and disordered and compacted lipid tails above the  $T_m$  (Figure 3-18).

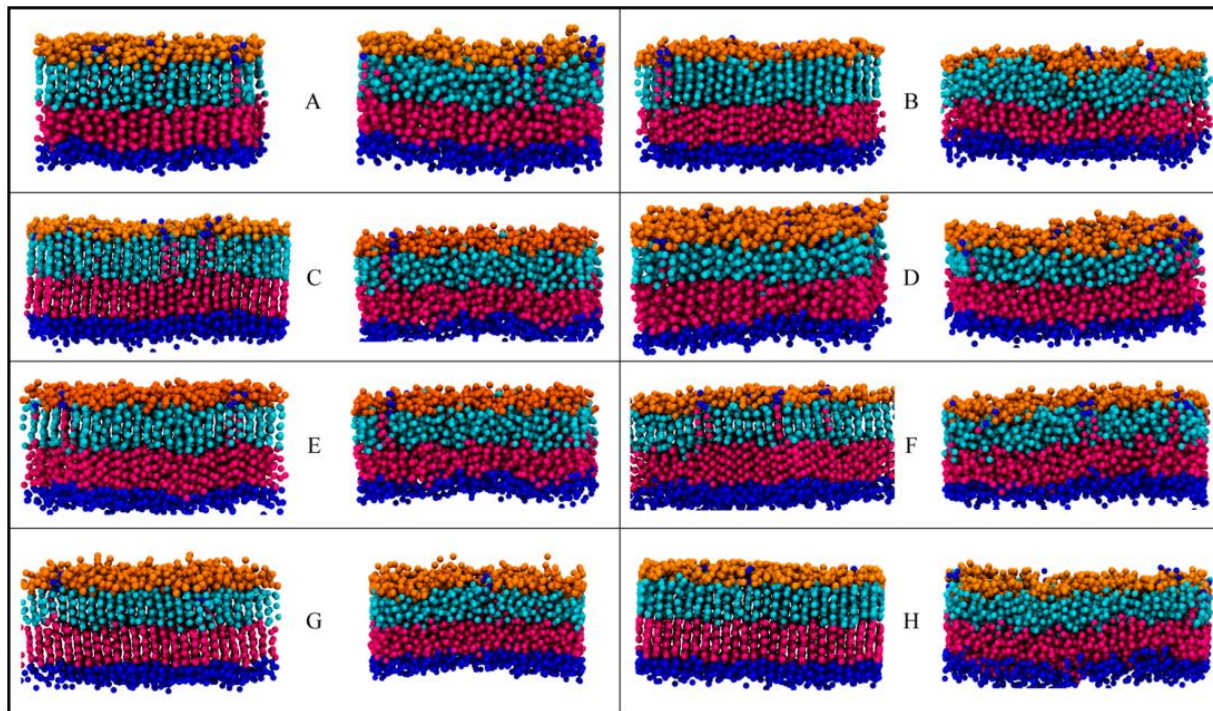


Figure 3-18. Snapshots of thermal phase transition of (A) *H. pylori*, (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitides*, and (H) *S. minnesota* in Set II. The panels show ordered phase (283 K, left) and disordered phase (350 K, right) membrane structure. Color scheme: Lipid A head groups (orange); Lipid A acyl chains (cyan); DPPE head group (blue); DPPE carbon tails (magenta).

Changing the lower leaflet composition (Set VI and VII) to include negatively charge POPG and cardiolipin lipid resulted in lipid A  $A_L$  values that were  $0.2 \text{ nm}^2$  larger than those in Set II. The slight increase in lipid A  $A_L$  is a direct consequence of the presence of charge in the lower leaflet, which causes increase in the bilayer cross-sectional area.

### 3.5.3 Membrane thickness

The  $D_M$  values were computed by measuring the perpendicular distance between the planes formed by the phosphate head groups of the top leaflet and the bottom leaflet. As expected  $D_M$  is larger at temperatures below  $T_m$  and smaller above the  $T_m$ . To capture this change in membrane structure, thickness was computed as a function of temperature for all eight membranes (Figure 3-19). All membranes, except *B. pertussis*, show  $\sim 0.51$  nm decrease in  $D_M$  after phase transition, which matches with the change in thickness observed experimentally in *S. minnesota* over a 30 K variation in temperature.<sup>26</sup> For *B. pertussis*, this decrease is only about 0.24 nm because of short 10-12 carbon acyl chains relative to others (Table 3-4).

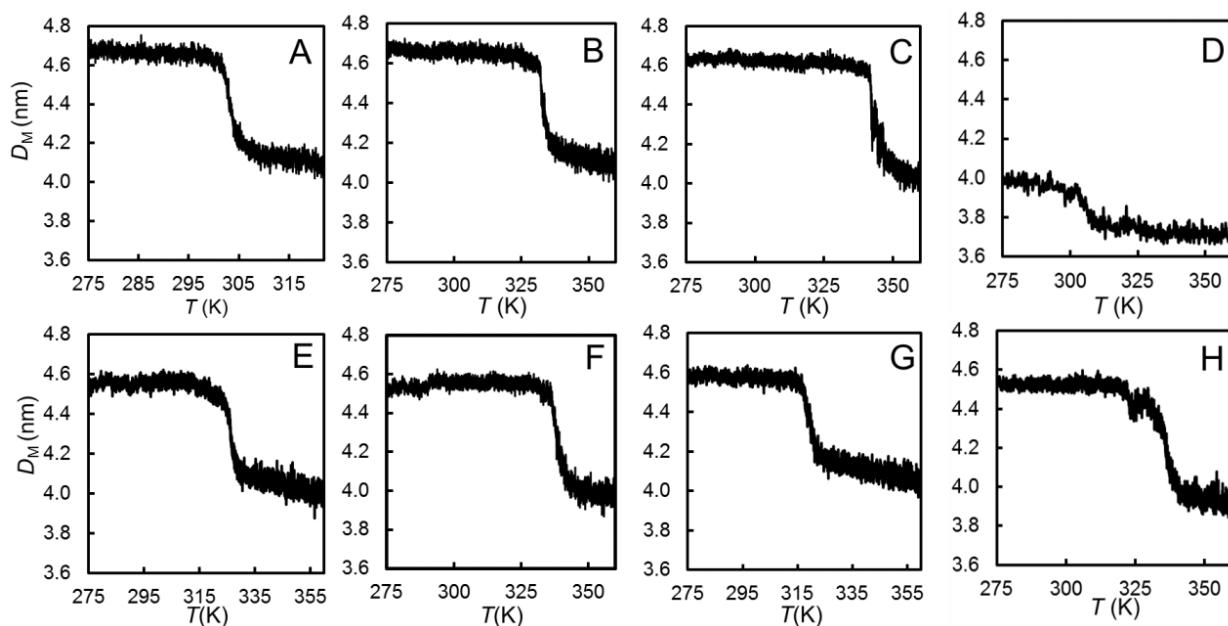


Figure 3-19. Membrane thickness (nm) as a function of  $T$  (K) for (A) *H. pylori*, (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitides*, and (H) *S. Minnesota* in Set II.

Table 3-4 Key Properties of the Membranes and Comparison of the Phase Transition

Temperatures from Sets II and VII with Available Experimental Data

bacteria	$A_L$ (nm <sup>2</sup> )		$D_M$ (nm)		$D_H$ (nm)	$T_m$ (K)			$D$ (cm <sup>2</sup> s <sup>-1</sup> )
	$T_m - 10$ K	$T_m + 10$ K	$T_m - 10$ K	$T_m + 10$ K	$T_m - 10$ K	set II	set VII	exptl	set I <sup>a</sup>
<i>H. pylori</i>	1.02	1.27	4.63	4.14	2.8	341 ± 4	299 ± 1	—	1.9 × 10 <sup>-7</sup>
<i>P. gingivalis</i>	1.27	1.54	4.64	4.11	2.9	338 ± 4	305 ± 2	—	3.0 × 10 <sup>-8</sup>
<i>B. fragilis</i>	1.25	1.55	4.59	4.09	3.0	349 ± 6	309 ± 1	—	1.3 × 10 <sup>-7</sup>
<i>B. pertussis</i>	1.33	1.52	3.96	3.74	2.6	309 ± 2	279 ± 2	—	1.6 × 10 <sup>-7</sup>
<i>C. trachomatis</i>	1.26	1.51	4.55	4.05	3.3	330 ± 3	289 ± 2	—	6.5 × 10 <sup>-8</sup>
<i>C. jejuni</i>	1.46	1.83	4.54	4.01	2.7	336 ± 4	296 ± 6	325 <sup>b</sup>	4.4 × 10 <sup>-8</sup>
<i>N. meningitidis</i>	1.49	1.80	4.55	4.12	2.5	324 ± 3	284 ± 2	318 <sup>c</sup>	7.3 × 10 <sup>-8</sup>
<i>S. minnesota</i>	1.71	2.08	4.31	3.93	2.8	337 ± 1	299 ± 8	310, 321 <sup>c</sup>	1.4 × 10 <sup>-8</sup>

<sup>a</sup>At  $T = 275$  K. <sup>b</sup>Reference 29. <sup>c</sup>Reference 30.

The simulated *S. minnesota*  $D_M$  was found 4.31 and 3.93 nm 10 K below and above the  $T_m$ , which is in good agreement with the electron density profile for rough mutant lipopolysaccharides Re (LPS Re) of *S. Minnesota* (strain R595). The experimentally reported upper leaflet head-group to lower leaflet head-group distance of the bilayer is 4.29 nm at 293 K and 3.87 nm at 323 K.

### 3.5.4 Density profile

The distribution of individual components within the lamellar asymmetrical bilayers was computed for all the membranes 10 K below their  $T_m$ . Because the lipids are in a thermal equilibrium, they adopt highly variable instantaneous molecular orientations; therefore, density profiles of all membrane components were calculated over 1  $\mu$ s of the simulation trajectory to account for ensemble averaging.

The density profiles computed for Set VI to determine the key feature of the membranes for comparison (Figure 3-20). At each membrane interface, the  $\text{Ca}^{2+}$  ions interact with the lipid headgroups and do not penetrate the hydrophobic tails of the outer and inner leaflets. The  $\text{Ca}^{2+}$  ion density is more pronounced in the lipid A headgroups of *N. meningitidis* because of the presence of additional phosphorylated residues that cap the phosphates at positions 1 and 4'. This also

explains the higher density of water surrounding the lipid A headgroups than phospholipids in the inner leaflet. The counterion peaks in the density profile were used as a measure of the membrane thickness. Additionally, the density profile of C1 beads (representing the acyl chains in both leaflets) was plotted as a function of the membrane normal (z coordinate) as a measure of the hydrophobic thickness ( $D_H$ ). The hydrophobic thickness lies in the 2.5 –3.3 nm range depending on the number of carbons in the acyl chains. The *C. trachomatis* membrane with an average of 17 carbons in the acyl chains has the highest hydrophobic thickness of 3.3 nm, which can be an important factor in determining the nature of the transmembrane porin proteins that can span the relatively thick outer membrane. Additionally, the high hydrophobic thickness in *C. trachomatis* also prevents the penetration of water deeper into the lipid A head groups.

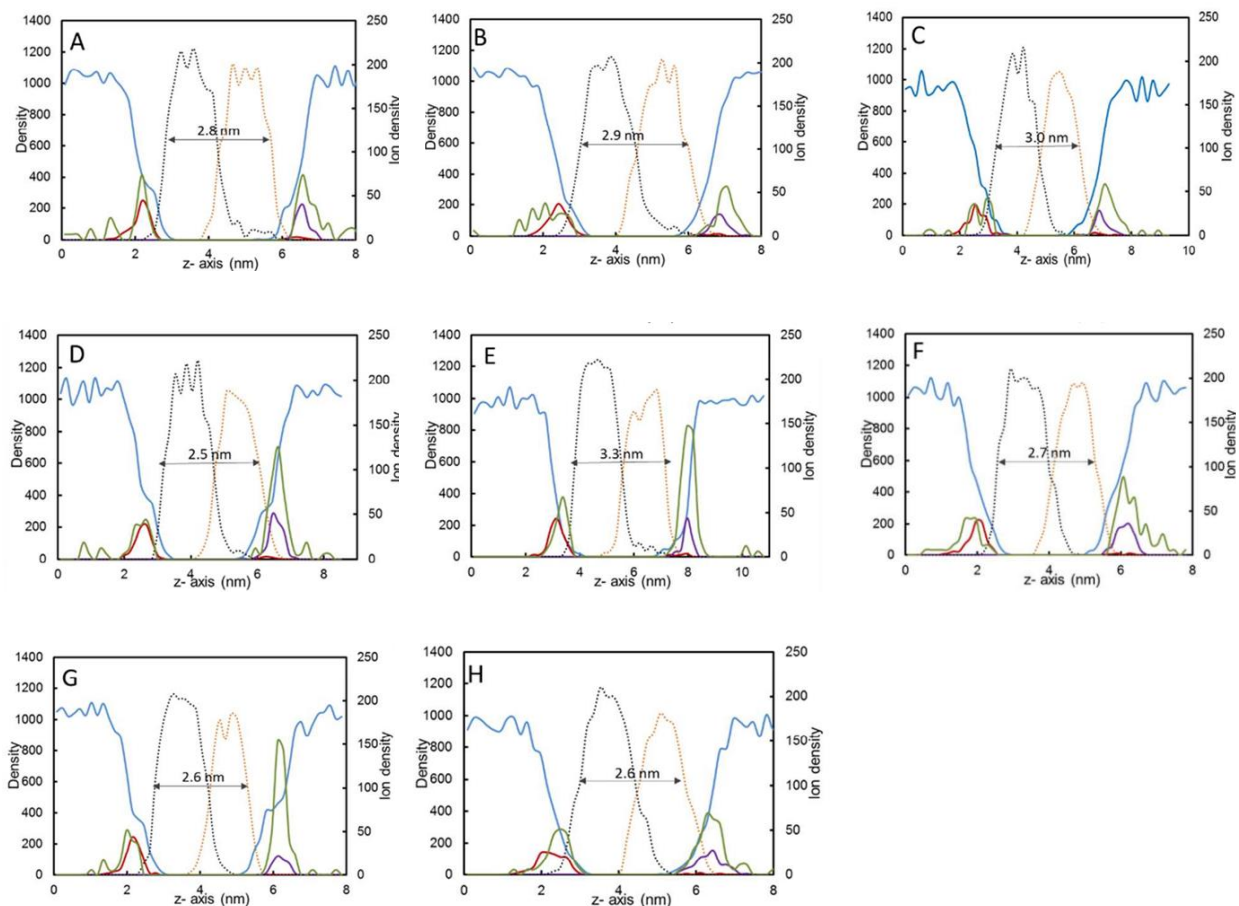


Figure 3-20. Density profile of key components of (A) *H. pylori*, (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis* (E) *C. trachomatis* (F) *C. jejuni*, (G) *N. meningitides*, and (H) *S. minnesota* membranes (Set VI). Color scheme: Water (blue); Ca<sup>2+</sup> counter ions (green); Lipid A phosphates (purple); DPPE head groups (red); DPPE carbon tails (black, dotted); and Lipid A carbon tails (orange, dotted). The Ca<sup>2+</sup>/counterion density is shown on the secondary y-axis. The trough between the DPPE carbon tails (black, dotted) and Lipid A carbon tails (orange, dotted) profiles marks the membrane mid-plane in each panel.

### 3.5.5 Radial distribution function

The radial distribution functions (RDF) of  $\text{Na}^+$  and  $\text{Ca}^{2+}$  ions interacting with the negatively charged phosphate and carboxyl groups of lipid A were calculated for all eight membranes (Figure 3-21). The curves for all membranes show similar trends, but notable is the peak for  $\text{Ca}^{2+}$ -carboxylate, which occurs at a longer distance ( $\sim 1.3$  nm) than the  $\text{Na}^+$ -carboxylate peak (at 0.51 nm). The peak positions imply that  $\text{Na}^+$  is able to penetrate deeper into the membrane and interact with carboxylate groups that lie below negatively charged phosphates.

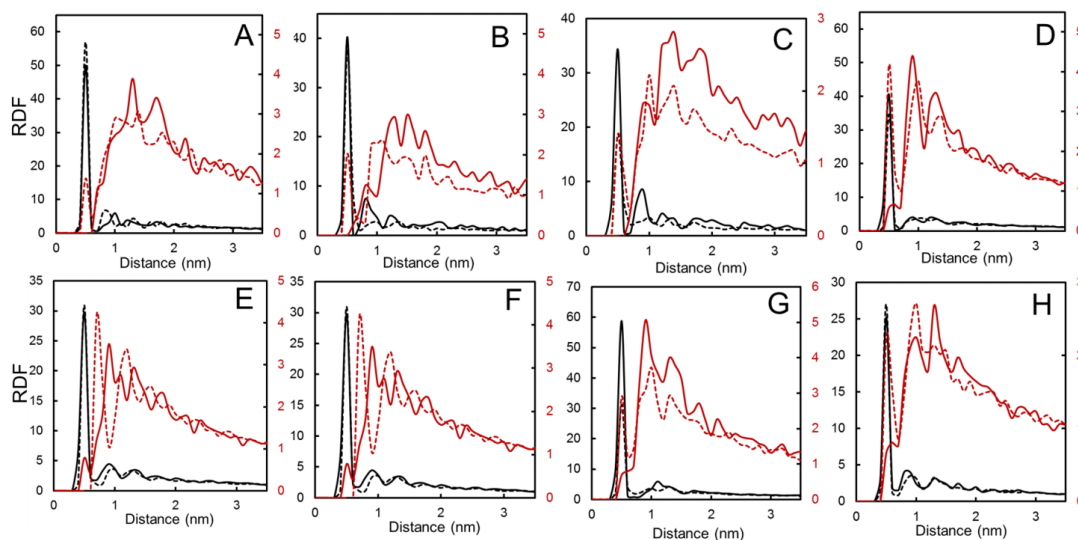


Figure 3-21. RDFs for set II ( $\text{Na}^+$ , dashed) and set VI ( $\text{Ca}^{2+}$ , solid) for phosphate (black) and carboxylate (red) for (A) *H. pylori*, (B) *P. gingivalis*, (C) *B. fragilis*, (D) *B. pertussis*, (E) *C. trachomatis*, (F) *C. jejuni*, (G) *N. meningitidis*, and (H) *S. minnesota*.

### 3.5.6 Diffusion coefficient

The dynamical properties of the membrane are sensitive to the lipids that constitute the membrane and the physiochemical aspects of the surrounding medium. Due to the inherent complexity of the membranes it is often difficult to parse through these contributing factors individually. To mitigate the variability among the membranes, the diffusion coefficients ( $D$ ) of Set I membrane systems

were computed, which have similar outer leaflet composition, pure DPPE inner leaflet, solvated in water with monovalent  $\text{Na}^+$  counterions, and at  $T=275$  K. The data shows the *S. minnesota* has the lowest diffusion coefficient, while *H. pylori* has the about an order of magnitude higher diffusion coefficient (Table 3-4). Although there is difference in the phosphorylation state of these two lipids, the difference in  $D$  value is attributed primarily to the number of lipids and difference in molecular weight. The hexa-acylated *C. jejuni* and *N. meningitidis* have  $D$  values in same order of magnitude. The trend in the  $D$  values for the penta-acylated lipid A membranes (*P. gingivalis*, *B. fragilis*, *B. pertussis*, and *C. trachomatis*), was less apparent, but the shortest acyl chain length *B. pertussis* lipid A has the highest diffusion coefficient (Figure 3-22). As is evident from the  $A_L$  and  $D$  data, acyl chain addition or deletion has a significant effect on the membrane properties. The acyl chain variability is an excellent example of a structure–property relationship showing how bacteria can employ this attribute to adapt to their habitats.

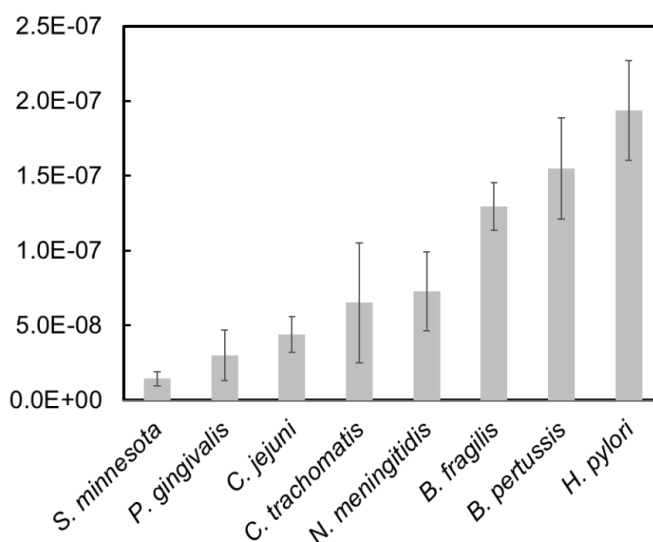


Figure 3-22. Diffusion coefficient ( $\text{cm}^2\text{s}^{-1}$ ) of for lipid A molecules (Set I) at 275 K.

### 3.6 DISCUSSION

The complexity of Gram-negative bacterial outer membranes has been a limiting factor in the computational modeling and characterization of these membranes. Until recently, the inherently asymmetric outer membranes (with LPS/phospholipid leaflets) were simplified as symmetric phospholipids in molecular simulations because of the lack of atomistic and coarse-grained force field parametrization. The development of LPS models is in its infancy, with models available for one or two bacterial species in atomistic and coarse-grained representations. There is therefore limited molecular-level understanding of the effects of the number of acyl chains, the length of acyl chains, and phosphorylation of a lipid on the membrane properties. The library of eight coarse-grained bacterial lipid models studied here will provide a systematic evaluation of the factors contributing to the membrane properties.

#### 3.6.1 Effect of number of acyl chains

Bacterial species adopt various acylation patterns to promote their survival by evading detection by the host innate immune system. It has been shown that penta-, tetra-, and tri-acylated lipid A surrogates stimulate a smaller immune response and lower cytokines levels compared to hexa-acylated lipid A. Some bacteria actively modify lipid A in response to changes in temperature of the host. An example is *Yersinia pestis* that produces hexa-acylated lipid A under ambient conditions but shifts to a tetra-acylated form at temperatures close to mammalian body temperature. In other cases, such as *S. typhimurium* and *P. aeruginosa*, acylation patterns are modified by enzyme activity in response to hostile chemical stimuli, such as depletion of cationic counterions, changes in pH, and presence of antimicrobial peptides, among others. The variability in acyl chains permits up and down regulation of outer membrane permeability and structural integrity, thus enhancing bacterial survival in harsh non-optimal environments.



The eight bacterial lipids studied here represent a range in lipid A structural diversity with 4–7 acyl chains. Analyzing the  $A_L$  data (Table 3-4) 10 K below the phase transition ( $T_m - 10$  K) shows that on an average each lipid tail contributes to  $\sim 0.24$  nm<sup>2</sup> to the area occupied by a lipid molecule. While this rule-of-thumb holds well for the membranes below their phase transition temperature, the contribution of the lipid tails increases to  $\sim 0.31$  nm<sup>2</sup> above the  $T_m$  in the disordered phase. Additionally, increasing the number of chains decreases the diffusivity of the lipid. Diffusion coefficient data from simulations of the eight lipid, under similar physiochemical conditions show that hepta-acylated lipid A is an order of magnitude lower than the tetra-acylated surrogate, and values of hexa- and penta-acylated lipid A are range between the two extremes (Figure 3-22).

### 3.6.2 Effect of acyl chain length

Membrane microstructure is dependent on the lipid-lipid interactions between adjacent molecules and is intimately tied to the length of the acyl chains and the average hydrophobic thickness. A membrane with a larger hydrophobic thickness experiences increased van der Waals attractions between neighboring lipids, resulting in lower area per lipid and a higher phase transition temperature. The data from the eight membranes studied here reflect the expected trend. Pair wise comparisons of *B. pertussis* and *C. trachomatis* membranes, which have the shortest and longest acyl chains of the group, show that *C. trachomatis* (with at least 2–6 additional carbons in the acyl chains) has a higher hydrophobic thickness ( $\Delta D_H = +0.6$  nm), lower area per lipid ( $\Delta A_L = -0.07$  nm<sup>2</sup>), and a higher phase transition temperature ( $\Delta T_m = + 27$  K) than does *B. pertussis*. Density profiles of the membranes (Figure 3-23) show that the difference in the total membrane thickness ( $\Delta D_M = +0.6$  nm) arises due to the hydrophobic thickness alone and not due to the disaccharide head groups. In addition, comparing density profiles (Figure 3-20) of *P. gingivalis* and *B. fragilis*

membranes, which have similar lipid A structures in terms of number of acyl chains, phosphorylation, but differ only by one carbon in two of its acyl chain, have similar values for  $A_L$ ,  $D_M$ ,  $D_H$ , and  $T_m$  (Table 3-4).

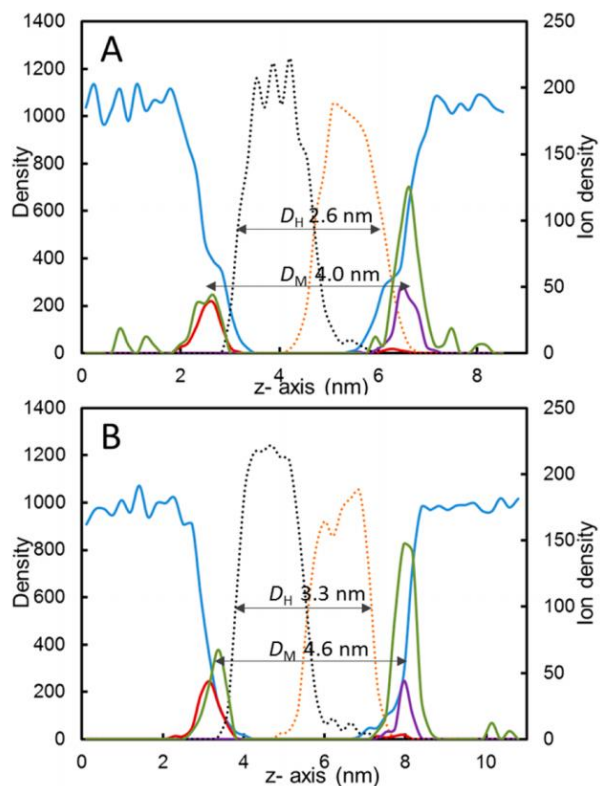


Figure 3-23. Density profiles of key components of (A) *B. pertussis* and (B) *C. trachomatis* membranes (set VI). Color scheme: water (blue); lipid A phosphates (purple); DPPE headgroups (red); inner leaflet carbon tails (black, dotted); outer leaflet carbon tails (orange, dotted). The  $\text{Ca}^{2+}$  counterion density (green) is shown on the secondary y-axis.

### 3.6.3 Effect of phosphorylation and counterions

The phosphorylation state of the disaccharide head group influences lipid A-mediated endotoxicity. Bacteria species with missing phosphates are resistant to antimicrobial peptides and are less active than the diphosphorylated lipids. For example, *H. pylori* consists of a tetra-acylated

lipid A that lacks the 4'-phosphate group to evade detection by Toll-like receptor 4 (TLR4) and resists action by antimicrobial peptides.<sup>27</sup> In *Salmonella typhimurium*, neutralizing the phosphates results in increased antimicrobial resistance and decreased immunogenic response.<sup>28</sup>

Furthermore, the negatively charged phosphates act as coordination sites for divalent ions to chelate adjacent LPS molecules. Comparison of phosphate- $\text{Ca}^{2+}$  radial distribution functions,  $g_{\text{P-Ca}^{2+}}(r)$  in *H. pylori*, *C. jejuni*, and *N. meningitidis*, all show a predominant peak at  $r = 0.5$  nm, irrespective of the number of phosphates on lipid A head groups (Figure 3-24). The separation distance of 0.5 nm is particularly important because it is the signature of the closest non-bonded distance between two CG beads. Unlike  $g_{\text{P-Ca}^{2+}}(r)$ , phosphate-phosphate (P-P) radial distribution functions  $g_{\text{P-P}}(r)$  clearly show differences in the phosphorylation states among these lipid A structures. *H. pylori* with one phosphate (at position 1 of the disaccharide head group) shows a low intensity P-P peak at 0.5 nm, mediated by the  $\text{Ca}^{2+}$  ions, but majority of the phosphates are less organized illustrated by the broader peak centered at 0.8 nm (Figure 3-24A). On the other hand, *C. jejuni* with two phosphate groups (at positions 1 and 4') shows well-defined peaks at 0.6 and 0.9 nm that correspond to the head-on intermolecular and the intramolecular P-P interactions, respectively (Figure 3-24B). Finally, *N. meningitidis* with four phosphates (bonded pair at positions 1 and 4') shows a bonded P-P peak at 0.3 nm, and broader peak centered at 0.9 nm. The differences in the ionic charge density at the lipid A-water interface is shown in the inset snapshots in Figure 3-24. As expected *H. pylori* snapshot shows lowest charge density, which is considered to be the cause of higher resistant to cationic antimicrobial peptides.

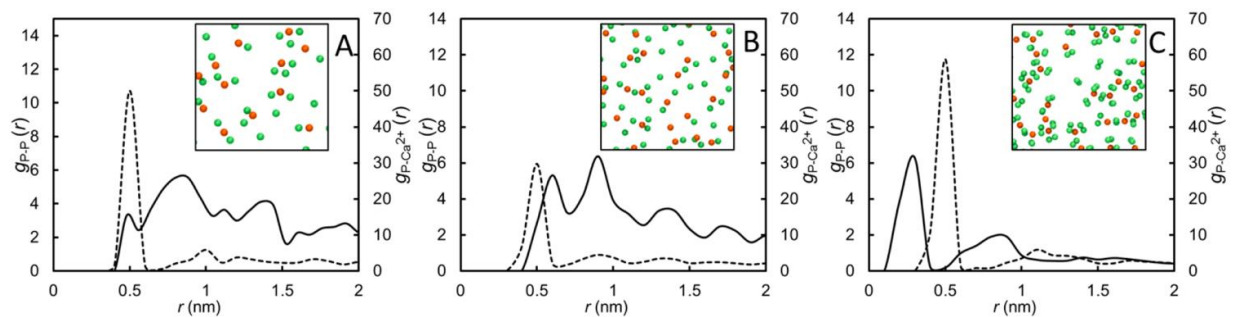


Figure 3-24. Lipid A headgroup phosphate-phosphate ( $g_{P-P}(r)$ ; solid lines) and phosphate- $Ca^{2+}$  counterion ( $g_{P-Ca^{2+}}(r)$ ; dashed lines) radial distribution functions for (A) *H. pylori*, (B) *C. jejuni*, and (C) *N. meningitidis*. The inset images show top views of the lipid A headgroup phosphates (green) and  $Ca^{2+}$  counterions (orange) from the surrounding medium.

### 3.7 CONCLUSIONS

This work provides thermodynamic and dynamical properties of a diverse set of eight bacterial membranes commensal or human pathogenic Gram-negative bacteria species: *Helicobacter pylori*, *Porphyromonas gingivalis*, *Bacteroides fragilis*, *Bordetella pertussis*, *Chlamydia trachomatis*, *Campylobacter jejuni*, *Neisseria meningitidis*, and *Salmonella minnesota*. This representative set of Gram-negative bacteria have lipid A domains that differ in the degree of phosphorylation presence of phosphate substituents, glucosamine head group as well as the nature, number, location, and length of acyl chains. After multiple independent simulations for all membranes, several key characteristics emerge. First, we find that on an average each lipid tail contributes  $\sim 0.24 \text{ nm}^2$  to the total area of the lipid, therefore  $A_L$  values of hepta-acylated *S minneosta* and tetra-acylated *H. pylori* lipid A are in 7:4 ratio. Second, the membranes composed of longer acyl chain lipid A have smaller  $A_L$  and a higher phase transition temperature compared to their shorter acyl chain counterparts. Third, membrane composition and charge of the inner leaflet can influence the phase transition temperature of the membrane by 20-30 K. Four, the monovalent ions bury themselves deeper in the membrane headgroups whereas the divalent ions are superficial and act as chelating agents binding to the phosphates on adjacent lipid A molecules. The insights from the work presented here coupled with the development of library of lipid A coarse-grained topology will facilitate advances in knowledge-based design on antimicrobial agents.

## REFERENCES

1. Nikaido, H., Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews* **2003**, 67 (4), 593-656.
2. Needham, B. D.; Trent, M. S., Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nature Reviews Microbiology* **2013**, 11 (7), 467-481.
3. Raetz, C. R. H.; Whitfield, C., Lipopolysaccharide endotoxins. *Annual Review of Biochemistry* **2002**, 71, 635-700.
4. Maeshima, N.; Fernandez, R. C., Recognition of lipid A variants by the TLR4-MD-2 receptor complex. *Frontiers in Cellular and Infection Microbiology* **2013**, 3.
5. Dixon, D. R.; Darveau, R. P., Lipopolysaccharide heterogeneity: Innate host responses to bacterial modification of lipid A structure. *Journal of Dental Research* **2005**, 84 (7), 584-595.
6. Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B. S.; Lee, H.; Lee, J. O., The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **2009**, 458 (7242), 1191-U130.
7. Munford, R. S., Sensing Gram-negative bacterial lipopolysaccharides: a human disease determinant? *Infection and Immunity* **2008**, 76 (2), 454-465.
8. Maldonado, R. F.; Sa-Correia, I.; Valvano, M. A., Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *Fems Microbiology Reviews* **2016**, 40 (4), 480-493.
9. Ma, H.; Irudayanathan, F. J.; Jiang, W.; Nangia, S., Simulating Gram-Negative Bacterial Outer Membrane: A Coarse Grain Model. *The Journal of Physical Chemistry B* **2015**, 119 (46), 14668-14682.

10. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H., The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* **2007**, *111* (27), 7812-7824.
11. Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J., Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *Plos Computational Biology* **2010**, *6* (6).
12. Kusters, J. G.; van Vliet, A. H. M.; Kuipers, E. J., Pathogenesis of *Helicobacter pylori* Infection. *Clinical Microbiology Reviews* **2006**, *19* (3), 449-490.
13. Moran, A. P.; Lindner, B.; Walsh, E. J., Structural characterization of the lipid A component of *Helicobacter pylori* rough- and smooth-form lipopolysaccharides. *J Bacteriol* **1997**, *179* (20), 6453-63.
14. Griffen, A. L.; Becker, M. R.; Lyons, S. R.; Moeschberger, M. L.; Leys, E. J., Prevalence of *Porphyromonas gingivalis* and Periodontal Health Status. *Journal of Clinical Microbiology* **1998**, *36* (11), 3239-3242.
15. Kumada, H.; Haishima, Y.; Umemoto, T.; Tanamoto, K. I., Structural Study on the Free Lipid-A Isolated from Lipopolysaccharide of *Porphyromonas-Gingivalis*. *Journal of Bacteriology* **1995**, *177* (8), 2098-2106.
16. Weintraub, A.; ZÄHringer, U.; Wollenweber, H.-W.; Seydel, U.; Rietschel, E. T., Structural characterization of the lipid A component of *Bacteroides fragilis* strain NCTC 9343 lipopolysaccharide. *European Journal of Biochemistry* **1989**, *183* (2), 425-431.
17. Basheer, S. M.; Bouchez, V.; Novikov, A.; Augusto, L. A.; Guiso, N.; Caroff, M., Structure activity characterization of *Bordetella petrii* lipid A, from environment to human isolates. *Biochimie* **2016**, *120*, 87-95.

18. Brunham, R. C.; Rey-Ladino, J., Immunology of Chlamydia infection: implications for a Chlamydia trachomatis vaccine. *Nat Rev Immunol* **2005**, 5 (2), 149-161.
19. Qureshi, N.; Kaltashov, I.; Walker, K.; Doroshenko, V.; Cotter, R. J.; Takayama, K.; Sievert, T. R.; Rice, P. A.; Lin, J. S.; Golenbock, D. T., Structure of the monophosphoryl lipid A moiety obtained from the lipopolysaccharide of Chlamydia trachomatis. *The Journal of biological chemistry* **1997**, 272 (16), 10594-600.
20. Erridge, C.; Bennett-Guerrero, E.; Poxton, I. R., Structure and function of lipopolysaccharides. *Microbes and Infection* **2002**, 4 (8), 837-851.
21. Haddad, N.; Burns, C. M.; Bolla, J. M.; Prévost, H.; Fédérighi, M.; Drider, D.; Cappellet, J. M., Long-Term Survival of Campylobacter jejuni at Low Temperatures Is Dependent on Polynucleotide Phosphorylase Activity. *Applied and Environmental Microbiology* **2009**, 75 (23), 7310-7318.
22. Reuter, M.; Mallett, A.; Pearson, B. M.; van Vliet, A. H. M., Biofilm Formation by Campylobacter jejuni Is Increased under Aerobic Conditions. *Applied and Environmental Microbiology* **2010**, 76 (7), 2122-2128.
23. Reinhardt, A.; Yang, Y.; Claus, H.; Pereira, C. L.; Cox, A. D.; Vogel, U.; Anish, C.; Seeberger, P. H., Antigenic potential of a highly conserved Neisseria meningitidis lipopolysaccharide inner core structure defined by chemical synthesis. *Chemistry & biology* **2015**, 22 (1), 38-49.
24. Kulshin, V. A.; Zahringer, U.; Lindner, B.; Frasch, C. E.; Tsai, C. M.; Dmitriev, B. A.; Rietschel, E. T., Structural Characterization of the Lipid-A Component of Pathogenic Neisseria-Meningitidis. *Journal of Bacteriology* **1992**, 174 (6), 1793-1800.



25. Janusch, H.; Brecker, L.; Lindner, B.; Alexander, C.; Gronow, S.; Heine, H.; Ulmer, A. J.; Rietschel, E. T.; Zahring, U., Structural and biological characterization of highly purified hepta-acyl lipid A present in the lipopolysaccharide of the *Salmonella enterica* sv. Minnesota Re deep rough mutant strain R595. *J Endotoxin Res* **2002**, 8 (5), 343-56.
26. Garidel, P.; Rappolt, M.; Schromm, A. B.; Howe, J.; Lohner, K.; Andra, J.; Koch, M. H.; Brandenburg, K., Divalent cations affect chain mobility and aggregate structure of lipopolysaccharide from *Salmonella minnesota* reflected in a decrease of its biological activity. *Biochimica et biophysica acta* **2005**, 1715 (2), 122-31.
27. Needham, B. D.; Trent, M. S., Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nat Rev Micro* **2013**, 11 (7), 467-481.
28. Kong, Q.; Six, D. A.; Liu, Q.; Gu, L.; Wang, S.; Alamuri, P.; Raetz, C. R.; Curtiss, R., 3rd, Phosphate groups of lipid A are essential for *Salmonella enterica* serovar Typhimurium virulence and affect innate and adaptive immunity. *Infect Immun* **2012**, 80 (9), 3215-24.
29. Brandenburg, K.; Andra, J.; Müller, M.; Koch, M. H. J.; Garidel, P. Physicochemical properties of bacterial glycopolymers in relation to bioactivity. *Carbohydr. Res.* **2003**, 338 (23), 2477-2489.
30. Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J. "Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, 6 (6), e1000810.

## **CHAPTER 4**

# **Dynamics of OmpF trimer formation in the bacterial outer membrane of Escherichia coli**

---

Ma, H.; Khan, A.; Nangia, S., Dynamics of OmpF trimer formation in the bacterial outer membrane of Escherichia coli. *Langmuir* **2017**, *34* (19), 5623-5634.

## 4.1 ABSTRACT

The self-assembly of outer membrane protein F (OmpF) in the outer membrane of *Escherichia coli* Gram-negative bacteria was studied using multiscale molecular dynamics simulations. To accommodate the long timescale required for protein assembly, coarse-grained parameterization of *E. coli* outer membrane lipids was first developed. The OmpF monomers formed stable dimers at specific protein-protein interactions sites, exactly as identified in earlier literature. The dimer intermediate was asymmetric but provided a template to form a symmetric trimer. Superposition analysis of the self-assembled trimer with the X-ray crystal structure of the trimer available in the protein data bank showed excellent agreement with global root-mean square deviation of less than 2.2 Å. The free energy change associated with dimer formation was  $-26 \pm 1$  kcal mol<sup>-1</sup>, and for a dimer to bind to a monomer and to form a trimer yielded  $-56 \pm 4$  kcal mol<sup>-1</sup>. Based on thermodynamic data, an alternate path to trimer formation via interaction of two dimers is also presented.

## 4.2 INTRODUCTION

Porins are barrel-shaped membrane proteins in the Gram-negative bacterial outer membrane; these hydrophilic nanochannels permit the diffusion-mediated influx of nutrients through the otherwise impermeable outer membrane.<sup>1-4</sup> Porins are classified as non-specific channels because their low affinity to bind to substrates results in the gradient-based diffusion of small substrates into the bacterial cell. In contrast, specific protein channels are restrictive and have binding sites for particular chemical substrates to facilitate selective transport. A typical Gram-negative bacterial species may express several outer membrane channel proteins that differ structurally (lumen diameter, number of strands, and oligomeric state) and channel functionally (specific or non-specific).

In the bacterial outer membrane, porins encounter an asymmetric lipid environment with lipopolysaccharide (LPS) in the outer leaflet and a mixture of phospholipids in the inner leaflet.<sup>5</sup> The LPS is a complex molecule that is composed of three domains—Lipid A, the core oligosaccharide, and the O-antigen.<sup>6</sup> Lipid A domain forms stabilizing non-bonded interactions with the exposed hydrophobic residues on the porin surface. The four to seven saturated fatty acid chains of Lipid A facilitate the tight packing of LPS molecules with protein channels to maintain the outer membrane's impermeability to hydrophobic substrates. The negatively charged core oligosaccharide domain is cross-linked via divalent counterions to enhance outer membrane stability and impermeability.<sup>7</sup> The O-antigen consists of many repeats of an oligosaccharide unit that extend outwards from the membrane into the bacterial surroundings.

In *Escherichia coli*, nutrient uptake is mediated by the non-specific outer membrane protein F (OmpF), which consists of a homotrimeric  $\beta$ -barrel assembly.<sup>3,8</sup> Each monomer has a cylindrical topology formed by a 16-stranded peptide backbone arranged in an antiparallel motif. Its

hydrophobic side-chain residues are exposed to membrane lipids, while the hydrophilic residues form the lumen of the channel. The  $\beta$ -strands are connected via short turns (T1–T8) on the intracellular side of the channel and longer loops (L1–L8) on the extracellular side. Loops interact with LPS core domain to provide stability.<sup>9-12</sup> In addition, L2 of each monomer participates in stabilizing non-covalent interactions with the adjoining monomers within the trimeric assembly, while the L3 buries into the lumen of the barrel to form a size-selective constriction zone for the channel.<sup>13</sup> The *E. coli* OmpF constriction zone is marked by positively and negatively charged residues on opposing sides of the lumen, which are important for diffusion of charged substrates.

The high-resolution X-ray crystal structures of OmpF provide atomistic-level resolution of the trimer,<sup>8</sup> but these static structures lack the mechanistic and dynamic details governing trimerization. Molecular dynamics (MD) simulations enable the extrapolation of static structures to physiological events that can then be compared to experimental results. In previous molecular dynamics (MD) simulations, the influence of the variability of the LPS environments on the structure and dynamics of OmpF trimer was reported using all-atom MD simulations.<sup>14,15</sup> Studies of trimeric OmpF revealed deviations of dynamical structure relative to the crystal structure and showed that L3 flexibility affected a change in pore cavity. Molecular dynamics simulations were also successfully used to observe solute behavior and passage through OmpF, and findings compared to experimental results.<sup>12,13</sup> The recent development of new algorithms has enabled research to simulate ion conduction directly using applied field MD simulations.<sup>16,17</sup>

Dimeric and monomeric states of OmpF also have been observed *in vivo* and *in vitro* experiments.<sup>18-22</sup> These findings suggest that the mechanism of OmpF formation proceeds in a stepwise manner from monomers to dimers to trimers. In a computational study involving dynamics of OmpF monomers in asymmetric phospholipid bilayers showed clustering and reduced

mobility of the OmpFs in a crowded lipid environments,<sup>23</sup> but did not specifically focus on the mechanism of OmpF trimer formation. Other studies focused on the formation of OmpF assembly revealed that the oligomerization occurs via specific protein-protein pair interactions.<sup>24-26</sup> Each OmpF monomer has two distinct patches centered at residues Glycine-19 (G19) and Glycine-135 (G135) that participate in oligomerization. A dimer is formed when G19 of an OmpF monomer interfaces with G135' patch of the OmpF' monomer. This dimer then interacts with the third OmpF'' monomer at the exposed G135 and G19' patches to form G135-G19'' and G19'-G135'' interfaces to complete the OmpF-OmpF'-OmpF'' trimer. Mutagenesis experiments have revealed that perturbing these interfacial residues results in loss of oligomerization.<sup>24</sup> In addition to thermodynamic stability, oligomerization of OmpF provides structural support for the extracellular loops. Niramitranon *et al.* showed that the OmpF trimer functions as a non-specific pore, but as a monomer it becomes anion-selective due to the dislocation of the D113 side chain on L3 loop, which blocks the cation pathway.<sup>15</sup> Other research indicates that the change of the location of side chains, the helices, or even the quaternary structure will result in porin malfunction.<sup>27</sup> The self-assembly simulation results in conjunction to previous literature indicates that being a trimer in *E.coli*'s outer membrane significantly decreases the structural flexibility of the OmpF and aids in maintaining the pore function by allowing exchange of both cations and anions, which is essential for *E.coli* survival.<sup>15</sup> Several reports of computational studies involving the stability and interactions of other trimeric bacterial porins based on from their trimeric X-ray crystallographic structure.<sup>10-12,14-15,28-30</sup> However, the stepwise assembly and molecular origins of OmpF trimerization have not yet been studied in detail due to the inability to perform long-time scale simulations with high fidelity to the underlying molecular structure.

Here we report the development of a coarse-grain force field parameter set for *E. coli* outer membrane lipids, which includes the Lipid A (LPA) and Lipid A with core (LPC). The force field parameters are similar in spirit to the recently reported coarse-grained parameter sets for bacterial lipids<sup>30-32</sup> that reduce the degrees of freedom of the system and provides a computationally affordable route to explore dynamics of molecular assemblies in microsecond timescales.<sup>33-40</sup> Using the developed parameter set, we computed structural properties of LPA and LPC lipid membranes, such as area per lipid, phase transition temperature, density profiles, hydrophobic thickness, bond distances and bond angles analysis, and compared our findings to available experimental and atomistic simulation results.

The outer membrane formed by the combination of the LPA in the outer leaflet and phospholipids in the inner leaflet was used to study the oligomerization process of OmpF monomers to more complex structures over tens of microseconds. We demonstrate the stepwise assembly of monomeric OmpF into a stable dimer and subsequent interaction with another monomer to form a stable trimer. The formation of OmpF trimer is a multibody interaction involving numerous protein-protein and protein-lipid binding and unbinding events. In general, protein self-assembly is a complex interplay of long-range (electrostatics, diffusivity, viscosity) and short-range (hydrogen bonding, hydrophobic and hydrophilic interactions, and van der Waal interactions) forces that drive the entire process from initial association to the final form via structural rearrangements.

## 4.3 METHODS

### 4.3.1 Parameterization

The *E. coli* lipid A head group (Figure 4-1) is a  $\beta(1\rightarrow6)$ -linked D-glucosamine disaccharide linked to C14 acyl carbon chains at positions 2 and 3 as well as 2' and 3' via amide or ester linkages.<sup>41</sup> The glucosamine residues (GlcN I and II) of the disaccharide head group are phosphorylated at positions 1 and 4'. Nonhydroxy saturated C12 and C14 secondary carbon chains further esterify the primary acyl chains at positions 2' and 3'. The primary hydroxyl at position 6' acts as the binding site of the core LPS oligosaccharide domain. The core oligosaccharide domain is branched and contains six to 10 hexoses (glucose and galactose units) with multiple anionic groups.<sup>6</sup> The divalent counterions act as chelating agents for the core anionic groups, diminish the electrostatic repulsion between neighboring LPS molecules, and provide stability to the membrane.<sup>116</sup> The outermost O-antigen domain is also an oligosaccharide consisting of 1–40 repeat units that extend into the surrounding medium.<sup>31,43</sup> Overall, there is high variability in the length of core and O-antigen oligosaccharides in *E. coli*; however, only lipid A and partial core oligosaccharides are essential for survival.



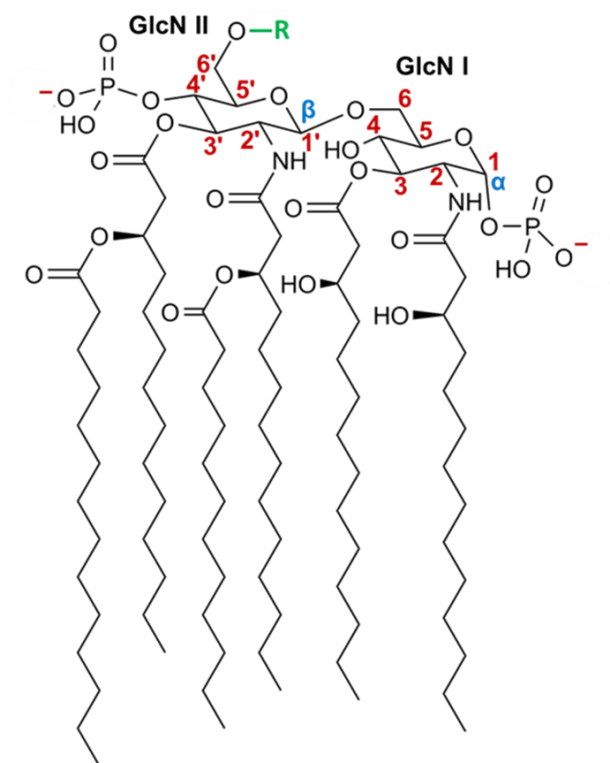


Figure 4-1. The chemical structure of *E. coli* lipid A domain. The head group is a  $\beta(1\rightarrow6)$ -linked D-glucosamine disaccharide linked to C14 acyl carbon chains at positions 2 and 3 as well as 2' and 3' via amide or ester linkages. The glucosamine residues (GlcN I and II) of the disaccharide head group are phosphorylated at positions 1 and 4'. Nonhydroxy saturated C12 and C14 secondary carbon chains further esterify the primary acyl chains at positions 2' and 3'. The primary hydroxyl at position 6' acts as the binding site of the core LPS oligosaccharide domain.

Here we develop the parameter set for lipid A and oligosaccharide core in the outer membrane of *E. coli* (Tables S1–S4) using the coarse-graining approach reported previously.<sup>36,38</sup> The coarse-grain parameterization of lipid A was developed based on MARTINI many-to-one mapping in which on average four or three heavy atoms are mapped into one bead (Figure 4-2).<sup>44-46</sup> Using the Lipid A template reported earlier,<sup>36,38</sup> each glucosamine unit was assigned four beads (P1, P2, P4, P5) and each phosphate was mapped into one Qa bead type with a unit negative charge (Figure 4-3). The acyl chain beads were assigned C1 bead type. The frequency distribution of the average bond lengths and angles of the acyl chains and glucose residues were computed (Figure 4-4 and

4-5). Sodium and calcium ions were used as counter ions to make the whole system electrically neutral. The *E. coli* parameter set adds to our library of nine bacterial species: *Pseudomonas aeruginosa*, *Helicobacter pylori*, *Porphyromonas gingivalis*, *Bacteroides fragilis*, *Bordetella pertussis*, *Chlamydia trachomatis*, *Campylobacter jejuni*, *Neisseria meningitidis*, and *Salmonella minnesota*.

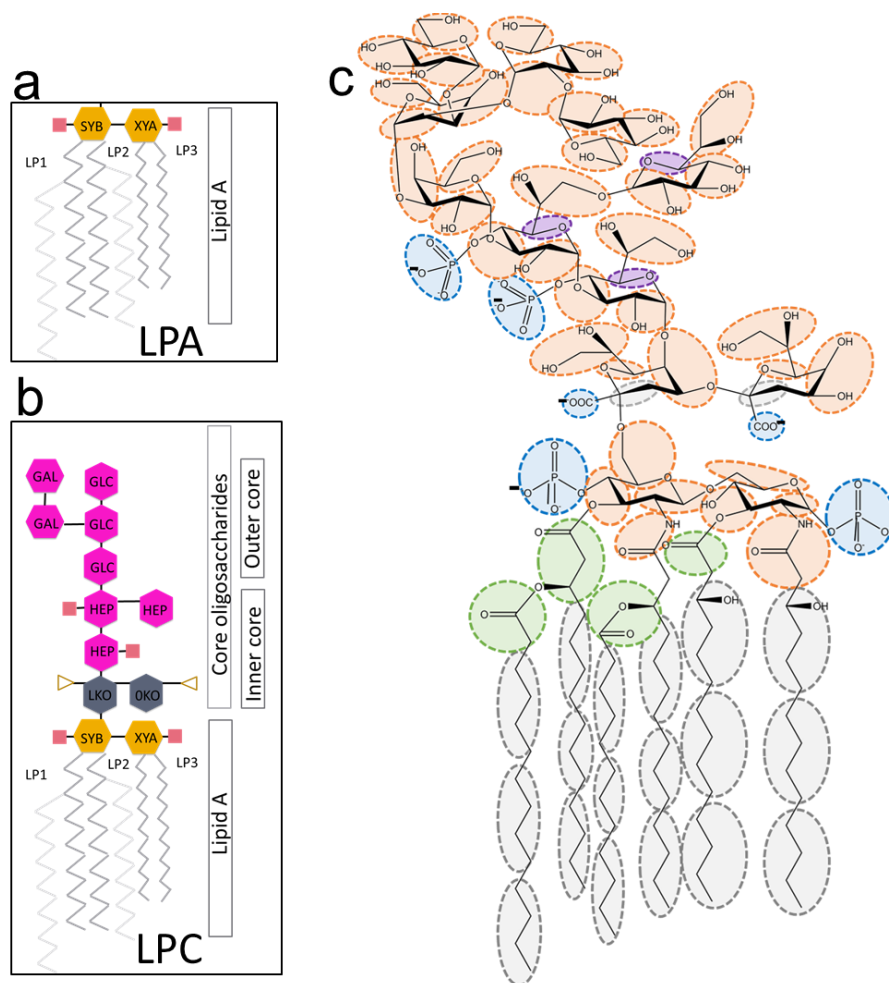


Figure 4-2. Representative chemical structure of *E.coli* (a) LPA, (b) LPC and (c) coarse grain mapping of LPC. Panels (a) and (b) shows LPA and LPC phosphates (peach) and carboxylate (orange; triangle) anionic groups, residue SYB and XYA (yellow; 3-(acetyl amino)-3-deoxy-D-glucose), LP1 (gray; 2-hydroxytetradecanoyl acid), LP2 (gray; dodecanoyl acid decyl ester) and LP3 (gray; tetraacyl acid decyl ester), and residues GAL (magenta; D-galactose), GLC (magenta; D-glucose) and HEP (magenta; L-glycero-D-manno heptose), LKO and OKO (gray; 3-deoxy-D-manno-oct-2-ulosonic acid). The coarse-grain mapping (dashed lines) show bead boundaries.

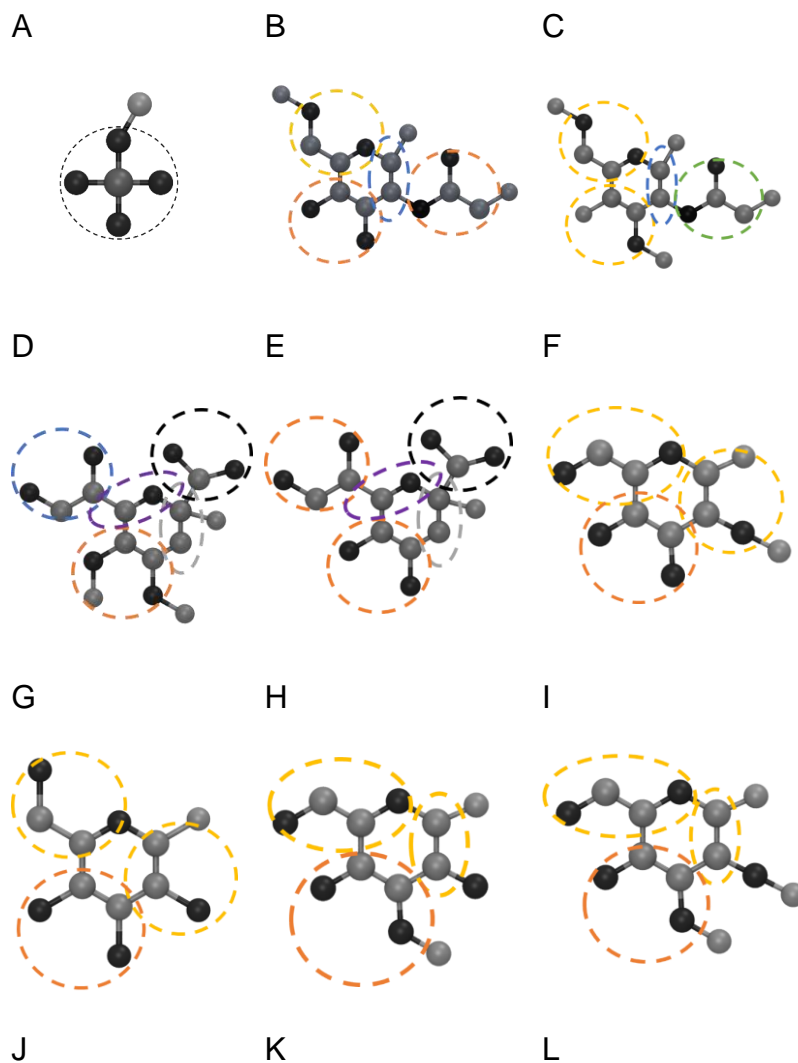


Figure 4-3. (A) PO4, (B) XYA, (C) SYB, (D) LKO, (E) OKO, (F) Gal 1, (G) Gal 2, (H) Glc1, (I) Glc 2, (J) Hep 1, (K) Hep 2, (L) Hep 3, (M) LP1, (N) LP2, (O) LP3. The coarse grained mapping scheme showing Martini beads of types Qa (black), P1 (yellow), P2 (blue), P4 (orange), P5 (green), N0 (purple), Na (light blue), and C1 (grey) and overlaid on the atomistic structure in ball and stick representation with carbon, oxygen, and linking bonds. Hydrogen atoms are not shown for clarity.

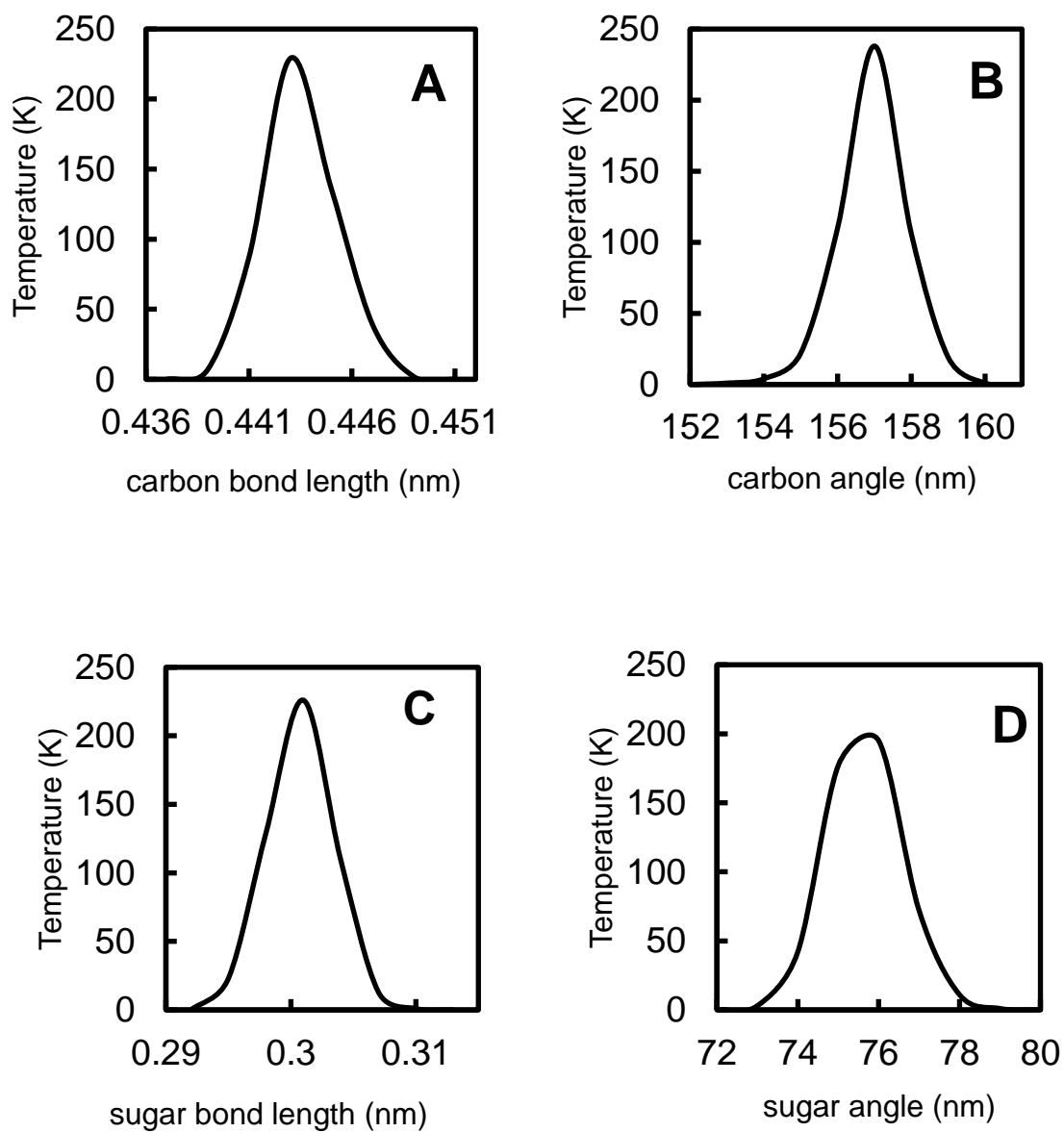


Figure 4-4. (A) Average acyl bond and B) angle frequency distribution. Average disaccharide head group bond (C) and angle frequency distribution (D).

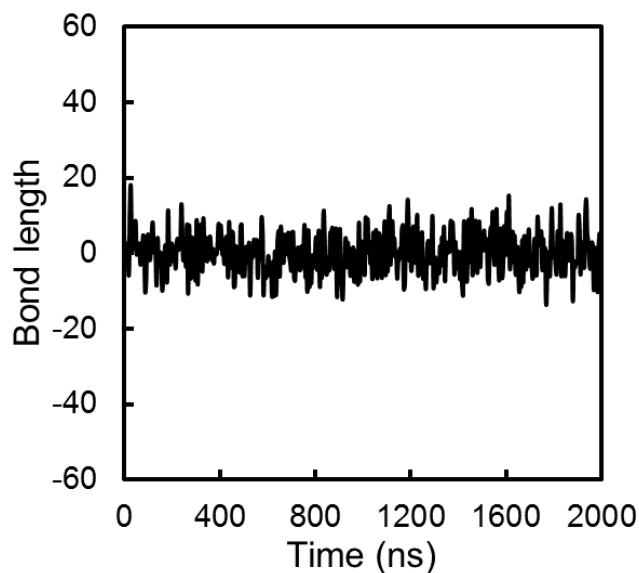


Figure 4-5. Dihedral angle fluctuations (degrees) as a function of simulation time.

#### 4.3.2 Simulation details

**Asymmetric Membranes:** We investigated three model membranes with varying level of lipid complexity to capture the asymmetric behavior of *E. coli* membrane and the dynamics of OmpF trimer self-assembly. The first membrane type (MT1) comprises of a Lipid A (LPA) in the outer leaflet and DPPE in the inner leaflet. The second membrane type (MT2) is a combination of Lipid A plus core (LPC) in the outer leaflet and DPPE in the inner leaflet. The third membrane type (MT3) comprises of LPA:DPPE (9:1) outer leaflet, and a ratio 7:2:1 mixture of 1-hexadecanoyl-2-(9Z-octadecenoyl)-sn-glycero-3 phosphoethanolamine (POPE), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPG), and cardiolipin (CDL2) in the inner leaflet, which most closely mimics to the composition of the *E. coli* membrane. The outer leaflet LPA:DPPE lipid ratio was maintained at 9:1 to the mimic the small concentrations of phospholipids that get recruited from the lower leaflet to stabilize the membrane under stress.<sup>47</sup>

The initial configuration of a coarse-grained membrane patches (10×10 nm<sup>2</sup>) were for built using a locally modified version of python script called *insane.py*.<sup>48</sup> The script uses a built-in library of lipids, ions and solvents to generate systems of the MT1–MT3 lipid membrane (Table 4-1). The *E. coli* LPA and LPC lipid templates were added, and the latest local version of the script is available to researchers. The membranes were solvated with standard MARTINI water, and the electroneutrality of the system was maintained using hydrated Na<sup>+</sup> or Ca<sup>2+</sup> counterions.

Table 4-1. Details of MT1–MT3 asymmetric membrane systems (without OmpF proteins).

Membrane Type	Outer leaflet			Inner leaflet			Counter Ion	No. of		<i>T</i> (K)	<i>t</i> (μs)
	LPA	LPC	DPPE	DPPE	POPG	CDL2		water	ions		
MT1 <sup>a</sup>	77	-	-	231	-	-	Na <sup>+</sup>	4358	154	323	2
MT1 <sup>b</sup>	77	-	-	231	-	-	Na <sup>+</sup>	4358	154	295-360	2
MT2 <sup>a</sup>	-	81	-	243	-	-	Na <sup>+</sup>	5831	486	323	2
MT2 <sup>b</sup>	-	81	-	243	-	-	Na <sup>+</sup>	5831	486	295-411	2
MT3 <sup>a</sup>	69	-	7	137	39	19	Ca <sup>2+</sup>	4607	107	310	2
MT3 <sup>b</sup>	69	-	7	137	39	19	Ca <sup>2+</sup>	4607	107	260-345	2

<sup>a</sup>Simulations to determine equilibrium properties of the membrane

<sup>b</sup>Simulations to determine the phase transition temperature

**Membrane Protein Systems:** The coarse-grained form of monomeric OmpF (pdb:4LSF) was downloaded from the online MemprotMD server,<sup>49</sup> which is an online repository for obtaining equilibrated protein structures in coarse-grained representation. The modified insane script was then used to insert the coarse-grained OmpF monomers in the desired membrane type (Table 4-2), as well as add water and counterions. The monomers were placed in a square grid (2×2 or 3×3), equidistant from each other and in random orientations (Figure 4-6a) to remove any conformational bias during self-assembly. Simulations were performed in two stages, where the OmpF monomers were initially position-restrained for 0.5 μs to equilibrate the surrounding lipids, followed by unrestrained self-assembly production run in the second stage.

Table 4-2. System details of OmpF-membrane simulations.

System	MT	Outer leaflet			Inner leaflet			Protein	No. of		$T$	$t$
		LPA	LPC	DPPE	DPPE	POPG	CDL2	OmpF	water	ions	(K)	( $\mu$ s)
1	MT1	124	-	12	384	-	-	4	31376	296	310	48
2	MT3	124	-	12	244	68	36	4	19663	436	325	40
3	MT3	81	-	9	162	45	18	9	13842	351	325	24
4	MT2	-	72	8	224	-	-	4	15629	480	310	48
5	MT2	-	81	9	252	-	-	9	15777	297	310	20
6	MT3	207	-	22	410	117	58	3	34520	683	325	16
7	MT2	-	225	25	445	129	63	3	23049	1641	310	2
8	MT1	234	-	26	728	-	-	2	25570	246	310	0.5 <sup>a</sup>
9	MT1	216	-	24	672	-	-	3	26370	234	310	0.5 <sup>a</sup>

<sup>a</sup>Simulation time for each umbrella sampling window

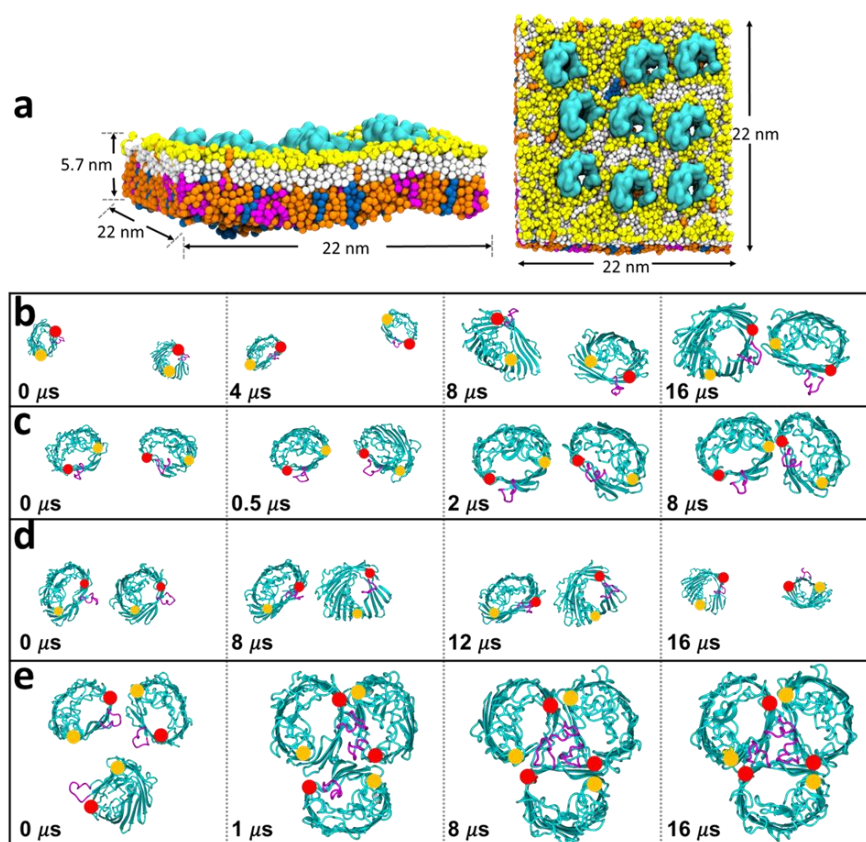


Figure 4-6. Simulation setup and self-assembly of OmpF trimer. (a) Side and top-view of a 3×3 grid of coarse-grained OmpF monomers (cyan, surface representation) in *E. coli* outer membrane MT3 (System 3) comprised of LPA (yellow and white beads) outer leaflet and a complex inner leaflet of DPPE (orange beads), POPG (blue beads), and CDL2 (magenta beads). Snapshots (top view) of (b) stable dimer (System 1), (c) dimer (System 3), (d) fleeting dimer interaction (System 2), and (e) formation of OmpF trimer (System 6). In panels (b–e), the position of G19 (red bead) and G135 (yellow bead) residues is highlighted; other components of the system are not shown for clarity.

**Simulation set-up:** The GROMACS molecular dynamics package<sup>50,51</sup> (5.1.2) was used to perform all simulations. Energy minimization was performed using the steepest-descent algorithm with a 20 fs time-step until the maximum force on any bead was below the tolerance parameter of 10 kJmol<sup>-1</sup>nm<sup>-1</sup>. Periodic boundary conditions were applied in all three dimensions. The *NVT* (constant number of particles, volume, and temperature) and *NPT* (constant number of particles, pressure, and temperature) equilibration runs were performed for 0.2  $\mu$ s. Semi-isotropic pressure coupling was used, and systems were maintained at 1 bar using the Berendsen barostat<sup>27</sup> with time constant,  $\tau_p = 4.0$  ps. Temperature was maintained at 310 or 325 K by independently coupling the lipids, proteins, and the solvent to an external velocity rescaling thermostat with  $\tau_T = 1.0$  ps. The neighbor list was updated every 25 steps using 1.4 and 1.2 nm for short-range van der Waals and electrostatic cutoffs, respectively. The production *NPT* simulations were performed for 2  $\mu$ s for the membrane without proteins (Table 4-1) and between 2–48  $\mu$ s for membranes with embedded OmpF proteins (Table 4-2). For membranes without proteins, annealing simulations were performed to obtain the phase transition temperatures; in brief, a short *NPT* was performed at 275 K followed by a heating scan from 275–360 K with a 15 K interval.

**Potential of Mean Force (PMF) Calculations:** To compute the PMF curves, we extracted the self-assembled dimer and trimer from the equilibrium simulations (Systems 3 and 6, respectively) and embedded them independently in 30 $\times$ 10 nm<sup>2</sup> membrane patches of simple MT1 bilayer; system 8 and 9, respectively. The MT1 lipids were equilibrated while keeping the proteins position-restrained for 0.5  $\mu$ s at 310 K. The restraints were then removed and pull simulations were performed along the reaction coordinate defined by separation of the center of mass (COM) of the interacting OmpFs while still embedded in MT1 bilayer. In case of both the dimer and the trimer, one OmpF was pulled with respect to the COM of its position-restrained oligomeric counterpart.



A harmonic potential with  $1000 \text{ kJmol}^{-1} \text{ nm}^2$  force constant was used for the pull. A total of 30-40 independent umbrella sampling windows were extracted along the reaction coordinate for both systems, and each window was simulated for 0.5  $\mu\text{s}$ . The weighted histogram analysis method (WHAM) along with bootstrapping was used to extract the PMF curves.

**General Analysis:** Post simulation analyses were performed using in-built GROMACS utilities, and external software suites such as YASARA<sup>52</sup> for protein alignment with X-ray crystal structure and CAVER<sup>41</sup> for analysis and visualization of porin channels. The use of YASARA and CAVER required reverse mapping of the OmpF from the CG representation to atomistic, which was achieved by using *backward.py* script that uses a library of mapping definitions to reconstruct the all-atom representation. Structural properties of the membranes such as area per lipid ( $A_L$ ), phase transition temperature ( $T_m$ ), membrane thickness ( $D_M$ ), and hydrophobic thickness ( $D_H$ ) were computed. Molecular visualization and graphics were generated using VMD<sup>43</sup> and YASARA.<sup>52</sup>

## 4.4 RESULTS AND DISCUSSION

### 4.4.1 Force field development and validation of membrane properties

Parameterization of *E. coli* outer membrane lipids is essential for examining OmpF assembly because the protein-protein interactions occur in the membrane milieu. Unlike most biological membranes that are symmetric, the asymmetric outer membrane comprising LPS-rich outer leaflets and phospholipid-rich inner leaflets characteristic of Gram-negative bacteria contribute to the membrane's striking properties. The presence of LPS in the outer leaflet provides a highly negatively charged hydrophilic nature to the bacterial membrane. Given that the complex outer membrane environment influences the *in vivo* OmpF assembly, the *E. coli* LPS force field parameter set should be accurately benchmarked before the *in silico* characterization of OmpF

assembly. A well-parameterized membrane force field should provide agreement with the available experimental data within the statistical uncertainty in the measurement.

#### 4.4.1.1 Area per lipid

The  $A_L$  quantifies the two-dimensional density of a membrane and captures the membrane phase behavior.<sup>29,30</sup> The correct prediction of  $A_L$  value enables us to assess the quality of the parameterization of the force field because the  $A_L$  value is sensitive to molecular-level interactions (lipid-lipid and lipid-water interface). The  $A_L$  is a highly averaged equilibrium property of a membrane and is often computationally expensive to calculate in an atomistic simulation due to slow diffusion of lipids in a bilayer.<sup>23</sup> Thus, a simplistic approach to compute  $A_L$  is to divide the cross-sectional area of the equilibrated membrane by the total number of lipids in each leaflet. Using the CG parameter set developed here, we were able to equilibrate the *E.coli* outer membrane and compute  $A_L$ .

The  $A_L$  values for MT1 and MT2 at 323 K after 2  $\mu$ s of equilibration are  $1.39 \pm 0.02$  and  $1.53 \pm 0.06$  nm<sup>2</sup>, respectively (Figure 4-7A). These values compare well to published  $A_L$  values of 1.38–1.56 nm<sup>2</sup> (Table 4-3).<sup>14,42,53</sup> Given that  $A_L$  is a function of membrane composition and temperature, it is not surprising that there is a 0.18 nm<sup>2</sup> variation in experimentally observed  $A_L$  values. The  $A_L$  for LPC-rich MT2 should be higher than MT1 because the charged core oligosaccharide domain has four additional negative charges per lipid A leading to higher electrostatic repulsion and larger lipid-lipid separation, whereas in the LPA-rich MT1, the lipid tails predominantly interact via van der Waals forces. Wu *et al.* using atomistic simulations also showed that  $A_L$  increases with the addition of core domain.<sup>14</sup> The  $A_L$  values for MT3 at 310 K after 2  $\mu$ s of equilibration is  $1.86 \pm 0.06$

$\text{nm}^2$ ; this value is higher than MT1 and MT2 (Table 4-3) because the heterogeneity of the mixture of lipids in the lower leaflets leads to looser packing and a higher area per lipid.<sup>38</sup>

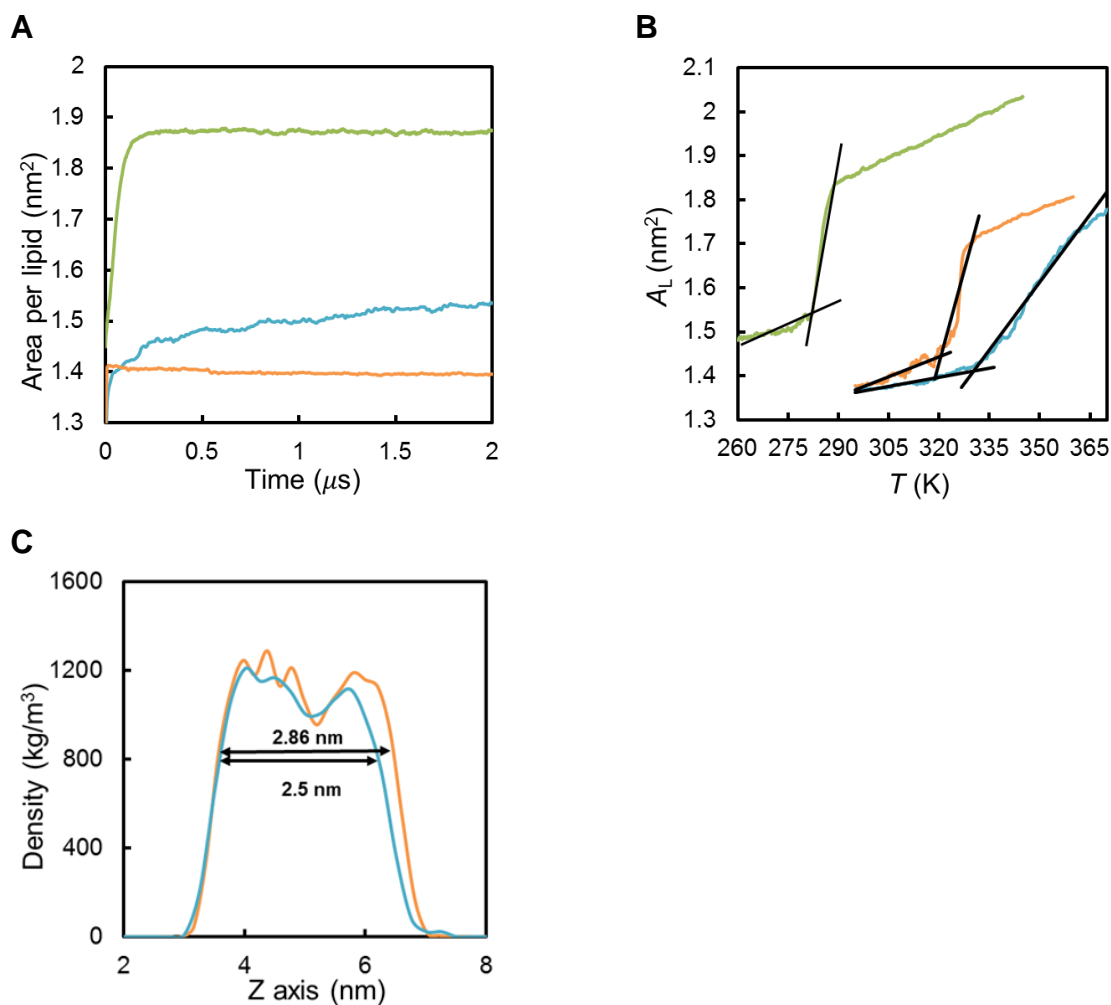


Figure 4-7. (A)  $A_L$  of outer leaflet lipids in MT1 (orange), MT2 (cyan), and MT3 (green) membranes in (B) Effect of temperature on  $A_L$  of outer membrane (same color scheme). (C) Hydrophobic thickness of membrane in Systems 1 and 2.

Table 4-3 Comparison of area per lipid of lipopolysachharides, phase transition temperature and membrane thickness, and hydrophobic thickness data of MT1, MT2, and MT3 models with available data in the literature.

Properties	MT1 <sup>a,b</sup>		MT2 <sup>a,b</sup>		MT3 <sup>a,b</sup>	
	This work	Ref.	This work	Ref.	This work	Ref.
$A_L$ (nm <sup>2</sup> )	1.39±0.02	1.38/1.51/1.56	1.53±0.03	1.8	1.86±0.06	
$T_m$ (K)	319±2	317.5/314.15	328±1	-	285±3	
$D_M$ (nm)	4.28±0.11	4	6.11±0.18	6.1/6.6	3.92±0.09	
$D_H$ (nm)	2.86±0.04	2.3	2.50±0.09	2.2	2.32±0.02	2.4

<sup>a</sup>Simulations to determine equilibrium properties  $A_L$ ,  $D_M$ , and  $D_H$

<sup>b</sup>Annealing simulations to determine  $T_m$

#### 4.4.1.2 Phase transition temperature

The bacterial outer membrane can be in a gel (ordered) or liquid (disordered) phase depending upon whether it's characteristic phase transition temperature ( $T_m$ ) is lower or higher than ambient temperature. The phase transition is a rapid physical change in the membrane's properties as a function of temperature; it depends on multiple factors such as membrane composition, charge per lipid, length and number of acyl chains, degree of unsaturation in lipid chains, and branching in core domain. Each bacterial species has a characteristic  $T_m$ .<sup>38</sup> Bacterial species are often able to manipulate the makeup of their membrane lipids to adapt to the surrounding temperature for survival.<sup>32,41</sup>

For an equilibrated membrane, the  $A_L$  for a lipid increases with increases in temperature; therefore,  $A_L$  of a disordered liquid phase membrane is significantly larger than in gel phase due to the weaker interactions between lipid A chains. The change in  $A_L$  is gradual with temperature until  $T_m$  is achieved, where gel-liquid phase change occurs. Once the temperature is above the  $T_m$ , a rapid increase of  $A_L$  occurs due to the phase change, which is captured from the change in the slope of

$A_L$  versus  $T$  plot (Figure 4-7B). The three model membranes (MT1-MT3), have  $T_m$  ranging from 285 to 328 K based on the composition of outer and inner leaflets (Table 4-3). These results are consistent with our earlier work, where  $T_m$  of LPA-complex or MT3 membrane of *N. meningitidis* (hexacyl Lipid A) was lower ( $284 \pm 2$ ) compared to LPA-DPPE or MT1 ( $324 \pm 3$  K).<sup>38</sup> The  $T_m$  of *E. coli* LPS-DPPE membrane was observed to be 317.15 K in atomistic simulations.<sup>53</sup> Naumann *et al.* estimated  $T_m$  to be 314.15 K via Fourier-transform infrared spectroscopy.<sup>54</sup> However, limited data were reported about the  $T_m$  of complete outer membrane of *E. coli* with core oligosaccharide.

#### 4.4.1.3 Other structural properties

Density profile of water, lipid tails, phosphate, and counter ions in MT1-MT3 systems were computed (Figure 4-8) to examine the effect of LPA and LPC in the outer leaflet, and the influence of a mixture of lipids in the lower leaflet. MT2 shows penetration of counter ions into the outer LPC leaflet to interact with the core oligosaccharide phosphates. The difference between the phosphate peaks of the outer and inner leaflets of the membrane was used to calculate the membrane thickness of the membrane, which is  $4.28 \pm 0.11$  and  $3.92 \pm 0.09$  nm for MT1 and MT3, respectively.

The MT2 membrane with the additional core domain in the outer leaflet lipids is considerably thicker ( $6.11 \pm 0.18$  nm) than the both MT1 and MT3, as expected. In comparison to the *Pseudomonas aeruginosa* LPC-DPPE membrane type the membrane thickness value lies within the previously reported 6.1-6.6 nm range.<sup>36</sup> The density profile of water showed that the membranes are impermeable to water as the density of water decreases to zero close to the membrane mid-plane between 4-6 nm, although water can penetrate the core domain of LPA and polar head groups of LPA and DPPE.

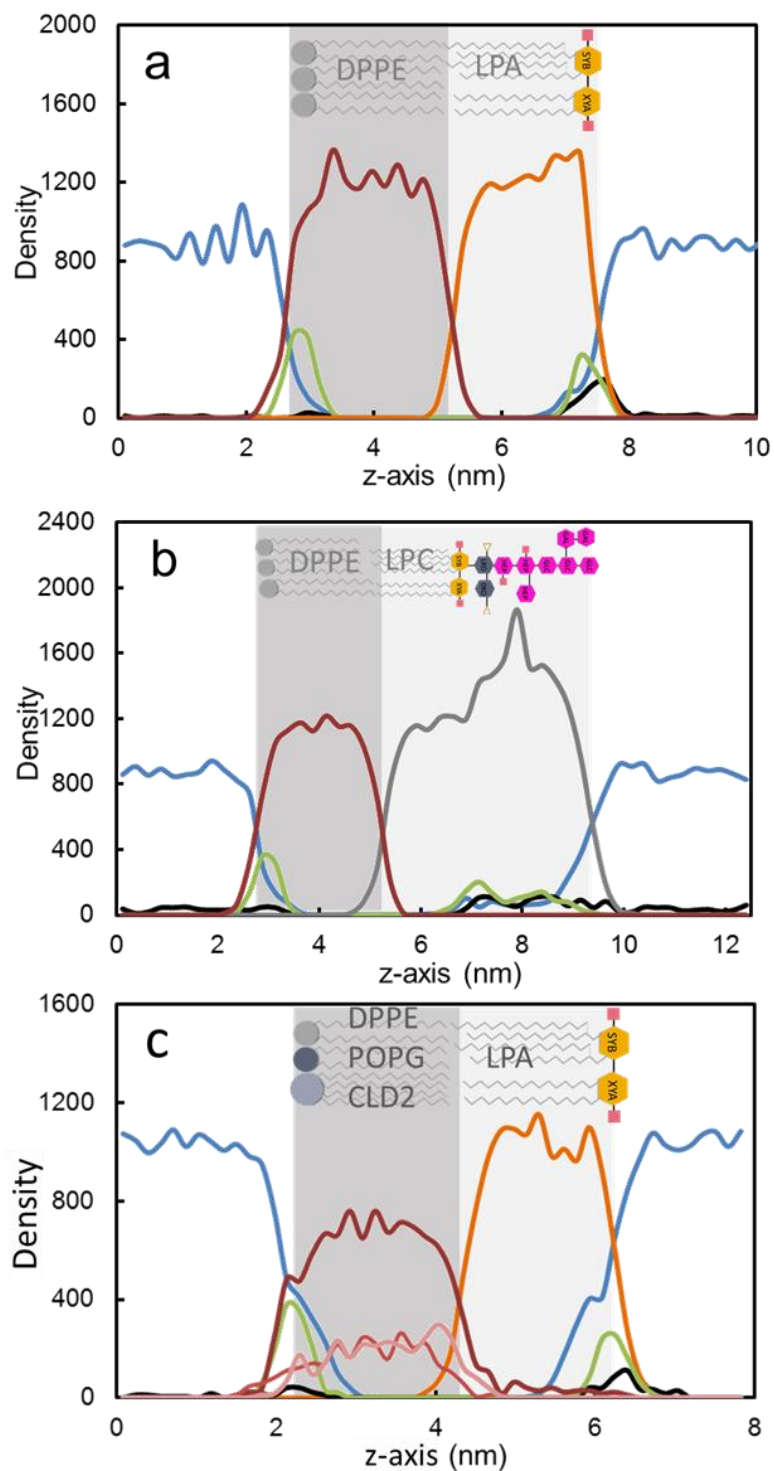


Figure 4-8. Density profile of (a) MT1, (b) MT2 and (c) MT3. Color scheme: LPA (orange line) and LPC (gray line); DPPE (red line); POPG (light red line); CDL2 (pink line); phosphates (green line); ions (black line); water (blue line); inner DPPE leaflet (dark gray, shaded area) and outer leaflet (light gray shaded area). The inset cartoon images of lipids in the shaded areas are provided as a guide.

The hydrophobic thickness of MT1 and MT2 bilayers was calculated from the density profile of the acyl chains (Figure 4-7C). The MT1 bilayer, having only the lipid A domain in the outer leaflet, has a larger (2.9 nm) hydrophobic thickness compared to the MT2 bilayer (2.5 nm) with the additional core domain linked to lipid A. This observation demonstrates the inverse correlation of the hydrophobic thickness with the  $A_L$ , because higher area per lipid leads to smaller hydrophobic thickness, as reported previously.<sup>14</sup>

#### **4.4.2 OmpF self-assembly simulations**

##### **4.4.2.1 Dimer and trimer formation**

The dynamics of OmpF assembly in the outer membrane revealed a two-step process involving formation of an intermediate dimer that leads to the formation of a trimer. Several events involving OmpF binding and unbinding were observed (Figure 4-6) as the proteins diffused through the MT1 and MT3 bilayers. The interactions in which monomers aggregated via the G19-G135 protein-protein interface resulted in stable dimers (Figure 4-6b and 4-6c). The lifetime of the OmpF dimer was dependent on the stability of the G19-G135 protein-protein interface. In most instances monomers OmpF and OmpF' had fleeting interactions and did not lead to stable dimers (Figure 4-6d). Note that the superscript on OmpF' has been used to differentiate the two monomers only for the clarity of the present discussion. To track formation of stable dimers, we computed the G19-G135 separation distances for all interacting OmpF and OmpF' pairs. System 1, 2, and 3 formed stable dimers and achieved an average minimum G19-G135 separation distance of  $2.05 \pm 0.05$  nm (Figure 4-9). Other OmpF-OmpF' dimer interactions in which association did not occur via G19-G135 interfaces, the G19-G135 pair separation remained larger than  $2.05 \pm 0.05$  nm (Figure 4-9) during the entire trajectory, and dimer association was short-lived (Figure 4-6d). In a series of studies involving OmpF oligomerization, the G19 and G135 patches were identified as weakly

stable regions of the OmpF  $\beta$ -barrel that interface with another monomer to acquire stability. Naveed *et al.* showed that site-directed mutagenesis of G19 with energetically unfavorable residues resulted in OmpF mutants with only monomers.<sup>24</sup> Although simulations of the self-assembly of mutated OmpF are beyond the scope of the present work, the observed G19-G135 dimer is consistent with these previously reported results.

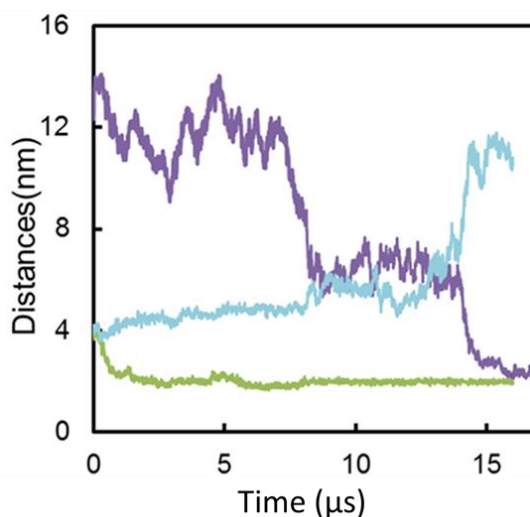


Figure 4-9. Distances between interfacial G19 and G135 residues in Fig.4-6b (purple), Fig.4-6c (green) and Fig. 4-6d (blue).

Analysis of the number density plots of equilibrated lipids in system 3, showed a high density corona of LPA molecules around the OmpF monomers (Figure 4-10a). The presence of LPA, however, did not prevent the oligomerization or arrest diffusion of OmpF in the membrane (Figure 4b–d). The plots exhibit asymmetry in the localization of upper and lower leaflet lipids and their adaptation to the OmpFs. The negatively charged head groups of LPA interact with the charged loop regions of the OmpFs, but do not appear to stabilize the weakly stable G19 and G135 patches, because in cases where a stable dimer is formed, the interacting OmpF interface gradually



delipidates (Figure 4-10a). The lower leaflet DPPE molecules do not form a corona around the OmpF monomer at the beginning of the simulation or after the dimer is formed (Figure 4-10b). Similarly, the negatively charged POPG and CDL2 also do not form a high density corona, which may be due to their low concentration in compared to DPPE in the lower leaflet (Figure 4-10 c–d). The counterions interact with the charged loop domains and form hot spots at the dimer interface due to localization of the OmpF loops.

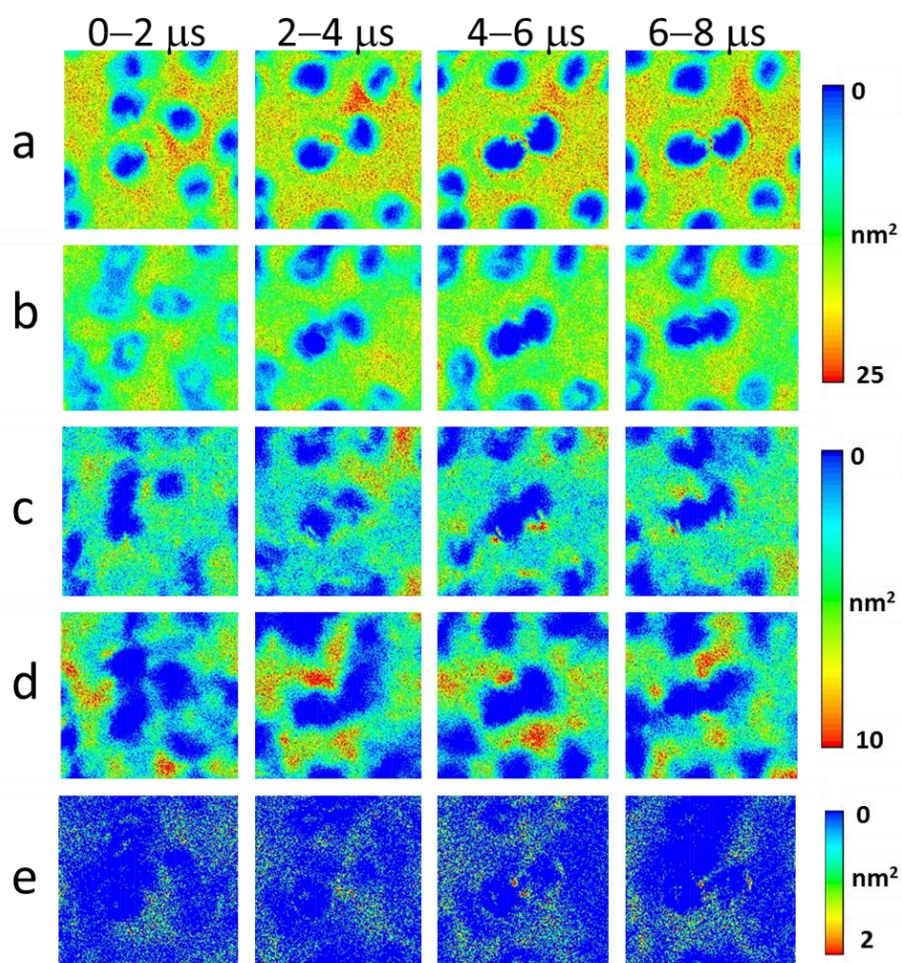


Figure 4-10. Average partial number density of membrane components (a) LPA (outer leaflet), (b) DPPE (inner leaflet), (c) POPG (inner leaflet), (d) CDL2 (inner leaflet), and (e) ions (membrane surface) during two microsecond intervals of trimer formation (System 3).

Structural analysis of the assembled dimer in System 3 shows lack of symmetry (Figure 4-11a). In addition to the obvious asymmetry caused by the G19-G135 interface, the OmpF and OmpF' channel lumens also have different diameters in the bottleneck region of 0.71 and 0.84 nm, respectively. The overall root-mean-square deviations (RMSD) of individual monomers are also different; the wider lumen monomer has higher flexibility (Figure 4-11b). Evidence of asymmetric conductance through the dimeric pore has been reported previously.<sup>22</sup> The asymmetry was observed in all dimers, including those formed in Systems 1 and 2.

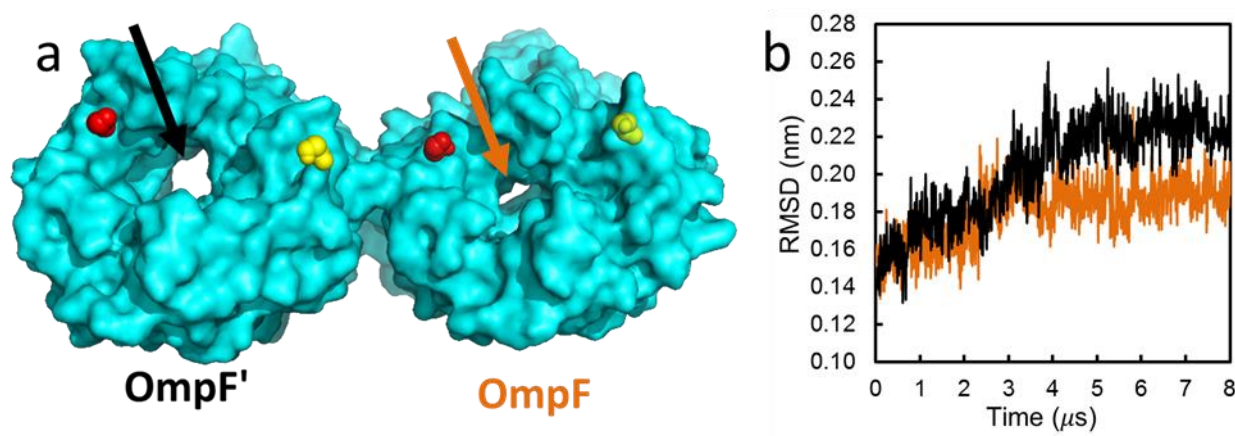


Figure 4-11. Reversed mapped snapshot of the self-assembled OmpF-OmpF' dimer (cyan; surface representation) in System 3. Difference in (a) pore lumens (arrows) and (b) the overall root-mean-square deviation of individual OmpF (orange) and OmpF' (black) units in the dimer as a function of time. The G19 (red) and G135 (yellow) residues are shown as beads.

Furthermore, the asymmetric orientation of the dimer leaves two exposed patches, G135 on OmpF and the G19 on OmpF', which together can accommodate a third OmpF chain to yield a trimer. Simulation results show trimer formation after 16 μs (Figure 4-6e), where the OmpF-OmpF' dimer gradually orients to interface with the OmpF'' monomer to form the trimer (OmpF-OmpF'-OmpF'') with C3 symmetry. The number density profile of the membrane lipids involved in trimer

formation (System 6) shows a thin LPA corona around the OmpFs throughout the assembly (Figure 4-12a). In contrast, the lower leaflet lipids do not show a preference to localize around the trimer (Figure 4-12 b–d). The lipids form a tight seal around the trimer, which is evident from the membrane mid-plane water number density profile that shows three distinct water filled channels formed by the lumens of the OmpFs (Figure 4-12 e). The exposed OmpF loops electrostatically attract counter ions and form a high density charge ring at the entrance of the trimer (Figure 4-12f).

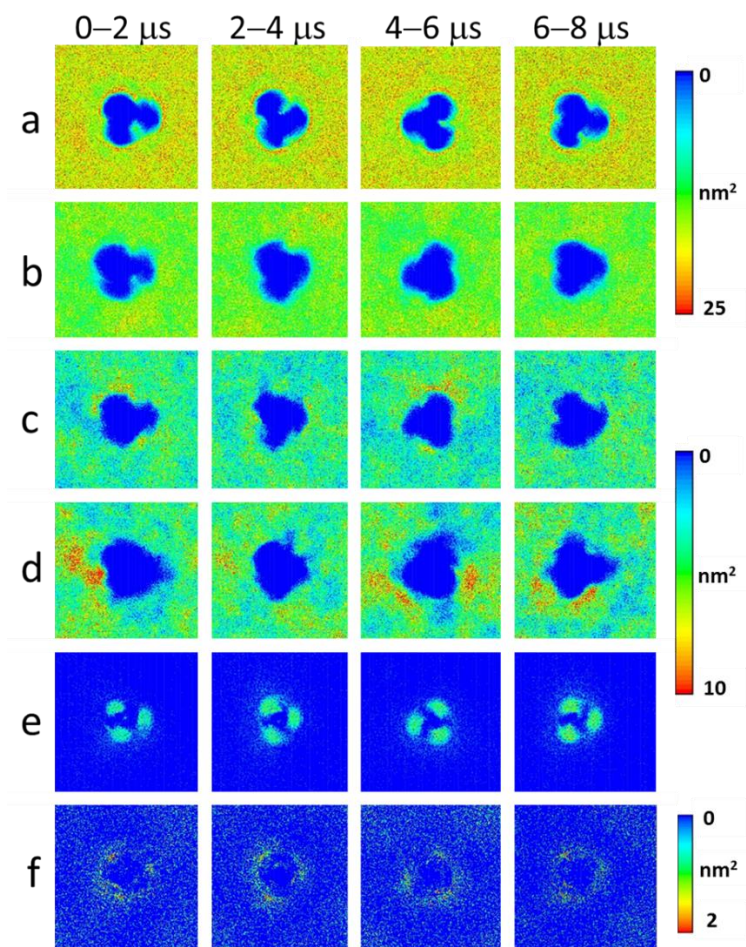


Figure 4-12. Average partial number density of membrane components (a) LPA (outer leaflet), (b) DPPE (inner leaflet), (c) POPG (inner leaflet), (d) CDL2 (inner leaflet), (e) water (membrane midplane), and (f) ions (membrane surface) during two microsecond intervals of trimer formation (System 6).

The root mean square fluctuations (RMSF) of the amino acid residues were computed over the ensemble of structures throughout the simulation (Figure 4-13a). The RMSF values provide unique information about the oligomeric state of a protein. For example, dynamic residues exhibit larger fluctuations from their reference structure and can be a signature of instability. Notably, the L2 residues in the OmpF dimer (Figure 4-13) show large fluctuations (0.4 nm) compared to the other loops and the  $\beta$ -strand regions. Since L2 has a significant role in latching one monomer to its neighbor, these large fluctuations become attenuated with trimer formation. Higher RMSF values were observed for a few residues (6, 52 and 304), in the OmpF turns did not diminish upon oligomerization. In the dimeric state, the OmpF-OmpF' contact is not fully established and the G19-G135 intermonomer distance averages 2.05 nm, whereas in the fully formed trimer, the separation distance reduces to 1.22 nm (Figure 4-13b). Evidence of high structural flexibility of the monomer was obtained computing the RMSD of the structures in each oligomeric state (Figure 7c). The average RMSD values for the trajectory were consistently higher for the monomer (0.25 nm), followed by the dimer (0.20-21 nm), and then the trimer (0.18-0.19 nm), indicating higher structure flexibility of monomer and a stable trimer, as expected. Another indicator of protein structural flexibility is the radius of gyration ( $R_g$ ), which refers to the distribution of the components of an object around a center of mass of the molecule. The  $R_g$  provides a measure of the compactness of OmpF porins in different oligomeric states (Figure 4-13d). The average  $R_g$  values per monomer for monomeric, dimeric, and trimeric forms are 2.12, 2.09, and 2.04 nm, respectively. The OmpF trimer with lowest  $R_g$  exhibits the tightest packing, consistent the RMSD data.



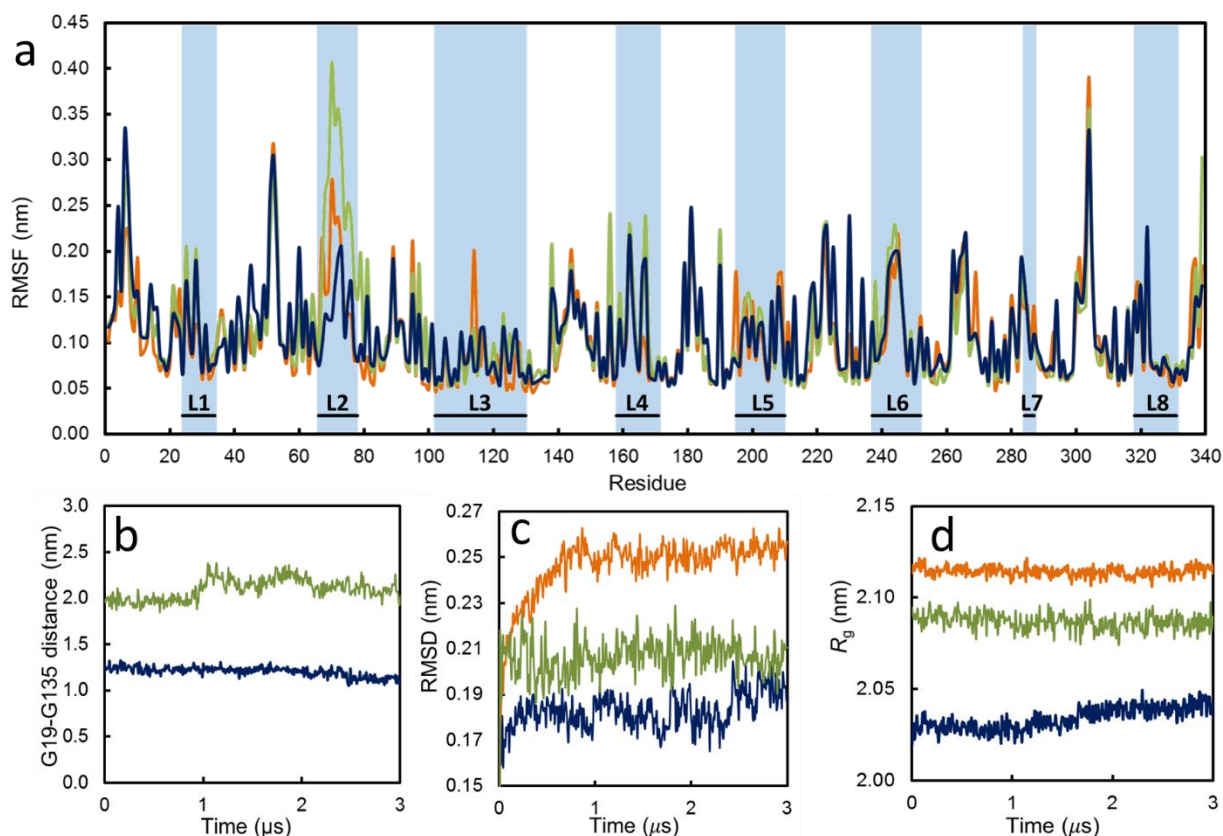


Figure 4-13. Dynamical properties of OmpF monomers<sup>a</sup> (orange, System 3), dimer<sup>b</sup> (green, System 3) and trimer<sup>c</sup> (blue, System 6) during self-assembly. (a) Root-mean-square fluctuations of backbone beads, (b) G19–G135 pair separation distance, (c) overall root-mean-square deviation, and (d)  $R_g$  of the protein as a function of time. In panel (a), the shaded (light blue) regions show the OmpF loop domains (L1–L8).

#### 4.4.2.2 Comparison of self-assembled OmpF trimer and native structure

The detailed analysis of the structural and thermodynamic properties of the self-assembled OmpF trimer with the X-ray crystal structure (pdb:4LSF) showed remarkable agreement. The global RMSD of the assembled OmpF relative to X-ray structure was 2.161 Å, which is a relatively small number indicating a very high similarity. Further comparison of the C $\alpha$ -backbone for individual monomers relative to the X-ray structure shows 1.99–2.11 Å RMSD (Table 4-4). The  $\beta$ -barrel, loops, and turn motifs of the assembled structure also show small deviations ranging from 1.12–1.89 Å. The extracellular loops are longer and have more flexibility than the periplasmic turns so

it is expected that their deviations are slightly larger than the turns. Further, the total and buried surface areas of the trimers were computed for comparison. The buried surface area provides a measure of the surface-to-surface contacts between the monomers within the trimer. The total surface area of the assembled trimer was 40483 Å<sup>2</sup>, which is within 0.4% of the crystal structure's surface area of 40296 Å<sup>2</sup>. Similarly, the buried surface area of the assembled trimer (8995 Å<sup>2</sup>) was also found to be within 1.5% of the native structure (8861 Å<sup>2</sup>).

Table 4-4. RMSD (Å) of self-assembled trimer (OmpF-OmpF'-OmpF'') relative to the X-ray crystal structure (pdb:4LSF).

Monomers <sup>a</sup>	Cα	β-barrel	Loops	Turns
OmpF	1.99	1.12	1.73	1.89
OmpF'	2.10	1.18	1.77	1.81
OmpF''	2.11	1.13	1.69	1.88

<sup>a</sup>Labels OmpF, OmpF', OmpF'' represent the three monomers that constitute the trimer.

#### 4.4.2.3 Mechanism and thermodynamics of OmpF oligomerization

The self-assembly simulations revealed the dynamics of the trimer formation, but not the thermodynamics of the process. To compute the thermodynamic stability of a dimer relative to well-separated monomers, umbrella-sampling simulations were performed to dissociate the dimer along the intermonomer separation coordinates. Similarly, the trimer was dissociated into a dimer and a monomer. The dissociation of the dimer into two well-separated monomers required 26±1 kcal/mol, whereas dissociating a trimer into a dimer and monomer required 56±4 kcal/mol (Figure 4-14). Both simulations were performed while the proteins were embedded in LPA-phospholipid membrane to capture the contribution from the asymmetric membrane environment. The importance of membrane asymmetry has been emphasized in prior computational and *in vivo* experimental studies involving protein folding and assembly.<sup>55-56</sup>

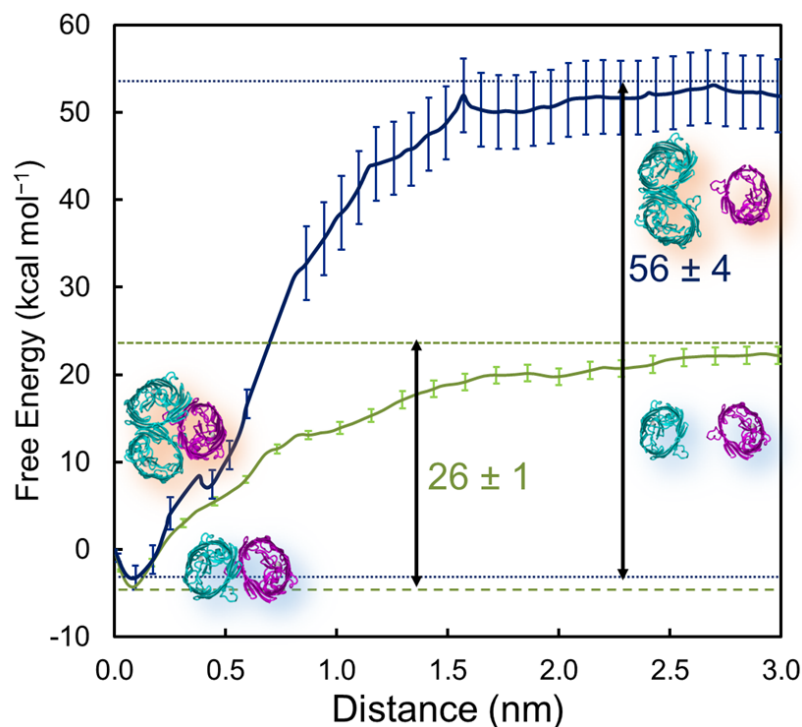


Figure 4-14. Potential of mean force profiles for unbinding of the dimer (System 8) and trimer (System 9) interfaces along the interseparation distance coordinate of the interacting pair. Error bars calculated by Bayesian bootstrapping method are shown in the same color.

The thermodynamic stability of the dimer relative to monomers in the LPA-phospholipid membrane indicates that the first step of the assembly process is the pairing of monomers. This finding is consistent with the *in vitro* and *in vivo* ability of OMPs to form oligomers in the outer membrane.<sup>22</sup> Our results indicate that two monomers diffusing independently in the outer membrane will form a stable dimer if they interact in the correct orientation. Interestingly, thermodynamic data suggests that the further oligomerization of dimer to trimer can occur via two different pathways: (a) Path I—the dimer interacts with a monomer to form the trimer (observed in the simulations), and (b) Path II—the dimer interacts with another dimer to form a trimer and a lone monomer (not observed in the limited simulation time). Although both paths are feasible (Figure 4-15), Path II involves interaction between two slowly diffusing dimers that need to

interact in the correct orientation to cause dissociation of one dimer in the process of forming a trimer. The transition state for dimer-dimer interaction is expected to have a higher barrier. It is, therefore, not entirely surprising that Path II was not observed in the 48  $\mu$ s of simulation time. However, another factor contributing to the choice of Path I or II would be the relative population of monomers versus dimers in the bacterial membrane during oligomerization.

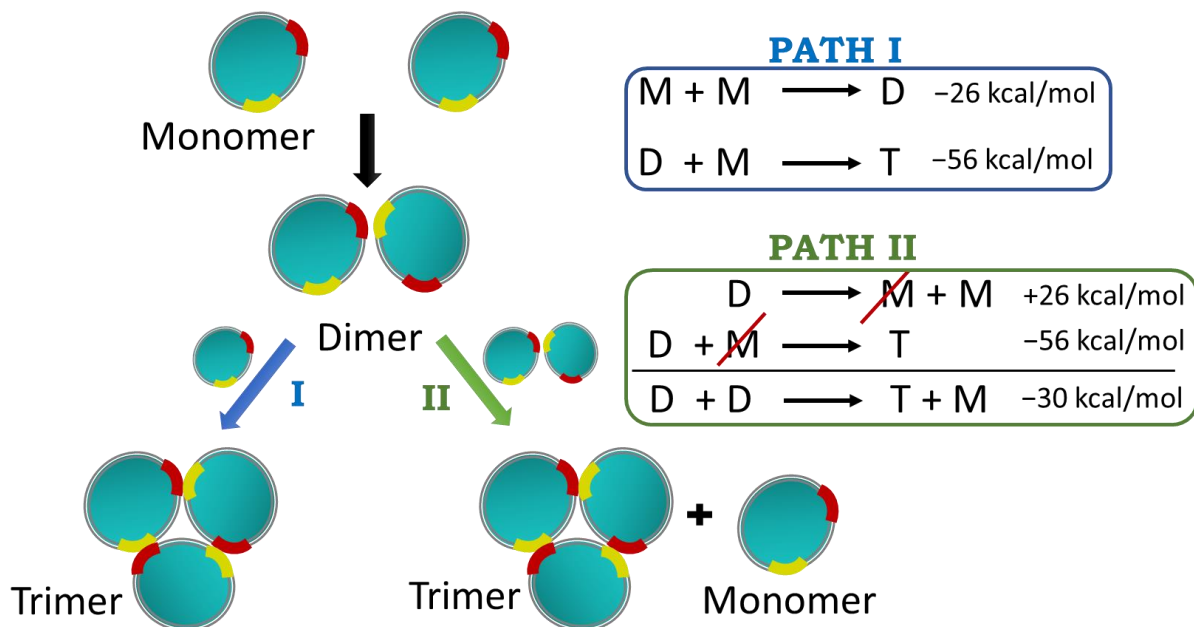


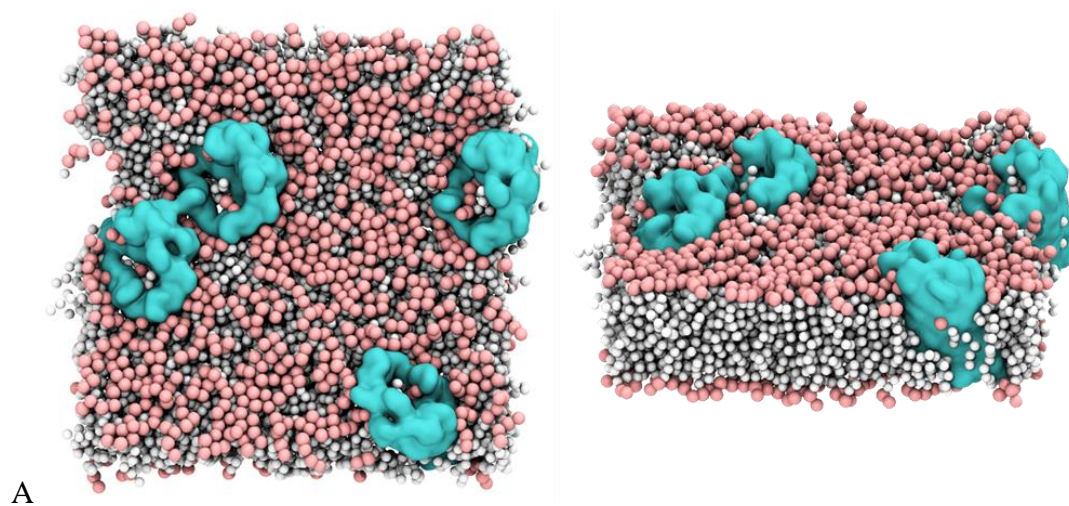
Figure 4-15. A schematic showing two possible pathways that lead to OmpF trimer formation.

#### 4.4.2.4 Interaction of OmpF timer with membrane lipids

The membrane lipids provide a tight seal around the OmpF to direct the passage of water-soluble nutrients through the trimeric nanochannels. The lipid A tails interact with the hydrophobic residues of the  $\beta$ -barrel and the phosphates in the lipid A head groups interact with the positively charged residues in porin loops. *In vitro* experiments have demonstrated that irrespective of the core and O-antigen domains, lipid A binds to the OmpF trimer. The cross-sectional view of the membrane midplane shows a water-filled OmpF lumen with water absent in the rest of membrane



(Figure 4-12). The water density profile shows the distinct footprint of homotrimeric water channels for LPC membrane (Figure 4-12e). Overall, we observe that like many integral membrane proteins, OmpF porins assemble to form oligomeric structures in lipid microenvironment. Despite the extensive hydrogen-bond network that maintains the  $\beta$ -barrel tertiary structure of the OmpF, specific weakly stable regions of the  $\beta$ -barrel remain, which drive OmpF oligomerization. OmpF has been observed as a trimer in both *in vivo* and *in vitro* experiments. The assembly of OmpF monomer into oligomers was observed in MT1 and MT3 bilayers (Systems 1, 2, 3, and 6). In systems with LPC outer leaflet (Systems 4, 5 and 7), the diffusion of OmpF monomers at 310 K was limited and even with 48  $\mu$ s of simulation time, stable dimers or trimers were observed. This implies that we will need to perform these simulations at a higher temperature ( $T > T_m$ ) and for longer times. However, formation of oligomers is not restricted to a specific membrane composition. In fact, OmpF oligomers were observed in symmetric phospholipids membranes that we tested (Figure 4-16). In addition, there have been reports of OmpF assembly in asymmetric lipopolysaccharide membranes,<sup>56</sup> vesicles, and detergents.<sup>22</sup> The lack of preference for a membrane environment suggests that the surrounding lipids do not provide the required stability to the OmpF monomer to thwart oligomerization.



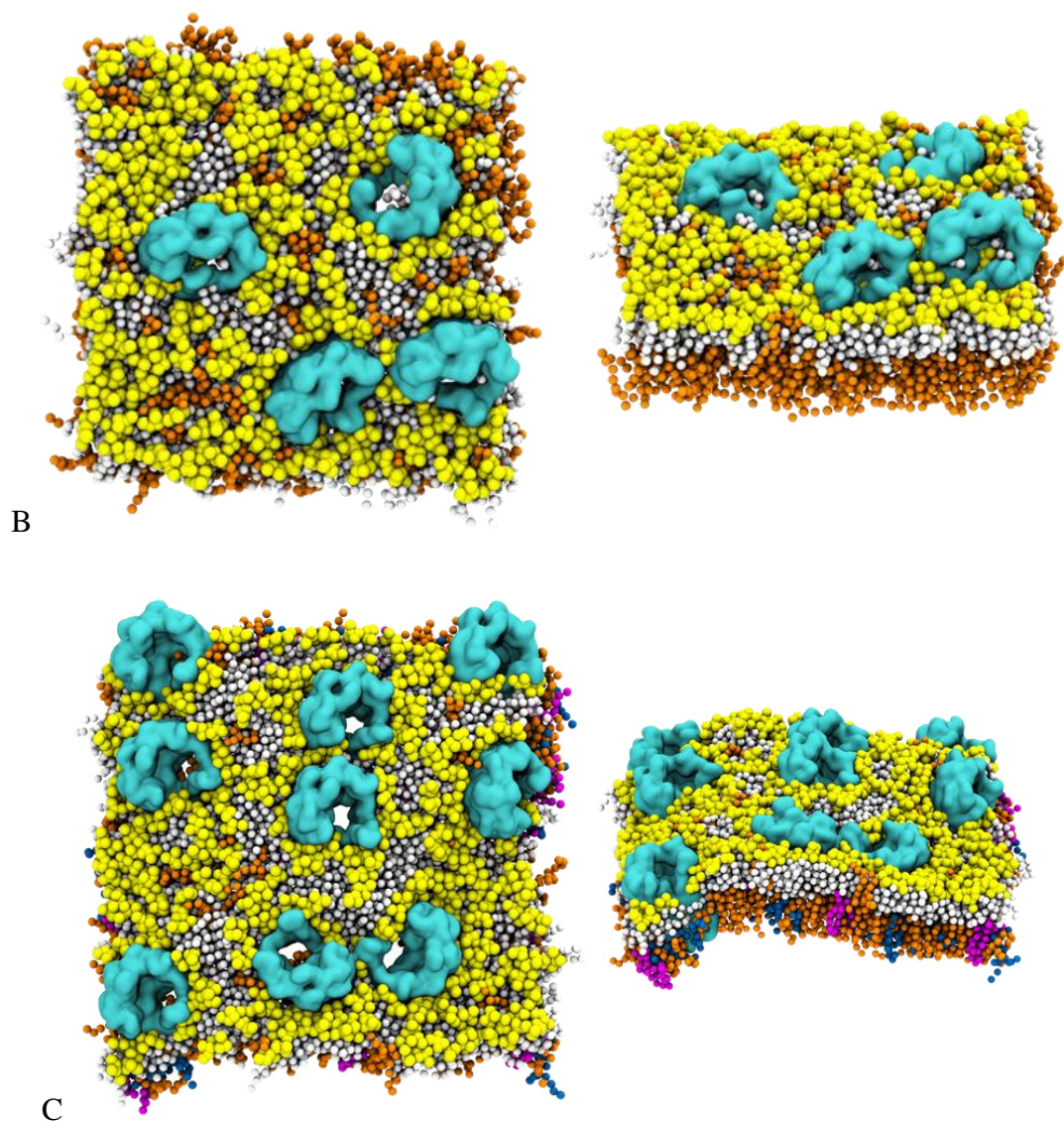


Figure 4-16. Dimer formation in different membranes. A). Dimerization of OmpF monomers in 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) bilayer (OmpF: cyan,  $\text{NH}_3^+$ ,  $\text{PO}_4^-$ : pink, acyl chains: white). B). Dimerization of OmpF monomers MT1 bilayer (OmpF: cyan, SYB, XYA: yellow, acyl chains: white, DPPE: orange). C). Dimerization of OmpF monomers in MT3 bilayer (OmpF: cyan, SYB, XYA: yellow, acyl chains: white, DPPE: orange, POPG: blue, CDL2: magenta).

## 4.5 CONCLUSIONS

The dynamics of OmpF trimer assembly in *E. coli* outer membrane were studied in coarse-grained molecular dynamics simulations. Motivated by the importance of bacterial membrane lipids in protein assembly, we developed a coarse-grained parameter set for *E. coli* membrane lipids. Membrane properties such as area per lipid, hydrophobic thickness, and phase transition temperature were benchmarked against available experimental and computational data. Development of the coarse-grained lipids was crucial in avoiding the computational bottleneck involved in long timescale self-assembly simulations. Multiple simulations starting from OmpF monomers embedded in asymmetric membrane were performed to determine the mechanism and thermodynamics of the OmpF assembly in bacterial outer membrane. Simulations revealed two key steps in OmpF trimer formation. In the first step, two monomers interact via specific complementary protein-protein interfaces to yield an asymmetric dimer, with a  $-26 \pm 1$  kcal mol<sup>-1</sup> free energy change. The root-mean-square fluctuations of the dimer residues show flexibility in loop regions, especially in the L2 latching loop, suggesting that the dimer is not fully structurally stabilized via the single protein-protein interface. In fact, the partially stable dimer acts as a template for the attachment of a third OmpF monomer that yields a C3 symmetric trimeric structure. Formation of the trimer from the dimer and a monomer is the second step of the OmpF oligomerization process and it is associated with a  $-56 \pm 1$  kcal mol<sup>-1</sup> free energy change. The self-assembled trimer showed excellent agreement with the X-ray crystal structure of OmpF trimer with a global root-mean square deviation of less than 2.2 Å. Based on thermodynamic data of the two-step assembly process, an alternate path to trimer formation is presented which involves interaction of two dimers in the second step. Although a dimer-dimer interaction did not yield a trimer in our current set of simulations, likely due to limited simulation time, such a step is

thermodynamically feasible with an associated  $-30 \text{ kcal mol}^{-1}$  free energy change. Overall, we observed that OmpF porins self-assemble to form dimeric and trimeric structures in lipid microenvironments, and oligomerization is not restricted to a specific membrane composition.

## REFERENCE

1. Cowan, S. W.; Schirmer, T.; Rummel, G.; Steiert, M.; Ghosh, R.; Pauptit, R. A.; Jansonius, J. N.; Rosenbusch, J. P., Crystal structures explain functional properties of two *E. coli* porins. *Nature* **1992**, *358*, 727-733.
2. Nikaido, H., Porins and specific channels of bacterial outer membranes. *Mol. Microbiol.* **1992**, *6*, 435-442.
3. Koebnik, R.; Locher, K. P.; Van Gelder, P., Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **2000**, *37*, 239-253.
4. Schulz, G. E., The structure of bacterial outer membrane proteins. *Biochim. Biophys. Acta, Biomembr.* **2002**, *1565*, 308-317.
5. Nikaido, H., Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 593-596.
6. Raetz, C. R. H.; Reynolds, C. M.; Trent, M. S.; Bishop, R. E., Lipid a modification systems in Gram-negative bacteria. In *Ann. Rev. Biochem.* **2007**, *76*, 295-329.
7. Dunton, T. A.; Goose, J. E.; Gavaghan, D. J.; Sansom, M. S. P.; Osborne, J. M., The Free Energy Landscape of Dimerization of a Membrane Protein, NanC. *PLOS Comp. Biol.* **2014**, *10*, e1003417.
8. Dhakshnamoorthy, B.; Ziervogel, B. K.; Blachowicz, L.; Roux, B., A structural study of ion permeation in OmpF porin from anomalous X-ray diffraction and molecular dynamics simulations. *J. Am. Chem. Soc.* **2013**, *135*, 16561-16568.
9. Tieleman, D. P.; Berendsen, H. J. C., A molecular dynamics study of the pores formed by Escherichia coli OmpF porin in a fully hydrated palmitoylcholine bilayer. *Biophys. J.* **1998**, *74*, 2786-2801.

10. Im, W.; Roux, B., Ion permeation and selectivity of OmpF porin: A theoretical study based on molecular dynamics, brownian dynamics, and continuum electrodiffusion theory. *J. Mol. Biol.* **2002**, *322*, 851-869.
11. Im, W.; Roux, B., Ions and counterions in a biological channel: A molecular dynamics simulation of OmpF porin from Escherichia coli in an explicit membrane with 1 M KCl aqueous salt solution. *J. Mol. Biol.* **2002**, *319*, 1177-1197.
12. Roux, B.; Allen, T.; Berneche, S.; Im, W., Theoretical and computational models of biological ion channels. *Q. Rev. Biophys.* **2004**, *37*, 15-103.
13. Danelon, C.; Suenaga, A.; Winterhalter, M.; Yamato, I., Molecular origin of the cation selectivity in OmpF porin: single channel conductances vs. free energy calculation. *Biophys.Chem.* **2003**, *104*, 591-603.
14. Wu, E. L.; Engstrom, O.; Jo, S.; Stuhlsatz, D.; Yeom, M. S.; Klauda, J. B.; Widmalm, G.; Im, W., Molecular Dynamics and NMR Spectroscopy Studies of *E. coli* Lipopolysaccharide Structure and Dynamics. *Biophys. J.* **2013**, *105*, 1444-1455.
15. Niramitrannon, J.; Sansom, M. S. P.; Pongprayoon, P., Why do the outer membrane proteins OmpF from *E. coli* and OprP from *P. aeruginosa* prefer trimers? Simulation studies. *J. Mol. Graph. Model.* **2016**, *65*, 1-7.
16. Kutzner, C.; Grubmüller, H.; de Groot, Bert L.; Zachariae, U., Computational electrophysiology: The molecular dynamics of ion channel permeation and selectivity in atomistic detail. *Biophys. J.* **2011**, *101*, 809-817.
17. Kutzner, C.; Köpfer, D. A.; Machtens, J.-P.; de Groot, B. L.; Song, C.; Zachariae, U., Insights into the function of ion channels by computational electrophysiology simulations. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 1741-1752.

18. Reid, J.; Fung, H.; Gehring, K.; Klebba, P. E.; Nikaido, H., Targeting of porin to the outer membrane of *Escherichia coli*. Rate of trimer assembly and identification of a dimer intermediate. *J. Biol. Chem.* **1988**, *263*, 7753-7759.
19. de Cock, H.; Tommassen, J., Lipopolysaccharides and divalent cations are involved in the formation of an assembly-competent intermediate of outer-membrane protein PhoE of *E.coli*. *The EMBO J.* **1996**, *15*, 5567-5573.
20. Surrey, T.; Schmid, A.; Jahnig, F., Folding and membrane insertion of the trimeric beta-barrel protein OmpF. *Biochemistry* **1996**, *35*, 2283-2288.
21. Watanabe, Y., Effect of various mild surfactants on the reassembly of an oligomeric integral membrane protein OmpF porin. *J. Protein Chem.* **2002**, *21*, 169-175.
22. Visudtiphole, V.; Thomas, Matthew B.; Chalton, David A.; Lakey, Jeremy H., Refolding of *Escherichia coli* outer membrane protein F in detergent creates LPS-free trimers and asymmetric dimers. *Biochem. J.* **2005**, *392*, 375-381.
23. Goose, J. E.; Sansom, M. S. P., Reduced lateral mobility of lipids and proteins in crowded membranes. *PLOS Computational Biology* **2013**, *9*, e1003033.
24. Naveed, H.; Jimenez-Morales, D.; Tian, J.; Pasupuleti, V.; Kenney, L. J.; Liang, J., Engineered oligomerization state of OmpF Protein through computational design decouples oligomer dissociation from unfolding. *J. Mol. Biol.* **2012**, *419*, 89-101.
25. Naveed, H.; Xu, Y.; Jackups, R.; Liang, J., Predicting three-dimensional structures of transmembrane fommains of beta-barrel membrane proteins. *J. Am. Chem. Soc.* **2012**, *134*, 1775-1781.
26. Naveed, H.; Liang, J., Weakly stable regions and protein-protein interactions in beta-barrel membrane proteins. *Curr. Pharm. Des.* **2014**, *20*, 1268-1273.

27. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.
28. Bond, P. J.; Sansom, M. S. P., The simulation approach to bacterial outer membrane proteins *Mol. Membr. Biol.* **2004**, *21*, 151-161.
29. Holdbrook, D. A.; Piggot, T. J.; Sansom, M. S.; Khalid, S., Stability and membrane interactions of an autotransport protein: MD simulations of the Hia translocator domain in a complex membrane environment. *Biochim. Biophys. Acta*, **2013**, *1828*, 715-23.
30. Hsu, P. C.; Bruininks, B. M. H.; Jefferies, D.; Cesar Telles de Souza, P.; Lee, J.; Patel, D. S.; Marrink, S. J.; Qi, Y.; Khalid, S.; Im, W., CHARMM-GUI Martini Maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides. *J. Comput. Chem.* **2017**, *38*, 2354-2363.
31. Hsu, P.-C.; Jefferies, D.; Khalid, S., Molecular dynamics simulations predict the pathways via which pristine fullerenes penetrate bacterial membranes. *J. Phys. Chem. B* **2016**, *120*, 11170-11179.
32. Van Oosten, B.; Harroun, T. A., A MARTINI extension for *Pseudomonas aeruginosa* PAO1 lipopolysaccharide. *J. Mol. Graph. Model.* **2016**, *63*, 125-133.
33. Marrink, S. J.; Tieleman, D. P., Perspective on the Martini model. *Chem. Soc. Rev.* **2013**, *42*, 6801-6822.
34. Shi, C. Y.; Yuan, D. K.; Nangia, S.; Xu, G. F.; Lam, K. S.; Luo, J. T., A Structure-property relationship study of the well-defined telodendrimers to improve hemocompatibility of nanocarriers for anticancer drug delivery. *Langmuir* **2014**, *30*, 6878-6888.



35. Jiang, W. J.; Luo, J. T.; Nangia, S., Multiscale approach to investigate self-assembly of telodendrimer based nanocarriers for anticancer drug delivery. *Langmuir* **2015**, *31*, 4270-4280.
36. Ma, H.; Irudayanathan, F. J.; Jiang, W.; Nangia, S., Simulating Gram-negative bacterial outer membrane: A coarse grain model. *J. Phys. Chem. B* **2015**, *119*, 14668-14682.
37. Jiang, W. J.; Wang, X. Y.; Guo, D. D.; Luo, J. T.; Nangia, S., Drug-specific design of telodendrimer architecture for effective doxorubicin encapsulation. *J. Phys. Chem. B* **2016**, *120*, 9766-9777.
38. Ma, H. L.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan, A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S., Modeling diversity in structures of bacterial outer membrane lipids. *J. Chem. Theory Comput.* **2017**, *13*, 811-824.
39. Nangia, S.; May, E. R., Influence of membrane composition on the binding and folding of a membrane lytic peptide from the non-enveloped flock house virus. *Biochim. Biophys. Acta, Biomembr.* **2017**, *1859*, 1190-1199.
40. Ward, M. D.; Nangia, S.; May, E. R., Evaluation of the hybrid resolution PACE model for the study of folding, insertion, and pore formation of membrane associated peptides. *J. Comput. Chem.* **2017**, *38*, 1462-1471.
41. Kozlikova, B.; Sebestova, E.; Sustr, V.; Brezovsky, J.; Strnad, O.; Daniel, L.; Bednar, D.; Pavelka, A.; Manak, M.; Bezdeka, M.; Benes, P.; Kotry, M.; Gora, A.; Damborsky, J.; Sochor, J., CAVER Analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *Bioinformatics* **2014**, *30*, 2684-2685.

42. Kirschner, K. N.; Lins, R. D.; Maass, A.; Soares, T. A., A Glycam-based force field for simulations of lipopolysaccharide membranes: parametrization and validation. *J. Chem. Theory Comput.* **2012**, *8*, 4719-4731.
43. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33-38.
44. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H., The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812-7824.
45. Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J., The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819-834.
46. de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J., Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theory Comput.* **2013**, *9*, 687-697.
47. Romantsov, T.; Guan, Z.; Wood, J. M., Cardiolipin and the osmotic stress responses of bacteria. *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788*, 2092-2100.
48. Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J., Computational Lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144-2155.
49. Stansfeld, Phillip J.; Goose, Joseph E.; Caffrey, M.; Carpenter, Elisabeth P.; Parker, Joanne L.; Newstead, S.; Sansom, Mark S. P., MemProtMD: Automated insertion of membrane protein structures into explicit lipid membranes. *Structure* **2015**, *23*, 1350-1361.

50. Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R., GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **1995**, *91*, 43-56.
51. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19-25.
52. Krieger, E.; Vriend, G., New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 996-1007.
53. Murzyn, K.; Pasenkiewicz-Gierula, M., Structural properties of the water/membrane interface of a bilayer built of the *E. coli* lipid A. *J. Phys. Chem. B* **2015**, *11*, 5846-5856.
54. Naumann, D.; Schultz, C.; Born, J.; Labischinski, H.; Brandenburg, K.; von Busse, G.; Brade, H.; Seydel, U., Investigations into the polymorphism of lipid A from lipopolysaccharides of *Escherichia coli* and *Salmonella minnesota* by Fourier-transform infrared spectroscopy. *Eur. J. Biochem.* **1987**, *164*, 159-169.
55. Hagge, S. O.; de Cock, H.; Gutschmann, T.; Beckers, F.; Seydel, U.; Wiese, A., Pore formation and function of phosphoporin PhoE of *Escherichia coli* are determined by the core sugar moiety of lipopolysaccharide. *J. Biol. Chem.* **2002**, *277*, 34247-34253.
56. Arunmanee, W.; Pathania, M.; Solovyova, A. S.; Le Brun, A. P.; Ridley, H.; Basle, A.; van den Berg, B.; Lakey, J. H., Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. *Proc. Natl. Acad. Sci. U S A* **2016**, *113*, E5034-43.

## **CHAPTER 5**

### **The mechanism of antimicrobial activity of choline-based ionic liquids (CAGE)**

---

\* Permission granted by Kelly Ibsen to use the text of published article below in this chapter.

Ibsen, K. N.; Ma, H.; Banerjee, A.; Tanner, E. E.; Nangia, S.; Mitragotri, S., Mechanism of Antibacterial Activity of Choline-Based Ionic Liquids (CAGE). *Acs Biomaterials Science & Engineering* **2018**, 4(7), 2370-2379.

## 5.1 ABSTRACT

The continued emergence of antibiotic-resistant organisms has severely depleted our arsenal of effective antimicrobials. Ionic liquids (ILs), or molten salts, and their close relatives, deep eutectic solvents (DESs), show great promise as antibacterial agents. Understanding the mechanism by which ILs and DESs attack bacterial cells is key to ensuring that design of IL-based biocides impart maximum efficacy with minimal toxicity, while also avoiding the potential for the target organisms to become resistant. Here we report the antibacterial attributes of a set of choline and geranic acid (CAGE)-based ILs and DESs and identify the mechanism by which they interact with the Gram-negative cell wall of *Escherichia coli*. Four CAGE variants with varying ratios of choline and geranic acid were synthesized and tested for their antibacterial activity (1:4, 1:2, 1:1 and 2:1 choline:geranic acid). The minimum bactericidal concentration required to kill *E. coli* correlated with the geranic acid content. Using molecular dynamics (MD) simulations, we identified the mechanism of CAGE action on the *E. coli* membrane, namely that choline is attracted to the negatively-charged cell membrane and consequently inserts geranic acid into the lipid bilayer. This study provides the fundamental mechanism of the action of choline-based ILs on bacteria, and demonstrates the promise of CAGE as a powerful antimicrobial agent.

## 5.2 INTRODUCTION

Ionic liquids (ILs), also known as molten salts, are a broad class of compounds most commonly described by their low melting points ( $<100\text{ }^{\circ}\text{C}$ ) and low volatility. Common IL cations (imidazolium, pyridinium, ammonium and phosphonium) can be combined with a variety of anions, all of which can be further functionalized, to create a diverse set of compounds. Over the last two decades, ILs have become popular as green alternatives to volatile organic solvents used in the chemical industry.<sup>1</sup>

ILs have been recognized as effective disinfectants for almost a century, but their systematic investigation as antibacterial and antifungal agents is a relatively new phenomenon. In 1996 Pernak and Skrzypczak reported a correlation between the concentration of an imidazolium chloride IL and its minimum inhibitory concentration against bacteria.<sup>2</sup> Other studies followed to confirm this relationship and provide a second mechanistic hypothesis, namely that an IL's antibacterial activity is correlated with the chain length of its alkyl chain.<sup>3-9</sup> Several studies postulated that aliphatic chains of ILs insert into the bacterial membrane with a mechanism similar to that used by surfactants or pesticides. Other studies attributed the functionality of ILs to the inhibition of acetylcholinesterase because of the cation.<sup>10, 11</sup>

Molecular dynamics (MD) simulations have been used to gain insights into the action of select imidazolium ILs on model lipid bilayers.<sup>12-14</sup> These simulations showed that imidazolium cations interact with the polar head groups of the lipids and insert their hydrophobic tails into the membrane. The interactions, however, are highly dependent on the charge and structure of the cation, the counter anion, as well as the complexity of the membrane lipids. Literature studies on ILs have focused on simple phospholipid bilayer models as surrogates for bacterial membranes due to the lack of available force fields for the Gram-negative bacterial outer membranes.

Moreover, the atomistic MD simulations have been limited to short timescales (hundreds of nanoseconds) that are unable to provide adsorption kinetics of an IL cation on bacterial membranes. In recent years, there have been advances in coarse-grained force field libraries for bacterial membranes that can aid in elucidating the IL-induced morphological reorganization of the bacterial membranes.<sup>15-19</sup>

Despite experimental and computational investigation into the interaction of ILs with lipid membranes, the exact molecular mechanism remains unknown, including the secondary effects of membrane disruption on cellular signaling and other cellular functions.<sup>20</sup> The lack of a complete mechanistic description hampers the effective development of antimicrobial ILs, especially as it pertains to avoiding imparting resistance. Combining a full mechanistic knowledge with the fact that IL properties (hydrophilicity, hydrophobicity, density, viscosity, conductivity, and polarity) can be widely and readily tuned could provide a wealth of new IL-based antimicrobials with maximum efficacy and minimum toxicity.

Using a combination of experimental and simulation techniques, we investigated the antibacterial activity and mechanism of a set of choline (or cholinium)-based ILs on a Gram-negative bacterium, *Escherichia coli*. While imidazolium and pyridinium-based ILs are widely used as solvents, choline, a quaternary ammonium cation, is generally regarded as more benign, and therefore a good choice for antibiotics. Studies have reported antibacterial properties of choline-based ILs and deep eutectic solvents (DESs, a mixture of charged and neutral species), using a variety of counterions or functionalizing the choline cation.<sup>6,7,21,22,23</sup> Using geranic acid, a highly hydrophobic molecule with an 8-carbon backbone, as a counterion, we synthesized 4 choline-geranic acid (CAGE) formulations, varying the choline bicarbonate and geranic acid ratio: 1:4, 1:2, 1:1 and 2:1. The 1:1 CAGE is a true IL, while the other 3 are DESs composed of a cation:anion,

choline:geranate pair plus additional neutral species, either geranic acid (1:4, 1:2) or choline bicarbonate (2:1). Each variant was tested to determine the minimum concentration required to kill *E. coli*, and MD simulations were performed to compute the interfacial properties of CAGE variants with *E. coli*, as well as choline bicarbonate, pure geranic acid, and sodium-substituted 1:4 CAGE.

### 5.3 METHODS

A cuboidal simulation box comprising of two coarse-grained membrane patches ( $10 \times 10 \text{ nm}^2$ ) in the *xy*-plane with 81 LPC (lipid A and core oligosaccharides without o-antigen) and 243 DPPE lipids each were built using a locally modified version of membrane generator script called *insane*.<sup>24</sup> The two membranes were stacked along the *z*-direction (4 nm spacing) with the LPC leaflets oriented towards the center of the box. The intermembrane space was filled with CAGE and the remainder of the simulation box was solvated with explicit coarse grained MARTINI water.<sup>25</sup> The CAGE components were coarse grained using the PyCGtool and the MARTINI four-to-one mapping protocol.<sup>26,27</sup> The electroneutrality of the system was maintained using hydrated  $\text{Ca}^{2+}$  counterions. The two-membrane setup was adopted to compartmentalize CAGE toward the outer LPS leaflet to mimic experiments and prevent issues that may arise due to periodic boundary conditions along the *z*-direction. The simulation setup was repeated for all seven CAGE variants.

The GROMACS molecular dynamics package (version 5.1.2) was used to perform all simulations.<sup>28,29</sup> Energy minimization was performed using the steepest-decent algorithm with a 20 fs time-step until the maximum force on any bead was below the tolerance parameter of  $10 \text{ kJmol}^{-1}\text{nm}^{-1}$ .<sup>30</sup> Periodic boundary conditions were applied in all three dimensions. The *NVT* (constant number of particles, volume, and temperature) and *NPT* (constant number of particles,



pressure, and temperature) equilibration runs were performed for 0.2  $\mu$ s. Semi-isotropic pressure coupling was used, and systems were maintained at 1 bar using the Berendsen barostat with time constant,  $\tau_p = 4.0$  ps.<sup>31</sup> Temperature was maintained at 335 K by independently coupling the lipids, and the solvent to an external velocity rescaling thermostat with  $\tau_T = 1.0$  ps. The neighbor list was updated every 25 steps using 1.4 and 1.2 nm for short-range van der Waals and electrostatic cutoffs, respectively. The production *NPT* simulations were performed for 8  $\mu$ s or 16  $\mu$ s for all systems.

Post simulation analyses were performed using in-built GROMACS utilities and in-house python scripts. To quantify the CAGE and the membrane interaction, we developed an in-house script to compute the number of contacts that CAGE components make with *E. coli* membrane. In this analysis, we defined the entire *E. coli* membrane as one unit, which includes  $\beta$ -(1 $\rightarrow$ 6)-linked D-glucosamine disaccharide head group linked to six acyl carbon chains, the core oligosaccharide domain and the lower phospholipid leaflet. A contact was counted when a CAGE component molecule was within a 1.1 nm cut-off distance with any part of the membrane. Molecular visualization and graphics were generated using visual molecular dynamics (VMD) software.<sup>32</sup>

## 5.4 RESULTS

The MD simulations were performed in the coarse-grained representation to provide a long timescale comparison of CAGE penetration and partitioning in the *E. coli* outer membrane. The MD simulation results show that the choline geranate pair has a unique cooperative penetration profile into *E. coli* membranes. The quaternary ammonium choline cation with its short hydroxyl alkyl chain is sufficiently small to penetrate the LPS domain and form stable ionic interactions with the negatively charged membrane. The presence of choline also facilitates the penetration of

geranate, which acts as short chain fatty acid chain and inserts itself into the lipid A tails. The negatively-charged head group of geranate remains above the hydrophobic tails and is stabilized by the embedded cholines.

Focusing on the simulation results of choline bicarbonate and 2 CAGE variants—choline bicarbonate, 2:1, and 1:1 (i.e. without free geranic acid), it is clear that the positively charged choline easily penetrates the *E. coli*'s membrane and binds to the negatively charged core and lipid A head groups. The simulation snapshots (Figure 5-1) of the three variants show that in each case, choline is trapped within the negatively charged core and lipid A head groups. A higher choline concentration in 2:1 variant leads to higher density in the core LPS head groups (Figure 5-1b). It is evident from the geranate density profiles of geranate that it penetrates the outer LPS leaflet (Figures 5-1 b-c). The contact plot shows how CAGE components penetrate the membrane as a function of time. In the beginning of the simulation, most of the CAGE components are on the surface so there are fewer contacts, and as time goes by, the CAGE components disperse into the membrane, creating more contacts. The CAGE-membrane contact analysis shows as choline continues to penetrate the membrane until equilibrium is achieved in 6-7  $\mu$ s. Compared to choline, the geranate contacts are 6 to 8 times lower in 2:1 and 1:1 CAGE, respectively. Unlike geranate, the bicarbonate ion, devoid of the alkyl chain, does not penetrate the LPS (Figure 3a), which explains the limited efficacy observed in experiments.

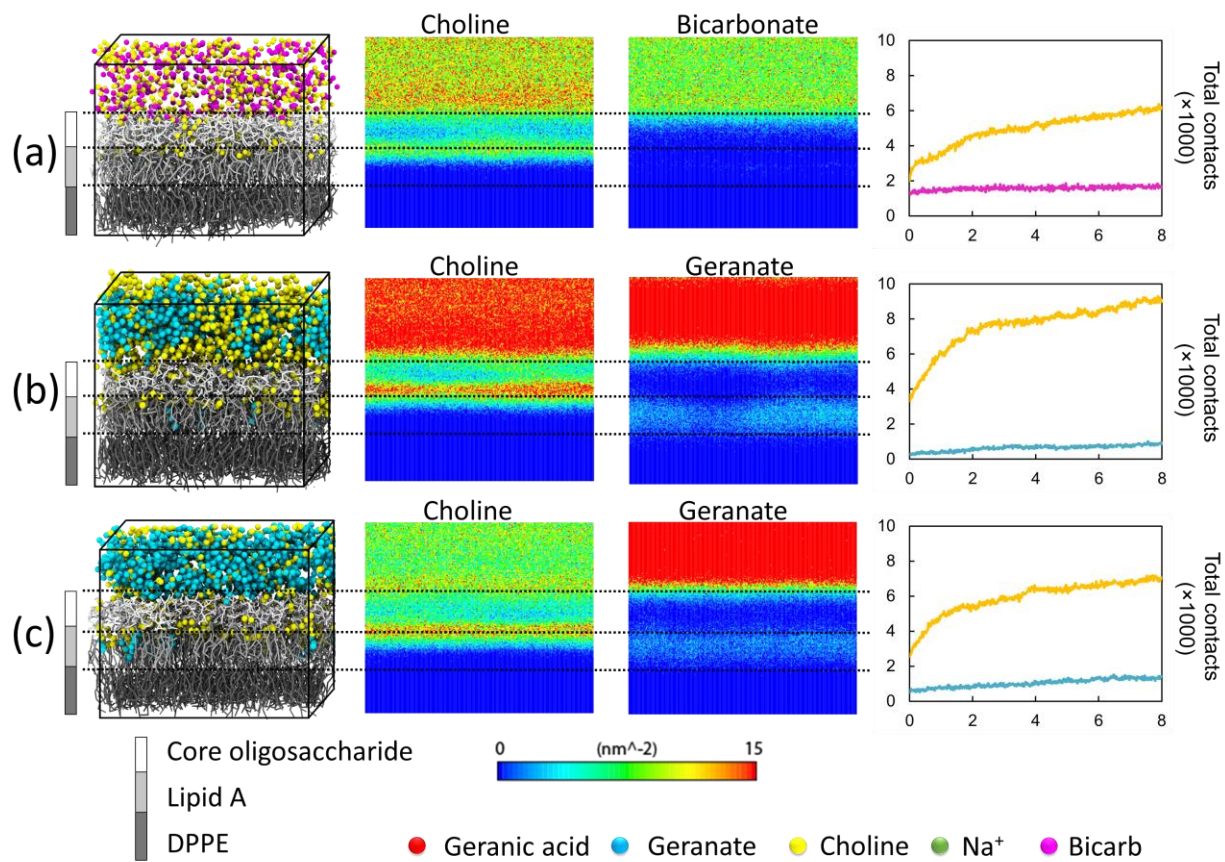


Figure 5-1. Comparison of CAGE variants (a) Choline bicarbonate (b) 2:1 and (c) 1:1. For each variant, the panels show the molecular simulation box, the number density of choline, the number density of bicarbonate or geranate, and the number of contacts as function of simulation time.

In the case of 1:2 and 1:4 CAGE variants, in addition to choline and geranate, the uncharged geranic acid molecules penetrate both the outer and inner leaflets of the membrane (Figures 5-2 a-b). The abundance of geranic acid in 1:4 CAGE results in very high membrane penetration, which explains the high efficacy observed in experiments (Figure 5-2b). The geranic acid penetration in 1:2 CAGE is only half of that observed in 1:4 CAGE (Figure 5-2a). However, in the absence of choline, the penetration of pure geranic acid is drastically reduced (Figure 5-2c).

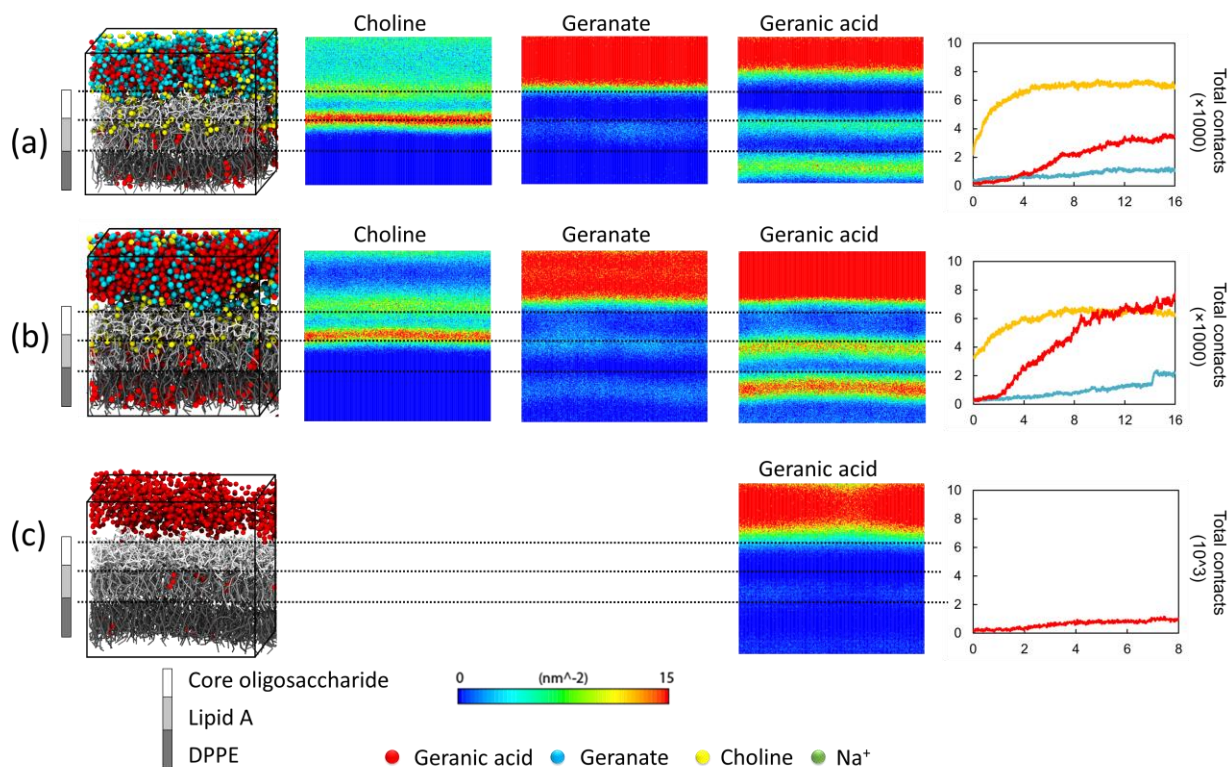


Figure 5-2. Comparison of CAGE variants (a) 1:2 (b) 1:4, and (c) pure geranic acid. For each variant, the panels show the molecular simulation box, the number density of choline, the number density of geranate, the number density of geranic acid, and the number of contacts as function of simulation time.

To further investigate the role of choline, we tested a 1:4 CAGE variant in which choline was replaced with Na<sup>+</sup> ions (Figure 5-3). Just as choline, the Na<sup>+</sup> ions penetrate the LPS core and make similar contacts with the membrane lipids, but being a hard cation, Na<sup>+</sup> ions do not interact cooperatively with softer anions such as geranate and geranic acid molecules to facilitate their penetration. The substitution of choline with Na<sup>+</sup> demonstrate that choline is vital for the penetration of geranate and geranic acid.

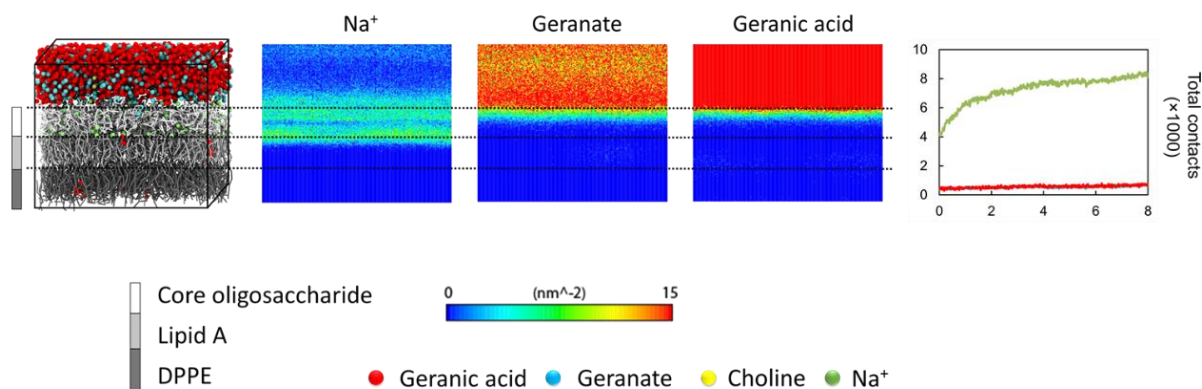


Figure 5-3. Comparison of Na<sup>+</sup> 1: 4 Sodium Geranate. The panels show the molecular simulation box, the number density of Na<sup>+</sup> ions, the number density of geranate, the number density of geranic acid, and the number of contacts as function of simulation time.

## 5.5 DISCUSSION

A full understanding of how a candidate antibiotic acts against a pathogen is of utmost importance to develop antibiotics with high efficacy and low potential to impart resistance. Since the properties of ILs can be finely tuned, a full mechanistic knowledge of their action of cell disruption can allow us to design a wealth of IL-based antimicrobials. CAGE ILs and DESs are easily synthesized via an ambient temperature salt metathesis reaction using commercially sourced and FDA-listed GRAS reactants choline bicarbonate and geranic acid. Choline is a water-soluble essential nutrient, made in the liver, and present in phospholipids that are abundant in cell membranes. Geranic acid, commonly used as a flavoring agent, is naturally occurring in lemongrass, which has reported antimicrobial activity itself.<sup>33</sup>

Mechanistic hypotheses for ILs' antibacterial activity most commonly include cell membrane disruption.<sup>4,8,9,20</sup> Some studies suggest additional signal interruptions as a result, but what actually causes cell death remains unknown. The *E. coli* outer membrane is comprised of a lipopolysaccharide-rich outer leaflet and a 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine

(DPPE)-rich inner leaflet. Each *E. coli* lipopolysaccharide molecule has a core oligosaccharide domain and lipid A domains with six and two negative charges, respectively.<sup>19</sup> The presence of a negatively charged outer leaflet makes the bacterial membrane unique compared to phospholipid bilayers and makes it susceptible to penetration by CAGE. In the MD simulations, the density profiles and the contact analysis of the seven compounds demonstrate that the negatively charged LPS core forms a barrier for geranate and geranic acid. In the presence of choline-containing CAGE variants, however, the LPS negative charge is effectively screened and choline is able to facilitate the geranate and geranic acid penetration into the membrane. Among the seven compounds simulated, 1:4 CAGE has the highest penetration, which explains the high toxicity observed in the experiments. Using the CAGE component penetration as a measure of their efficacy, the simulation results show the following order of CAGE variants toxicity: 1:4 > 1:2 > 1:1 > 2:1 > choline bicarbonate > pure geranic acid > Na<sup>+</sup>1:4. The order corroborates with the experimentally observed CAGE toxicity.

There are no previous studies on cholinium-geranic acid salts, however some groups have investigated other choline-based ILs for their bacterial activity, and our results are generally consistent with these studies. Petkovic, synthesized a group of ILs using a choline cation paired with a range of linear alkanoate anions ( $[C_nH_{2n+1}CO_2]^-$ , where  $n=1-9$ ) and found that the longer anion chains resulted in lower MFC values. Choline chloride, tested as a proxy for the choline cation alone, showed the lowest toxicity.<sup>7</sup> Zhao synthesized a variety of choline-based DESs using choline chloride and several different types of hydrogen-bond donors including organic acids, amines, alcohols and sugars; only the organic acid-containing DES showed bacterial inhibition.<sup>23</sup> The CAGE mechanism of membrane attraction and insertion, while similar to those suggested for cation-substituted ILs, has a unique feature – the hydrophobic long chain can dissociate from the

more bulky, hydrophilic cation and penetrate deeper into the membrane. This dissociation ability may prove useful in developing highly effective antimicrobial ILs. It is also interesting to note that geranic acid is structurally similar to free fatty acids, which have demonstrated bioactivities related to chain length and degree of saturation, but poor solubility.<sup>34</sup> Combining choline, a hydrophilic molecule, with hydrophobic geranic acid may improve its ability to contact cells in aqueous environments like wounds.

## 5.6 CONCLUSIONS

By varying the ion ratios in CAGE, we were able to show that increasing the geranic acid content increases the biocidal activity. Through MD simulations we identified cell membrane disruption via choline attraction to the negatively-charged cell membrane and geranic acid insertion as a disrupting mechanism. Overall, this study provides the basic mechanism for choline-based IL activity on the cell membrane of Gram-negative *E. coli*. CAGE is a promising new antibacterial that kills *E. coli* with low mM concentrations and exhibits no evidence of imparting vertical evolution-based resistance.

## REFERENCE

1. Pendleton, J.N.; Gilmore, B.F. The antimicrobial potential of ionic liquids: A source of chemical diversity for infection and biofilm control. *Int J Antimicrob Agents* **2015**, 46(2), 131-139. DOI: 10.1016/j.ijantimicag.2015.02.016.
2. Pernak, J.; Skrzypczak, A. 3-alkylthiomethyl-1-ethylimidazolium chlorides. Correlations between critical micelle concentrations and minimum inhibitory concentrations. *Eur J Med Chem* **1996**, 31, 901-903.
3. Pernak, J.; Sobaszekiewicz, K.; Mirska, I. Anti-microbial activities of ionic liquids. *Green Chem* **2003**, 5, 52-56, 10.1039/b207543c.
4. Docherty, K.M.; Kulpa Jr, C.F. Toxicity and antimicrobial activity of imidazolium and pyridinium ionic liquids. *Green Chem* **2005**, 7, 185-189, DOI: 10.1039/b419172b.
5. Pernak, J.; Feder-Kubis, J.; Cieniecka-Roslonkiewicz, A.; Fischmeither, C.; Griffin, S.T.; Rogers, R.D. Synthesis and properties of chiral imidazolium ionic liquids with a (1R,2S,5R)-(-)-menthoxymethyl substituent. *New J Chem* **2007**, 31, 879-892, DOI: 10.10139/b616215k.
6. Pernak, J.; Syguda, A.; Mirska, I.; Pernak, A.; Nawrot, J.; Pradzynska, A.; Griffin, S.T.; Rogers, R.D. Choline-derivative-based ionic liquids. *Chem Eur J* **2007**, 13, 6817-6827, 10.1002/chem.200700285.
7. Petkovic, M.; Ferguson, J.L.; Nimal Gunaratne, H.Q.; Ferreira, R.; Leitão, M.C.; Seddon, K.R.; Rebelo, L.P.N.; Silva Pereira, C. Novel biocompatible cholinium-based ionic liquids-toxicity and biodegradability. *Green Chem* **2010**, 12, 643-649, DOI: 10.1039/b922247b.
8. Yu, Y.; Nie, Y. Toxicity and antimicrobial activities of ionic liquids with halogen anion. *J Environ Prot* **2011**, 2, 298-303.



9. Jeong, S.; Ho Ha, S.; Han, S.; Lim, M.; Kim, S.M.; Kim, Y.; Koo, Y.; So, J.; Jeon, T.  
Elucidation of molecular interactions between lipid membranes and ionic liquids using model cell membranes. *Soft Mater* **2012**, 8, 5501-5506, DOI: 10.1039/c2sm25223d.
10. Arning, J.; Stolte, S.; Bösch, A.; Stock, F.; Pitner, W.; Welz-Biermann, U.; Jastorff, B.; Ranke, J. Qualitative and quantitative structure activity relationships for the inhibitory effects of cationic head groups, functionalized side chains and anions of ionic liquids on acetylcholinesterase. *Green Chem* **2008**, 10, 47-58, DOI: 10.1039/b712109a.
11. Torrecilla, J.S.; García, J.; Rojo, E.; Rodríguez, F. Estimation of toxicity of ionic liquids in leukemia Rat Cell Line and Acetylcholinesterase enzyme by principal component analysis neural networks and multiple linear regressions. *J Hazard Mater* **2009**, 164, 182-194. DOI: 10.1016/j.jhazmat.2008.08.022.
12. Bingham, R.J.; Ballone, P. Computational study of room-temperature ionic liquids interacting with a POPC phospholipid bilayer. *J Phys Chem B* **2012**, 116(36), 11205-11216, DOI: 10.1021/jp306126q.
13. Yoo, B.; Jing, B.; Jones, S.E.; Lamberti, G.A.; Zhu, Y.; Shah, J.K.; Maginn, E.J. Molecular mechanisms of ionic liquid cytotoxicity probed by an integrated experimental and computational approach. *Sci Rep* **2016**, 6, 19889, DOI: 10.1038/srep19889.
14. Klähn, M.; Zacharias, M. Transformations in plasma membranes of cancerous cells and resulting consequences for cation insertion studied with molecular dynamics. *Phys Chem Chem Phys* **2013**, 15, 14427-14441, DOI: 10.1039/c3cp52085d.
15. Ma, H.; Irudayanathan, F.J.; Jiang, W.; Nangia, S. Simulating Gram-negative bacterial outer membrane: A coarse grain model. *J Phys Chem B* **2015**, 119, 14668–14682, DOI: 10.1021/acs.jpcb.5b07122.

16. Hsu, P.C.; Jefferies, D.; Khalid, S. Molecular dynamics simulations predict the pathways via which pristine fullerenes penetrate bacterial membranes. *J. Phys. Chem. B* **2016**, 120, 11170–11179, DOI: 10.1021/acs.jpcb.6b06615.
17. Van Oosten, B.; Harroun, T.A. (2016). A MARTINI extension for *Pseudomonas aeruginosa* PAO1 lipopolysaccharide. *J Mol Graph Model* **2016**, 63, 125–133, DOI: 10.1016/j.jmgm.2015.12.002.
18. Ma, H.L.; Cummins, D.D.; Edelstein, N.B.; Gomez, J.; Khan, A.; Llewellyn, M.D.; Picudella, T.; Willsey, S.R.; Nangia, S. Modeling diversity in structures of bacterial outer membrane lipids. *J Chem Theory Comput* **2017**, 13, 811–824, DOI:10.1021/acs.jctc.6b00856.
19. Ma, H.; Khan, A.; Nangia, S. Dynamics of OmpF trimer formation in the bacterial outer membrane of *Escherichia coli*. *Langmuir* **2017**, ASAP, DOI: 10.1021/acs.langmuir.7b02653.
20. Bhattacharya, G.; Giri, R.P.; Saxena, H.; Agrawal, V.V.; Gupta, A.; Mukhopadhyay, M.K.; Ghosh, S.K. X-ray reflectivity study of the interaction of an imidazolium-based ionic liquid with a soft supported lipid membrane. *Langmuir* **2017**, 33, 1295-1304. DOI: 10.1021/acs.langmuir.6b03192.
21. Zakrewsky, M.; Lovejoy, K.S.; Kern, T.L.; Miller, T.E.; Le, V.; Nagy, A.; Gourmas, A.M.; Iyer, R.S.; Del Sesto, R.E.; Koppisch, A.T.; Fox, D.T.; Mitragotri, S. Ionic liquids as a class of materials for transdermal delivery and pathogen neutralization. *Proc Natl Acad Sci* **2014**, 111, 13313. DOI: 10.1073/pnas.140395111.
22. Siopa, F.; Figueiredo, T.; Frade, R.F.M.; Neto, I.; Meirinhos, A.; Reis, C.P.; Sobral, R.G.; Afonso, C.A.M.; Rijo, P. Choline-based ionic liquids: Improvement of antimicrobial activity. *ChemistrySelect* **2016**, 1, 5909-5916, DOI: 10.1002/slct.201600864.

23. Zhao, B.; Xu, P.; Yang, F.; Wu, H.; Zong, M.; Lou, W. Biocompatible deep eutectic solvents based on choline chloride: Characterization and application to the extraction of rutin from *Sophora japonica*. *ACS Sustain Chem Eng* **2015**, 3, 2746-2755, DOI: 10.1021/ACSSUSCHEMENG.5B00619.
24. Wassenaar, T.A.; Ingólfsson, H.I.; Böckmann, R.A.; Tieleman, D.R.; Marrink, S.J. Computational lipidomics with *insane*: a versatile tool for generating custom membranes for molecular simulations. *J Chem Theory Comput* **2015**, 11(5), 2144-2155, DOI: 10.1021/acs.jctc.5b00209.
25. Marrink, S.J.; Risselada, H.J.; Yefimov, S.; Tieleman, D.P.; de Vries, A.H. The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* **2007**, 111(27), 7812-7824, DOI: 10.1021/jp071097f.
26. Graham, J.A.; Essex, J.W.; Khalid, S. PyCGTOOL: Automated Generation of Coarse-Grained Molecular Dynamics Models from Atomistic Trajectories. *J Chem Inf Model* **2017**, 57(4), 650-656, DOI: 10.1021/acs.jcim.7b00096.
27. de Jong, D.H.; Singh, G.; Bennett, W.F.D.; Arnarez, C.; Wassenaar, T.A.; Schäfer, L.V.; Periole, X.; Tieleman, D.P.; Marrink S.J. Improved parameters for the martini coarse-grained protein force field. *J Chem Theory Comput* **2012**, 9(1):687-697, DOI: 10.1021/ct300646g.
28. Abraham, M.J.; van der Spoel, D.; Lindahl, E.; Hess, B.; and the GROMACS development team, GROMACS User Manual version 5.1. 2, **2016**. There is no corresponding record for this reference.
29. Berendsen, H.J.C.; van der Spoel, D.; van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun* **1995**, 91(1-3), 43-56, DOI: 10.1016/0010-4655(96)00042-E.

30. Armijo, L. Minimization of functions having Lipschitz continuous first partial derivatives. *Pac J Math* **1966**, 16(1), 1-3.
31. Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J Chem Phys* **1984**, 81(3684), 10.10, DOI: 10.1063/1.448118.
32. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J Mol Graph Model* **1996**, 14(1), 33-38, DOI:
33. Friedman, M.; Henika, P.R.; Levin, C.E.; Mandrell, R.E. Antibacterial activities of plant essential oils and their components against *Escherichia coli* O157:H7 and *Salmonella enterica* in apple juice. *J Agric Food Chem* **2004**, 52, 6042-6048, DOI: 10.1021/jf0495340.
34. Desbois. A.P.; Smith, V.J. Antibacterial free fatty acids: activities, mechanisms of action and biotechnological potential. *Appl Microbiol Biotechnol* **2010**, 85, 1629-1642, DOI: 10.1007/s00253-009-2355-3.

## **CHAPTER 6**

### **A computational platform for accelerating antibiotics discovery**

## 6.1 ABSTRACT

Antibiotic resistance has become one of the greatest challenges. Finding new class of antibiotics is becoming more urgent when Gram-negative bacteria are becoming more resistance to most available antibiotics. The cell envelope of Gram-negative bacteria contains porin proteins, which are specific to a few molecules. Understanding molecules transport is urgently needed for the rational design of existing and new antibiotics. To quickly and accurately obtain the transport pathway of a large set of small molecules, we built a high throughput computational automated screening platform. We used *P.aeruginosa* as the first test of our platform. *P.aeruginosa* has an even narrower outer membrane porins, which make penetration of antibiotics harder. The transport of Carbapenem across *P.aeruginosa*'s major channel, the OcrD1, was studied. The detailed transport process of Carbapenem was revealed and compared with some reported results. Based on the fast and accurate information acquisition ability, the computational platform, which can be used to process large numbers of small molecules and extended to more bacterial membrane as well as their all identified porins. These results and the platform will help understand the permeability of drug candidates and facilitate the drug discovery process.

## 6.2 INTRODUCTION

Antibiotic resistance has become one of the biggest threats to human health. Even though the occurrence of antibiotic resistance is a natural process, the overuse and misuse are becoming the primary reasons for accelerating this process to causes higher patient mortality and treatment expenses.<sup>1,2</sup> However, the development of new antibiotics cannot catch up with the emergence of resistant bacteria. Since the last discovery of a new class of antibiotics was on 1987, there is a huge void in the history of antibiotics development.<sup>3</sup> Many major pharmaceutical companies are dropping antibiotics development programs now, the low success rate makes the cost higher than expectation, antibiotics resistance develops fast after new antibiotics being introduced, which makes antibiotics a short-term drug and profits are also reduced. The development of new antibiotics needs at least 10 years and over 1 billion dollars, potential antibiotics need to be screened out from more than thousands of small molecules, there is a lack of efficient screening technique, which can help lower the investment.<sup>4</sup> This technique should be able to easily acquire the antibiotics transport pathway through the bacterial membrane in a fast manner because the membrane has been proved to be the main barrier for most small molecules. It is worth noting that Gram-negative bacteria are becoming resistant to most available antibiotic drugs. The complicated outer membrane of Gram-negative bacteria was found to be able to help them develop antibiotic resistance easily and quickly.<sup>2</sup>

The outer membrane of Gram-negative bacteria comprises lipopolysaccharides (LPS) as outer leaflet, mixture of phospholipids as inner leaflet and outer membrane proteins (OMPs) as pores to allow transport of small molecules. The rigid outer membrane forms the first line of defense, which makes the Gram-negative bacteria much harder to treat than Gram-positive bacteria.<sup>5-7</sup> Understanding the interactions between this unique membrane and small molecules is essential to

guide the development process of new antibiotics. To achieve this goal, many potential candidates need to be tested from some small molecules libraries, e.g. Microsource SPECTRUM Collection, ChemBridge Diversity Set Library, etc. To facilitate this process, computational approaches should be employed to comprehensively understand the details at the molecular level and screen out the most promising antibiotics from these drug candidates to boost the drug discovery pipeline.

Recently, with the development of Computer-aided drug design (CADD), computational approaches have been widely used to guide and accelerate the early-stage development of new compounds and reduce the cost.<sup>8, 9</sup> Molecular dynamics (MD), which is a computer-based simulation method to study chemical systems and provide physical behavior of each atom or molecule in the system for nanoseconds to microseconds, can be employed to gain insights into the actions of small molecules on bacterial outer membrane or membrane proteins in molecular level to help develop robust antibiotics.<sup>5-7, 10</sup>

All-atom MD and Coarsed-grained MD are being widely used to perform simulations. All-atom MD provides us detailed interactions between each atom but needs more computing resources and time, so it is too expensive for All-atom MD to achieve long simulation time and simulate complicated systems. MD in coarse-grained (CG) level could help achieve longer time scale and larger system sizes by merging several atoms into one bead to reduce the degree of freedom so that people can explore more complex systems.<sup>11-14</sup> Martini force field is a popular used CG force field that provides us a variety of parameters of LPS, membrane lipids, amino acids, proteins, solvents and ions.<sup>15-21</sup> We built our CG representation of simulation systems using Martini force field, comprising Gram-negative bacterial outer membrane with or without OMPs embedded depending on the specific bacteria, target molecule and solvents to explore the exhaustive transport behavior of target molecule through the membrane or OMPs into the periplasmic space. The



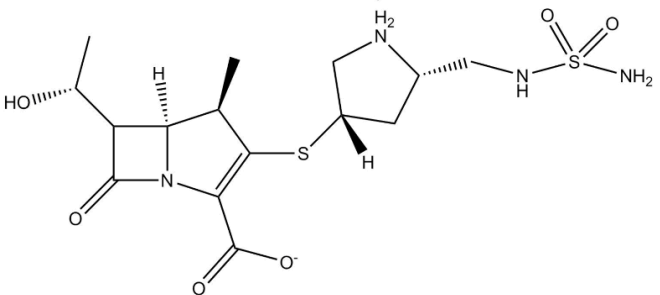
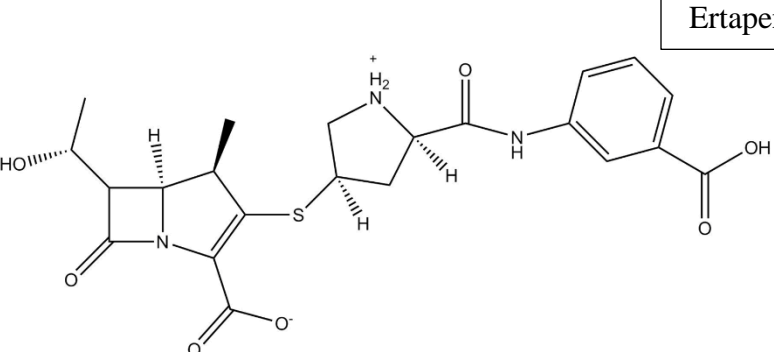
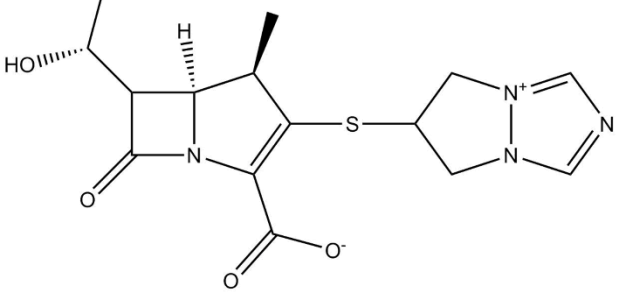
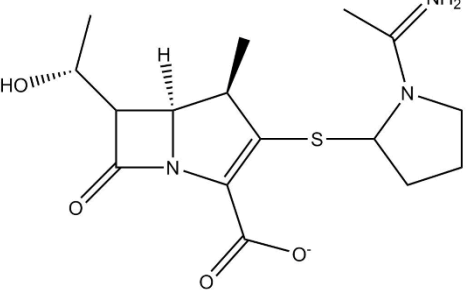
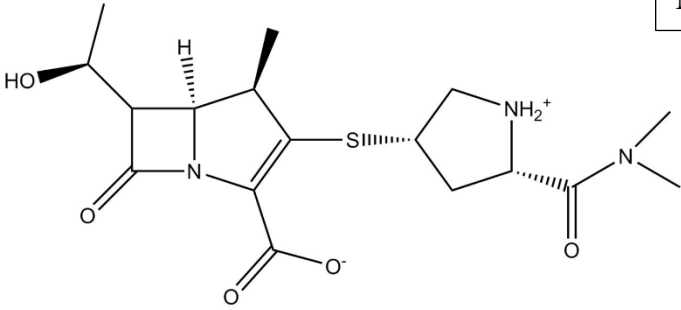
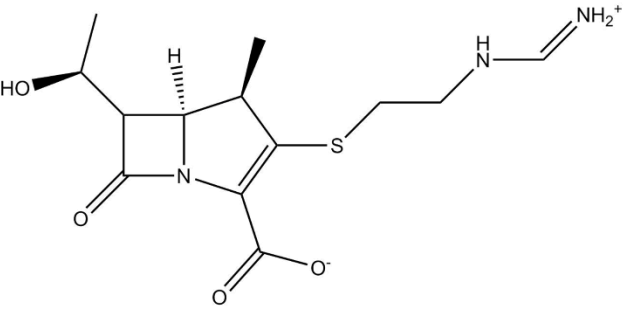
thermodynamic and kinetic data of the transport can be obtained by our CG molecular simulations as well as the molecule-membrane/molecule-protein interactions. To alleviate the burden of building new systems with different bacteria membranes, OMPs and small molecules, we made an automatic simulation control algorithm to easily achieve our goals, which is a computational automated screening platform that can quickly generate the data for further analysis. However, to validate the reliability of the computational platform, we had to prove that the simulation results could reproduce bench experimental or atomistic simulation data.

What we noticed was, the outer membrane carboxylate channel D (OccD), the largest family of substrate-specific proteins in *P. aeruginosa*, has been well studied in recent years.<sup>22-27</sup> OccD, known as the main channel for majority of small molecules such as basic amino acids, is a barrel protein with 18  $\beta$ -strands connected by large extracellular loops and short turns and its x-ray structure has been determined. Additionally, OccD1 is the important gateway for carbapenem antibiotics, which are being widely used to treat infections caused by *P. aeruginosa*.<sup>28, 29</sup> Carbapenems belong to  $\beta$ -lactam antimicrobial agents, which are able to bind the penicillin binding proteins (PBPs) to inhibit *P. aeruginosa*'s cell wall synthesis. The advantage of using Carbapenems is their excellent stability against most  $\beta$ -lactamases.<sup>30, 31</sup> However, *P. aeruginosa* acquires resistance against carbapenems by changing their PBPs' structure, expressing efflux pumps and lower their OccD protein's permeability or stopping expressing these specific porin proteins.<sup>32-34</sup> Understanding the mechanism of Carbapenems' transport is imperative to reduce the threat of *P. aeruginosa*, which has been listed as the top 3 pathogenic bacteria by World Health Organization in 2017. Recently, the penetration of Carbapenems across *P. aeruginosa*'s OccD1 protein has been studied in detail, elaborating the specific penetration process, which could be used as benchmark to validate the accuracy, reliability and feasibility of our computational platform

when CG force field parameters of *P. aeruginosa* are already available, which was done by our previous work.<sup>5, 25, 31</sup>

In this work, we reported the development of a novel computational automated screening platform (CLASP) for small molecules screening. Six carbapenems including doripenem, ertapenem, biapenem, panipenem, meropenem, imipenem and OccD1 porin protein of *P. aeruginosa* were chosen as the first test for our newly built CLASP (Table 6-1). After 6 independent simulations, the 6 carbapenems showed excellent behavior of penetration profiles, which reproduced the experimental findings and provided more insight into the penetration process within a very short time. Based on our performance tests, each simulation can be finished within 45 minutes using our available computer resources and could be reduced when adopting high performance computing. Our results provided an automated computational platform for exploring penetration process of small molecules, which can be applied to all CG lipid bilayer and available bacterial outer membranes, atomistic models are applicable by mapping to CG modeling using martinize.py<sup>18</sup> and PyCGTOOL.<sup>35</sup> This platform developed a new method of optimizing lead compounds and breaking through the obstacle of finding effective antibiotics.

## Background Information

 <p>Chemical structure of Doripenem, showing the penam nucleus and the (2S,4S)-2-((S)-sulfamoylamino)pyrrolidin-4-ylthio side chain.</p>	Doripenem
 <p>Chemical structure of Ertapenem, showing the penam nucleus and the (2S,4S)-2-((S)-((S)-1-aminopropyl)carbamoyl)pyrrolidin-4-ylthio side chain.</p>	Ertapenem
 <p>Chemical structure of Biapenem, showing the penam nucleus and the (2S,4S)-2-(1H-imidazol-2-ylthio)pyrrolidin-4-ylthio side chain.</p>	Biapenem
 <p>Chemical structure of Panipenem, showing the penam nucleus and the (2S,4S)-2-((S)-1-aminopropyl)pyrrolidin-4-ylthio side chain.</p>	Panipenem
 <p>Chemical structure of Meropenem, showing the penam nucleus and the (2S,4S)-2-((S)-((S)-1-((dimethylamino)carbonyl)ethyl)amino)pyrrolidin-4-ylthio side chain.</p>	Meropenem
 <p>Chemical structure of Imipenem, showing the penam nucleus and the (2S,4S)-2-((S)-1-((S)-1-aminopropyl)carbamoyl)pyrrolidin-4-ylthio side chain.</p>	Imipenem

## 6.3 METHODS

### 6.3.1 Simulation details

Simulation set-up: The GROMACS molecular dynamics package (5.1.2) was used in the present work.<sup>36</sup> Energy minimization simulations were performed using the steepest-decent algorithm with a 20 fs time-step until the maximum force on any bead was below the tolerance parameter of 10 kJ mol<sup>-1</sup> nm<sup>-1</sup>.<sup>37</sup> Periodic boundary conditions were applied in all three dimensions. The *NVT* (constant number of particles, volume, and temperature) and *NPT* (constant number of particles, pressure, and temperature) equilibration runs were performed for 0.2  $\mu$ s.<sup>38, 39</sup> Semi-isotropic pressure coupling was used and systems were maintained at 1 bar using the Berendsen barostat with time constant,  $\tau_p = 4.0$  ps.<sup>40, 41</sup> Temperature was maintained at 325 K by independently coupling the lipids, proteins, and the solvent to an external velocity rescaling thermostat with  $\tau_T = 1.0$  ps. The neighbor list was updated every 25 steps using 1.4 and 1.2 nm for short-range van der Waals and electrostatic cutoffs, respectively. The production *NPT* simulations were performed for 2  $\mu$ s.

Potential of Mean Force (PMF) Calculations: To compute the PMF curves, we built the *P. aeruginosa* lipopolysaccharide with OccD1 porin protein embedded in 10 $\times$ 10 nm<sup>2</sup> membrane patches. The membrane was equilibrated while keeping the proteins position-restrained for 0.5  $\mu$ s at 310 K. The restraints were then removed and pull simulations were performed along the reaction coordinate defined by separation of the center of mass (COM) of the interacting OccD1 while still embedded in the membrane. A single diglycine was pulled with respect to the COM of its position-restrained OccD1 protein. A harmonic potential with 5000 kJmol<sup>-1</sup> nm<sup>2</sup> force constant was used for the pull. A total of 100 independent Umbrella Sampling (US) windows were extracted along

the reaction coordinate for both systems, and each window was simulated for 0.2  $\mu$ s. The weighted histogram analysis method (WHAM) along with bootstrapping was used to extract the PMF profiles.<sup>42</sup>

Table 6-1 System details of Carbapenem-OccD1 simulations

Antibiotic	Inner	Outer		No. of				$T$ (K)	$t$ (s)
	DPPE	LPA	DPPE	carbapenem	OccD 1	water	ions		
Doripenem	144	55	6	1	1	6507	127	325	200
Ertapenem	144	55	6	1	1	6507	127	325	200
Biapenem	144	55	6	1	1	6507	127	325	200
Panipenem	144	55	6	1	1	6507	127	325	200
Meropenem	144	55	6	1	1	6507	127	325	200
Imipenem	144	55	6	1	1	6507	127	325	200

### 6.3.2 Method development

The CLASP algorithm has been developed to accelerate the simulation process for calculating PMF using Martini Force Field and subsequent data processing and analysis. The aim is to make the tedious Umbrella Sampling simulations automated and quickly obtain some important biological data. Although the CLASP was designed for CG simulations, atomistic models are also supported by converting to CG model using Martinize or PyCGTOOL. The work flow is shown in Fig.6-1.

Atomistic or Coarse grain structure of proteins and molecules as well as the topology are taken as the input files, different kinds of membrane can be built by insane.py with interested proteins embedded,<sup>43</sup> 10 Gram-negative bacterial outer membrane can be also generated using in-house modified insane.py called BOB.py (Bacterial Outer membrane Builder). A short Energy minimization, NVT and NPT are needed after building the membrane-protein-solvent system followed by a 2  $\mu$ s production NPT simulation to establish the starting configuration. The pre-

prepared molecules input files are inserted into the well-designed N positions across the protein channel to generate N configurations by gmx insert-molecules. Each one will be taken as the starting configuration for the US window and will be assigned an independent directory to perform US run. Finally, the PMF curves will be generated and trajectories will be used for further analysis to obtain useful biological information.

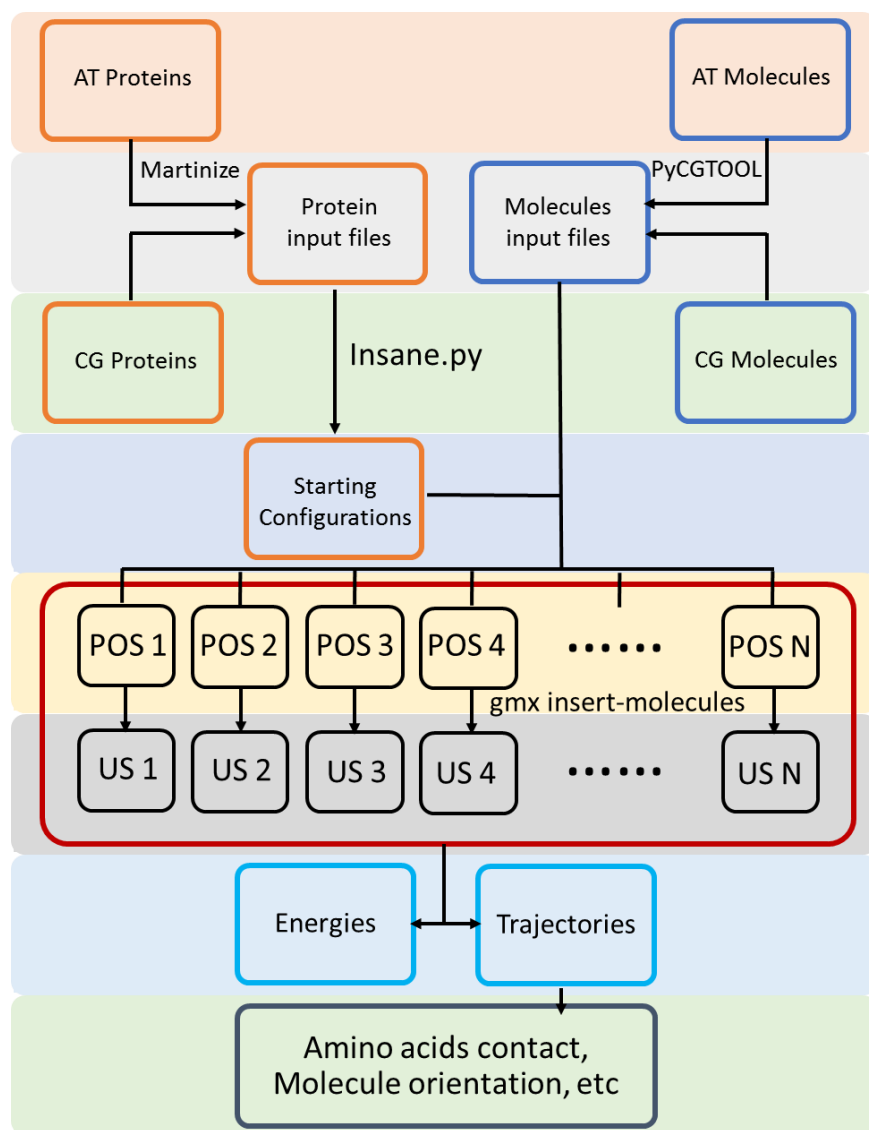


Figure.6-1 CLASP workflow

### 6.3.3 CLASP performance

In this work, all simulations were performed in GROMACS as mentioned in the method. The performance is provided for The Academic Virtual Hosting Environment (AVHE) computer clusters provided by Syracuse University. Table 2 shows the performance for different number of nodes and openMP threads used for the umbrella sampling simulations. The best performance we calculated was about 3200 ns/day and only 45 mins are needed for the simulations. With enough nodes provided, our platform can be used in a very fast manner.

Table 6-2 Simulation Performance

# of Nodes	# of OpenMP	Rate (ns/day)	Time (mins)
2	2	637.775	225
4	2	889.276	161
8	2	1875.672	76
16	2	1847.037	77
24	1	2346.15	60
24	2	3202.917	45
24	3	2207.56	65

### 6.3.4 Carbapenem coarse graining

The coarse grain parametrization of carbapenem was developed based on MARTINI many to-one mapping in which on average four heavy atoms are mapped into one bead. The detailed mapping scheme of six carbapenem were shown in Fig.6-2. The  $\beta$  lactam ring was assigned four beads (SP1, SP3, SC5, SQa) and the side chain linked to  $\beta$  lactam ring through a sulfur was assigned different beads types based on each one's specific side group. PyCGTOOL was used to generate the initial system coordinates and topologies for CG simulations. The validation was also performed to

compare the CG model with the atomistic structure using PyCGTOOL tutorial to ensure the high quality of the CG model.

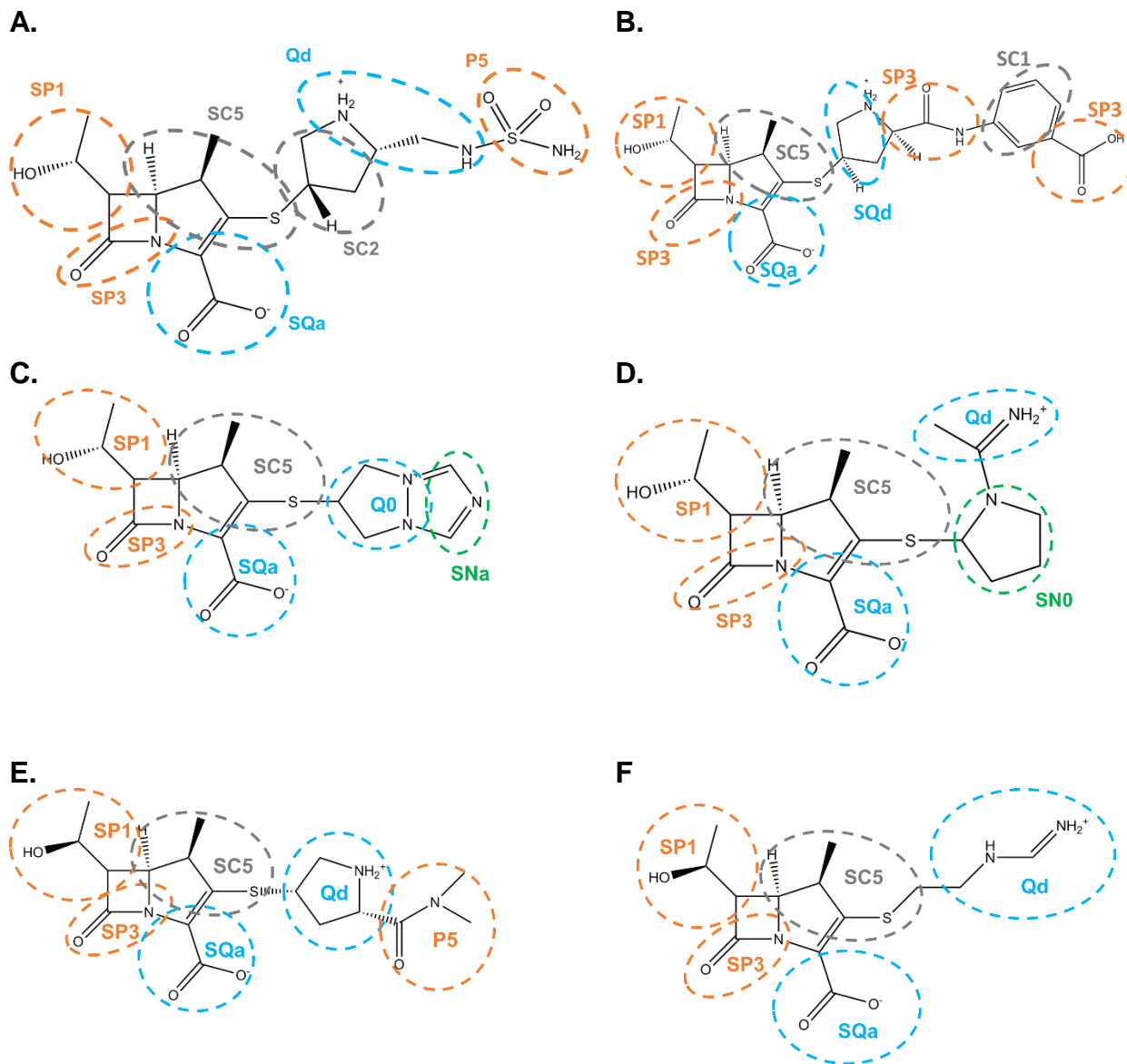


Figure.6-2 The coarse grained mapping scheme showing Martini beads of types Q/SQ (cyan), P/SP (orange), C/SC (grey), N/SN (green), A) Doripenem B) Ertapenem C) Biapenem D) Panipenem E) Meropenem F) Imipenem



### 6.3.5 Analysis

Post simulation analyses were performed using CLASP data analysis scripts including some in-built GROMACS utilities, and external software suites such as YASARA for protein alignment with X-ray crystal structure and CAVER for analysis and visualization of porin channels. Molecular visualization and graphics were generated using VMD, PyMol and YASARA.<sup>44-47</sup>

## 6.4 RESULTS

In this work, we showed the ability of CLASP to obtain the transport barrier of small molecules, interactions between small molecules and channel proteins, key residues involved during transport and the orientation of small molecules during the whole transport process. We used our selected 6 carbapenem as examples. Based on the size and coarse grain mapping of these 6 carbapenem, I categorize them as 1) Doripenem and Ertapenem, which have 7 and 8 beads, respectively 2) Panipenem, Meropenem and Biapenem, which contain 6 martini beads and 3) Imipenem, which only has 4 beads.

### Case study 1: Doripenem and Ertapenem

The PMF profiles of the six-selected carbapenems were obtained using the automated umbrella-sampling simulations implemented in CLASP. One hundred of configurations were generated along the OccD1's channel and run in independent simulations. All simulations were performed while the protein was embedded in *P. aeruginosa*'s lipidA-phospholipid membrane. The free energy profiles of Doripenem and Ertapenem carbapenem were shown in Fig.6-3. There is a common binding site around Displacement (D) = 2.3 nm for Doripenem and Ertapenem, which is the landing site of carbapenem in the entrance of OccD1's channel that right above loop 7. The region from D= 2.5 nm to D= 4.7 nm was found to be the constriction zone of OccD1, which has narrower cavity compared with other regions of the protein. Doripenem and Ertapenem showed transport barrier about 52 and 48 kJ/mol to go through this highly confined region.

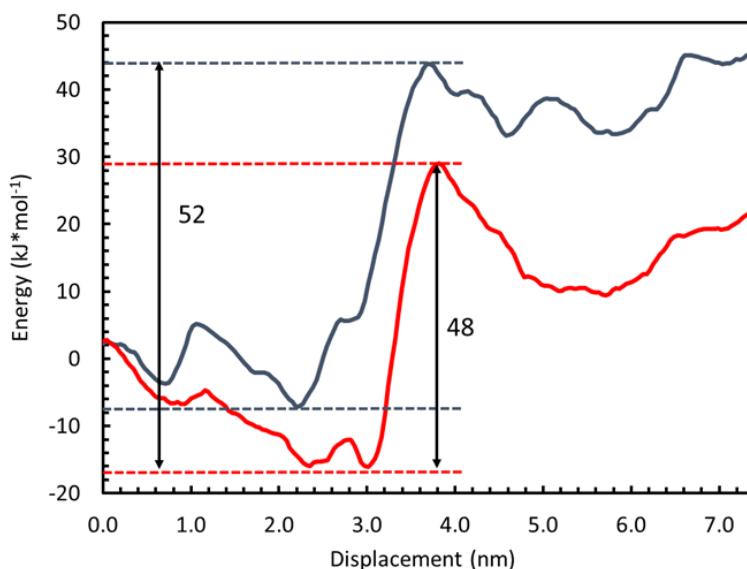


Figure.6-3 PMF of Doripenem (Navy blue) and Ertapenem (red)

The PMF profile of Doripenem and Ertapenem showed us two interesting sites inside the protein, D= 2.3 nm and D= 4 nm. At D = 2.3 nm, both of them had favorite binding sites due to the

formation of some intermolecular interactions. Doripenem and Ertapenem have energy well-depth of  $\sim 8$  kJ/mol and  $\sim 17$  kJ/mol, respectively, which shows that Ertapenem has stronger binding affinity towards this landing site of OccD1.

The detailed transport process of carbapenem have been studied over the past few years. It is well known that carbapenem need the assistant of some specific amino acids to successfully transport through the OccD1, the functions of arginine ladder has been studied carefully recently to reveal the importance of it to mediate the penetration of carbapenem. Arginine would interact with the carboxylate of carbapenem to stabilize the binding of carbapenem and OccD1. The arginine ladder includes Arg 30, Arg 39, Arg 319, Arg 337, Arg 389, Arg 391 and Arg 410, which extend through the constriction zone, would guide the permeation of carbapenem toward the inside of OccD1, leading to a successful penetration. In order to prove the importance of the arginine ladder during the carbapenem's penetration, CLASP was designed to be able to catch the interactions between small molecules and amino acids of the protein channel in a dynamic way and give us a comprehensive view of the transport process, which is an important function to find out the most important residues that can facilitate the transport process or block the penetration so that we can modify or design small molecules on a basis of these results. Here we report the application of CLASP to capture how small molecules contact with all amino acids of the protein to reveal the participation of arginine ladder in carbapenem penetration.

Firstly, the trajectories of Doripenem and Ertapenem penetrating the OccD1 were generated and then analyzed by CLASP to calculate the contact or collision between them and the amino acids of OccD1. The contact was calculated based on the center of mass distance between the them and the specific amino acids, when the distance was smaller than van der Waals force cut-off (1.2 nm), they would be recognized as having contact. By using this method, the total number of contacts

between the Doripenem and Ertapenem molecule and all amino acids of OccD1 were recorded. The most notable six amino acids that have the highest number of contact were shown in Fig.6-4 for Doripenem and Ertapenem. It showed that Tyr, Arg and Phe had the most significant contact with Doripenem and Ertapenem. During penetration, doripenem and Ertapenem molecule would make significant contacts with tyrosine first followed by very frequent contacts with arginine. Phenylalanine also showed high contact because most phenylalanine within OccD1 are adjacent to arginine so they showed similar contact. These results highlighted the important amino acids involved during the penetration, modifications or mutations of these amino acids may lead to different pathway and permeability.

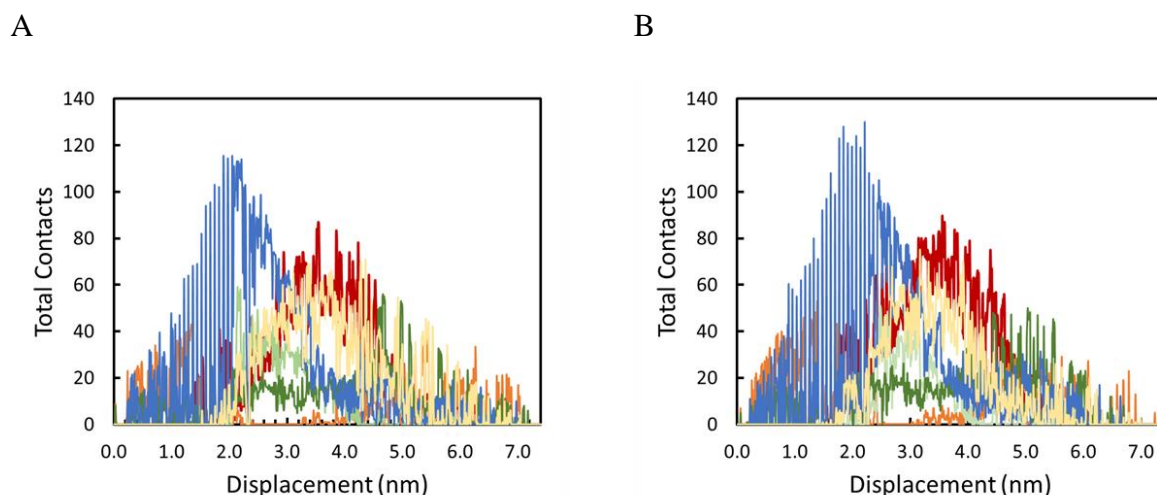


Figure.6-4 Amino acid – Carbapenem contacts A) Doripenem B) Ertapenem. Tyrosine (blue), Arginine (red), Phenylalanine (yellow), Threonine (green), Histidine (light green), Lysine (orange)

However, this was still not enough to understand which Tyrosine and Arginine made contributions to the contact. Even though it was notable that tyrosine and arginine were critical during the penetration process. However, OccD1 includes 26 tyrosine and 16 arginine, not all of them were involved in this process. In order to identify the specific tyrosine and arginine involved, CLASP

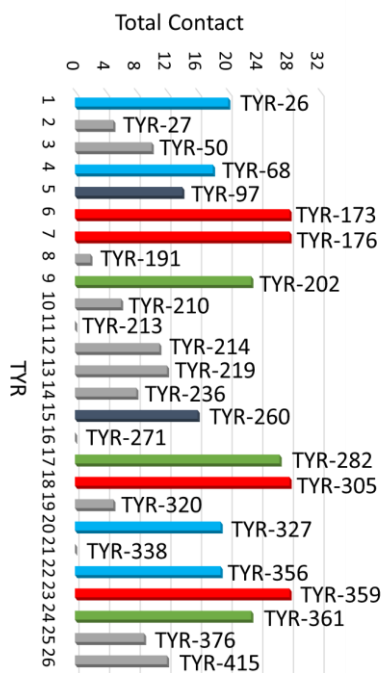
was designed to be able to obtain the contact between the carbapenem molecule and tyrosine, arginine of OccD1, respectively using similar method in last paragraph. This result would provide us more accurate targets if any modifications or mutations are needed to be done to change OccD1's selectivity or permeability. Fig. 6-5 shows the total contact between Doripenem and Ertapenem molecule with all tyrosine in OccD1, the assignment of colors was based on the number of contacts, tyrosine that has the highest contact with Doripenem or Ertapenem was assigned red, then we showed gradually decreasing concentrations in green, cyan, navy blue and gray. Tyr 173, Tyr 176, Tyr 305 and Tyr 359 were the most significant tyrosine to interact with both Doripenem and Ertapenem. The positions of these 4 red-labeled tyrosine were highlighted in OccD1 in Fig. 6-5 C. Not surprisingly, all red-labeled tyrosine were found at the entrance of OccD1. And the fact is the entrance of OccD1 shows high density of tyrosine.

More importantly, Fig. 6-4 revealed the significant participation of arginine. Similar to the tyrosine-Doripenem/Ertapenem contact map above, the 16 arginine's contacts with Doripenem/Ertapenem were also recorded. Arg 30, 39, 131, 319, 389, 391 were found to be the most significant residues to interact with Doripenem and Ertapenem (Fig. 6-6). These residues are located on or near the basic ladder or opposite. This finding matched the aforementioned reported results about the arginine basic ladder and confirmed the importance of them.<sup>41</sup> Ertapenem showed higher contacts with tyrosine and arginine due to its bigger size, it contains 8 beads that had higher probability to have contact with these amino acid residues.

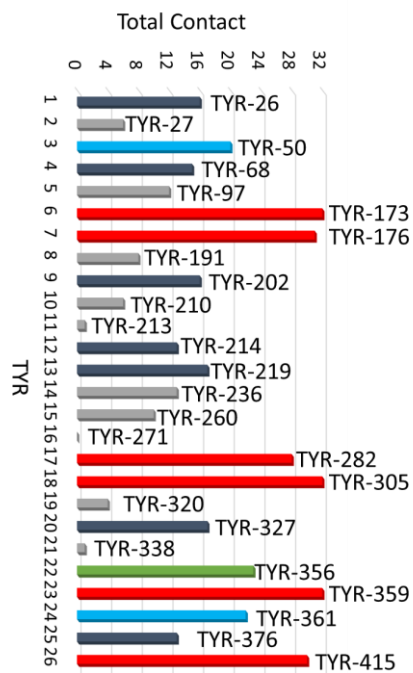
The orientation of Doripenem and Ertapenem during the transport were recorded using CALSP (Fig. 6-7). The  $D_{ij}$  is the z-component of the norm of the interatomic vector connecting the first SP1 bead on the  $\beta$  lactam ring and the last bead on the side chain (P5 of Doripenem and SP3 of Ertapenem). The magnitude of this vector is about 1 nm and the direction is from SP1 to Qd for

panipenem. The magnitude was also calculated to ensure that it would not change during the process so that its z-component can be used as a good mark for the orientation. Fig. 6-7 showed the changes of  $D_{ij}$ , it was clear that the molecule can move freely in the wide and open regions ( $D=0 - 2.5$  nm and  $D > 5$ nm), but the orientation was restrained in the constriction zone ( $D= 2.5 - 5$  nm). Moreover, for Doripenem, when  $D= 3$  nm, the  $D_{ij}$  was mostly negative, indicating the SP1 bead was mostly above the Qd bead, when  $D$  was close to 3.8 nm,  $D_{ij}$  was close to +1 nm indicating the Doripenem was almost parallel to the z axis of the protein with SP1 bead ahead, it looks like that a flip-flop happened around  $D= 3.7$  nm . However, after  $D= 4$ nm, there was a sudden change of the orientation to a negative  $D_{ij}$ . When  $D= 4$ nm, another flip-flop happened again so the  $D_{ij}$  became negative again. Ertapenem had one flip-flop happened around  $D= 3.3$  nm.

A



B



C

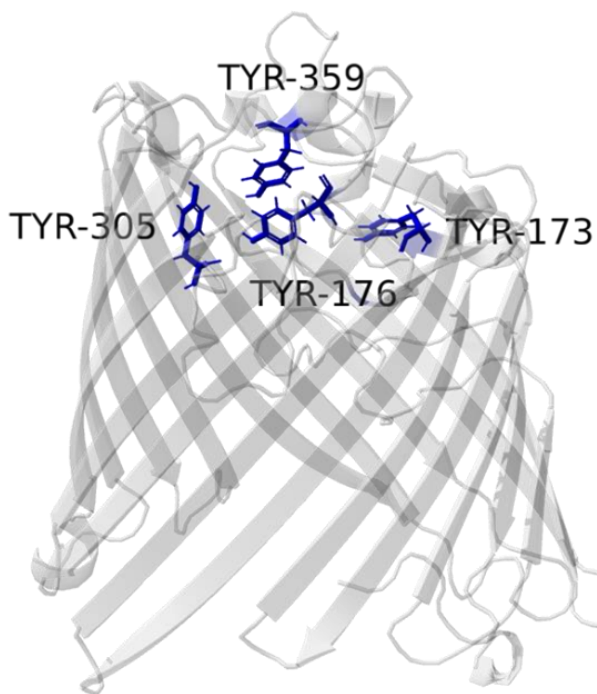
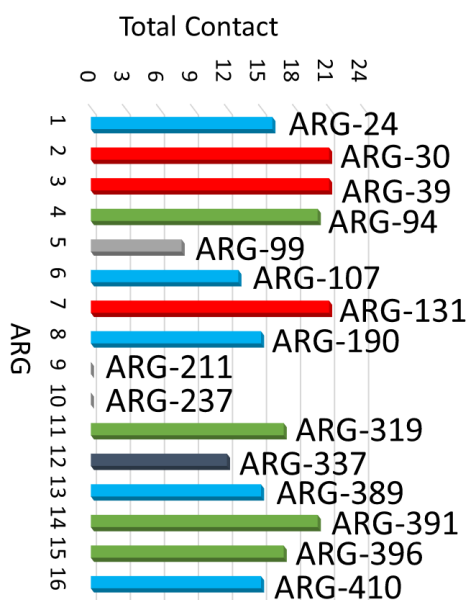


Figure. 6-5 Tyrosine - Carbapenem contact A) Tyrosine – Doripenem contact B) Tyrosine – Ertapenem contact C) Positions of tyrosine with highest contact

A



B

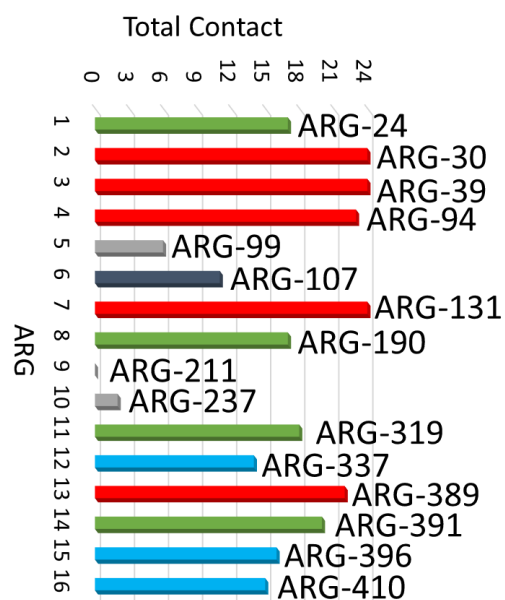
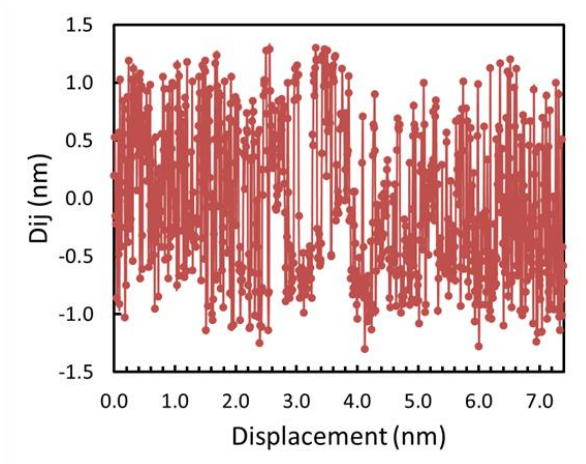


Figure.6-6 Arginine - Carbapenem contact A) Arginine – Doripenem contact B) Arginine – Ertapenem contact

A



B

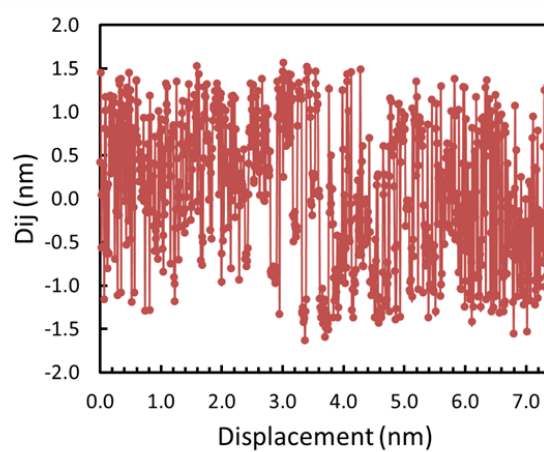


Figure. 6-7 Dij of A) Doripenem and B) Ertapenem



## Case study 2: Panipenem, Biapenem and Meropenem

The free energy profiles of Panipenem, Biapenem and Meropenem were shown in Fig. 6-8. These three carbapenem have similar binding site around  $D = 2.3$  nm. They have intermediate transport barrier about 28, 23 and 22 kJ/mol, respectively, which are significantly smaller than Doripenem and Ertapenem. This difference may arise from the bulky side groups that Doripenem and Ertapenem have. It is also interesting that all of them showed similar depth of the binding sites about 15 kJ/mol compared with 29 kcal/mol of panipenem, indicating a possible stronger binding of panipenem toward OccD1's binding site at  $D = 2.4$  nm.

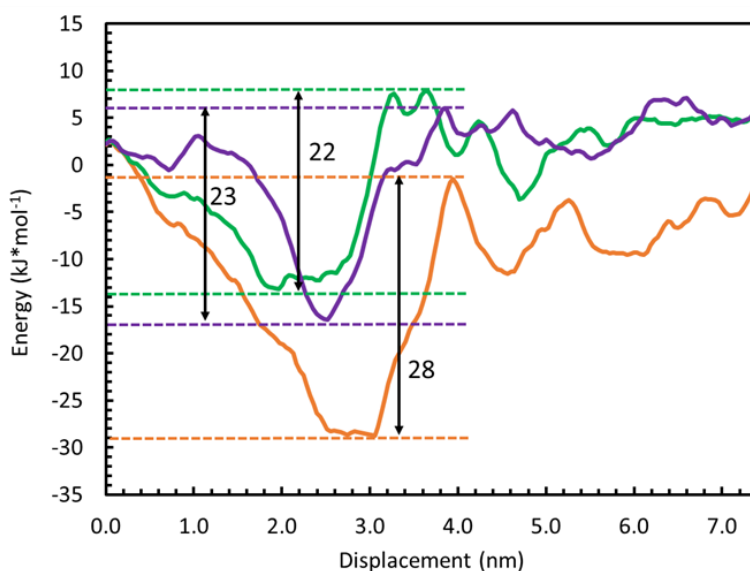


Figure. 6-8 PMF of Panipenem (orange), Biapenem (purple) and Meropenem (green)

In order to catch the important residues, the total number of contacts between the Panipenem, Biapenem and Meropenem molecule and all amino acids of OccD1 were computed as described in Case study 1 (Fig. 6-9). Same color scheme was used here. It is very clear that tyrosine and arginine were also the most significant ones as Doripenem and Ertapenem. Because of the smaller size compared with Doripenem and Ertapenem, less frequent contacts were observed for these 3

carbapenem. For tyrosine, Tyr 173, Tyr 176, Tyr 305 and Tyr 359 were still found to have the highest contacts (Fig. 6-10), which also means most carbapenem have similar preferred “landing site” at the entrance of OccD1. For arginine, Arg 30, 39, 131, 319, 389, 391 that from the arginine ladder had the highest contact with Panipenem, Biapenem and Meropenem (Figure 6-11).

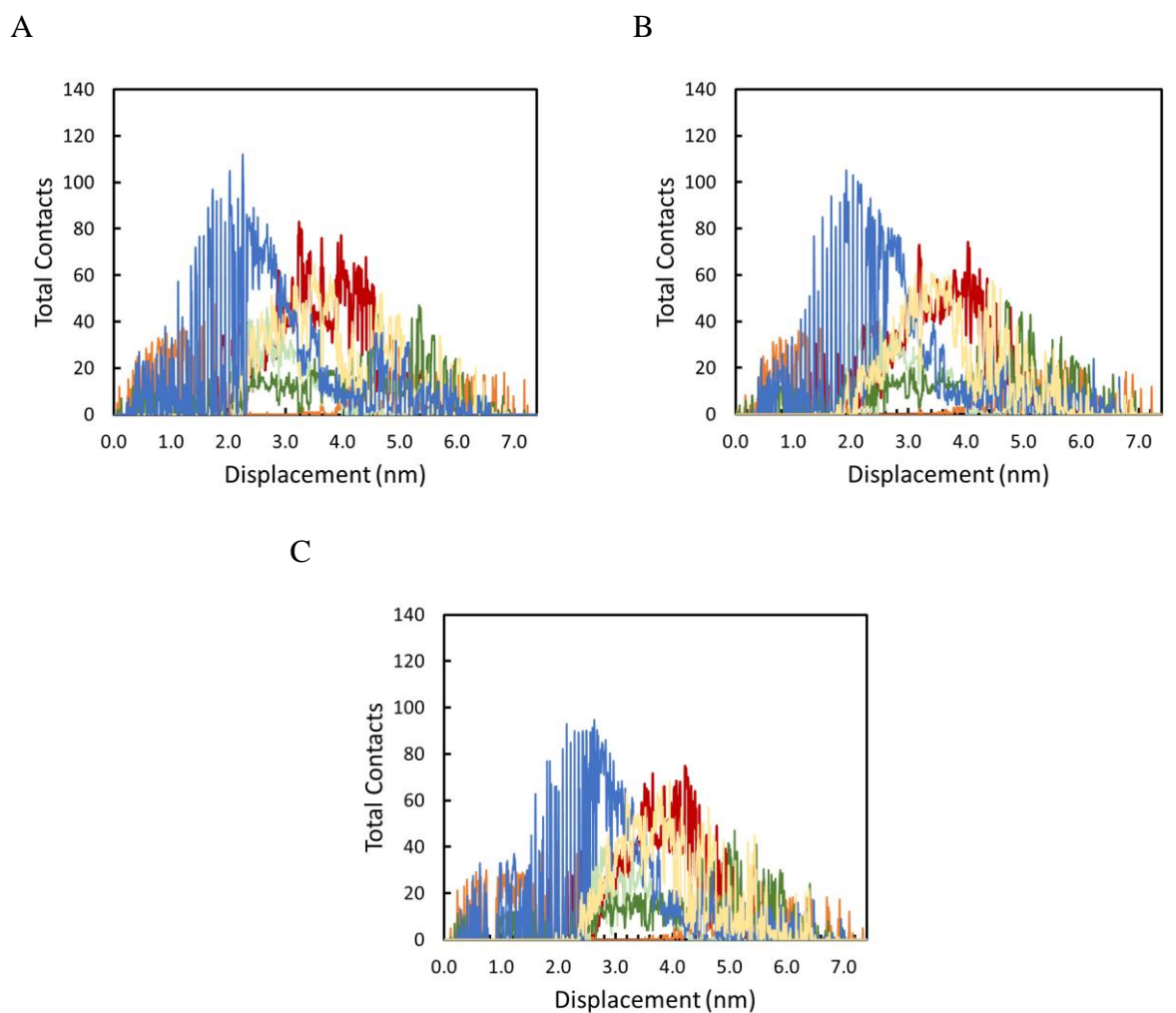
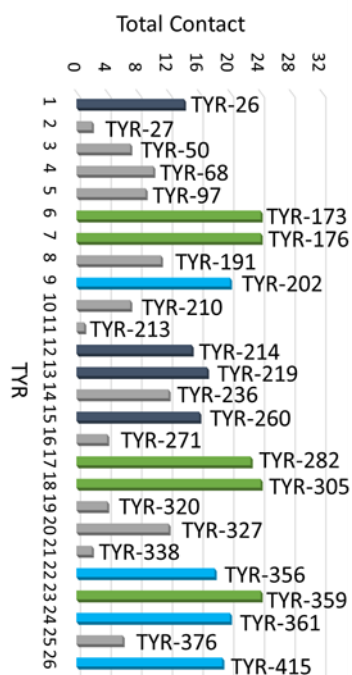
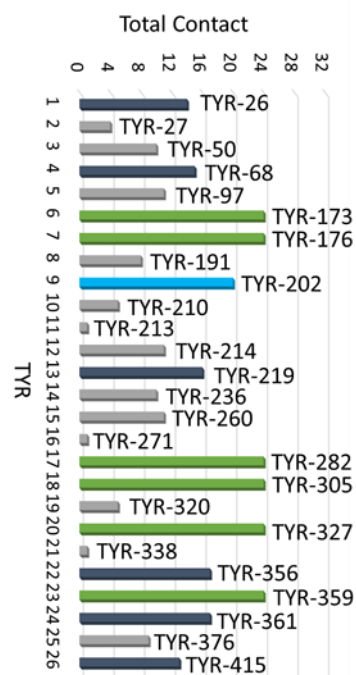


Figure.6-9 Amino acid – Carbapenem contacts A) Panipenem B) Biapenem C) Meropenem Tyrosine (blue), Arginine (red), Phenylalanine (yellow), Threonine (green), Histidine (light green), Lysine (orange)

A



B



C

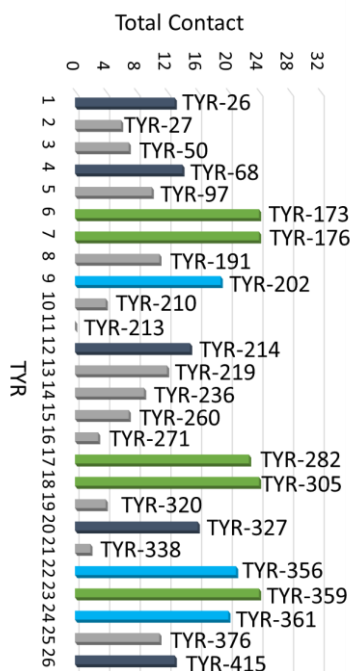
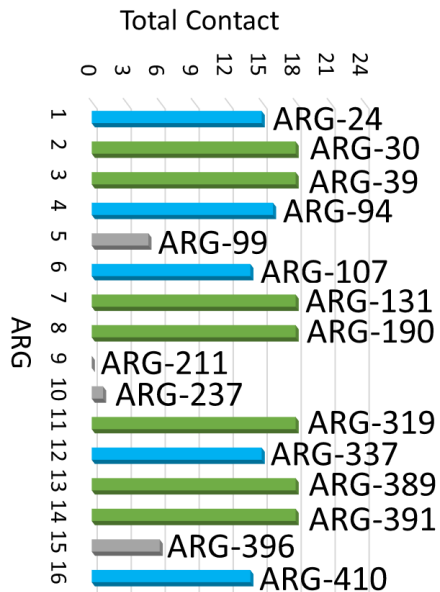
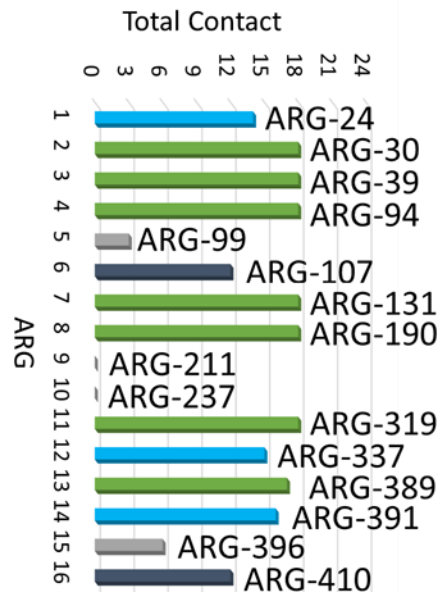


Figure. 6-10 Tyrosine - Carbapenem contact A) Tyrosine – Panipenem contact B) Tyrosine – Biapenem contact C) Tyrosine – Meropenem contact

A



B



C

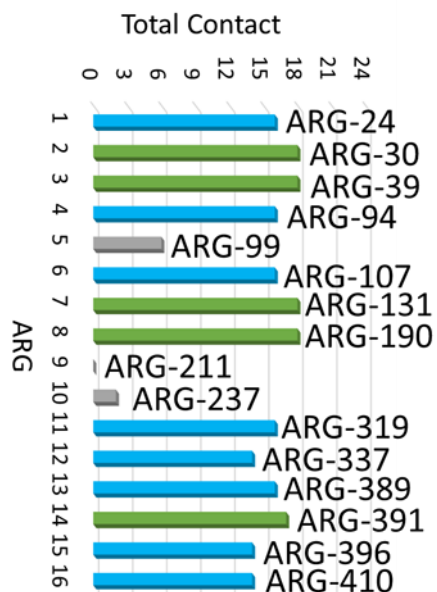


Figure.6-11 Arginine - Carbapenem contact A) Arginine –Panipenem contact B) Arginine – Biapienem contact C) Arginine – Meropenem contact

The orientation of Panipenem, Biapenem and Meropenem were also computed (Fig. 6-12). The flip-flop effect was observed in all three carbapenem here. For example, panipenem showed lumped  $D_{ij} = 1$  around  $D = 3.6 - 3.8$  nm and suddenly below  $-0.5$  around  $D = 4 - 4.6$  nm, indicating a sudden flip-flop happened during the penetration.

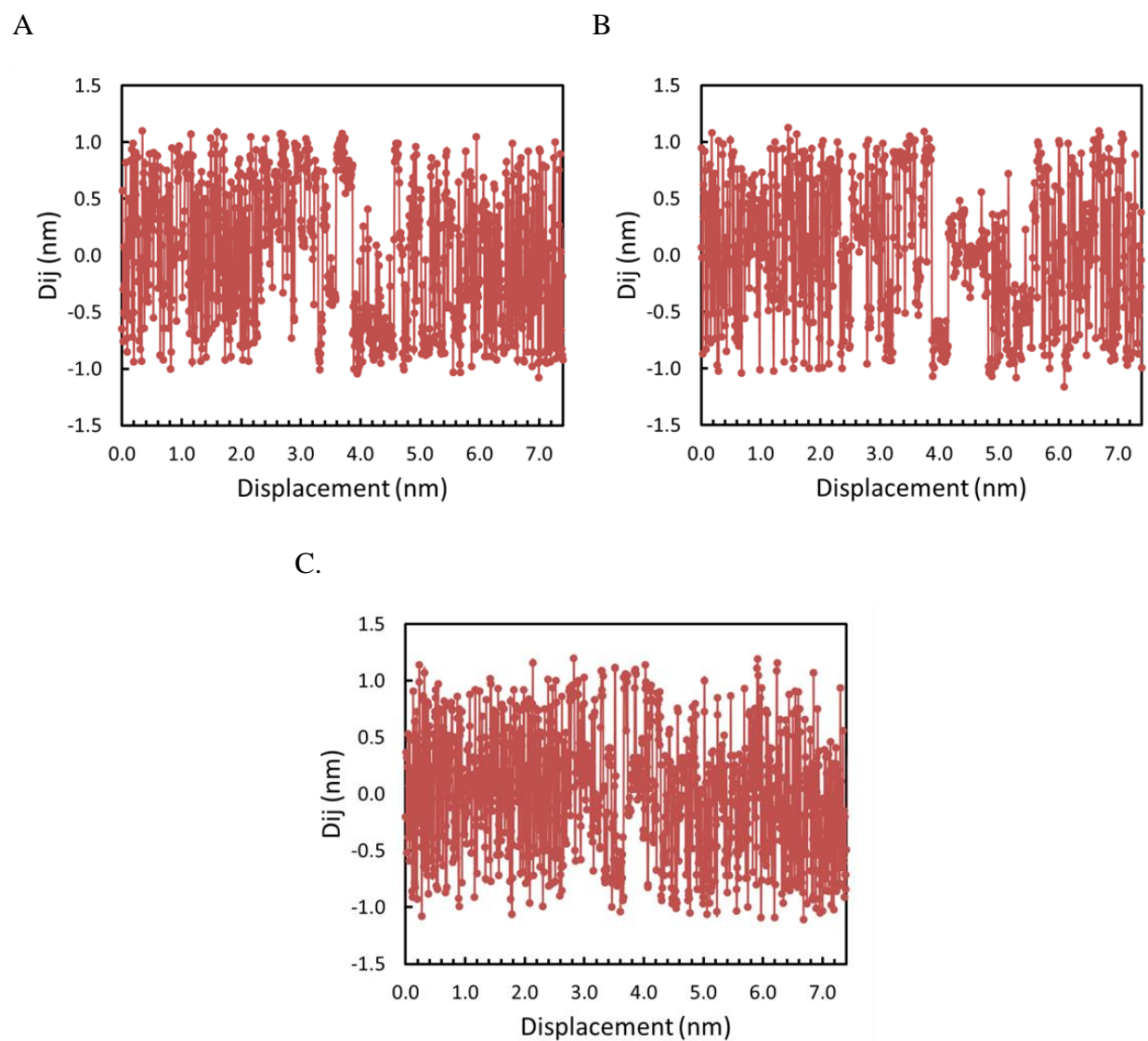


Figure. 6-12  $D_{ij}$  of A) Panipenem and B) Biapenem C) Meropenem

In order to catch the orientation change, we extracted the frames from our simulations to show how panipenem transported through  $D = 3.6 - 4.6$  region. Panipenem molecule, its SP1 bead and Qd bead were colored skyblue, red and yellow, respectively. The carboxylate group of carbepenem (SQa bead) was colored magenta and the positively charge nitrogen in arginine's guanidino group was colored green to better show the positions of the functional groups. The penetration sequence was also labeled as 1-5 in Fig. 6-13. When the panipenem molecule was close the constriction zone when  $D > 3.6$ , it maintained the orientation like step 1 and 2 with carboxylate group pointing to the guanidino group of Arg 131, now the  $D_{ij}$  is close to +1. When it was moving deeper, the opposite arginine would take over the guide function and the panipenem would turn around to face to these arginine with carboxylate group pointing to their guanidino group (step 3-4). Then they would still maintain this tailfirst shape due to the restrain of the narrow constriction zone until it reached the wide region and finally got out of the protein, now the  $D_{ij}$  was about -1. The Arg 131 was working as a transfer station when panipenem was guided by previous arginine in the ladder to the later arginine of the ladder in the narrowest region of the constriction zone. Arg 131 is not part of the arginine ladder, so its function was rarely investigated in previous studies. By studying the whole penetration process of carbepenem, its function was partly revealed.

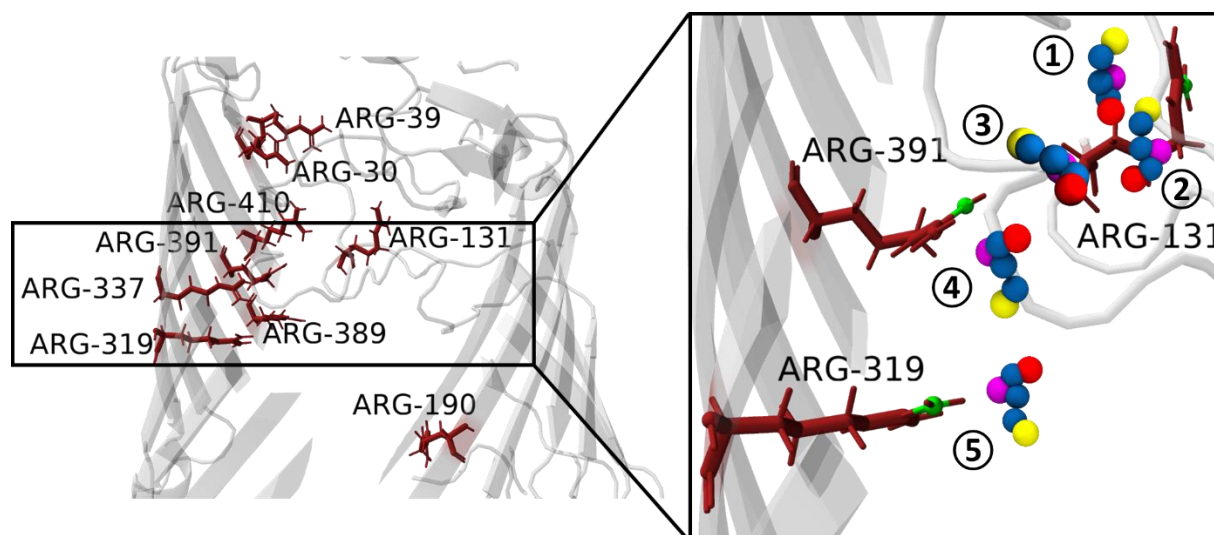


Figure.6-13 Panipenem penetrating OccD1 constriction zone process. Major arginine involved in this process are colored red.

### Case study 3: Imipenem

Imipenem is the smallest carbapenem among the 6 selected carbapenem. It showed the lowest barrier of 15.5 kJ/mol (Fig. 6-14). This was correlated with the size of the six carbapenem, based on Martini Force Field's mapping scheme, Doripenem and Ertapenem were mapped to 7 and 8 beads, respectively, Biapenem, Panipenem and Meropenem were mapped to 6 beads, imipenem was mapping to 5 beads (Fig. 6-2). All of them showed no significant barrier after  $D > 4.7$  nm, indicating a wide and open protein channel after the constriction zone and the main barrier of OccD1 was at  $D = 4$  nm.

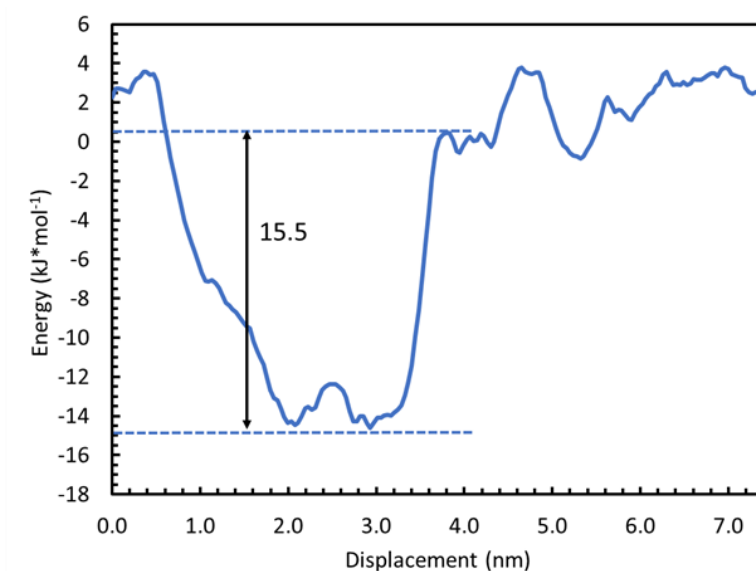


Figure.6-14 PMF of Imipenem

From Fig. 6-15, it is obvious that tyrosine and arginine were still the leading amino acids that made the most contact with imipenem. Tyr 173, Tyr 176, Tyr 305 and Tyr 359 had the highest contact with imipenem (Fig. 6-16 A) and arginine ladder also had high contact with imipenem (Fig. 6-16 B), which remains the same as other 5 carbapenem.

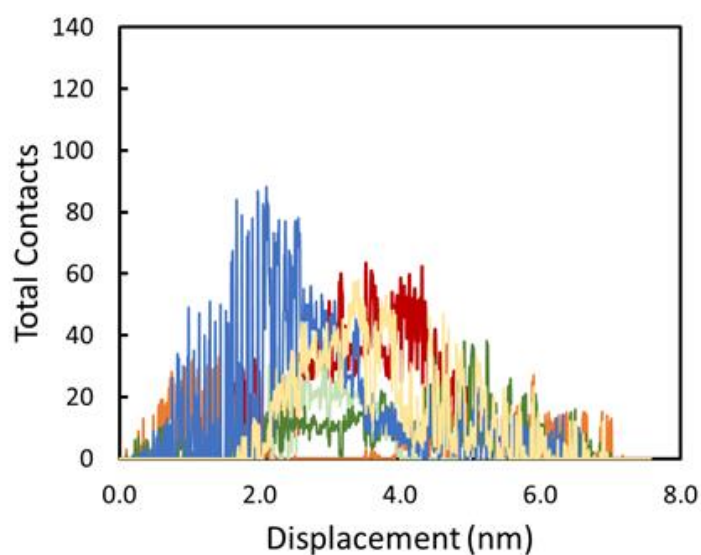
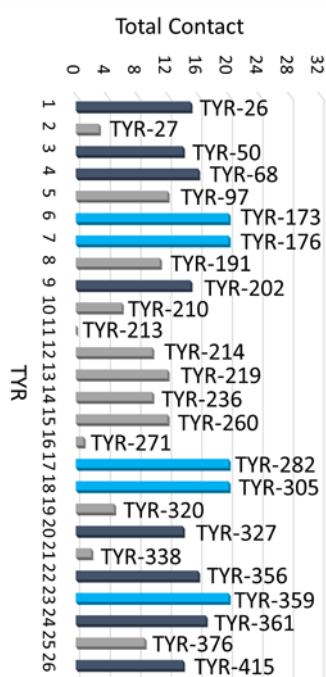


Figure.6-15 Amino acid – Imipenem contacts



A



B

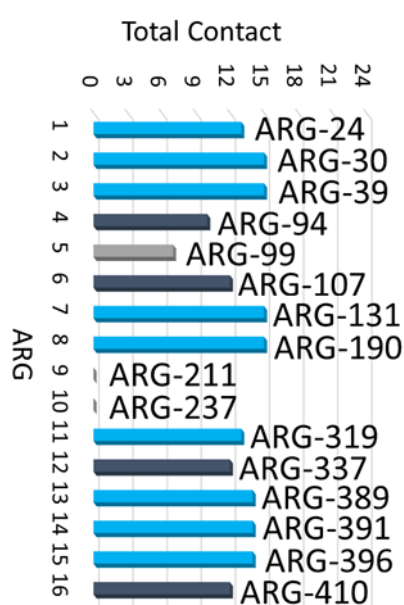


Figure. 6-16 A) Tyrosine - Imipenem contact B) Arginine – Imipenem contact

The orientation of imipenem during the transport still had less frequent changes in the constriction zone, but not as obvious as other 5, which may be because of the size of imipenem that allowed it to move more freely in the highly confined constriction zone. Flip-flop of  $D_{ij}$  were also observed as  $D=3.4-3.6$  nm,  $3.6-3.8$ nm and  $D_{ij}$  were more negative around  $D= 4.0$  nm.

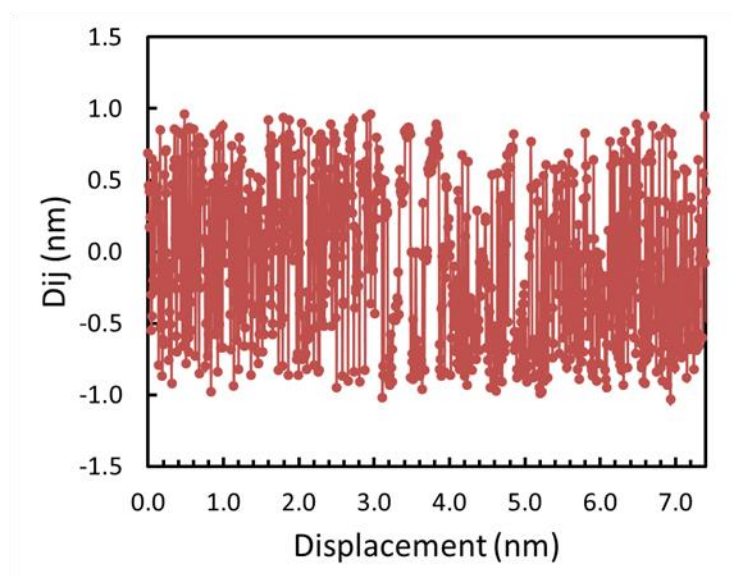


Figure. 6-17  $D_{ij}$  of Imipenem

## 6.5 DISCUSSION

The major applications of CLASP were revealed by the 3 case studies. Transport barriers of six selected carbapenem were computed using automated Umbrella Sampling simulations conducted by CLASP. The transport barriers showed positive correlation with the size of the molecule. The smallest imipenem had the lowest barrier. As  $D=2.5$  nm all carbapenem have a strong binding site toward OccD1. All carbapenem showed similar depth of the binding sites below 20 kJ/mol compared with -29 kJ/mol of panipenem, indicating a possible stronger binding of panipenem toward OccD1's landing site of entrance.

We computed the contacts between all amino acids of OccD1 and the six carbapenem. All of them showed similar contact map without any exception. Furthermore, we noticed that tyrosine showed significant higher contact with carbapenem when they started entering the channel because of tyrosine's high density in the entrance. Even though there is few research on the functions of tyrosine, the importance of tyrosine during the transport should attract more attention.

Arginine ladder plays important role in guiding the transport of carbapenem, arginine from the arginine ladder formed stable interactions with carbapenem and had high contact with them based on hydrophobic interactions/hydrogen bonding plots and contact map. Besides the arginine ladder, some other arginine also had high contact with all carbapenem. Arg 131, which is not part of the arginine ladder, had high contact with all carbapenem and its function was partly studied in this paper. However, all amino acid residues that are involved during the transport should be studied carefully. And CLASP is able to catch these important residues accurately and easily. Suitable mutations of these identified important residues would provide comprehensive understanding of their features

All 6 carbapenem showed the similar turning behavior, which was achieved with the assistance of Arg 131, either from tailfirst to headfirst, or from headfirst to tailfirst depending on their initial orientation before penetrating this area. This finding confirmed the guide function of arginine ladder as well as the importance of Arg 131. The detailed penetration steps of carbapenem were also revealed here, which would be important for amino acid mutation study, modification of available antibiotics and design of new antibiotics.

## 6.6 CONCLUSIONS

The spread of resistant bacteria species is becoming uncontrollable in the absence of efficient platforms for antibiotic discovery. To facilitate the process of finding more promising lead compounds, CLASP was developed to quickly acquire information about the transport process of small molecules. *P. aeruginosa* and its special outer membrane protein OccD1 were chosen as the first test of this platform. Six carbapenem were used in this test. CLASP was able to obtain the barrier for these carbapenem transport through OccD1 within 1 hour. During the transport process, tyrosine and arginine were found to be the most dominant amino acids involved and specific residues of two amino acids were also identified, further study on these residues would help understand the functions of them. In order to successfully go through the constriction zone of OccD1, all carbapenem were guided by the arginine ladder. In this region, the frequency of molecular directional changes is significantly reduced. Arg 131, as a transfer stop, could cause sudden directional change, even though it is not part of the arginine ladder, its function cannot be ignored. These information would guide rational design and modification of antibiotic molecules. Overall, we successfully designed the CLASP, and it is a promising computational platform for screening of promising small molecules and facilitating antibiotic discovery.

## REFERENCE

1. Landers, T. F.; Cohen, B.; Wittum, T. E.; Larson, E. L., A review of antibiotic use in food animals: perspective, policy, and potential. *Public health reports* **2012**, *127* (1), 4-22.
2. Ventola, C. L., The antibiotic resistance crisis: part 1: causes and threats. *Pharmacy and Therapeutics* **2015**, *40* (4), 277.
3. Silver, L. L., Challenges of antibacterial discovery. *Clinical microbiology reviews* **2011**, *24* (1), 71-109.
4. Fox, J. L., The business of developing antibacterials. *Nature biotechnology* **2006**, *24* (12), 1521.
5. Ma, H.; Irudayanathan, F. J.; Jiang, W.; Nangia, S., Simulating Gram-negative bacterial outer membrane: a coarse grain model. *J. Chem. Phys* **2015**, *119* (46), 14668-14682.
6. Ma, H.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan, A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S., Modeling diversity in structures of bacterial outer membrane lipids. *Journal of chemical theory and computation* **2017**, *13* (2), 811-824.
7. Ma, H.; Khan, A.; Nangia, S., Dynamics of OmpF trimer formation in the bacterial outer membrane of Escherichia coli. *Langmuir* **2017**, *34* (19), 5623-5634.
8. Yang, L.; Rybtke, M. T.; Jakobsen, T. H.; Hentzer, M.; Bjarnsholt, T.; Givskov, M.; Tolker-Nielsen, T., Computer-aided identification of recognized drugs as Pseudomonas aeruginosa quorum-sensing inhibitors. *Antimicrobial agents and chemotherapy* **2009**, *53* (6), 2432-2443.
9. Marshall, G. R., Computer-aided drug design. *Annual review of pharmacology and toxicology* **1987**, *27* (1), 193-213.

10. Ibsen, K. N.; Ma, H.; Banerjee, A.; Tanner, E. E.; Nangia, S.; Mitragotri, S., Mechanism of Antibacterial Activity of Choline-Based Ionic Liquids (CAGE). *ACS Biomaterials Science & Engineering* **2018**.
11. Jiang, W.; Wang, X.; Guo, D.; Luo, J.; Nangia, S., Drug-specific design of telodendrimer architecture for effective doxorubicin encapsulation. *J. Chem. Phys* **2016**, *120* (36), 9766-9777.
12. Irudayanathan, F. J.; Trasatti, J. P.; Karande, P.; Nangia, S., Molecular architecture of the blood brain barrier tight junction proteins—a synergistic computational and in vitro approach. *J. Chem. Phys* **2015**, *120* (1), 77-88.
13. Irudayanathan, F. J.; Wang, N.; Wang, X.; Nangia, S., Architecture of the paracellular channels formed by claudins of the blood–brain barrier tight junctions. *Annals of the New York Academy of Sciences* **2017**, *1405* (1), 131-146.
14. Irudayanathan, F. J.; Wang, X.; Wang, N.; Willsey, S. R.; Seddon, I. A.; Nangia, S., Self-Assembly Simulations of Classic Claudins—insights into the Pore Structure, Selectivity and Higher Order Complexes. *J. Chem. Phys* **2018**.
15. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H., The MARTINI force field: coarse grained model for biomolecular simulations. *J. Chem. Phys* **2007**, *111* (27), 7812-7824.
16. Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J., The MARTINI coarse-grained force field: extension to proteins. *Journal of chemical theory and computation* **2008**, *4* (5), 819-834.
17. Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J., Polarizable water model for the coarse-grained MARTINI force field. *PLoS computational biology* **2010**, *6* (6), e1000810.

18. de Jong, D. H.; Singh, G.; Bennett, W. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J., Improved parameters for the martini coarse-grained protein force field. *Journal of Chemical Theory and Computation* **2012**, 9 (1), 687-697.
19. López, C. A.; Rzepiela, A. J.; De Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J., Martini coarse-grained force field: extension to carbohydrates. *Journal of Chemical Theory and Computation* **2009**, 5 (12), 3195-3210.
20. López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J., Martini force field parameters for glycolipids. *Journal of chemical theory and computation* **2013**, 9 (3), 1694-1708.
21. Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J., Martini coarse-grained force field: extension to DNA. *Journal of chemical theory and computation* **2015**, 11 (8), 3932-3945.
22. Parkin, J.; Khalid, S., Atomistic molecular-dynamics simulations enable prediction of the arginine permeation pathway through OccD1/OprD from *Pseudomonas aeruginosa*. *Biophysical journal* **2014**, 107 (8), 1853-1861.
23. Trias, J.; Nikaido, H., Outer membrane protein D2 catalyzes facilitated diffusion of carbapenems and penems through the outer membrane of *Pseudomonas aeruginosa*. *Antimicrobial agents and chemotherapy* **1990**, 34 (1), 52-57.
24. Biswas, S.; Mohammad, M. M.; Patel, D. R.; Movileanu, L.; Van Den Berg, B., Structural insight into OprD substrate specificity. *Nature Structural and Molecular Biology* **2007**, 14 (11), 1108.

25. Eren, E.; Parkin, J.; Adelanwa, A.; Cheneke, B.; Movileanu, L.; Khalid, S.; Van den Berg, B., Towards understanding the outer membrane uptake of small molecules by *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* **2013**, jbc. M113. 463570.
26. Chevalier, S.; Bouffartigues, E.; Bodilis, J.; Maillot, O.; Lesouhaitier, O.; Feuilloley, M. G.; Orange, N.; Dufour, A.; Cornelis, P., Structure, function and regulation of *Pseudomonas aeruginosa* porins. *FEMS microbiology reviews* **2017**, 41 (5), 698-722.
27. Eren, E.; Vijayaraghavan, J.; Liu, J.; Cheneke, B. R.; Touw, D. S.; Lepore, B. W.; Indic, M.; Movileanu, L.; Van den Berg, B., Substrate specificity within a family of outer membrane carboxylate channels. *PLoS biology* **2012**, 10 (1), e1001242.
28. Köhler, T.; Michea-Hamzehpour, M.; Epp, S. F.; Pechere, J.-C., Carbapenem activities against *Pseudomonas aeruginosa*: respective contributions of OprD and efflux systems. *Antimicrobial agents and chemotherapy* **1999**, 43 (2), 424-427.
29. Shah, P. M.; Isaacs, R. D., Ertapenem, the first of a new group of carbapenems. *Journal of Antimicrobial Chemotherapy* **2003**, 52 (4), 538-542.
30. Pillar, C. M.; Torres, M. K.; Brown, N. P.; Shah, D.; Sahm, D. F., In vitro activity of doripenem, a carbapenem for the treatment of challenging infections caused by Gram-negative bacteria, against recent clinical isolates from the United States. *Antimicrobial agents and chemotherapy* **2008**, 52 (12), 4388-4399.
31. Iyer, R.; Sylvester, M. A.; Velez-Vega, C.; Tommasi, R.; Durand-Reville, T. F.; Miller, A. A., Whole-cell-based assay to evaluate structure permeation relationships for carbapenem passage through the *Pseudomonas aeruginosa* porin OprD. *ACS infectious diseases* **2017**, 3 (4), 310-319.



32. Subedi, D.; Vijay, A. K.; Willcox, M., Overview of mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*: an ocular perspective. *Clinical and Experimental Optometry* **2018**, *101* (2), 162-171.
33. Nordmann, P., Mechanism of resistance to betalactam antibiotics in *Pseudomonas aeruginosa*. *Annales Francaises D Anesthesie Et De Reanimation* **2003**, *22* (6), 527-530.
34. Lambert, P. A., Mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*. *Journal of the Royal Society of Medicine* **2002**, *95*, 22-26.
35. Graham, J. A.; Essex, J. W.; Khalid, S., PyCGTOOL: Automated Generation of Coarse-Grained Molecular Dynamics Models from Atomistic Trajectories. *Journal of Chemical Information and Modeling* **2017**, *57* (4), 650-656.
36. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25.
37. Armijo, L., MINIMIZATION OF FUNCTIONS HAVING LIPSCHITZ CONTINUOUS FIRST PARTIAL DERIVATIVES. *Pacific Journal of Mathematics* **1966**, *16* (1), 1-&.
38. Nose, S., A MOLECULAR-DYNAMICS METHOD FOR SIMULATIONS IN THE CANONICAL ENSEMBLE. *Molecular Physics* **1984**, *52* (2), 255-268.
39. Evans, D. J.; Morriss, G. P., THE ISOTHERMAL ISOBARIC MOLECULAR-DYNAMICS ENSEMBLE. *Physics Letters A* **1983**, *98* (8-9), 433-436.
40. Andersen, H. C., MOLECULAR-DYNAMICS SIMULATIONS AT CONSTANT PRESSURE AND-OR TEMPERATURE. *Journal of Chemical Physics* **1980**, *72* (4), 2384-2393.

41. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., MOLECULAR-DYNAMICS WITH COUPLING TO AN EXTERNAL BATH. *Journal of Chemical Physics* **1984**, *81* (8), 3684-3690.
42. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., THE WEIGHTED HISTOGRAM ANALYSIS METHOD FOR FREE-ENERGY CALCULATIONS ON BIOMOLECULES .1. THE METHOD. *Journal of Computational Chemistry* **1992**, *13* (8), 1011-1021.
43. Wassenaar, T. A.; Ingolfsson, H. I.; Bockmann, R. A.; Tieleman, D. P.; Marrink, S. J., Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *Journal of Chemical Theory and Computation* **2015**, *11* (5), 2144-2155.
44. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics & Modelling* **1996**, *14* (1), 33-38.
45. Seeliger, D.; de Groot, B. L., Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of Computer-Aided Molecular Design* **2010**, *24* (5), 417-422.
46. DeLano, W. L.; Lam, J. W., PyMOL: A communications tool for computational models. *Abstracts of Papers of the American Chemical Society* **2005**, *230*, U1371-U1372.
47. Krieger, E.; Vriend, G., New Ways to Boost Molecular Dynamics Simulations. *Journal of Computational Chemistry* **2015**, *36* (13), 996-1007.

## **CHAPTER 7**

# **CONCLUSIONS AND FUTURE WORK**

## 7.1 CONCLUSIONS

In this work, force field parameterization of ten Gram-negative bacterial outer membranes has been completed. Coarse grained computational models of these ten membranes showed great agreement with experimental data. Force field parameters can be downloaded online now for free either from our group website or Martini force field website. As an integral part of the cell membrane, behavior of OmpF porin protein were also studied. The results highlighted the necessity of OmpF to remain trimeric state. The process of OmpF to form trimer were observed by simulations. The mechanism of forming trimer was also confirmed.

After successfully modeling Gram-negative bacterial outer membrane, we applied our coarse grained *E.coli*'s outer membrane to the study of ionic liquids. Molecular dynamic simulations were used to reveal the effect of CAGE variants on *E.coli*'s outer membrane. Combined with experimental data from our collaborator, basic mechanism of choline-based IL activity was illuminated. This also confirmed the reliability and applicability of our models.

To best use these models, CLASP was developed to achieve automated screening of antibiotics. 6 carbapenem were chosen as the first test of CLASP. The results basically confirm the previously published conclusions and provide more valuable information. The easy extension of CLASP makes it a promising tool to investigate more potential small molecules and help rational design of available antibiotics. The process of developing antibiotics will benefit from the information provided by CLASP.

## 7.2 FUTURE WORK

Gram-negative and Gram-positive bacteria both have peptidoglycan in their cell wall, which is not included in my work. Coarse grained parameterization of bacterial peptidoglycan has been partly done in our group. Mechanism of how antibiotics transport through the entire cell wall of bacteria can be achieved in the future.

A library of coarse grained bacterial outer membranes are needed to improve the applicability of CLASP especially for those pathogenic species. Based on available models and identified porin proteins, CLASP can already be extended to 10 bacterial as well as their associated porin proteins. We only tested one porin protein of *P. aeruginosa*, whose major porin proteins are more than 10 types.

## Appendix: Tables of parameters for coarse grained *E.coli* lipid A and core oligosaccharides

Table S4-1. Equilibrium bond length and force constant values for the coarse grained lipid A

model of *E. coli*

Residue	Bonds	$R_{\min}(\text{nm})$	$K_{\text{bond}}(\text{kJ mol}^{-1} \text{ nm}^{-2})$
XYA	L1-L3	0.32	5000
XYA	L1-L5	0.32	5000
XYA	L1-L2	0.32	1250
XYA-LP1	L2-L6	0.47	1250
XYA-LP1	L3-L9	0.47	1250
XYA	L4-L1	0.32	5000
XYA	L4-L3	0.32	5000
LP1	L6-L7	0.47	1250
LP1	L7-L8	0.47	1250
LP1	L9-L10	0.47	1250
SYB	L13-L15	0.3	5000
SYB	L13-L16	0.47	5000
SYB-XYA	L13-L4	0.47	5000
SYB	L14-L15	0.47	5000
SYB	L14-L13	0.3	5000
SYB	L15-L17	0.32	5000
SYB-LP3	L15-L28	0.47	1250
SYB-LP2	L16-L21	0.47	1250
LP2	L18-L19	0.47	1250
LP2	L19-L20	0.47	1250
LP2	L21-L18	0.47	5000
LP2	L21-L22	0.47	1250
LP2	L22-L23	0.47	1250
LP2	L23-L24	0.47	1250
LP3	L25-L28	0.47	1250
LP3	L25-L26	0.47	1250
LP3	L26-L27	0.47	1250
LP3	L28-L29	0.3	1250
LP3	L29-L30	0.47	1250
LP3	L30-L31	0.47	1250
LP3	L31-L32	0.47	1250

**LP1** (2-hydroxytetradecanoyl acid 14:0); **LP2** (dodecanoyl acid decyl ester) ; **LP3** (tetracanoyl acid decyl ester); **SYB** (3-(acetyl amino)-3-deoxy-D-glucose); **XYA**(3-(acetyl amino)-3-deoxy-D-glucose);

Table S4-2. Equilibrium angle and force constant values for the coarse grained Lipid A model of *E. coli*

Angle	$\theta_{eq}$ (degrees)	$K_{angle}(\text{kJ mol}^{-1})$
L4-L3-L1	120	50
L3-L1-L5	120	50
L2-L6-L7	180	25
L6-L7-L8	180	25
L9-L10-L11	180	25
L10-L11-L12	180	25
L13-L4-L3	107	50
L14-L15-L17	87	50
L14-L13-L15	59	50
L14-L15-L13	68	50
L13-L16-L21	180	25
L18-L19-L20	180	25
L18-L21-L22	120	50
L21-L22-L23	180	25
L22-L23-L24	180	25
L25-L26-L27	180	25
L13-L15-L28	101	50
L15-L28-L29	180	25
L28-L29-L30	180	25
L30-L31-L32	180	25

Table S4-3. Equilibrium bond length and force constant values for the coarse grained core oligosaccharide model of *E. coli*

Residue	Bonds	$R_{\min}$ (nm)	$K_{\text{bond}}$ (kJ mol <sup>-1</sup> nm <sup>-2</sup> )
GAL	L1-L2	0.368	17000
GAL	L1-L3	0.291	17000
GAL-2GA	L1-L4	0.321	17000
GAL	L2-L3	0.298	17000
2GA	L4-L5	0.274	17000
2GA	L4-L6	0.372	17000
2GA-2GL	L4-L10	0.33	17000
2GA	L5-L6	0.268	17000
GLC	L7-L8	0.286	17000
GLC	L7-L9	0.372	17000
GLC	L8-L9	0.275	17000
GLC-2GL	L7-L11	0.352	17000
2GL	L10-L12	0.369	17000
2GL	L10-L11	0.281	17000
2GL-3GL	L10-L14	0.341	17000
2GL	L11-L12	0.33	17000
3GL	L14-L15	0.301	17000
3GL	L14-L13	0.365	17000
3GL	L13-L15	0.282	17000
3GL-2HP	L15-L23	0.365	17000
HEP	L16-L18	0.319	17000
HEP	L16-L19	0.234	17000
HEP-2HP	L16-L21	0.323	17000
HEP	L17-L18	0.287	17000
HEP	L18-L19	0.227	17000
2HP	L20-L22	0.308	17000
2HP	L20-L23	0.272	17000
2HP-3HP	L20-L28	0.306	17000
2HP	L21-L22	0.334	17000
2HP	L22-L23	0.387	17000
2HP	L23-L24	0.393	17000
3HP	L25-L27	0.267	17000
3HP	L25-L28	0.309	17000
3HP-LKO	L25-L34	0.293	17000
3HP	L26-L27	0.259	17000
3HP	L27-L28	0.316	17000
3HP	L28-L29	0.289	17000
OKO	L36-L35	0.258	17000
OKO-3HP	L35-L27	0.262	17000



3HP	L27-L25	0.248	17000
3HP	L27-L28	0.258	17000
3HP	L28-L29	0.289	17000
LKO	L30-L31	0.239	17000
3HP	L28-L29	0.289	17000
LKO	L30-L31	0.239	17000
LKO	L30-L33	0.216	17000
LKO	L30-L34	0.248	17000
LKO	L32-L33	0.216	17000
LKO	L33-L34	0.258	17000
LKO-OKO	L34-L35	0.349	17000
OKO	L35-L36	0.233	17000
OKO	L35-L38	0.21	17000
OKO	L37-L38	0.284	17000
OKO	L37-L39	0.238	17000
OKO	L38-L39	0.242	17000

---

**LP1** (2-hydroxytetradecanoyl acid 14:0); **LP2** (dodecanoyl acid decyl ester) ; **LP3** (tetracanoyl acid decyl ester); **SYB** (3-(acetyl amino)-3-deoxy-D-glucose); **XYA**(3-(acetyl amino)-3-deoxy-D-glucose); **LKO** and **OKO** (3-deoxy-D-manno-oct-2-ulosonic acid) ; **HEP** or **HP** (L-glycero-D-manno heptose); **GAL** or **GA** (D-galactose); **GLC** or **GL** (D-glucose).

Table S4-4. Equilibrium angle and force constant values for the coarse -grained core oligosaccharide model of *E. coli*

Angle	$\theta_{eq}$ (degrees)	$K_{angle}$ (kJ mol <sup>-1</sup> )
L5-L1-L4	122	50
L3-L1-L2	52	50
L1-L3-L2	77	50
L2-L1-L3	52	50
L1-L4-L6	116	50
L1-L4-L5	132	50
L4-L5-L6	67	50
L9-L8-L7	79	50
L9-L7-L8	63	50
L4-L10-L12	140	50
L10-L11-L12	73	50
L11-L12-L10	47	50
L11-L7-L8	105	50
L14-L15-L13	77	50
L14-L13-L15	47	50
L23-L22-L20	83	50
L23-L20-L22	63	50
L23-L22-L21	131	50
L15-L23-L20	180	50
L19-L18-L16	70	50
L19-L18-L17	131	50
L19-L16-L21	125	50
L22-L20-L28	122	50
L28-L27-L25	58	50
L28-L27-L26	110	50
L28-L25-L33	112	50
L27-L28-L29	134	50
L26-L27-L25	134	50
L25-L34-L33	91	50
L34-L33-L32	117	50
L34-L30-L31	140	50
L33-L30-L31	106	50
L33-L34-L30	50	50
L25-L34-L35	108	50
L37-L39-L38	52	50
L39-L37-L38	63	50
L38-L35-L36	105	50
L39-L38-L37	110	50

## **VITA**

**NAME OF AUTHOR:** Huilin Ma

**PLACE OF BIRTH:** Zhenjiang, Jiangsu, China

### **GRADUATION AND UNDERGRADUATE SCHOOLS ATTENDED:**

Bachelor of Engineering, 2013;

Department of Bioengineering, China Pharmaceutical University

Nanjing, Jiangsu, China.

Master of Science, 2015;

Department of Biomedical and Chemical Engineering, Syracuse University

Syracuse, New York, USA.

### **POSSESSIONAL EXPERIENCE**

Department of Biomedical and Chemical Engineering, Syracuse University

Teaching Assistant, 2015-2016

Research Assistant, 2016 - present