UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Chemistry

Complexity Matters: Polymyxin Antibiotics and the Bacterial Cell Envelope

by

Iain Peter Shand Smith
MPhys
ORCiD: 0000-0002-1562-3361

A thesis for the degree of
Doctor of Philosophy

November 2024

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Chemistry

Doctor of Philosophy

Complexity Matters: Polymyxin Antibiotics and the Bacterial Cell Envelope

by Iain Peter Shand Smith

The rise of antibiotic resistance in bacteria, such as *Escherichia coli*, poses a significant threat to global public health. One class of antibiotics, the polymyxins, are seeing a surge in use as a treatment against infection by multi-drug resistant Gram-negative bacteria due to their highly potent antimicrobial activity. Whilst these drugs are effective in their purpose, reports of severe nephro- and neurotoxicity have limited their use to only the most serious clinical cases in which all other treatment options have failed. This toxicity, in combination with recent reports of polymyxin resistance observed within certain bacterial strains, provides an immediate motivation to gain a mechanistic understanding of how these drugs perform each stage of their antimicrobial function so as to better inform the development of novel therapeutics. To do so in a rational manner demands a thorough understanding of the behaviour of the polymyxins within the region of the bacterial cell in which they perform their function.

To this end, the work presented in this thesis uses the computational technique of molecular dynamics for the purpose of simulating polymyxin antibiotics within the *E. coli* cell envelope. The nature of the molecular interactions between the polymyxins and the various structures of the cell envelope are studied in depth. Through varying the composition of the proteins, osmolytes and ions within the periplasm it is shown that the duration and biochemical nature of polymyxin binding with the cell wall and Braun's lipoprotein are modulated by interactions with the constituents of the surrounding environment.

The energetics and conformational dynamics of polymyxin permeation into the inner membrane are also investigated using a variety of enhanced sampling methods. It is shown that the polymyxins spontaneously insert into the headgroup region of the inner membrane through the adoption of a folded amphipathic conformation. The similarities of this mechanism to the insertion of other antimicrobial and designed peptides into lipid bilayers, reported in literature, are also discussed.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: Iain Peter Shand Smith, Conrado Pedebos, and Syma Khalid. Molecular crowding alters the interactions of polymyxin lipopeptides within the periplasm of e. coli: insights from molecular dynamics. *The Journal of Physical Chemistry B*, 128(11):2717–2733, 2024
   Conrado Pedebos, Iain Peter Shand Smith, Alister Boags, and Syma Khalid. The hitchhiker's guide to the periplasm: Unexpected molecular interactions of polymyxin B1 in E. coli. *Structure*, 29(5):444–456, May 2021. ISSN 1878-4186.

Signed:.. Date:..................

Acknowledgements

I would like to first thank my supervisor, Prof. Syma Khalid, for her support and guidance throughout my PhD studies. Your extensive subject expertise, and ability to deliver high impact research year upon year have been, and continue to be, a true inspiration.

Also my thanks go to Dr Conrado Pedebos. The foundations that you laid, upon which this thesis has sought to develop, were invaluable. Your kindness and patience when mentoring me at the start of my studies will never be forgotten.

To all members of the Khalid research group, past and present, and in particular Kahlan Newman and Callum Waller. These years have been made all the more enjoyable for the conversation and camaraderie that you each have provided.

To my father, Peter John Shand Smith, for providing a guiding light throughout each step of this journey. You inspire me today as you have done forever.

To my mother, Elisabeth Marjorie Glazener, for your endless love and support. Needless to say I would not be where I am today without you. There are no words strong enough to express my love and gratitude.

To my siblings, Rob and Catriona, no matter the distance you are always close to my heart.

And finally to all members of Southampton University Mountaineering Club, you have given me peace and joy through the brightest and darkest moments alike. I will treasure our memories for the rest of my life.

To my Opa - Hendrik Adriaan Glazener.

Het verdriet bestaat uit twee delen.

De eerste is verlies.
De tweede is het herscheppen van het leven.

Definitions and Abbreviations

AMP	Antimicrobial peptide
AA	All-atom
BLP	Braun′s lipoprotein
CG	Coarse grain
CoM	Center of mass
CMAP	Correction map
Conc	Concentrated
DAB	Diaminobutyric acid
DOPE	Dioleoylphosphatidylethanolamine
DOPC	Dioleoylphosphatidylcholine
DPPE	Dipalmitoylphosphatidylethanolamine
DPPC	Dipalmitoylphosphatidylcholine
EM	Energy minimisation
E. coli	*Escherichia coli*
IACT	Integrated autocorrelation time
IM	Inner membrane
IMP	Inner membrane protein
KDE	Kernel density estimate
LPS	Lipopolysaccharide
MD	Molecular dynamics
MSD	Mean squared displacement
meso-DAP	Meso-diaminopimelic acid
NAG	N-acetylglucosamine
NAM	N-acetylmuramic acid
Neut	Neutralised
NVT	Isothermal isochoric ensemble
NPT	Isothermal isobaric ensemble
OM	Outer membrane
OMP	Outer membrane protein
OPG	Osmoregulated periplasmic glycan
PBC	Periodic boundary condition
PGN	Peptidoglycan

PMF	Potential of mean force
PMB	Polymyxin B (admixture)
PMB1	Polymyxin B1
PME	Polymyxin E (colistin)
POPC	Palmitoyloleoylphosphatidylcholine
POPE	Palmitoyloleoylphosphoethanolamine
POPG	Palmitoyloleoylphosphoglycerol
RDF	Radial distribution function
ROG	Radius of gyration
RMSD	Root mean squared deviation
UBQ	Ubiquitin
SPD	Spermidine
SMD	Steered MD
SASA	Solvent accessible surface area
US	Umbrella sampling
vdW	van der Waals
WHAM	Weighted histogram analysis method

Chapter 1

Introduction

> *"If we fail to act, we are looking at an almost unthinkable scenario where antibiotics no longer work and we are cast back into the dark ages of medicine"* (Cameron (2014))

A major UN study Antimicrobial Resistance Collaborators (2022) reported that an estimated 1.27 million deaths in 2019 were attributable to bacterial antimicrobial resistance. A separate study commissioned by the UK government O'Neill (2016) predicted that this number may rise to 10 million by 2050 if proactive solutions were not found to slow the rise of drug resistance; putting at risk a cumulative 100 trillion USD of economic output. In order to combat a threat of such magnitude, a determined response from the medical sciences is to be demanded. However, due to the large upfront costs of developing novel antimicrobial agents and the comparatively small revenue generated from their sale, many major pharmaceutical companies have stopped funding research in this field altogether Plackett (2020).

There is thus an immediate need for academics and charities to fill this gap in the research and to develop novel therapeutic agents to combat multi-drug resistant bacteria. To do so in a rational manner demands a thorough understanding of the challenges posed to the motion of antibiotics as they interact with the complex environment of the bacterial cell.

1.1 The Cell

All forms of cellular life may be classified into 3 so-called *domains*; bacteria, archaea and eukarya. Eukarya, or eukaryotes, are uni- or multi-cellular organisms whose cells contain membrane-bound organelles such as nuclei, mitochondria, chloroplasts, Golgi apparatus or the endoplasmic reticulum (ER). These organelles are specialised subunits of the cell that perform some particular function critical to cell survival and, in eukaryotic cells, they are almost always separated from the rest of the cellular space by a membrane. Conversely, bacteria and archaea are strictly unicellular with organelles that are not partitioned from the surrounding cytoplasm; these features are used to define them as the constituent members of a broader group called the prokaryotes. The work presented in this thesis focuses on the bacterial domain, and so we shall restrict our discussion to the particulars of the bacterial cell.

The bacterial cell comprises two parts; the cytoplasm and the cell envelope. The cytoplasm is a gel-like matrix of water, enzymes, nutrients, wastes and gases that forms the interior of the cell and is encapsulated by the cell envelope; providing refuge for the ribosomes, chromosomes and plasmids that all contribute to the proper function of the cell. The cell envelope, as mentioned, encapsulates the cytoplasm; providing a barrier between the interior of the cell and the often hostile external environment, thus acting as the gatekeeper for molecules entering the cell.

In order to combat bacterial infection, antibiotics must either kill the cell directly or suppress cell growth and reproduction sufficiently to enable the host immune system to overcome the infection. This delineation defines the two categories of antibiotics, bactericidal (induce cell lysis) and bacteriostatic (suppress cell growth) Bernatová et al. (2013); Baquero and Levin (2021); Loree and Lappin (2019). Both of these modes of action, however, require the antibiotic to first gain entry into the cell; thus understanding the nature of molecular interactions that occur between these antimicrobial agents and the various components of the bacterial cell envelope is vital to further our understanding of how they perform their function.

1.2 The Bacterial Cell Envelope

The bacterial cell envelope is a complex, multilayered structure that, as mentioned, wholly encapsulates the cytoplasmic region of the cell. In the case of Gram-negative bacteria, the cell envelope comprises two lipid bilayers; the inner membrane (IM) and outer membrane (OM) that form the boundaries of a crowded aqueous compartment known as the periplasm. Within the periplasm lies the cell wall, a mesh-like structure composed of cross-linked strands of peptidoglycan (PGN) polymers, along with a wide variety of proteins, osmolytes and ions Goemans et al. (2014); Weiner and Li (2008).

The only protein known to provide a covalent link with the cell wall is Braun's lipoprotein (BLP, also known as ”Lpp” or murein lipoprotein) Braun and Rehn (1969); Braun and Bosch (1972). BLP is anchored in the OM *via* a lipidated N-terminus, with the C-terminus covalently bound to the cell wall. BLP is one of the most abundantly expressed proteins in Gram-negative bacteria, with an estimated 3×10^5 copies per

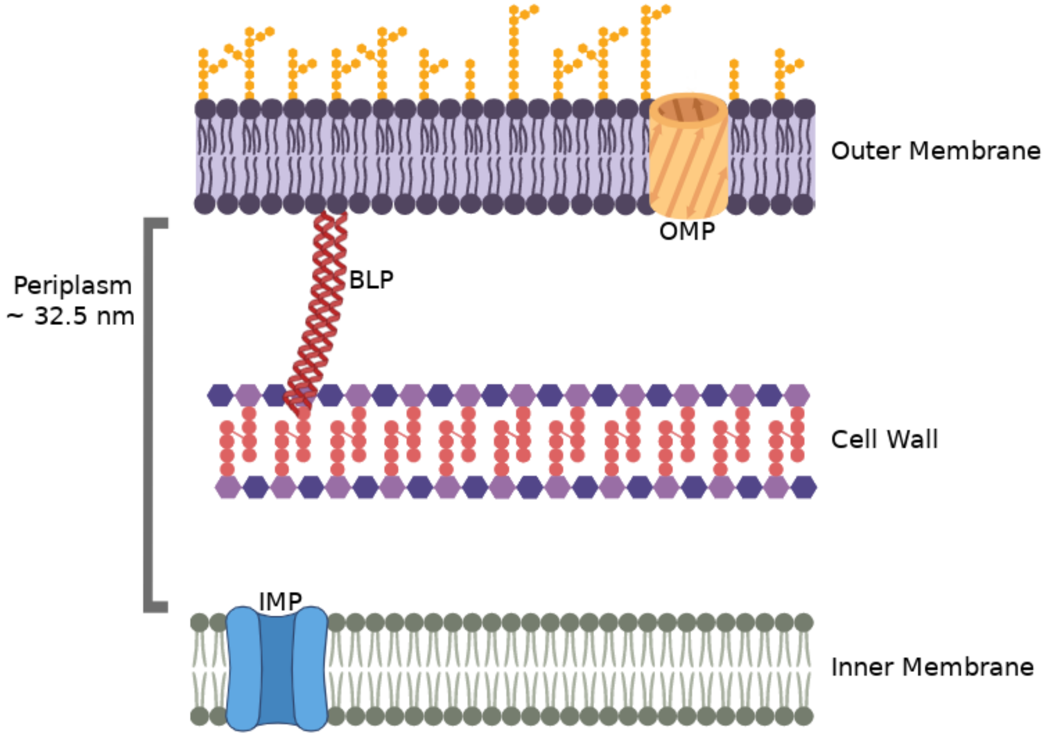


FIGURE 1.1: Schematic representation of the Gram-negative cell envelope. The periplasmic width of *E. coli* is estimated experimentally to be approximately 32.5 nm Matias et al. (2003).
Figure created with BioRender.

cell Braun (1975). This abundance, along with the connection that BLP provides between the OM and cell wall, enables the protein to act as a scaffold between the two structures, modulating their separation Cohen et al. (2017) and facilitating the non-covalent interaction of other OM proteins with the cell wall Samsudin et al. (2017).

The cell envelope of Gram-positive bacteria is comparatively simple, lacking an outer membrane and thus comprised of a single cytoplasmic membrane, encapsulated by a thick cell wall composed of numerous layers of the PGN mesh Silhavy et al. (2010).

1.2.1 The Peptidoglycan Cell Wall

One of the defining features distinguishing bacteria from eukarya is the presence of the PGN cell wall that envelopes the cell, maintaining its structural integrity and giving rise to the many varied morphologies observed across the different bacterial species Vollmer et al. (2008); van Teeseling et al. (2017). The cell wall also provides a barrier to the passage of molecules into and out of the cell; sequestering vital metabolites and excluding cytotoxic molecules whilst providing resistance to the effects of osmotic and turgor pressures that would, in its absence, result in lysis Silhavy et al. (2010).

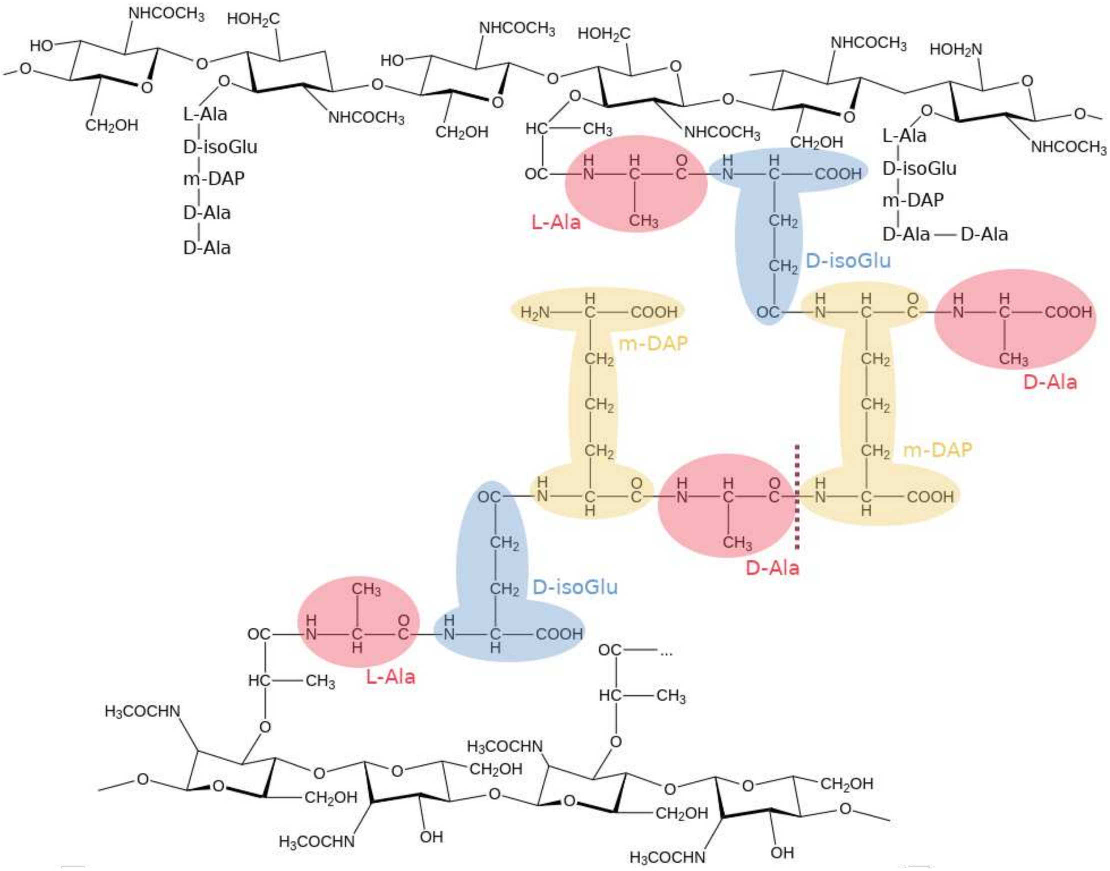


FIGURE 1.2: Structure of the peptidoglycan cell wall of *E. coli*. The structure of the peptide cross-link between adjacent glycan strands is shown in the center, with the peptide bond between adjacent PGN monomers highlighted by the dotted red line. The L-Ala and D-Ala residues of the cell wall peptide stems are highlighted with red bubbles, D-isoGlu residues are highlighted with blue bubbles and the meso-diaminopimelic acid (meso-DAP) residues are highlighted by yellow bubbles.

Since the cell wall is not present in eukaryotes, and bacteria are dependent upon its proper function for their survival; the cell wall presents an obvious target for the development of antimicrobial agents that express selective toxicity to bacteria while posing a limited threat to the mammalian cell. Indeed, cell wall biosynthesis inhibitors (CBIs) such as penicillins, cephalosporins and vancomycin have historically been one of the most effective classes of antibiotics Sarkar et al. (2017), and make up the majority of prescribed antibiotics throughout the developed West Center for Disease Control et al. (2020); Bruyndonckx et al. (2021).

The structure of the cell wall is largely conserved between Gram-negative and Gram-positive bacteria; and is composed of linear glycan strands, cross-linked by short peptides (Figure 1.2). The glycan strands of both Gram-negative and Gram-positive bacteria are made up of alternating $\beta - (1,4)$ linked N-Acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues. The D-lactoyl group of each NAM residue is substituted by a peptide stem, typically composed of 5 amino acids. The exact sequence of the peptide stem varies between species Schleifer and Kandler (1972), but is most often composed of (1) L-Ala - (2) D-isoGlu - (3) meso-diaminopimelic acid (meso-DAP) or L-Lys - (4) D-Ala - (5) D-Ala Vollmer et al. (2008); Rajagopal and Walker (2017). Adjacent glycan strands are typically cross-linked *via* peptide bonds between the carbonyl group of the D-Ala residue at position (4) of the peptide stem and the amino group of the meso-DAP (or L-Lys) residue at position (3); in the case of *E. coli*, roughly 50% of peptide stems form such cross-links Glauner et al. (1988). PGN monomers are polymerised in this manner on the external side of the cytoplasmic membrane; forming the mesh-like cell wall structure that wholly encapsulates the cell, also known as the sacculus.

The cell wall of Gram-negative bacteria is composed of a single layer of the PGN mesh, resulting in a thin sacculus ($\sim$ 4 nm) Huang et al. (2008); Gumbart et al. (2014). In contrast to this, the cell wall of Gram-positive bacteria comprise numerous layers of the PGN mesh, resulting in a much thicker sacculus ($\sim$ 20 - 80 nm) Salton et al. (1996). The distinction between these two bacterial sub-divisions was first discovered by Christian Gram in 1884 through application of the eponymous Gram staining procedure Gram (1884).

1.2.2 Gram Staining and Bacterial Subdivisions

The Gram staining procedure involves treating a cell with a form of dye, typically crystal violet Tripathi and Sapra (2023), that is known to be absorbed into the cell envelope of all bacteria; saturating both the lipid bilayer and the pores within the cell wall. The dye is then fixed using iodine, and a solvent is subsequently applied to the cell; dissolving the lipid membrane and dehydrating the cell wall, causing the pores in its structure to contract. In Gram-positive bacteria, contraction of the pores in the thick multilayered cell wall traps the dye within its structure and thus the cell remains stained even after dissolution of the lipid membrane, giving rise to a ′positive′ staining result. In Gram-negative bacteria, the single layered cell wall lacks the volume and internal structure required to retain the dye molecules upon application of the solvent and thus, once the lipid membrane dissolves, the cell does not remain stained and returns a ′negative′ staining result.

Gram-negative bacteria are considered to be among the most significant threats to global public health due to their high resistance to antibiotics Oliveira and Reygaert (2019). They are responsible for a wide range of food-borne and nosocomial infections such as pneumonia, meningitis and bloodstream, wound or surgical site infections Oliveira and Reygaert (2019); Ramirez and Giron (2023); thus they pose a significant risk within healthcare settings, particularly towards immuno-compromised and intensive care patients.

Whilst the threat posed by Gram-positive bacteria is still significant Bouza and Finch (2001), the structural characteristics of their cell envelope make them an easier target for clinical intervention compared to their Gram-negative counterparts. Their thick cell wall presents an obvious target for selective antibiotics, such as the aforementioned CBIs, whilst the lack of an OM encapsulating the cell wall leaves it easily accessible to such agents.

1.2.3 The Outer Membrane

The outer membrane of Gram-negative bacteria is an asymmetric lipid bilayer with an outer leaflet composed primarily of the glycolipid lipopolysaccharide (LPS) Funahara and Nikaido (1980); Caroff and Karibian (2003); and an inner leaflet composed of a mixture of phospholipids, primarily with phosphoethanolamine (PE), phosphoglycerol (PG) and cardiolipin headgroups (Figure 1.4). Both the length and saturation of the acyl tails on these phospholipids can vary Ernst et al. (2016); however, the typical length is between 12 to 18 carbons Rilfors and Lindblom (2002). Amongst these lipids, the OM is interspersed with various integral outer membrane proteins (OMPs), such as porins and efflux pumps, which regulate nutrient uptake, drug resistance, and virulence factor secretion Silhavy et al. (2010); Koebnik et al. (2000); Blair et al. (2014); Soto (2013); Green and Mecsas (2016).

The structure of LPS (Figure 1.3) may be broken down into 3 main components; lipid A, the core oligosaccharides, and the O-antigen polysaccharide Bertani and Ruiz (2018). Lipid A serves as the hydrophobic anchor for LPS insertion into the OM, whilst the core oligosaccharides form a bridge between the lipid A and O-antigen regions. The O-antigen polysaccharide projects radially outward from the OM and exhibits remarkable structural diversity between Gram-negative species, contributing to antigenic variation and evasion of the host immune system Reeves (1995).

When packed together in the OM, the polysaccharide chains of neighbouring LPS molecules interact with one another, forming strong electrostatic interactions that result in exceptionally slow lateral diffusion rates within the outer leaflet; increasing the structural integrity of the membrane and forming a formidable barrier to chemical attack Jefferies et al. (2019). This permeability barrier, formed by LPS in the outermost layer of the cell envelope, is a main contributor to the innate resistance that Gram-negative bacteria display against many antimicrobials and its absence in Gram-positive bacteria is a key factor in the relative ease with which clinicians are able to treat Gram-positive infections Bertani and Ruiz (2018); Miller (2016); Exner et al. (2017).

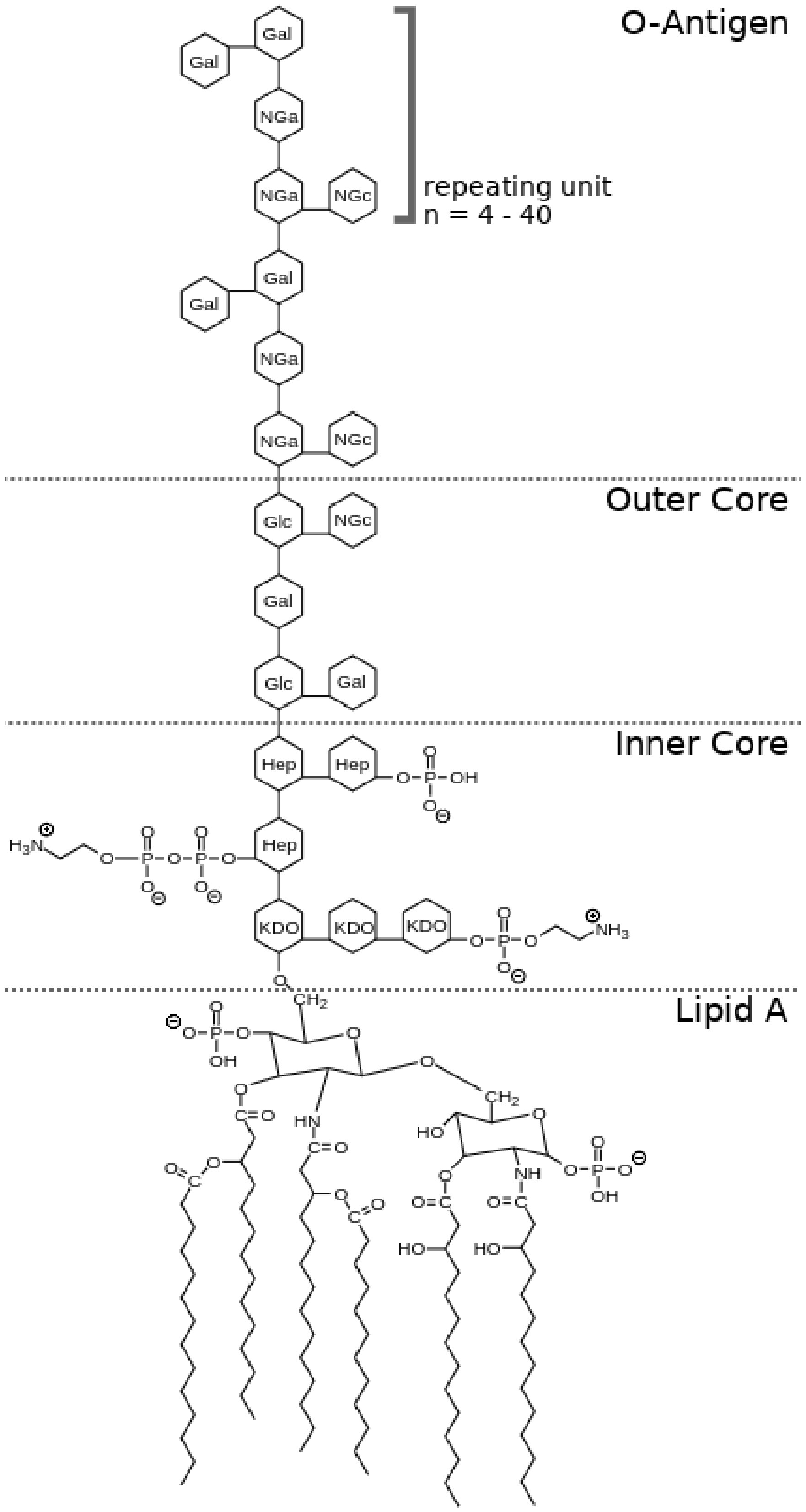


FIGURE 1.3: Chemical structure of the full-length R3 core LPS from *E. coli* Rietschel et al. (1994).

1.2.4 The Inner Membrane

One of the defining traits of eukaryotic cells is the presence of membrane-bound, intracellular organelles. These organelles perform various cellular processes necessary for the proper function of the cell: mitochondria produce energy for the cell through oxidative phosphorylation; the smooth ER synthesizes lipids whilst the rough ER is responsible for protein secretion. Since bacteria lack these intracellular organelles, all of the membrane-associated functions of the eukaryotic organelles must be performed by the bacterial IM Silhavy et al. (2010).

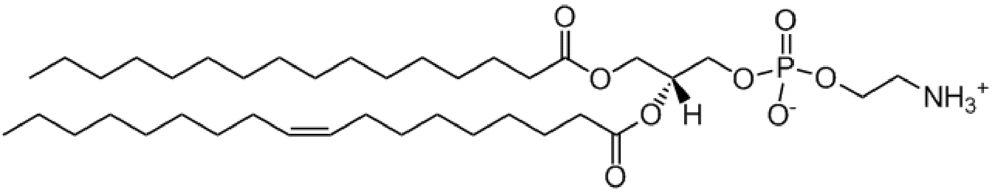
The IM also serves an important role as a selective permeability barrier to the interior of the cell; excluding the entry of large molecules, whilst allowing certain smaller molecules to pass through. To aid the membrane in this function, many of the IM transmembrane proteins take the form of pores or ion channels; providing direct pathways for the uptake and removal of small organic molecules and ions between the cytoplasmic and periplasmic regions of the cell Papanastasiou et al. (2013).

The Gram-negative IM is a phospholipid bilayer composed primarily of the same PE, PG and cardiolipin phospholipids that form the inner leaflet of the OM. Whilst the exact distribution of these phospholipids across the Gram-negative IM is unknown, it has recently been reported that rod-shaped *E. coli* cells exhibit a 75%/25% distribution of PE lipids in the cytoplasmic/periplasmic leaflets, respectively Bogdanov et al. (2020). Despite this, it is still the norm for contemporary simulation studies to consider the IM as a simplified symmetric bilayer Corey et al. (2021); Waller et al. (2023); Chakraborty et al. (2020).

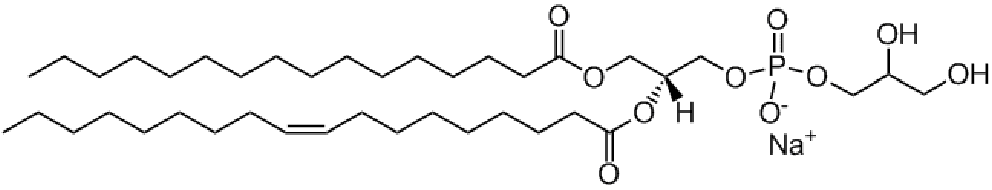
1.3 The Model Bacterium, *E. coli*

One particular species of Gram-negative bacterium is *E. coli*, belonging to the family *Enterobacteriaceae* which accounts for approximately 80% of Gram-negative infections identified in clinic Oliveira and Reygaert (2019). *E. coli* is the leading cause of AMR related mortality worldwide Antimicrobial Resistance Collaborators (2022), and the second largest contributor to mortality due to bacterial infection in high income countries Ikuta et al. (2022). Thus, understanding the interactions of this bacterium with antimicrobial agents is of key clinical interest. To this end, *E. coli* has been used as a model to study Gram-negative bacteria for well over a century due to its clinical relevance and the abundance of non-pathogenic, hardy strains that are easily cultured on a variety of nutrient media. For these reasons, the simulations presented in this thesis used *E. coli* as a model organism to investigate the interactions of antibiotics within the Gram-negative cell envelope.

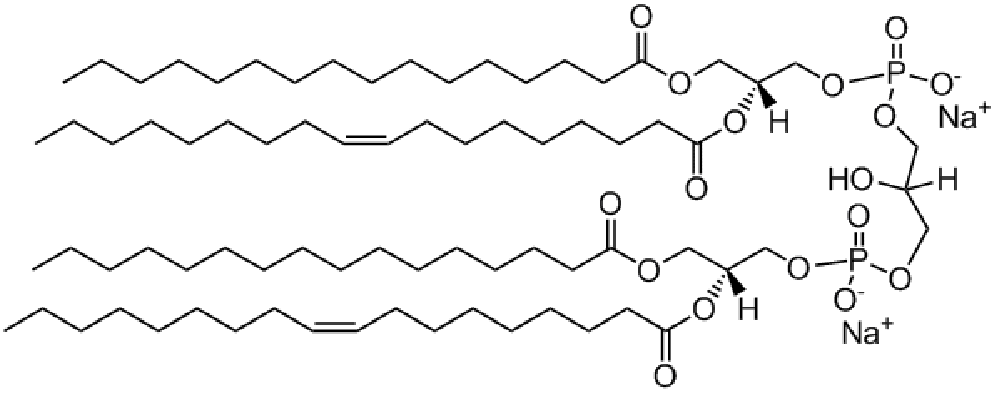
1.3.1 Composition of the *E. coli* Cell Envelope Model



(A) 16:0,18:1 phosphoethanolamine (POPE)



(B) 16:0,18:1 phosphoglycerol (POPG)



(C) 18:1,18:1 / 18:1,18:1 Cardiolipin

FIGURE 1.4: Chemical structures of common *E. coli* phospholipids. The length and saturation of the aliphatic tails may vary Sinensky (1971); Rowlett et al. (2017).

The work presented in Chapters 4 and 5 of this thesis utilised an asymmetric model of the outer membrane; with an inner leaflet lipid composition of 90% 16:0,18:1 phosphoethanolamine (POPE, Figure 1.4a), 10% 16:0,18:1 phosphoglycerol (POPG, Figure 1.4b) and 10% 18:1,18:1 / 18:1,18:1 cardiolipin (Figure 1.4c) and an outer leaflet composed entirely out of LPS. This composition of OM lipids and the corresponding levels of tail saturation were based on experimental studies of *E. coli* Lugtenberg and Van Alphen (1983) and have been validated by previous simulation studies performed by the Khalid research group Piggot et al. (2011); Ortiz-Suarez et al. (2016); Samsudin et al. (2016).

Since the focus of this work was primarily on the behaviour and interactions of molecules within the periplasmic region of the cell envelope, a model of LPS was chosen that simplified the complex environment on the exterior of the OM outer leaflet. This was achieved through the use of an LPS model that included only the lipid A region of each molecule. This choice excluded the entirety of the core oligosaccharide and O-antigen regions, thus removing a vast region of complex electrostatic interactions from the system. This allowed for the simulations to be performed within a smaller periodic domain involving fewer atoms, thus leading to a faster rate of simulation. Furthermore, the absence of the core oligosaccharide and O-antigen regions directly sped up the correct packing of acyl tails within the OM during the various stages of equilibration; as the lateral movement of outer leaflet lipids was not slowed down by the abundance of strong electrostatic interactions between the external polysaccharide chains.

A model of the cell wall was utilised that consisted of a single layer of the PGN mesh. Each PGN monomer comprised one pair of $\beta - (1, 4)$ linked N-Acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues and a short peptide stem. The full five residue sequence of the peptide stem was (1) L-Ala - (2) D-isoGlu - (3) meso-DAP - (4) D-Ala - (5) D-Ala. Each PGN monomer was polymerised in the manner discussed in Section 1.2.1. A further detailed discussion of the construction of the *E. coli* cell envelope model and its constituent components can be found in Chapter 3.

1.4 Antimicrobial Peptides

Antimicrobial peptides (AMPs) are small, cationic, membrane-active peptides. They are found in most living organisms, playing an important role in the innate immune response of their hosts Izadpanah and Gallo (2005); Kimura et al. (1992); Kraemer et al. (2011). These peptides exhibit broad-spectrum antimicrobial activity against bacteria, fungi and viruses Lei et al. (2019) and thus they are of biomedical interest for use as therapeutic agents themselves or as the inspiration for other novel antibacterial agents.

1.4.1 Polymyxin B and Colistin

One class of highly potent bactericidal AMPs are the polymyxins; a family of cyclic lipopeptides originally derived from the bacterial species *Paenibacillus polymyxa* Storm et al. (1977). There are five chemically distinct compounds within the family, namely polymyxins A-E, however, only polymyxin B (PMB) and polymyxin E (PME or “colistin”) have been used in clinical practice. First approved for clinical use in the 1950s, their use was limited by the 1970s due to reports of severe nephro- and neurotoxicity Avedissian et al. (2019). In recent decades, however, the emergence of multi-drug resistant Gram-negative “superbugs” and their associated threat to global public health Jasovský et al. (2016), coupled with improvements in clinical application Dubashynskaya and Skorik (2020) and reports of lower levels of toxicity Falagas and

Kasiakou (2006), has led to the revival of their use as a last-resort intervention when all other treatment options have failed Falagas et al. (2021); Vaara (2019).

1.4.2 Polymyxin Structure

Polymyxin B is an admixture of polymyxins B1, B1-I, B2, B3 and B6; with polymyxins B1 and B2 considered to be the major components. These compounds are structurally identical, aside from slight variations to the acyl tail at the terminus of the branched fatty acid region. Given these structural similarities, and its dominance over the other admixture components, the work presented in this thesis focuses on polymyxin B1 (PMB1) (Figure 1.5A). The full sequence of PMB1 is DABC − Thr − DAB − DAB − DLeu − DPhe − DAB − DABC − DAB − Thr − DAB − $CO(CH_2)4CH(CH_3)CH_2CH_3$, where DABC represents the cyclised diaminobutyric acid (DAB) residue. The 5 non-cyclised DAB residues each contain a protonated amine group carrying charge +1 *e*, conferring a total charge of +5 *e* to the molecule.

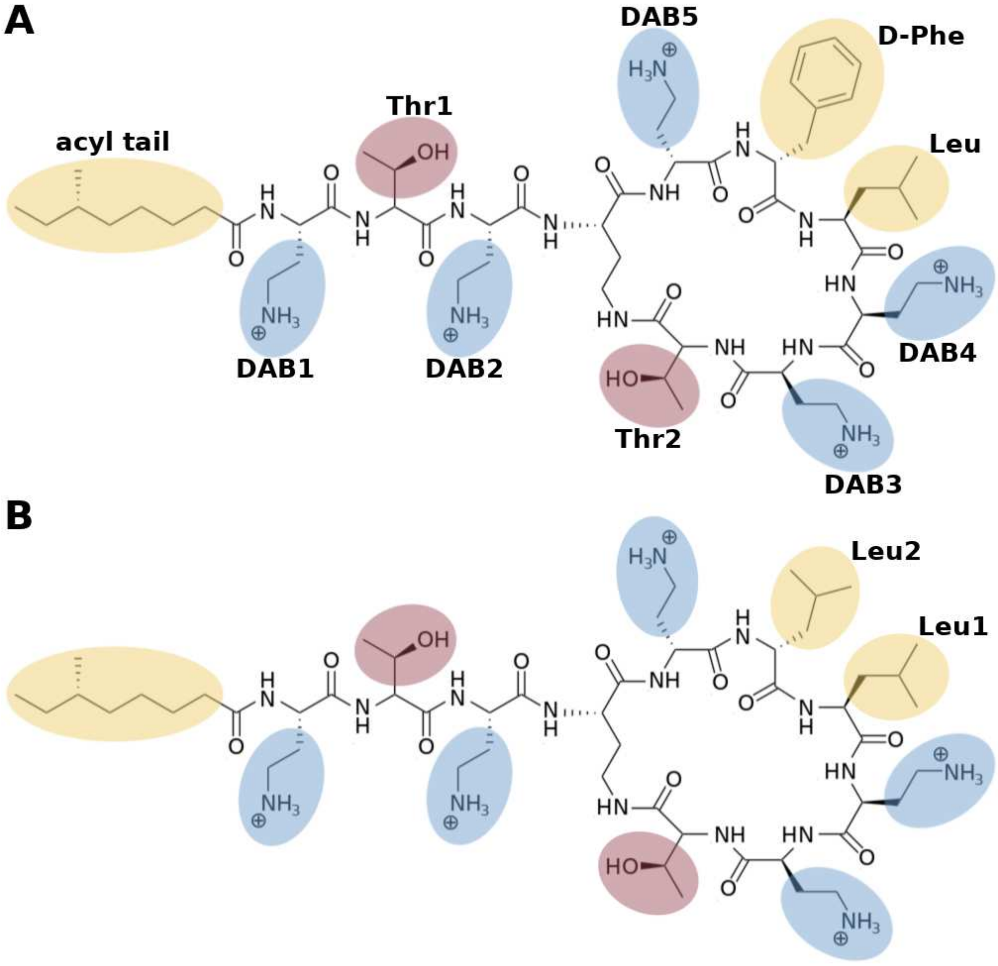


FIGURE 1.5: Chemical structures of A) PMB1 and B) PME. Hydrophobic residues are highlighted in yellow, basic DAB residues in blue and polar Thr residues in red.

Colistin (PME) is structurally similar to PMB1, differing only in the substitution of the D-Phe residue on the heptapeptide ring of PMB1 for a D-Leu residue in PME (Figure 1.5B). The full sequence of PME is thus DABC − Thr − DAB − DAB − DLeu − DLeu − DAB − DABC − DAB − Thr − DAB − $CO(CH_2)4CH(CH_3)CH_2CH_3$. Similar to PMB1, the 5 non-cyclised DAB residues within the structure of PME each contain a protonated amine group carrying charge +1 e, thus conferring a total charge of +5 e to PME.

The amphipathic nature of the polymyxins, arising from the cationic DAB residues and hydrophobic D-Phe, Leu and acyl tail residues, enables the polymyxins to disrupt both bacterial and mammalian cell membranes Jiang et al. (2020c); Velkov et al. (2010) and is highly likely the origin of both their potent antimicrobial activity and clinical toxicity.

1.5 The Importance of Crowding

Throughout the field of biomolecular simulation it has been common practice to investigate the functional dynamics of a given biophysical system in a highly idealised chemical environment, often with just a few species of interest in solution with ions alone Dieudonné et al. (2023); Park et al. (2023); McDowell et al. (2023); Li et al. (2023). This approach, however, ignores the vast complexity of the biological environments within which these molecules exist *in vivo*.

It has long been known that the cytoplasmic region of living cells is densely packed with macromolecules at volume fractions of up to 40% Zimmerman and Minton (1993); such macromolecular crowding necessarily affects the behaviour of the molecules within this region through excluded volume effects and the addition of a multitude of varied electrostatic interactions between constituents. These effects are most pronounced when considering proteins with intrinsically disordered fragments, or those that undergo significant conformational transitions as part of their function Ostrowska et al. (2019). Whilst a majority of proteins behave in this way, examples of contemporary studies of such proteins can be found that make no attempt to consider the potential impact of the biological environment on their results Robustelli et al. (2020); Shrestha et al. (2021); Herrera-Nieto et al. (2023).

The slow rate at which protein folding takes place *in vivo* has typically limited the study of these systems using MD simulations, often necessitating the use of enhanced sampling techniques and/or a reduction of the model complexity Gershenson et al. (2020); Durrant and McCammon (2011). However, the presence of external factors during the *in vivo* folding process will no doubt impact the direct biological relevance of the *in silico* folding mechanisms obtained in an idealised environment Devi et al. (2022); Macdonald and Johnson JR (2001); Xu et al. (2021).

These limitations further extend to the study of protein-ligand binding. It has long been known known that complex formation between a protein and ligand implies some structural rearrangement within that protein Betts and Sternberg (1999) and, more recently, that even small conformational changes may impact the accuracy of ligand-docking algorithms Hospital et al. (2015). Various methods have been implemented to account for these conformational changes during docking Meng et al. (2011); Pinzi and Rastelli (2019), however the choice of method used has been shown to significantly affect the accuracy of the predicted binding pose Elokely and Doerksen (2013). The formation of complexes between proteins, drugs and the vast array of biomolecules that comprise the *in vivo* environment is therefore highly likely to significantly affect the conformational dynamics and thus the function of these molecules.

Aside from protein folding and docking, considerable work has been done to quantify the effects that molecular crowding has on both small and large molecular diffusion rates von Bülow et al. (2019); Kekenes-Huskey et al. (2016); Pedebos et al. (2021), lipid diffusion, sorting and membrane deformation Guigas and Weiss (2016); Nawrocki et al. (2019). Whilst these studies have begun to reveal the importance of performing simulations with biologically accurate solute compositions, they still do not account for the presence of the complete cell envelope. It is therefore still largely unknown how the presence of structural macromolecules, such as BLP or the cell wall, may affect the dynamics and interactions of small molecules within the crowded periplasm (or *vice versa*).

A study previously published by the Khalid research group, on which I am a co-author, investigated the behaviour of PMB1 within the crowded environment of the *E. coli* cell envelope using the GROMOS54a7 united-atom forcefield Pedebos et al. (2021). It was found that PMB1 formed both transient and long-lived interactions with the proteins and osmolytes within the periplasmic space, as well as with BLP; greatly reducing the observed translational diffusion coefficients as the crowding volume fraction was increased. The diffusion coefficients for the proteins, however, closely matched those measured in simulations that did not include the cell wall or membrane and thus further investigation is required into what effect the various structures of the bacterial cell envelope may have on molecular interactions within the periplasm.

1.6 Outline of Work

The work presented in this thesis therefore aims to shed light on the nature of polymyxin interactions with the various components of the *E. coli* cell envelope and how they may be impacted by biomolecular crowding and chemical complexity. Thus, Chapter 3 focuses on the construction of a new model of the *E. coli* cell envelope using the all-atom CHARMM36m forcefield Huang et al. (2017). Analysis of this model is presented in Chapters 4 & 5.

Simulations of the *E. coli* cell envelope were performed under three different periplasmic fluid compositions of varying complexity with the aim of characterising whether the various molecular interactions experienced by PMB1 and PME during translocation between the OM and IM facilitate or hinder their movement. The task of constructing a biologically accurate model of the periplasm is not trivial, with sparse information regarding the exact composition, concentration and distribution of molecular components. The model periplasm presented in this thesis was therefore constructed with a subset of periplasmic osmolytes that were both known to be abundant within the periplasm and that represented a range of diverse chemical behaviours. The model periplasm included osmoregulated periplasmic glucans (OPG), an oligosaccharide that plays a prominent role in regulating osmotic pressure and virulence Cayley et al. (2000). The polyamine spermidine was also included, whose function includes supporting bacterial growth, incorporation into the cell wall and biosynthesis of siderophores Cohen (1997). Finally, glycerol was also included, a small molecule that is metabolized in *E. coli*, both aerobically and anaerobically, for a range of functions Wang et al. (2019); Shah and Swiatlo (2008).

Whilst by no means is this a comprehensive set of molecular constituents comparable with the true complexity of the biological environment, this choice allows us to probe the impacts of various different chemical moieties on the nature of polymyxin interactions within the cell envelope whilst minimizing the computational cost and complexity of the analysis. In this manner, the work presented in this thesis provides a much needed initial step towards the fully atomistic study of the impact of biological complexity on the interactions of AMPs within the Gram-negative periplasm.

Thus, through analysis of the aforementioned simulations, Chapter 4 highlights the impacts of biomolecular crowding on polymyxin interactions with the cell wall, whilst Chapter 5 focuses on how such crowding also impacts the nature of polymyxin interactions with BLP.

To complete our picture of polymyxin behaviour throughout the cell envelope, Chapter 6 focuses on the nature of polymyxin interactions with the IM using the coarse-grain MARTINI2.2 forcefield de Jong et al. (2013). We investigate the energetic barriers to polymyxin permeation into the IM, highlighting how the initial insertion of polymyxin into the IM is driven by the adoption of the folded amphipathic binding mode at the membrane surface. The nature of PMB1 aggregation at the surface of the IM is also studied, illustrating the role that the clustering of hydrophobic D-Phe residues plays in the formation of such aggregates.

Chapter 2

Methods

> *"Thermodynamics is one of those words best avoided in a book with any pretence to be popular"* (Lane (2010))

In this chapter we discuss the range of computational techniques used in this work. The mathematical framework of Molecular Dynamics (MD) simulation is described in depth, along with the theory of statistical mechanics that allows for information to be extracted from these simulations. We provide an overview of all the analysis tools used and discuss the implementation of the new analysis scripts that have been written for this thesis.

2.1 Molecular Dynamics

In classical MD, Newton's second law of motion (Equation 2.1) is applied to a system of atoms in order to generate a so-called 'trajectory' - a chronologically ordered set of system configurations, separated by a small and constant time interval, that specifies how the positions and velocities of the atoms in the system vary with time.

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \tag{2.1}$$

Where $\mathbf{F}_i$ is the force applied on particle i, m_i its mass, $\mathbf{a}_i$ its acceleration, $\mathbf{r}_i = (x_i, y_i, z_i)$ its Cartesian coordinates and t the time.

The resultant force exerted upon each individual particle is comprised of contributions from its interactions with all nearby surrounding atoms. This dependence of the force calculations on each particle's position relative to all other particles in the system leads to a set of high dimensional coupled differential equations governing the motion of the system that are unable to be solved analytically. This therefore necessitates the use of numerical procedures to integrate the equations of motion, usually in the form of a finite difference method. In order to evaluate these equations of motion, however, we must first provide the potential energy function for all types of interaction present within the system and their corresponding parameters. This is achieved through the provision of a so-called ‘force field’.

2.1.1 Force Fields

In order to calculate the trajectory of each particle in the system, according to Equation 2.1, we must be able to evaluate the individual forces exerted on each particle within the system. To achieve this we utilise a ‘force field’; a global potential energy function, U, with an accompanying set of parameters defining all possible particle types, their corresponding properties (*e.g.* charge, mass and radius), the bonded interaction strengths of all possible particle pairs, triplets and quadruplets and the non-bonded interaction strengths of all particle pairs. These parameters are typically derived from comparison to experimental data and quantum mechanical calculations.

There are a variety of force fields available for public use, each optimised for a particular use-case, that fall into three broad categories: all atom (AA), united atom (UA) and coarse grain (CG). Each of these categories represents a different level of resolution at which a molecular system may be modelled. AA force fields provide the greatest level of detail, modelling every atom within the system as an individual particle; this approach provides explicit detail of interatomic interactions at the cost of considerable computational expense. In contrast, CG force fields group multiple atoms together into a single particle (or ’bead’); reducing the computational complexity of the system at the cost of the atomic-level detail provided by AA models. UA force fields provide a compromise between these two paradigms, whereby aliphatic (non-polar) hydrogens are ignored whilst all other atoms are treated as individual particles.

The work presented in Chapters 4 and 5 of this thesis used the AA CHARMM36m force field Huang et al. (2017), whilst the work presented in Chapter 6 used the CG MARTINI2.2 de Jong et al. (2013) force field. The precise details of the composition of these force fields will be discussed in the following sections.

Despite the differences in resolution, the general form of the global potential energy function, U, is similar amongst all force fields; defined as the sum of all bonded and non-bonded potential energy terms (Equation 2.2).

$$U = U_{bonded} + U_{non-bonded} \tag{2.2}$$

Where the bonded potential energy term, U_{bonded}, includes contributions from all bonds (Section 2.1.1.1), angles (Section 2.1.1.2), dihedrals (Section 2.1.1.3) and impropers (Section 2.1.1.4) present within the system (Equation 2.3). The non-bonded potential energy term, $U_{non-bonded}$, includes contributions from all van der Waals (Section 2.1.1.5) and electrostatic (Section 2.1.1.6) interactions between non-bonded particle pairs (Equation 2.4). Notably, the treatment of angles and dihedrals within AA and CG force fields varies slightly, and will be discussed in more detail in the following sections.

$$U_{bonded} = \sum U_{bonds} + \sum U_{angles} + \sum U_{dihedrals} + \sum U_{impropers} \tag{2.3}$$

$$U_{non-bonded} = \sum U_{vdW} + \sum U_{Coulomb} \tag{2.4}$$

Through expressing the force on a given particle as the gradient of the global potential energy function at its coordinate position, we can recast Newton's equation of motion (Equation 2.1) as a function of the known potential energy of the system:

$$\mathbf{F}_i = -\nabla_i U \tag{2.5}$$

$$-\frac{dU}{d\mathbf{r}_i} = m_i \frac{d^2\mathbf{r}_i}{dt^2} \tag{2.6}$$

This enables the calculation of particle accelerations, and thus trajectories, according to the positions and parameters of the surrounding particles. The practical implementation of this procedure is achieved through the application of an MD integrator, discussed in more detail in Section 2.15.

2.1.1.1 Bond Interactions

The first term in Equation 2.3 represents the potential energy of explicitly bound pairs of particles. Both the CHARMM36m and MARTINI2.2 force fields implement this term in the form of a harmonic restraint on the separation between the two bonded particles (Figure 2.1) and thus each bonded pair contributes the following potential energy component to the global potential energy function (Equation 2.2):

$$U_{bonds} = \sum_{bonds} \frac{1}{2} k_{ij} (r_{ij} - r_{0,ij})^2 \tag{2.7}$$

Where k_{ij} is the force constant associated with a bond between particles i and j, r_{ij} is the distance between the particles and $r_{0,ij}$ is their equilibrium bond length.

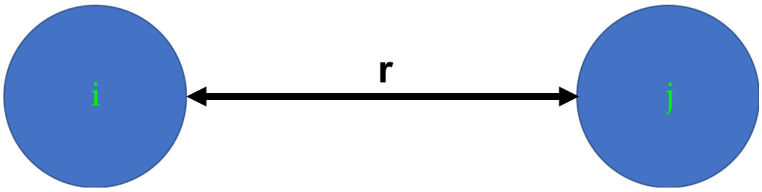


FIGURE 2.1: Schematic representation of the bond interaction (Equation 2.7) between two particles.

2.1.1.2 Valence Angle and Urey-Bradley Interactions

The second term in Equation 2.3 represents the potential energy of bonded particle triplets. In the CHARMM36m force field, this term takes the form of a harmonic restraint on the angle separating two adjacent bond vectors for which the central particle is shared, referred to as the valence angle (Figure 2.2). Each bonded triplet therefore contributes the following potential energy component to the global potential energy function (Equation 2.2):

$$U_{angles} = \sum_{angles} \frac{1}{2} k_{ijk} (\theta_{ijk} - \theta_{0,ijk})^2 \tag{2.8}$$

Where k_{ijk} is the valence angle force constant associated with particles i, j and k, θ_{ijk} is the valence angle between the particles and $\theta_{0,ijk}$ is their equilibrium valence angle.

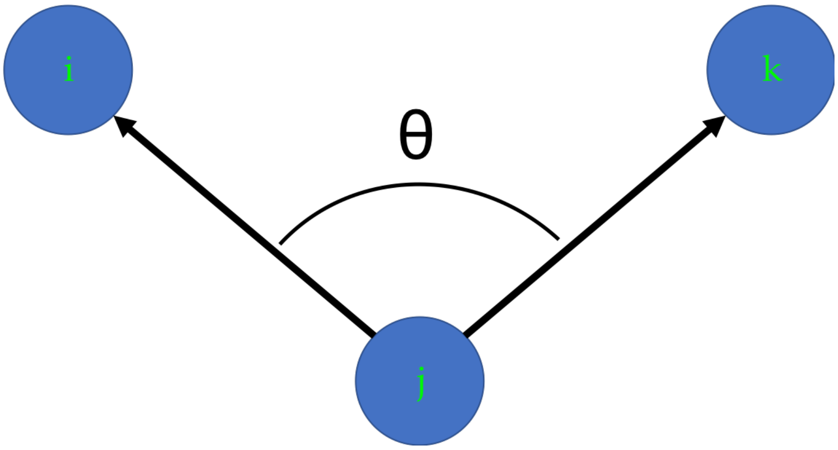


FIGURE 2.2: Schematic representation of the valence angle interaction (Equation 2.8) between three particles.

The CHARMM36m force field includes a further correction term to the potential energy component of each valence angle in the system; referred to as the Urey-Bradley correction. The Urey-Bradley correction limits the vibrational modes of each valence triplet through the application of a harmonic restraint on the distance between the two non-bonded particles, i and k (Figure 2.2). The additional contribution of the

Urey-Bradley correction to the global potential energy function (Equation 2.2) of the CHARMM36m force field therefore takes the following form:

$$U_{Urey-Bradley} = \sum_{\text{Urey-Bradley}} \frac{1}{2} k_{UB,ijk} (r_{ik} - r_{0,ik})^2 \tag{2.9}$$

Where $k_{UB,ijk}$ is the Urey-Bradley force constant associated with particles i, j and k, r_{ik} is the distance between particles i and k, and $r_{0,ik}$ is their equilibrium separation.

In the MARTINI2.2 force field, the valence angle potential energy term takes a similar form to that of the CHARMM36m force field; however, the harmonic restraint is applied to the cosine of the valence angle, as opposed to the angle itself. The potential energy contribution of valence angles to the global potential energy function (Equation 2.2) of the MARTINI2.2 force field therefore takes the following form:

$$U_{angles} = \sum_{angles} \frac{1}{2} k_{ijk} (\cos(\theta_{ijk}) - \cos(\theta_{0,ijk}))^2 \tag{2.10}$$

Where k_{ijk} is the valence angle force constant associated with particles i, j and k, θ_{ijk} is the valence angle between the particles and $\theta_{0,ijk}$ is their equilibrium valence angle.

2.1.1.3 Dihedral Interactions

The third term in Equation 2.3 represents the potential energy of dihedral angles between sequentially bound particle quadruplets. These dihedral restraints model the torsional forces that restrict the rotation of each quadruplet around their central bond. Given the positions of four sequentially bound particles, i, j, k and n; the dihedral angle is defined as the angle formed between the ijk and jkn planes (Figure 2.3).

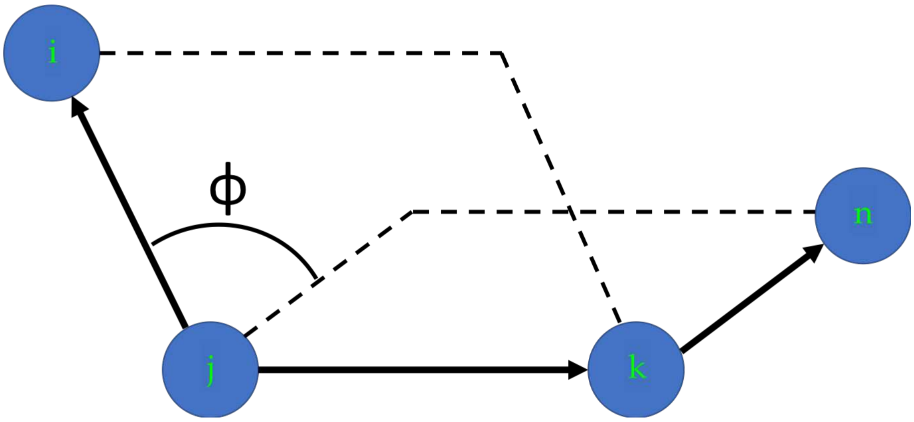


FIGURE 2.3: Schematic representation of the dihedral angle (Equation 2.11) between four particles.

The contribution of dihedral angles to the global potential energy function (Equation 2.2) of the CHARMM36m force field takes the following form:

$$U_{dihedrals} = \sum_{dihedrals} k_{ijkn}[1 + \cos(m\phi_{ijkn}) - \delta] \tag{2.11}$$

Where k_{ijkn} is the force constant associated with a dihedral between particles i, j, k and n, ϕ_{ijkn} is the dihedral angle, m is the periodicity of the potential and δ is the phase factor.

In the CHARMM36m force field, an additional correction map (CMAP) term is included between adjacent ϕ, ψ dihedral angles along the backbone of proteins. This term improves the accuracy of protein conformational sampling in general MacKerell Jr et al. (2004) and, in particular, resolves the spurious oversampling of left-handed α-helix secondary structures observed in simulations using the preceding CHARMM36 force field Huang et al. (2017). The potential energy contribution of the CMAP term is not implemented as a continuous function, rather it takes the form of a 2D grid of energy corrections in ϕ, ψ space.

The MARTINI2.2 force field does not include potential energy contributions from these ′proper′ dihedral angles, instead it only applies restraints on ′improper′ dihedral angles, discussed further in the following section.

2.1.1.4 Improper Dihedral Interactions

The fourth term in Equation 2.3 represents the potential energy of improper dihedral angles. These ‘impropers’ are used to restrict the rotation of particle quadruplets about a central atom. This is particularly useful in order to enforce a specific configuration within groups of four bonded atoms; such as maintaining the planarity of aromatic rings or restricting chiral centers from flipping to their mirror image. These improper dihedral angles are applied to particle quadruplets that are not bound in the sequential *i*-*j*-*k*-*n* manner illustrated in Figure 2.3; but are instead joined to one another *via* a single central atom (Figure 2.4).

Both the CHARMM36m and MARTINI2.2 force fields treat these improper dihedrals using a simple harmonic restraint applied to the dihedral angle formed between the *ijk* and *jkn* planes. Thus, in both force fields, each improper dihedral contributes a potential energy component of the following form to the global potential energy function (Equation 2.2):

$$U_{impropers} = \sum_{impropers} \frac{1}{2} k_{ijkn} (\phi_{ijkn} - \phi_{0,ijkn})^2 \tag{2.12}$$

Where k_{ijkn} is the improper dihedral force constant associated with particles i, j, k and n. ϕ_{ijkn} is the improper dihedral angle defined by the ijk and jkn planes, and $\phi_{0,ijkn}$ is the corresponding equilibrium angle.

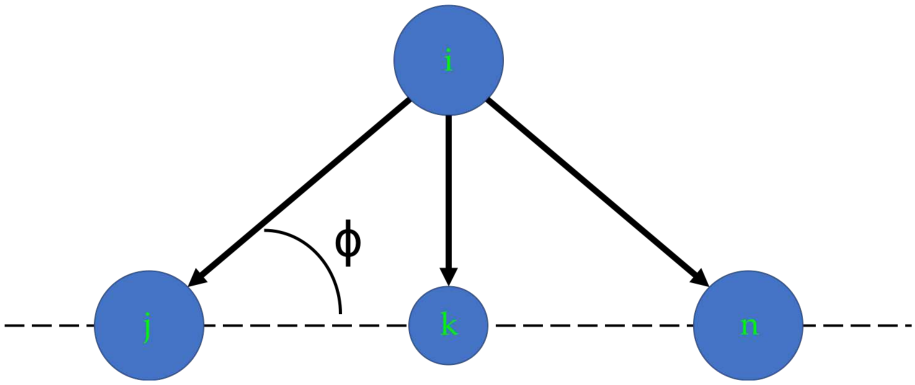


FIGURE 2.4: Schematic representation of the improper dihedral interaction (Equation 2.12) between four particles.

2.1.1.5 Van der Waals Interactions

The first term in Equation 2.4 represents the potential energy of short range interactions between non-bonded pairs of particles that arise from van der Waals (vdW) forces. Due to the Pauli exclusion principle, these forces are strongly repulsive when the particles' separation is smaller than their combined radii, thus stopping multiple particles from occupying the same position in space. The formation of instantaneous multipoles leads to a modestly attractive force when the particles' separation exceeds their combined radii; known as the London dispersion force. A Lennard-Jones potential (Figure 2.5) of the following form is used in both the CHARMM36m and MARTINI2.2 force fields to model these two dominant components of the vdW force:

$$U_{vdW} = \sum_{non-bonded} 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right) \tag{2.13}$$

Where r_{ij} is the separation between non-bonded particles i and j; σ_{ij} is the value of r_{ij} for which $U_{vdW,ij} = 0$ and is equivalent to the sum of the effective radii of the two particles; ϵ_{ij} represents the strength of the interaction between the two particles and defines the maximum depth of the potential well.

2.1.1.6 Electrostatic Interactions

The second term in Equation 2.4 represents the potential energy of electrostatic interactions between pairs of nearby charged particles that are not directly bound to

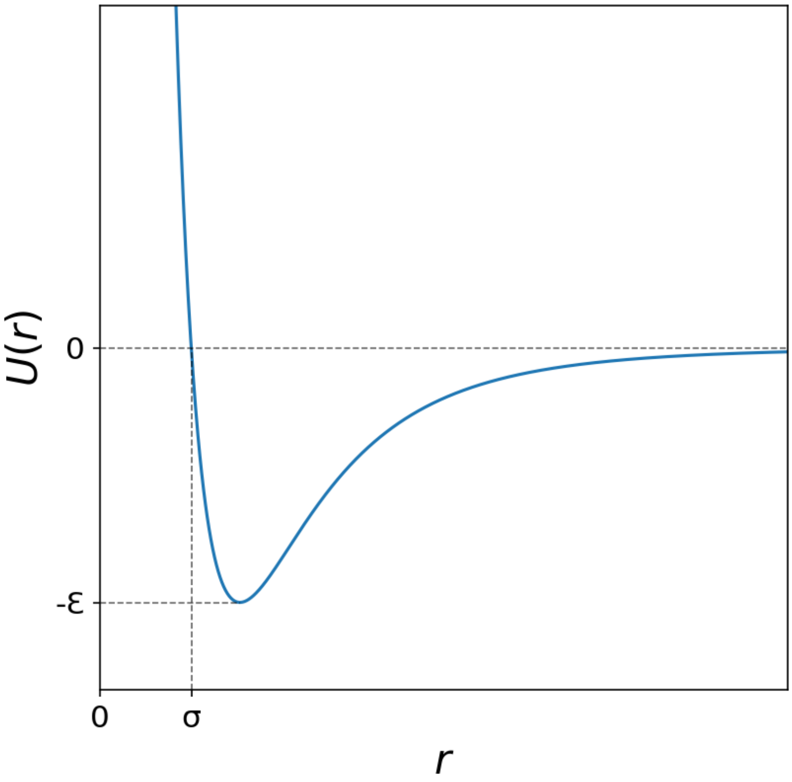


FIGURE 2.5: The Lennard-Jones potential (Equation 2.13).

each other. Both the CHARMM36m and MARTINI2.2 force fields implement these contributions *via* the standard form of the Coulombic potential:

$$U_{\text{Coulomb}} = \frac{1}{2} \sum_{non-bonded} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{2.14}$$

Where q_i, q_j are the charges of non-bonded particles i and j respectively, r_{ij} is their separation and ϵ_0 is the permittivity of free space.

In the case of large systems, the explicit calculation of Coulombic potentials between all possible pairs of charged particles quickly becomes computationally infeasible. Thus, the exact form of the electrostatic potential (Equation 2.14) is evaluated only for pairs of particles that are ’nearby’ to one another. Particles are deemed to be nearby when their separation is smaller than a given cut-off distance. The parameters of each force field are typically optimised for use with a specific value of this cut-off distance; both the CHARMM36m and MARTINI2.2 force fields use an electrostatic cut-off distance of 1.2 nm.

Electrostatic interactions between charged particles that are separated by distances greater than this threshold value are instead treated using an approximation to the long range electrostatic potential. The work presented in this thesis uses the particle mesh Ewald (PME) approximation Darden et al. (1993) with the CHARMM36m force field and the reaction field approximation Tironi et al. (1995) with the MARTINI2.2 force field.

2.1.1.7 Particle Mesh Ewald and Reaction Field Electrostatics

Evaluating the total sum resulting from all long range charged particle pairs in Equation 2.14 is complicated by the necessary inclusion of interactions between particles in the central simulation cell and particles in neighbouring periodic images, a further discussion of periodic boundary conditions can be found in Section 2.1.2. Depending on the size of the system, this can lead to a potentially intractable number of individual contributions and so attempting to directly sum all of the corresponding wave vectors, as is required by the classical Ewald method, leads to prohibitively slow convergence when considering large systems.

To avoid this issue, simulations using the AA CHARMM36m force field treat long range electrostatics using the Particle Mesh Ewald (PME) method Darden et al. (1993). PME avoids the issue of slow convergence through replacing the sum over point charges with summations over two series of Gaussian charge distributions in real and reciprocal space. This results in much faster convergence than in the classical Ewald approach; with the computational cost of Ewald scaling as $\mathcal{O}(N^{\frac{3}{2}})$ and PME scaling as $\mathcal{O}(N \log N)$.

In CG systems using the MARTINI2.2 force field, the aggregation of multiple atoms into single beads leads to an oversimplification of molecular electrostatic surfaces and thus leads to inherently poor modelling of electrostatics. The use of accurate PME electrostatics is therefore typically avoided when using CG force fields due to its computational expense, and the fundamental limitations of CG electrostatics. The reaction field approximation is a robust choice for CG simulations that provides a improvement on a simple cut-off treatment, with negligible additional computational cost Tironi et al. (1995). The reaction field approximation uses the analytical solution of the Poisson Boltzmann equation to model the long range electrostatics under the assumption of a uniform dielectric constant.

2.1.2 Periodic Boundary Conditions

Given the necessarily finite size of any MD simulation, careful consideration must be given to how particles are treated on the edges of the simulated region. If the boundaries of the system are treated as hard barriers to particle motion, then it is clear that the particles residing at the edge of the simulated domain will experience significantly different forces to those that reside close the center. In order to avoid this problem we can imitate an infinite system by introducing periodic boundary conditions (PBCs). A representation of PBCs in two dimensions is provided in Figure 2.6.

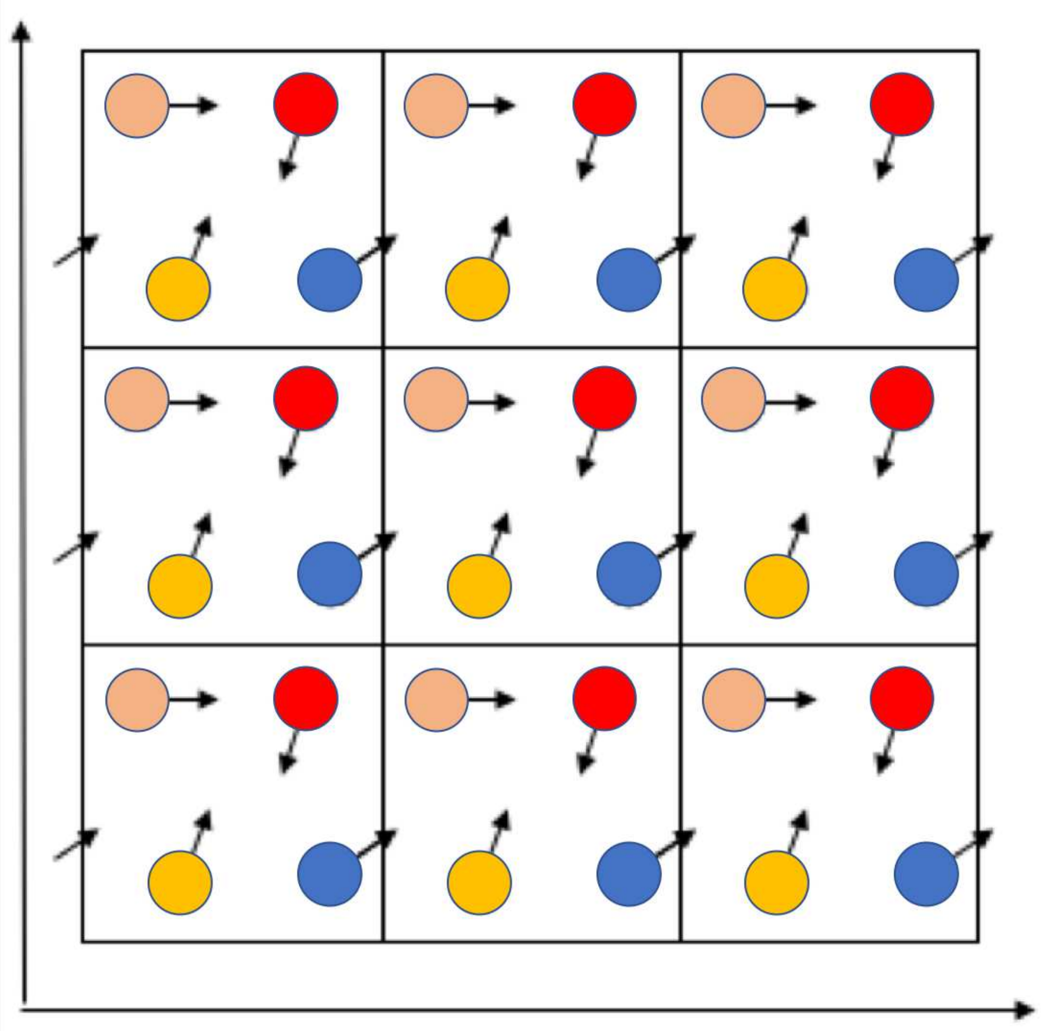


FIGURE 2.6: Representation of periodic boundaries in two dimensions. Circles of matching colour represent periodic copies of the same particle. Arrows show the direction of motion and grid lines represent the periodic boundaries.

In the two dimensional scheme, the central box is surrounded by images of itself. Each image contains copies of all the atoms within the central box along with their corresponding coordinates. Each time a particle moves within the central box, all periodic images of that particle move concordantly. If a particle was to move across one of the boundaries of the periodic domain then one of its images would enter the box from the opposite face, ensuring the number of particles within each periodic box remains constant. Furthermore, particles within the central box may interact with particles from neighbouring images, ensuring bulk behaviour at the edges of the periodic domain. It is important, however, to ensure that no individual particle is able to interact with its own periodic image as this would result in non-physical artefacts in the system dynamics; this issue is resolved by ensuring that all dimensions of the central box are larger than twice the electrostatic cut-off distance discussed in Section 2.1.1.6.

The scheme represented in Figure 2.6 can easily be extended to three dimensions through the replacement of the central square box with a cube. Whilst the cubic cell is the simplest periodic boundary regime to visualise and implement, in principle any shape of cell may be used provided that the regular tessellation of the cell fills the entirety of space. In three dimensions this ensures that there is no set of Cartesian coordinates that does not reside within a periodic image of the central cell.

2.1.3 The Leap-Frog Integrator

As discussed previously, the analytic evaluation of the Newtonian equations of motion for each individual particle in a system is not feasible due to the coupled dependence of the force calculation on the positions, types and binding of all other particles in the system. The solution to the resulting set of high dimensional coupled differential equations must therefore be estimated through the use of a numerical procedure, known as an integrator.

The work presented in this thesis utilises the leap-frog integrator (Equation 2.15); an algorithm based on the Verlet integrator, which can be derived *via* a Taylor expansion of Newton's equations of motion.

The algorithm uses the position of a particle, $\mathbf{r}$, at time t and the corresponding velocity of the particle, $\mathbf{v}$, at time $t - \frac{1}{2}\Delta t$ to determine the new set of atomic coordinates for the particle at time $t + \Delta t$, for some small time interval Δt. The particle velocity is then updated using the force, $\mathbf{F}(t)$, calculated using the global potential energy function provided by the force field (Equations 2.5 and 2.6). The following equations describe the iterative procedure:

$$
\begin{aligned}
\mathbf{v}(t + \frac{1}{2}\Delta t) &= \mathbf{v}(t - \frac{1}{2}\Delta t) + \frac{\Delta t}{m}\mathbf{F}(t) \\
\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}(t + \frac{1}{2}\Delta t)
\end{aligned}
\tag{2.15}
$$

These equations of motion are then modified to account for the specific temperature and pressure coupling regimes chosen by the user to ensure proper sampling of the relevant thermodynamic ensemble. These modifications are discussed in more detail in Sections 2.1.5 and 2.1.6.

2.1.4 Integrator Timestep and LINCS Constraints

The choice of timestep, Δt, in Equation 2.15 has an obvious impact on the speed at which one can perform MD simulations of any system. Whilst it would be preferable to choose as large a value as possible to maximise the rate of simulation, we are limited to choosing a value smaller than the period of the highest frequency motion present within the system. When using AA or UA force fields that include hydrogens (as opposed to CG force fields that do not), the highest frequency motion is typically the bond stretching vibrations between hydrogen atoms and their much heavier parent atoms Hopkins et al. (2015) with an oscillation period of roughly 10 fs. Correct sampling of these vibrations, however, requires a quantum-mechanical treatment and thus for classical MD calculations they are better represented by a constraint.

To this end, the work presented in this thesis uses the LINCS algorithm Hess et al. (1997) to replace the bonds between hydrogens and their heavy parent atoms with constraints. Once the hydrogen bond stretching vibrations have been accounted for, the next fastest motion in the system is the vibration of bond angles between hydrogen atoms and their heavy parent atoms; with an oscillation period of roughly 13 fs. Despite the robust nature of the LINCS algorithm as compared to its predecessors, it is unable to handle the highly connected restraints that result from constraining both bonds and angles simultaneously and, as such, having already constrained the bond stretching vibrations, the MD integrator must be able to correctly sample the bond angle vibrations without the use of constraints.

The leap-frog integrator requires a minimum of five numerical integration steps per period of harmonic oscillation to ensure correct sampling; accurate treatment of bond angle vibrations thus requires an integration time step of approximately 2.6 fs Feenstra et al. (1999). The simulations in this thesis use an integration time step of 2 fs where possible to abide by this threshold, whilst a shorter time step of 1 fs is used when a more careful treatment of system dynamics is required for stable equilibration.

2.1.5 Temperature Coupling

Whilst direct use of the leap-frog integrator (Equation 2.15) allows for the simulation of systems within the microcanonical-fixed ensemble (NVE), many quantities that we may wish to calculate require use of the canonical ensemble (NVT). This assertion of constant temperature can be implemented within the MD algorithm through the inclusion of a temperature coupling scheme. GROMACS includes a variety of temperature coupling schemes by default, each with their individual benefits; however, the work presented in this thesis utilises the velocity-rescaling algorithm Bussi et al. (2007); equivalent to a Berendsen coupling scheme with an additional stochastic term that ensures a correct kinetic energy distribution.

2.1.5.1 The Berendsen Regime

The Berendsen algorithm Berendsen et al. (1984) effectively couples the kinetics of the system to an external heat bath of temperature T_0. This acts to suppress fluctuations in the kinetic energy of the system by slowly correcting the system temperature as it deviates from T_0, according to:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \tag{2.16}$$

The above equation leads to the exponential decay of temperature deviations with a time constant τ. The suppression of kinetic energy fluctuations does, however, mean that one does not sample the true canonical ensemble. As such, calculation of quantities that require a canonical ensemble will not produce exactly correct results. The error in these quantities, however, scales with $\frac{1}{N}$ and so for sufficiently large systems most ensemble averages will not be significantly affected.

Practically, Equation 2.16 is implemented through scaling the velocities of each particle at every (or every ${n_{TC}}^{\text{th}}$) step by a time-dependent factor λ, given by:

$$\lambda = \left[1 + \frac{n_{TC}\Delta t}{\tau_T}\left\{\frac{T_0}{T(t - \frac{1}{2}\Delta t)} - 1\right\}\right]^{\frac{1}{2}} \tag{2.17}$$

The parameter τ_t is close, but does not exactly equal, the time constant τ of the temperature coupling regime given in Equation 2.16. These two quantities are instead related *via* the following expression:

$$\tau = \frac{2C_V\tau_T}{N_{df}k} \tag{2.18}$$

Where C_V is the total heat capacity of the system, k is the Boltzmann constant and N_{df} is the total number of degrees of freedom. The kinetic energy of the system is then modified at each scaling step by the following factor:

$$\Delta E_k = (\lambda - 1)^2 E_k \tag{2.19}$$

2.1.5.2 The Velocity-Rescale Regime

As mentioned in the previous section, the suppression of kinetic energy fluctuations during simulation leads to the sampling of an ensemble that is not an exact canonical ensemble. The velocity-rescale algorithm Bussi et al. (2007) is able to reproduce the correct ensemble by ensuring that the system exhibits the correct kinetic energy distribution. This is achieved through the inclusion of an additional stochastic term when updating the kinetic energies at each scaling step (Equation 2.19). This term takes the following form:

$$dK = (K_0 - K)\frac{dt}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau_T}} \tag{2.20}$$

Where K is the kinetic energy, N_f the number of degrees of freedom and dW an increment of a Wiener process. The Wiener process is responsible for the stochastic nature of this expression and is equivalent to a random walk generated by Brownian motion in one dimension. The term dW represents an incremental step along this random walk and can thus take a positive or negative, finite value taken from a normal distribution with a mean of 0 and variance u.

2.1.6 Pressure Coupling

Similar to temperature coupling, pressure coupling can be thought of as a coupling of the kinetics of the system to a pressure bath. GROMACS allows for the use of various pressure coupling regimes, however the simulations presented in this thesis primarily utilise the Parrinello-Rahman scheme Parrinello and Rahman (1981), with the Berendsen scheme Berendsen et al. (1984) used during certain equilibration procedures when a weaker coupling scheme is necessary to maintain system stability.

2.1.6.1 The Berendsen Regime

The Berendsen algorithm Berendsen et al. (1984) rescales the coordinates and box vectors at every (or every ${n_{PC}}^{\text{th}}$) step with a matrix μ. Similar to the Berendsen temperature coupling scheme, this has the effect of suppressing fluctuations in the pressure by relaxing the system kinetics towards a given reference pressure $\mathbf{P}_0$, according to:

$$\frac{d\mathbf{P}}{dt} = \frac{\mathbf{P}_0 - \mathbf{P}}{\tau_p} \tag{2.21}$$

The scaling matrix μ is given by:

$$\mu_{ij} = \delta_{ij} - \frac{n_{PC}\Delta t}{3\tau_p}\beta_{ij}\{P_{0ij} - P_{ij}(t)\} \tag{2.22}$$

Whilst the particle velocities are neither scaled nor rotated by the action of this matrix, the coordinates and box vectors, and thus the equations of motion, are. It is thus necessary to update the conserved energy quantity of the system at each scaling step. The value of this conserved energy is obtained by subtracting the work done by the coupling algorithm on the system from the total energy according to the following:

$$E_{mod} = E - \sum_{i,j}(\mu_{ij} - \delta_{ij})P_{ij}V = \sum_{i,j} 2(\mu_{ij} - \delta_{ij})\Xi_{ij} \tag{2.23}$$

Where E_{mod} is the modified conserved energy quantity, E is the total system energy, δ_{ij} is the Kroenecker delta and Ξ is the virial. It is important to note that although the Berendsen scheme will produce a simulation with the correct average pressure, similar to its temperature coupling equivalent, it will not sample the exact isothermal-isobaric (NPT) ensemble; as such, a more rigorous pressure coupling scheme is required if physically relevant information about the system's thermodynamic properties is to be extracted.

2.1.6.2 The Parrinello-Rahman Regime

In theory, use of the Parrinello-Rahman scheme Parrinello and Rahman (1981) reproduces the true isothermal-isobaric ensemble and is thus a more desirable choice than Berendsen for simulations that will be subjected to thermodynamic analysis. Under this regime, the box vectors, represented by the matrix **b**, are governed by the following matrix equation of motion:

$$\frac{d\mathbf{b}^2}{dt^2} = V\mathbf{W}^{-1}\mathbf{b}'^{-1}(\mathbf{P} - \mathbf{P}_{\mathrm{ref}}) \tag{2.24}$$

Where V is the volume of the box and the matrices $\mathbf{P}$ and $\mathbf{P}_{\mathrm{ref}}$ are the current and reference pressures, respectively. $\mathbf{W}$ is a mass parameter matrix that determines both the strength of the coupling and how the box can be deformed, its inverse elements are defined by the following equation:

$$(\mathbf{W}^{-1})_{ij} = \frac{4\pi^2\beta_{ij}}{3\tau_p^2 L} \tag{2.25}$$

Where β are the approximate isothermal compressibilities, τ_p is the pressure time constant and L is the largest box matrix element.

Just as in the Berendsen scheme, the modifications made to the box vectors of the system during simulation result in changes to each particle's equations of motion and, as such, the conserved energy quantity of the system will also be affected. The modified, conserved energy Hamiltonian takes the following form:

$$H_{\text{mod}} = E_{\text{pot}} + E_{\text{kin}} + \sum_i P_{ii} V + \sum_{i,j} \frac{1}{2} W_{ij} \left(\frac{db_{ij}}{dt} \right)^2 \tag{2.26}$$

The resulting equations of motion obtained from this Hamiltonian are as follows:

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \mathbf{M}\frac{d\mathbf{r}_i}{dt} \tag{2.27}$$

$$\mathbf{M} = \mathbf{b}^{-1} \left[\mathbf{b}\frac{d\mathbf{b}'}{dt} + \frac{d\mathbf{b}}{dt}\mathbf{b}' \right] \mathbf{b}'^{-1} \tag{2.28}$$

The extra term in Equation 2.27 takes the form of a frictional component. It is important to note that this is a fictitious force, and is an artefact of the particle coordinates being defined relative to the box vectors within the equations of the Parrinello-Rahman algorithm; this is in contrast to the accompanying leap frog integrator which utilises absolute particle positions.

2.2 System Equilibration

Once the choice of force field, integrator and accompanying temperature and pressure coupling schemes has been made; we can consider how they may be applied to a system in order to generate a novel trajectory.

The first step in this process is to obtain an initial configuration for the system of choice. Online repositories such as the RCSB Protein Data Bank Berman et al. (2000) provide an invaluable resource for obtaining the structure data for many individual proteins and their complexes. Web-based platforms such as CHARMM-GUI Jo et al. (2008) provide useful functionality for generating the initial structures of more complex systems involving membranes and membrane proteins.

These initial structures often occupy highly idealised conformations that would not readily exist in equilibrium physiological conditions. It is often the case that individual protein structures are determined using X-ray crystallography Shi (2014) or cryo-EM Lyumkis (2019) techniques and thus the obtained structure data is that of the crystalline or frozen protein. If algorithmic procedures such as CHARMM-GUI, or other specialised scripts, are used to generate the initial structures this may lead to unnaturally ordered structures that bear little resemblance to their *in vivo* forms. It is therefore necessary to perform a set of equilibration steps to bring the system to a more biologically relevant equilibrium state before any meaningful information can be extracted from its dynamics. This so-called equilibration phase of MD simulation is split into three distinct stages; energy minimisation, NVT and NPT.

2.2.1 Energy Minimisation

The first equilibration stage is energy minimization (EM); the purpose of this stage is to ensure that no excessively large forces are present within the system. Such forces are usually caused by improper bond geometries and steric clashes between molecules; removing these artefacts is the primary goal of the EM phase.

GROMACS provides a number of algorithms for performing EM, namely; the steepest descent, conjugate gradient and limited-memory Broyden-Fletcher-Goldfarb-Shanno quasi Newtonian minimizer (l-bfgs) methods. The steepest descent algorithm, whilst not the most efficient, is robust and easily implemented. Since the EM phase is the fastest of the equilibration stages by many orders of magnitude, the robust nature of the steepest descent algorithm when dealing with systems that are far from equilibrium is of much greater utility than the minimal gain in time efficiency obtained when using a less robust method. As such, the work presented in this thesis utilises the steepest descent algorithm and so we shall focus our discussion on this method.

First, the potential energy, U, and thus the forces, $\mathbf{F}$, within the system are evaluated. The forces are then used to calculate a new set of atomic positions:

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \frac{\mathbf{F}_n}{\max(|\mathbf{F}_n|)} h_n \tag{2.29}$$

Where h_n is the maximum displacement and $\mathbf{F}_n$ is the force, or the negative gradient of the potential U. The term $\max(|\mathbf{F}_n|)$ represents the largest scalar force acting on any atom in the system. Once these new positions are obtained, they are used to compute the new set of forces and potential energy. If the new potential energy is less than that of the previous iteration, then the new positions are accepted and the displacement of the next step is increased: $h_{n+1} = 1.2h_n$. If the new potential energy is greater than or equal to that of the previous iteration, then the new positions are rejected and the displacement of the next step is decreased: $h_n = 0.2h_n$. This process is then repeated until a specified number of steps have been completed, or until the maximum force present in the system has reached a suitably low, user-defined threshold.

2.2.2 NVT Equilibration

The second equilibration stage is NVT equilibration; the purpose of this phase is to impart velocities to the atoms within the static struture generated by EM, raising the system temperature to a physiologically relevant equilibrium value. During NVT, pressure coupling is turned off whilst a temperature coupling regime is introduced into the system's equations of motion. As such, the box dimensions are held constant whilst the atomic velocities are slowly increased until the velocity distribution corresponds to an average system temperature equal to a user-defined value. This user-defined value is the temperature of the heat bath, T_0, coupled to the system; further discussion of how this is implemented can be found in our previous section on Berendsen temperature coupling (Section 2.1.5.1).

Typically, any large biomolecules in the system are held in place using position restraints during this stage so as to ensure that no undesirable conformational changes occur to the molecules whilst the surrounding solvent molecules are allowed to correctly orient themselves around the larger structures.

2.2.3 NPT equilibration

The final equilibration phase is NPT equilibration; the purpose of this phase is to take the system generated during NVT equilbiration and raise or lower its pressure to physiological levels. During NPT, both a temperature and pressure coupling scheme are applied to the system's equations of motion and the system's dynamics are simulated using the chosen MD integrator. As such, the previously equilibrated temperature is held constant while the box dimensions and particle positions are scaled repeatedly until the system reaches a roughly constant average pressure of 1 bar. Further discussion of how this stage is implemented can be found in our previous sections on Berendsen and Parrinello-Rahman pressure coupling (Sections 2.1.6.1 & 2.1.6.2).

This is arguably the most important and nuanced stage in the equilibration process; it is often the case with more complex systems that numerous NPT stages are required with varied position restraints in order to allow all of the large biomolecules in the system to relax into their equilibrium conformations without causing non-physical deformations to the other structures in the system. This is especially the case when dealing with large structures that directly interact with many components of the system; *e.g.* BLP molecules that are inserted into the outer membrane whilst also covalently bound to the cell wall.

Once the NPT phase is complete, the equilibrated system is ready for unrestrained production MD, the stage in which meaningful dynamical data can be extracted from the system. Even so, when working with membranes, it is wise to ignore the first $\sim$ 100 ns of production MD when performing any analysis on membrane properties. This is required because the high density of lipids within the hydrophobic region of lipid bilayers means that it often takes a considerable amount of time for the lipids within this region to find the correct organisation and orientation with respect to each other and any external structures embedded within the membrane.

2.3 Statistical Mechanics

Molecular dynamics simulations provide us with the ability to observe the time evolution of complex biochemical systems at atomic scale resolution. In order to make use of such detailed data, however, we must ascertain the link between the simulated microscopic behaviour of individual particles and the observable macroscopic properties of the system. For this we turn to the field of statistical mechanics and its fundamental postulate which states that the observed properties of a system are the average of the microscopic properties of all possible microstates consistent with the thermodynamic state of the system Allen and Tildesley (1987).

In a classical system, each N atom microstate is defined by $6N$ component variables; with each atom contributing 3 Cartesian coordinates, $\mathbf{r}$, and 3 corresponding components of momentum, $\mathbf{q}$. The $6N$ dimensional space defined by all allowed combinations of these $6N$ position and momentum components constitutes the global phase space of the system, Γ:

$$\Gamma = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{q}_{n-1}, \mathbf{q}_n) \tag{2.30}$$

Each of these microstates represents a unique system configuration and thus the global phase space can be seen as a representation of all possible configurations of the given system. It thus follows that the time evolution of any individual microstate must be able to be depicted as a series of transitions between individual microstates.

Suppose now that we may write the instantaneous value of some property, A, as a function of this phase space, $A(\Gamma)$. As the system evolves through time and thus moves along some trajectory through its allowed phase space, the value of $A(\Gamma)$ will vary. Through reference to the fundamental postulate of statistical mechanics, one can reasonably assume that the experimentally observable, macroscopic form of this property, A_{obs}, is simply the time average of $A(\Gamma)$ evaluated over a long time interval:

$$A_{obs} = \langle A \rangle_{time} = \langle A(\Gamma(t)) \rangle_{time} = \lim_{t_{obs} \to \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A(\Gamma(t))dt \tag{2.31}$$

Clearly it is not possible to perform simulations of infinite length and so we must approximate the above integral numerically. This can be done by taking an average over a sufficiently long finite time, t_{obs}. Typically, MD simulations are performed as linear iterative processes and, as such, produce trajectories comprised of a chronological set of microstates each separated in time by a small, finite interval δt. Provided that a sufficiently large number of these microstates are sampled, *i.e.* that the system configuration is re-evaluated for a sufficiently large number of time steps, τ_{obs}, of length $\delta t = t_{obs}/\tau_{obs}$; then the integral in Equation 2.31 can be rewritten as the following:

$$A_{obs} = \langle A \rangle_{time} = \frac{1}{t_{obs}} \sum_{\tau=1}^{\tau_{obs}} A(\Gamma(\tau)) \tag{2.32}$$

Currently, typical MD studies involve $\mathcal{O}(10^6)$ atoms and time scales of $\mathcal{O}(10^{-7})$ s. Many biological processes, however, take place over much longer time scales; *e.g.*, the rate of transbilayer lipid motion (known as lipid flip-flopping) is $\mathcal{O}(10^{-15})$ s^{-1}, corresponding to one flip-flop event every 24 hours. It is therefore evident that there may exist regions of a system's allowed phase space that are simply inaccessible given the initial configuration and feasible simulation time. This poses an issue for calculating systemic properties using Equation 2.32 since we cannot be certain that any individual simulation is sufficiently long enough to sample the entire of the system's phase space and thus the equivalence of Equation 2.32 and Equation 2.31 cannot be guaranteed.

In order to mitigate this sampling problem, we may replace the time average with an ensemble average. An ensemble is defined as a collection of points in phase space satisfying the conditions of a particular thermodynamic state, distributed according to a probability density, $\mathbf{P}(\Gamma)$, with each of these points corresponding to a unique system configuration, or microstate, as discussed previously. As each of these microstates evolves over time, the phase space density $\mathbf{P}(\Gamma)$ will also change; however, if $\mathbf{P}(\Gamma)$ represents an equilibrium ensemble then its time dependence vanishes. This can be understood through the principle of ergodicity, a property of certain dynamical systems which is upheld by most many-bodied systems in nature. Explicitly, an ergodic system is a dynamical system which, given any initial configuration and in the limit of infinite time, will explore all regions of its phase space. Thus, the vanishing time dependence of an equilibrium ensemble $\mathbf{P}(\Gamma)$ is accounted for by the fact that each time a system leaves a particular state, $\Gamma(\tau)$, another system in the ensemble arrives to replace it. For these systems we are able to replace the time average with an average over all members of the ensemble, 'frozen' at a particular point in time, this *ensemble average* takes the following form:

$$\langle A \rangle = \sum_{\Gamma} A(\Gamma) P(\Gamma) \tag{2.33}$$

Provided that the ensemble average is calculated using simulations of much greater duration than the time scale of any relaxation of the system, and that a sufficiently large number of microstates are extracted from the trajectories; the ensemble average and time average are equivalent:

$$\langle A \rangle = A_{obs} \tag{2.34}$$

The three main ensembles in statistical mechanics and their fixed macroscopic variables are:

- Microcanonical-fixed (N, V, E)
- Canonical-fixed (N, V, T)
- Isothermal-isobaric-fixed (N, P, T)

Where N is the number of particles, V is the volume, E is the energy, T is the temperature and P is the pressure of the system.

2.4 Enhanced Sampling

In conventional equilibrium MD simulations, large energy barriers may separate regions of a system's potential energy landscape. Such energetic barriers can result, for example, from macromolecular structures physically separating regions of the simulation domain, or an energetically unfavourable transition between two molecular conformations that may affect the system's observed behaviour. These barriers can lead to certain regions of the system's allowed phase space remaining inaccessible from a given initial configuration, leaving these regions either poorly sampled or completely neglected. When the form of the potential energy landscape hinders ergodicity in this way, biased simulation techniques can be utilised to enhance the sampling of otherwise inaccessible system configurations.

Biased simulation methods involve applying external forces to a system during simulation so as to drive it towards, or to maintain, a desired configuration. These external forces simply take the form of additional potential functions that may be added to the energy Hamiltonian of the system; allowing for the original integrators, thermostats and barostats to be applied as normal. Whilst there are a multitude of enhanced sampling methods that may be used, the work presented in Chapter 6 of this thesis utilises one in particular; umbrella sampling.

2.4.1 Umbrella Sampling

It is often the case that one may wish to investigate the behaviour a system as it transitions between two thermodynamic states of interest, *e.g.* the docking of a drug to a target molecule, the membrane-insertion of a biomolecule or the folding of a protein. One popular enhanced sampling method for studying a system along such a reaction pathway is umbrella sampling (US). Using this method, one can obtain the probability of finding the system at any position along the given reaction pathway, and thus calculate the potential of mean force (PMF) along the reaction coordinate.

The US method involves the simulation of a number of system configurations, called "windows", along a reaction coordinate, ξ. The reaction coordinate is chosen so as to map the reaction pathway between two states of interest and can be a simple measure of distance between two molecules, a particular bond or dihedral angle, or a mathematical expression that transforms the system from one state to another. The initial configurations for the windows are generated at intervals of the reaction coordinate, so that each window starts from a different position along the reaction pathway.

In each window, i, the reaction coordinate, ξ is restrained by an additional bias potential, U_i. Typically a harmonic restraint of the following form is used to model each bias potential:

$$U_i(\xi) = \frac{k_i}{2}(\xi - \xi_i) \tag{2.35}$$

Where k_i is the strength of the harmonic restraint applied on window i, ξ is the value of the reaction coordinate at each frame of the window simulation and ξ_i is the equilibrium value of the reaction coordinate in window i. This restraint allows the system to sample the phase space perpendicular to the reaction coordinate, whilst remaining in proximity to the specified reference value of the reaction coordinate, ξ_i.

From each window, an umbrella histogram $h_i(\xi)$ is recorded, representing the probability distribution $P_i^b(\xi)$ along the reaction coordinate biased by the umbrella potential $U_i(\xi)$. The overall probability distribution along the reaction coordinate with respect to the biased simulations, P_{bias}, can then be determined. However, the addition of these bias potentials to the energy Hamiltonian of the system affects the potential energy landscape, and thus calculation of the PMF profile along the reaction coordinate requires use of the unbiased probability distribution, $P(\xi)$. One widely used technique which may be used to obtain the unbiased distribution is the Weighted Histogram Analysis Method (WHAM).

2.4.1.1 Weighted Histogram Analysis Method

The Weighted Histogram Analysis Method (WHAM) obtains the unbiased probability distribution *via* the following equations;

$$P(\xi) = \frac{\sum_{i=1}^{N_w} g_i^{-1} h_i(\xi)}{\sum_{j=1}^{N_w} n_j g_j^{-1} exp[-\beta(U_j(\xi) - f_j)]} \tag{2.36}$$

$$exp(-\beta f_j) = \int exp[-\beta U_j(\xi)] P(\xi) \tag{2.37}$$

Here, $h_i(\xi)$ denotes the histogram of the biased window, N_w the total number of windows, n_j the number of data points in the j^{th} window and $\beta = 1/k_B T$ where k_B is the Boltzmann constant and T the temperature. The statistical inefficiency of a window, g_i, is given by $g_i = 1 + 2\tau_i$. τ_i denotes the integrated autocorrelation time and is given by:

$$\tau_i = \sum_{\Delta t} R_i(\Delta t) \tag{2.38}$$

Where $R_i(\Delta t)$ denotes the normalized autocorrelation function of umbrella window i and is given by:

$$R_i(\Delta t) = \frac{\langle (\xi_i(t) - \langle \xi \rangle)(\xi(t + \Delta t) - \langle \xi_i \rangle) \rangle}{\sigma^2_{\xi,i}} \tag{2.39}$$

where $\xi_i(t)$ denotes the time-evolving value of the reaction coordinate during window i, $\sigma^2_{\xi,i} = \langle (\xi_i(t) - \langle \xi_i \rangle)^2 \rangle$ is the respective variance and $\langle \dots \rangle$ represents the average of a given quantity over all simulation frames.

2.5 Methods of Analysis

All analyses in this thesis were performed with Python Van Rossum et al. (1995) and BASH Ramey et al. (1998) scripts written using MDAnalysis Gowers et al. (2016); Michaud-Agrawal et al. (2011), NumPy Harris et al. (2020), Gromacs utilities Abraham et al. (2015), VMD Humphrey et al. (1996) and custom Tcl scripts Welch et al. (2003). Kernel density estimates (KDE) were calculated using the Seaborn python package Waskom (2021). All data presented in this thesis was plotted using Matplotlib Hunter

(2007). Various BASH pipelines were developed to save and load data efficiently and to analyse trajectories in tandem or remotely without requiring the transfer of large data files from the HPC infrastructure.

2.5.1 Radius of Gyration

The radius of gyration (ROG) of a body with total mass M about an axis is defined as the radial distance between the axis and a point mass, also of mass M, for which the moment of inertia is equivalent to that of the body. The radius of gyration of a molecule about it's center of mass takes the following form:

$$R_g^2 = \sum_i m_i \frac{|\mathbf{r}_i - \mathbf{r}_C|^2}{M} \tag{2.40}$$

Where R_g is the radius of gyration, m_i the mass of atom i, $\mathbf{r}_i$ the position of atom i, $\mathbf{r}_C$ the position of the molecule's center of mass and M the total molecular mass.

R_g gives us a measure of the molecule's 'compactness' in space; with smaller values correlated both with smaller molecules (*i.e.* those comprised of less atoms) and higher numbers of atomic contacts per residue Lobanov et al. (2008), an indicator of the density of atoms within a molecule's structure. The value of R_g may be calculated for a chosen molecule using the GROMACS *gyrate* function.

2.5.2 Mean Squared Displacement

The mean squared displacement (MSD) of a molecule is a measure of its deviation from a reference position over time. It is defined mathematically as the following:

$$\mathrm{MSD} \equiv \langle |\mathbf{x}(t) - \mathbf{x}_0|^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{x}_i(t) - \mathbf{x}_{i,0}|^2 \tag{2.41}$$

Where the middle expression takes the form of an ensemble average with parameters $\mathbf{x}(t)$ representing the set of atomic coordinates for the given molecule at time t and $\mathbf{x}_0$ representing the set of reference atomic coordinates for the molecule at time $t = 0$. The expression on the right hand side recasts the ensemble average as a sum over individual atoms with N representing the total number of atoms in the molecule, $\mathbf{x}_i(t)$ the position of atom i at time t and $\mathbf{x}_{i,0}$ the reference position of atom i at time $t = 0$.

By calculating the MSD of a chosen molecule at each frame of a simulation trajectory one may obtain a corresponding MSD curve; the gradient of which provides the value of the molecule's average drift velocity and can thus be used to determine the molecule's translational diffusion rate *via* the Einstein relation Frenkel and Smit (2002):

$$\frac{\partial(\mathrm{MSD})}{\partial t} = 2dD \tag{2.42}$$

Where d is the dimensionality of the system (*i.e.* 3 spatial dimensions) and D is the translational diffusion coefficient. Accurate calculation of D requires that the MSD curve be linear since non-linearity would imply a rate of change within the drift velocity; indicating that advective forces are present within the system and, as such, the molecule is not undergoing pure diffusive motion.

2.5.3 Solvent Accessible Surface Area

The solvent accessible surface area (SASA) of an atom is defined as the portion of a spherical shell centered on the atom and with radius $r_{vdW} + r_{sol}$, upon which the center of a spherical solvent molecule with radius r_{sol} can be placed in contact with the atomic van der Waals sphere of radius r_{vdW} without penetrating any other neighbouring atom. The sum of SASA values for all atoms in a given molecule provides a measure of the extent of that molecule's surface that may be exposed to the bulk solvent. The implementation of this algorithm is far more nuanced than that of the MSD and ROG calculations Eisenhaber et al. (1995) but is easily performed using the GROMACS *sasa* function.

2.5.4 Residue Interaction Analysis

Analysis of the specific residue-residue interactions between polymyxins and the various solutes and surfaces of the periplasmic environment was performed using bespoke python scripts written using the MDAnalysis package. Atom groups were defined for the two molecules (or molecule groups) under investigation, and matrices containing the contacts (separation < 4 Å) between each residue of the two groups were generated at every frame of the simulation. The concatenation of these matrices provided a contact tensor comprising all of the interaction data for the two atom groups. Analysis of this contact tensor allowed for the construction of residue interaction distributions, highlighting the percentage contribution of each individual residue of one atom group to all observed interactions with the opposing atom group.

2.5.5 Binding Duration Analysis

The contact tensor generated during the aforementioned residue interaction analysis was also used to analyse the binding durations between the polymyxins and the cell wall / BLP. Regions of zero values in the contact tensor corresponded to the absence of any interactions between the two atom groups and could thus be used to determine the number and duration of distinct binding events observed in each simulation. Molecules were considered to be "bound" to one another when contacts were formed between them for 5 consecutive frames (equivalent to 2.5 ns). Conversely, bound molecules were considered to have dissociated from one another when no contacts were observed between them for 5 consecutive frames.

Chapter 3

Developing the CHARMM36m Cell Envelope Model

3.1 Introduction

Previous work published by the Khalid research group, on which I was a co-author, investigated the molecular interactions of PMB1 within the periplasm of *E. coli* using the united atom GROMOS54a7 forcefield Pedebos et al. (2021). In order to investigate the behaviour of polymyxins within this space further, in atomistic detail, it was necessary to develop a new model of the periplasm using the all-atom CHARMM36m forcefield Huang et al. (2017). This chapter outlines the implementation of the various procedures required to construct this model. The use of the CHARMM-GUI web server Jo et al. (2008) to insert BLP into the OM is discussed, along with the importance of equilibration of individual system components prior to the construction of such complex systems. Detail is also provided regarding the implementation of bespoke scripts used to bind BLP to the cell wall.

3.2 Equilibration of the Outer Membrane & BLP

Initial attempts to insert the BLP trimer into the outer membrane using the CHARMM-GUI membrane builder Lee et al. (2019) routinely produced membranes with interlocking lipid tails from the opposing leaflets. This indicated that the membrane had been constructed in the gel-phase, resulting in dynamical freezing of the membrane during simulation. To avoid the long equilibration times and temperature changes required to return the membrane to the desired liquid disordered phase, a manual insertion procedure was used instead that utilised an equilibrated model of the OM that had been validated by previous studies performed by the

Khalid research group Khalid et al. (2019); Ortiz-Suarez et al. (2016); Piggot et al. (2011); Samsudin et al. (2016). VMD Humphrey et al. (1996) was used to position the acyl tails of the BLP trimer within the hydrophobic region of the OM and any lipids that substantially overlapped with the newly inserted BLP trimer were removed.

During the insertion of a single BLP trimer, 13 POPE residues were removed from the inner leaflet of the OM; since POPE is the most abundant lipid type in this region, representing 145 of the 161 total inner leaflet lipids prior to the insertion of BLP, the effect on the overall lipid composition may largely be ignored. However, the removal of these lipids and the insertion of BLP posed a significant structural change to the previously equilibrated membrane and as such necessitated further equilibration of the combined system to return it to an equilibrium state.

To this end, the standard equilibration scripts generated by the CHARMM-GUI membrane builder Lee et al. (2019) were utilised. These scripts comprised 2 × NVT and 4 × semi-isotropic NPT phases; with a combination of z-position restraints on the phosphate atoms in the lipid head groups and dihedral restraints on the branched and double-bonded regions in the lipid tails. The strengths of these position restraints were reduced in each subsequent phase and are given in Table 3.1.

After equilibration, the surrounding lipids were correctly packed around the acyl tails of the BLP trimer and the overall dimensions of the membrane had adjusted to account for the newly inserted membrane protein. A further 40 ns of production MD was then performed in order to account for any longer timescale membrane relaxation. A frame was then extracted from the final 5 ns of the resulting trajectory that exhibited a BLP orientation roughly parallel to the bilayer normal; thus providing a structure that could be easily bound to the PGN cell wall without imposing any excessively large forces due to further large-scale system relaxation.

Equilibration Phase	F_z (kJ.mol^{-1}.nm^{-2})	F_{dih} (kJ.mol^{-1}.nm^{-2})
NVT1	1000	1000
NVT2	1000	400
NPT1	400	200
NPT2	200	200
NPT3	40	100
NPT4	0	0

TABLE 3.1: Position restraint strengths during the various equilibration stages of the inserted BLP - OM structure.

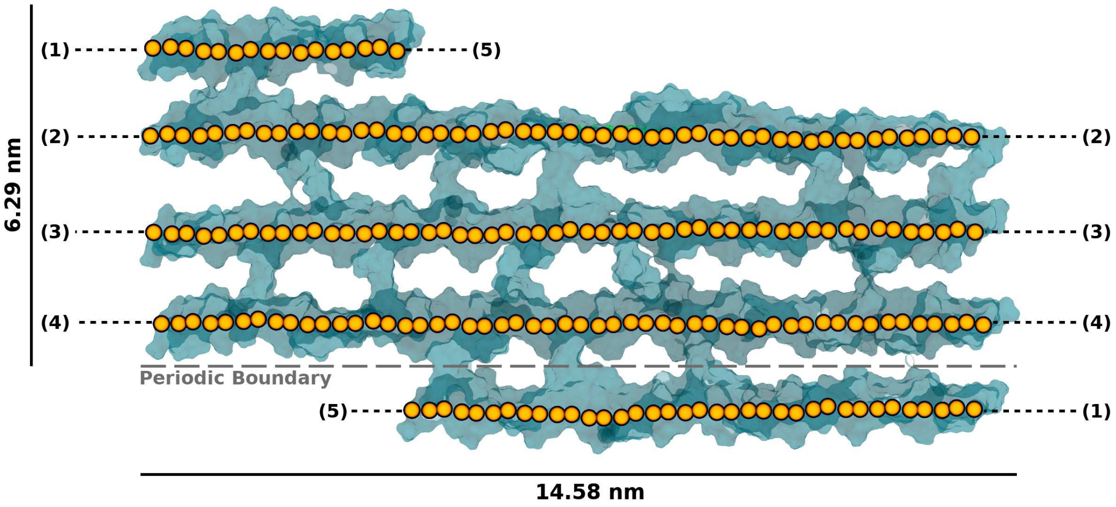


FIGURE 3.1: Unequilibrated cell wall structure. The carbon backbone of each glycan strand is highlighted in orange. A surface plot of the entire cell wall is shown in cyan. Glycan strands are explicitly bound to themselves across the periodic boundaries, the positions of each of these bonds are shown by pairs of bracketed numbers.

3.3 Construction of the Peptidoglycan Cell Wall

The cell wall model used for this project was generated using a script originally developed by previous members of the Khalid research group Samsudin et al. (2017); Boags et al. (2019). The work performed for this thesis, however, involved the contribution of significant updates and modifications to this script that have enabled the robust, scalable construction of cell wall models in both the GROMOS54a7 and CHARMM36m forcefields. This updated script was used throughout the work presented in this thesis.

The script incrementally builds up the cell wall by forming covalent links between individual PGN monomers to form glycan strands of a user-specified length. During this step, PGN monomers were bound together *via* $\beta - (1,4)$ glycosidic linkages between the N-Acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues to form the glycan strands that serve as the main structural scaffold of the cell wall. These glycan strands were subsequently bound together *via* the peptide residues of the PGN monomers. Explicitly, the ϵ-amino group of the meso-DAP residue on one PGN monomer was bound to the carbonyl group of the penultimate D-Ala residue on a PGN monomer from an adajcent glycan strand. Approximately half of the monomers were cross-linked in this way so as to match the 50 % cross-linking observed in experimental studies of *E. coli* Glauner et al. (1988). The chemical structure of this cross-linked mesh has been previously shown in Figure 1.2. A short energy minimisation step was performed after each bond is added to allow for the local structure of the new bond to relax into the most accessible energy minimum; ensuring proper bond geometry. This step alone was insufficient, however, to ensure that the cell wall as a whole existed in or near to its equilibrium state. Asymmetries in the placement of the peptide cross-links and the highly ordered structure of the repeating monomer units (Figure 3.1) necessitated further equilibration before the cell wall occupied such a state.

3.4 Construction of the Combined BLP, OM & PGN System

Since a model of the amide linkage between BLP and PGN did not exist prior to this work, it was therefore necessary to construct a new model of the bound PGN-BLP structure and to obtain accurate bond parameters for the region surrounding the amide linkage. Avogadro Hanwell et al. (2012) was used to ’draw’ the structure of the amide linkage along with 5 backbone atoms (and their corresponding bound hydrogens or functional groups) on either side of the linkage. This structure was then passed through the CHARMM-GUI ligand reader Kim et al. (2017) to generate a complete set of bond parameters for the local binding region.

VMD Humphrey et al. (1996) was then used to manually position the cell wall beneath the OM-inserted BLP trimer, such that the NZ atom of the terminal Lys residue of the second BLP helix lay within roughly 10 Åof a maximally exposed O13 atom from the terminal carboxyl group of a non-crosslinked meso-DAP residue on the cell wall. This distance, while not an essential requirement, was chosen as a standard to be small enough that energy minimization could easily obtain the correct bond geometry whilst not causing excessive overlap between BLP and nearby regions of the cell wall.

The coordinates of the positioned cell wall were then inserted into the BLP-OM structure file, beneath the entries for the second BLP helix. A bespoke script was then implemented to generate a complete system structure and topology from the binding region parameters and the correctly positioned and formatted PGN-BLP-OM structure file. The procedure involved creating the topology for a single molecule comprised of the cell wall bound to a single BLP monomer; this was achieved through mapping the atom indices from the local binding region, cell wall and BLP monomer topologies to the indices of the equivalent atoms in the bound molecule. Two hydrogens were removed from the terminal NZ atom of BLP along with the removal of the O13 atom from the terminal carboxyl group of the cell wall in order to make way for the additional amide linkage (Figure 3.2).

Once the indices had been updated and atoms removed, the parameters were appended into the relevant sections of a combined topology file, ensuring that no bonds, pairs, angles, dihedrals or impropers had multiple entries due to the overlap between the local binding region topology and the individual cell wall and BLP topologies. Notably, the CHARMM36m forcefield assigns the same parameters to both oxygen atoms in a carboxyl group; this approximation accounts for the readily occurring tautomerization that shifts the $C = O$ double bond between the two

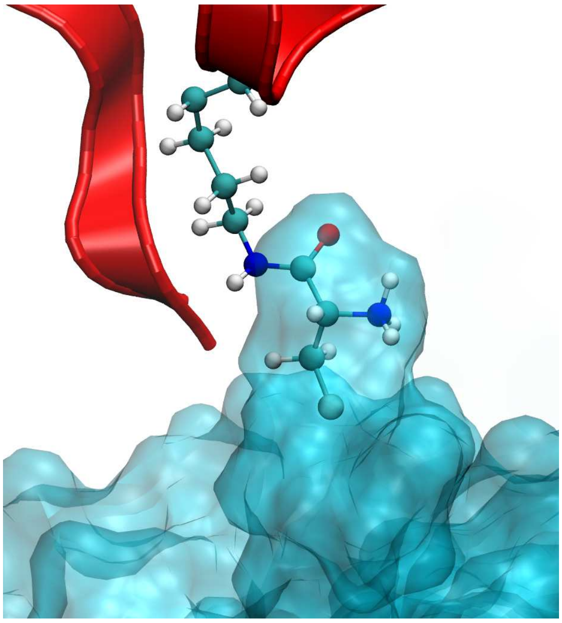
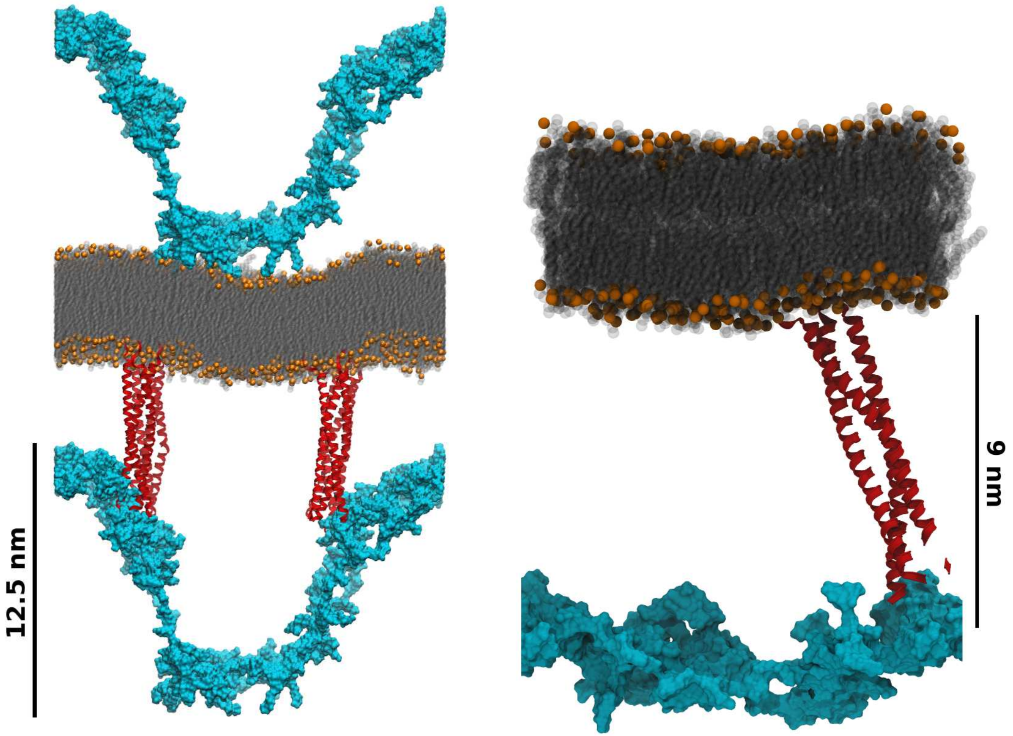


FIGURE 3.2: Amide linkage between BLP (red ribbons) and the cell wall (transparent cyan surface) after 40 ns NPT and 20 ns production MD of the complete system. The amide linkage is shown in diffuse CPK.

oxygens. Upon removal of the O13 atom it was therefore necessary to update the parameters to represent a permanent double bond between the remaining oxygen and carbon atoms. Further refinements of the script were then made to allow for the binding of multiple BLP monomers to the cell wall at the same time, giving the user the option to generate much larger systems with ease.

Simulations of the cell wall alone in solution invariably resulted in an undulating structure presumed to be caused by the lack of structural constraints usually applied to the cell wall by external proteins such as BLP. With this in mind, it was thought that the careful application of position restraints during whole system equilibration, along with the known elasticity of PGN, would allow for the cell wall to relax into a biologically relevant conformation without any curvature. Initial attempts to construct the combined system using this procedure therefore utilised an unequilibrated cell wall model that had been generated to match the dimensions of the pre-equilibrated OM model as closely as possible. This approach however, led to massive deformations in the cell wall and/or the OM (Fig. 3.3a) in all cases.

It has been shown by both *in vivo* and *in silico* studies that the elasticity of the *E. coli* cell wall differs when measured along the glycan strands or in the direction of the peptide cross-links Yao et al. (1999); Gumbart et al. (2014). In both studies, a much larger modulus of elasticity was measured in the direction of the glycan strands compared to in the direction of the peptide cross-links; *i.e.* the glycan strands were found to be much stiffer than the cross-links. Given this, any pressure coupling applied during an equilibration phase involving the cell wall must be anisotropic to



(A) Final structure after 20 ns production MD of large (4 x BLP) whole system simulation using an unequilibrated cell wall model. A section of the periodic copy of the cell wall (cyan) is shown to highlight its interaction with the OM across the periodic boundary. Lipid head group phosphates are shown in orange, BLP in red and the lipid tails of the OM hydrophobic region in grey.

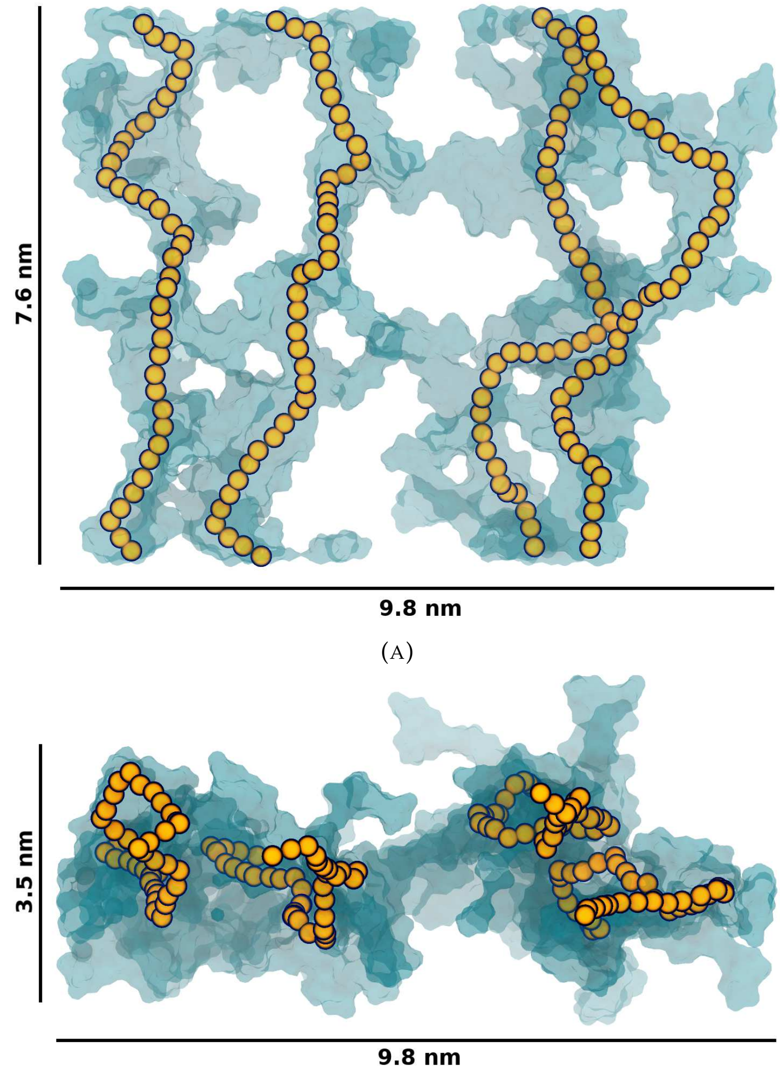
(B) Final structure of the complete system after 40 ns NPT and 20 ns production MD. Head group phosphates are shown in orange (vdW spheres), BLP in red (ribbons) and the cell wall in cyan (surface plot). The gap in one of the BLP monomers is an artefact of the visualisation software and does not represent a discontinuity in the molecule.

FIGURE 3.3: Final structures of cell envelope before (A) and after (B) correcting for self interactions across the periodic z boundary.

allow for asymmetric relaxation; without doing so, the tension in either direction would be incorrectly coupled to the tension in the orthogonal direction through the assertion of equal x and y box dimensions.

Through extensive testing of various equilibration procedures and independent simulation of the system components it was eventually found that, when equilibrated in isolation with anisotropic pressure coupling, the cell wall consistently contracted by roughly 30 % along the axis of the glycan strands and elongated by roughly 50 % in the orthogonal, peptide linkage axis. This motivated subsequent attempts at construction of the system to utilise a cell wall model with dimensions that were correspondingly larger/smaller in the glycan/peptide axes, respectively, such that the cell wall would relax into a conformation that better matched the dimensions of the equilibrated BLP-OM model. This new cell wall (Figure 3.1) was then passed through 2 ns NVT and 40 ns anisotropic NPT equilibration phases with z-axis position restraints of strength 500 kJ.mol^{-1}.nm^{-2} applied on the glycan backbone carbon atoms to minimise any curvature induced during simulation. The resulting equilibrated cell wall was then passed through 40 ns of unrestrained production MD after which it matched the dimensions of the equilibrated OM-BLP model to within 1 nm in either direction, exhibited the desired disordered glycan strand configuration and expressed an acceptable level of curvature (Figure 3.4).

The equilibrated cell wall was then bound to the equilibrated BLP-OM model using the previously described binding script before being passed through whole system equilibration and a subsequent 20 ns production run. The equilibration procedure is outlined in Table 3.2 and the final structure of the system after the 20 ns production run is shown in Figure 3.3b.



(A)

(B)

FIGURE 3.4: Top-down (A) and side-on (B) views of the equilibrated cell wall isolated in solution after 40 ns production MD. The carbon backbone of each glycan strand is highlighted in orange. Surface plots of the entire cell wall are shown in cyan.

Phase	Description
EM	• Integrator = steep • EM tolerance = 10 $kJ.mol^{-1}.nm^{-2}$
NVT	• Simulation time = 200 ps • Time step = 0.001 ps • Position restraints = 1000 $kJ.mol^{-1}.nm^{-2}$ on z coordinates of cell wall and lipid head group phosphates
NPT1	• Simulation time = 20 ns • Anisotropic pressure coupling • Time step = 0.001 ps • Position restraints = 1000 $kJ.mol^{-1}.nm^{-2}$ on z coordinates of lipid head group phosphates only
NPT2	• Simulation time = 20 ns • Anisotropic pressure coupling • Time step = 0.001 ps • Position restraints = None

TABLE 3.2: Key simulation parameters for each stage of whole system equilibration.

3.5 Crowding Regimes

In order to investigate the effects of crowding on the behaviour of PMB1 and PME within the periplasm of *E. coli*, three crowding regimes of increasing complexity were studied. These regimes describe different compositions of the periplasmic fluid added to the previously discussed cell envelope model during the simulations performed throughout this work and are outlined in the following sections.

3.5.1 Polymyxins

The first and simplest model of the periplasm comprised 8 polymyxin molecules (all PMB1 or all PME) in solution with either a neutralising or 150 mM concentration of KCl. This number of polymyxin molecules was chosen to match the number density used in the aforementioned complementary study of this system Pedebos et al. (2021). A snapshot taken from a simulation of the cell envelope in this crowding regime is presented in Figure 3.5a.

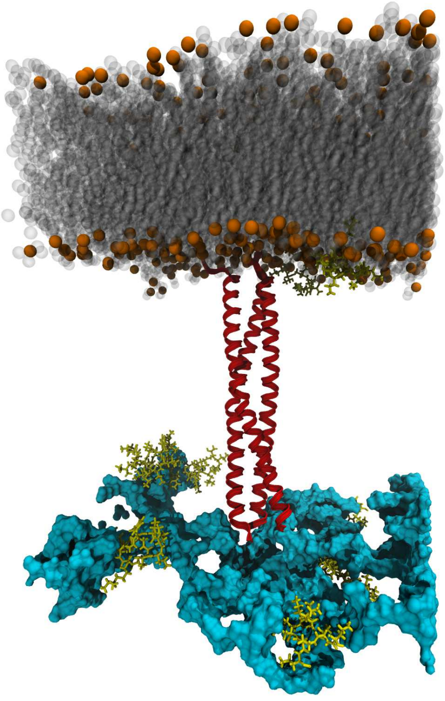
3.5.2 Polymyxins & Osmolytes

The next model of the periplasm studied in this work also contained 8 polymyxin molecules, with the addition of 3 types of osmolyte to increase the density and complexity of biomolecules within the system. The chosen osmolytes were glycerol, osmoregulated periplasmic glycan (OPG) and spermidine. Since spermidine is positively charged, OPG negatively charged and glycerol neutral; these osmolyte provided a rough sampling of polymyxin interactions with molecules of all net charge types. In total, 1 spermidine, 9 OPG and 17 glycerol molecules were added in this regime. These numbers were chosen to match the experimentally determined concentrations of these species within the periplasmic region of *E. coli* Wang et al. (2019); Cayley et al. (2000); Shah and Swiatlo (2008); Bontemps-Gallo et al. (2017). A snapshot taken from a simulation of the cell envelope in this crowding regime is presented in Figure 3.5b.

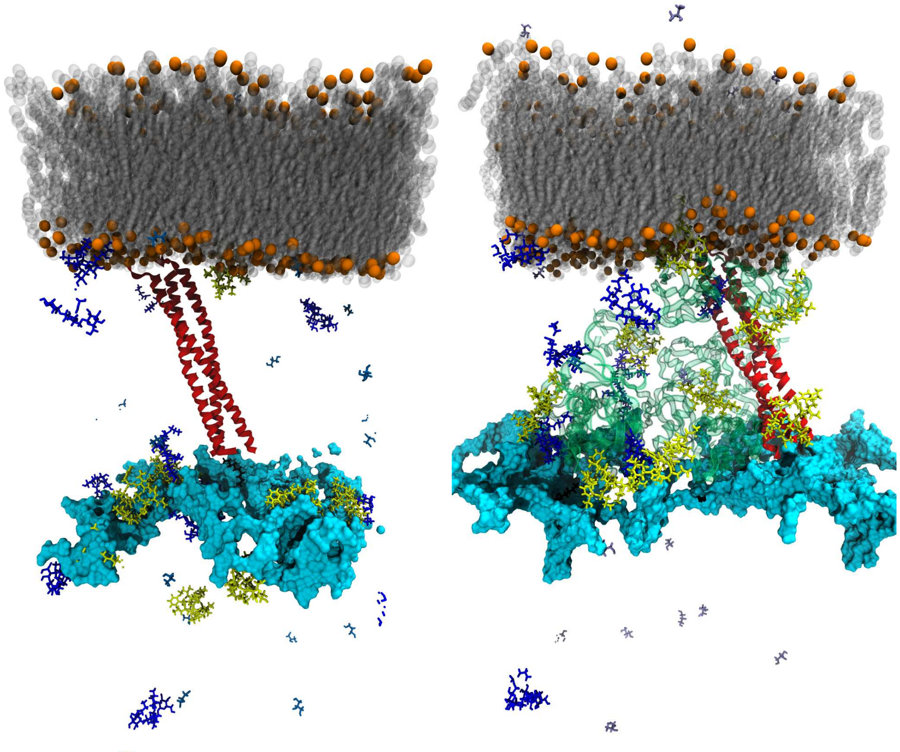
3.5.3 Polymyxins, Osmolytes & Ubiquitin

The final crowding regime studied in this work included polymyxins, glycerol, OPG, spermidine and ubiquitin. The addition of a crowding protein, ubiquitin, allowed for the periplasm to be packed with biomolecules to a crowding fraction of $\phi \simeq 0.21$. This value was chosen to match that used in the complementary study previously published by the Khalid research group Pedebos et al. (2021) and resulted in the addition of 11 ubiquitin molecules. A snapshot taken from a simulation of the cell envelope in this crowding regime is presented in Figure 3.5c.

Ubiquitin was chosen due to its abundance across all eukaryotic cells, with functional analogs also found in prokaryotes Darwin (2009), making it a biologically reasonable choice for a representative crowding protein *in lieu* of the full range of proteins present within the cell envelope. Furthermore, another study into the diffusive motion of proteins in crowded media utilised ubiquitin as a representative protein von Bülow et al. (2019), thus the use of ubiquitin within this work would allow for the future comparison of results with existing literature.



(A) Polymyxins



(B) Polymyxins & osmolytes

(C) Polymyxins, osmolytes & ubiquitin

FIGURE 3.5: Snapshots taken during simulations of the three crowding regimes. Diffuse orange vdW spheres represent OM head group phosphate atoms. Transparent grey vdW spheres represent OM carbon atoms. The PGN cell wall is represented by the diffuse cyan surface plot. BLP is represented by red ribbons. PMB1 shown in yellow licorice. Glycerol shown in ice-blue licorice (small molecules). OPG shown in true-blue licorice (large molecules). Spermidine shown in black licorice. Ubiquitin is represented by transparent green ribbons.

3.6 Conclusions

This chapter has outlined each step of the procedure through which the model of the *E. coli* cell envelope used in Chapters 4 and 5 was constructed. The implementation of the scripts used to generate the cell wall and combined PGN-BLP structure and topology files have been described in detail and the importance of independent equilibration of individual system components when constructing complex, multifaceted systems has been discussed.

Particular focus has been given to how asymmetries in the chemical structure of the PGN cell wall necessitate the use of anisotropic pressure coupling during both equilibration and production MD phases. This requirement is in contrast to typical membrane simulation studies that utilise semi-isotropic pressure coupling due to the approximate rotational symmetry of these systems about the membrane normal Waller et al. (2023); Shearer et al. (2019); Webby et al. (2022). The requirement of anisotropic pressure coupling in cell wall simulations is thus of interest to other researchers seeking to perform membrane simulations in the presence of the cell wall.

Finally, the crowding regimes utilised in Chapters 4 and 5 have been discussed, with motivation provided for the choice of constituent molecular species and their respective concentrations.

The CHARMM36m model of the *E. coli* periplasm presented in this chapter builds upon previously published work by our group, on which I am a co-author, that developed a model with a similar composition using the united-atom GROMOS54a7 force field Pedebos et al. (2021). Whilst existing studies may be found that utilise the CHARMM36m force field to study bacterial membranes Wu et al. (2013); Daison et al. (2021); Sharma and Ayappa (2022); Waller et al. (2023) or the PGN cell wall Gumbart et al. (2014); Vaiwala et al. (2022) independently, the system presented in this chapter is the only CHARMM36m model published to date that combines these structures in one model, linking them *via* the periplasm spanning biomolecule, BLP Smith et al. (2024). As such, this work marks a notable step forwards in the accuracy with which we may study this complex environment using molecular dynamics simulation.

Chapter 4

The nature of Polymyxin interactions with the *E. coli* cell wall

4.1 Introduction

The bacterial cell envelope is a complex, multilayered structure that serves as a barrier between the interior of the cell and the often hostile external environment. In the case of Gram-negative bacteria such as *E. coli*, the cell envelope is composed of an inner membrane (IM) and outer membrane (OM) that form the boundaries of the crowded aqueous compartment known as the periplasm. Within the periplasm lies the cell wall, a mesh-like structure composed of cross-linked strands of peptidoglycan (PGN) polymers, along with a wide variety of proteins, osmolytes and ions Goemans et al. (2014); Weiner and Li (2008). The only protein known to provide a covalent link with the cell wall is Braun's lipoprotein (BLP, also known as "Lpp" or murein lipoprotein). BLP is anchored in the OM *via* a lipidated N-terminus, with the C-terminus being covalently bound to the peptide stem of a PGN monomer. With an estimated 10^5 copies per cell, BLP is the most abundant protein in *E. coli* and acts as a structural scaffold linking the cell wall to the OM, maintaining their separation and facilitating the non-covalent interaction of other OM proteins with the cell wall Samsudin et al. (2017). The cell envelope is thus a complex, crowded environment and the considerable challenges posed to the movement of small molecules, such as antibiotics, throughout this space are starting to be elucidated through both *in vitro* and *in silico* studies Kuznetsova et al. (2014); Pedebos et al. (2021).

The recent emergence of bacterial strains resistant to both PMB and PME Gales et al. (2011); Li et al. (2019), coupled with the associated toxicities of their clinical use, necessitates either their modification or the development of completely novel antibiotics and this need is reflected by the ongoing search for polymyxin derivatives Kaguchi et al. (2023a). As such, it is of immediate interest to establish a molecular-level understanding of each stage of the process *via* which they bring about cell death and, in particular, to begin to fill the gap in our understanding of how polymyxins are transported across the periplasm, towards the target of their antimicrobial function, the IM.

A previous paper published by the Khalid research group, on which I am a co-author, highlighted the promiscuity of polymyxin interactions within the crowded periplasm of *E. coli* using the united-atom GROMOS54a7 forcefield Pedebos et al. (2021). In this study, PMB1 was seen to interact profusely with the cell wall, proteins, osmolytes and outer membrane lipids present within the periplasm. Indeed, PMB1 was rarely uncomplexed throughout the simulations. In order to better understand how such interactions may facilitate, hinder or leave unimpeded the translocation of polymyxins from the OM to the IM, it is prescient to investigate the comparative interactions of PMB1 and PME within models of the periplasm crowded to varying extents.

To this end, a model of a portion of the *E. coli* cell envelope (Figure 4.1A) has been constructed using the all-atom CHARMM36m forcefield, according to the protocol described in Chapter 3. A series of molecular dynamics simulations (Table 4.1) of this model have been performed in the presence of either PMB1 or PME under a range of periplasmic fluid compositions of differing complexity, these varied compositions are referred to as "crowding regimes" from here on.

The simulations presented in this chapter show that in the absence of a diverse chemical environment both PMB1 and PME tend to bind rapidly and irreversibly to the cell wall; predominantly *via* polar interactions between the cationic diaminobutyric acid (DAB) residues of the polymyxins and the various carboxylate groups on the peptide residues of the cell wall. These interactions are shown to be disrupted by the presence of physiological salt concentrations, or increased biomolecular crowding, allowing for the dissociation of the polymyxins from the cell wall and their subsequent interaction with the other components of the cell envelope. Evidence is provided that shows how certain cations, osmolytes and proteins contribute to the disruption of polymyxin-cell wall interactions by forming competing interactions with the carboxylate groups on the peptide stems of the cell wall, reducing the number of such interaction sites available to nearby polymyxin molecules.

4.2 Methods

4.2.1 Envelope model construction

The *E. coli* cell envelope model was based on a composition validated in previous work published by the Khalid research group Samsudin et al. (2017); Boags et al.

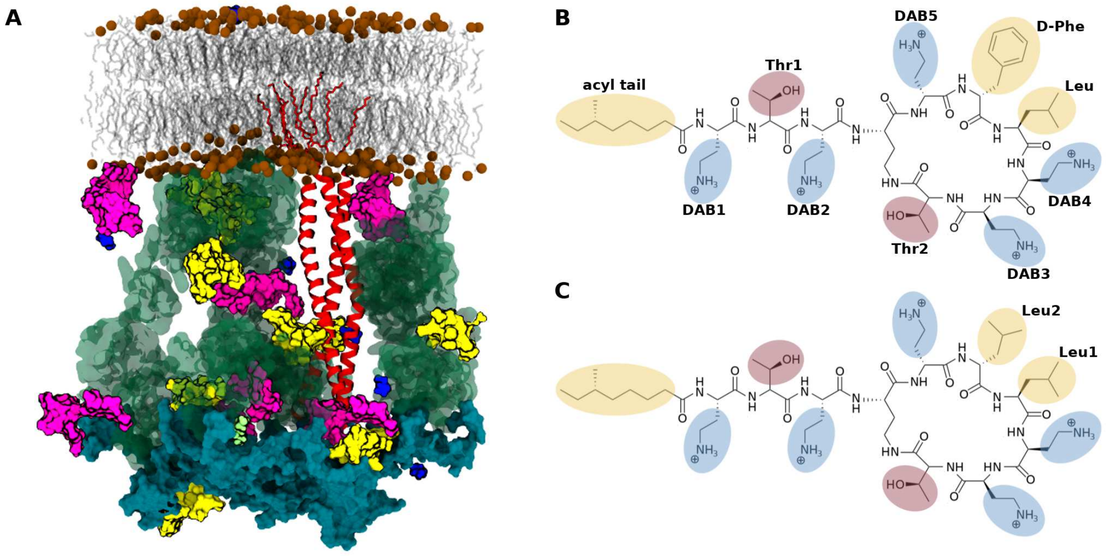


FIGURE 4.1: A) Image of the cell envelope in the Ubiq crowding regime, water and ions are omitted to aid visualization. Polymyxins shown in yellow, OPG in magenta, glycerol in blue, spermidine in lime, ubiquitin in transparent green, BLP in red ribbons and PGN in diffuse cyan. B/C) Chemical structures of PMB1/PME. Hydrophobic residues are highlighted in yellow, DAB in blue and Thr in red.

(2019). Details of the protocol used to construct and equilibrate the model can be found in Chapter 3.

An asymmetric model of the OM with an outer leaflet composed entirely of the lipid A region of LPS and an inner leaflet composed of 90% 1-palmitoyl 2-cis-vaccenic phosphatidylethanolamine (POPE), 5% 1-palmitoyl 2-cis-vaccenic phosphatidylglycerol (POPG) and 5% 1-palmitoyl 2-cis-vaccenic 3-palmitoyl 4-cis-vaccenic diphosphatidylglycerol (PVCL2, also known as cardiolipin) was constructed. This OM model has been validated in previous studies published by the Khalid research group Ortiz-Suarez et al. (2016); Samsudin et al. (2016); Piggot et al. (2011); Khalid et al. (2019).

The 1.9 Å crystal structure of the BLP homotrimer (1EQ7) from *E. coli* was used Shu et al. (2000). All three BLP helices were acylated at their C-termini using the CHARMM-GUI membrane builder tool Lee et al. (2019). The resulting acylated homotrimer was manually inserted into the inner leaflet of the OM using VMD Humphrey et al. (1996). The combined OM/BLP system was then equilibrated to ensure correct lipid packing around the newly inserted acyl tails.

A single layer model of the peptidoglycan (PGN) cell wall was generated as previously discussed in Section 3.3 Samsudin et al. (2017); Boags et al. (2019); this model was equilibrated alone in solution before being covalently bound to the aforementioned membrane-inserted BLP molecule. The N-terminus of one BLP monomer within the homotrimer was covalently bound to the C-terminus of a meso-DAP residue located on a non-crosslinked peptide stem within the cell wall.

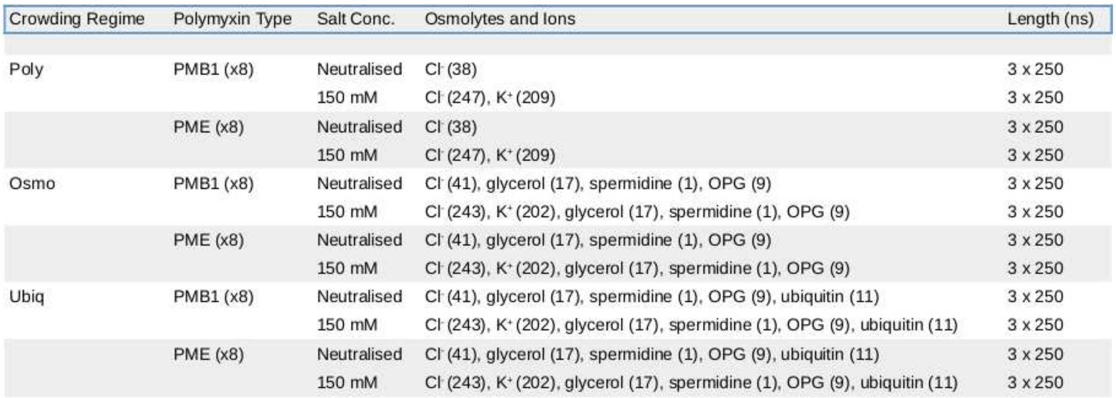
Avogadro Hanwell et al. (2012) was used to generate a structure of the region surrounding the amide linkage connecting BLP and PGN. This structure was then passed through the CHARMM-GUI ligand reader tool Kim et al. (2017) to generate bond parameters for the covalently bound region. Bespoke scripts were then implemented to combine these files into complete topology and structure files for the combined system.

4.2.2 Envelope system preparation

Three cell envelope crowding regimes were used in this work; namely the Poly, Osmo and Ubiq regimes (Section 3.5). The simplest of these, the Poly regime, contained polymyxin molecules alone in the periplasmic region between the OM and cell wall. In addition to these polymyxin molecules, the Osmo regime contained a range of small osmolytes; namely, spermidine, glycerol and osmoregulated periplasmic glucans (OPG). These osmolytes were selected based on their chemical diversity and abundance within the *E. coli* cell envelope. The concentrations of these molecules within the periplasm are either documented or estimated in the literature Wang et al. (2019); Cayley et al. (2000); Shah and Swiatlo (2008); Bontemps-Gallo et al. (2017), and are reproduced in the model used in this chapter: glycerol (36 mM), OPG (20 mM) and spermidine (3 mM). The Ubiq regime was the most compositionally complex system studied in this work and, along with the polymyxin molecules and osmolytes, included ubiquitin proteins. The number of proteins added into the periplasm was chosen to reproduce a crowding volume fraction of $\phi \sim 0.21$, as estimated from experimental studies Cayley et al. (2000).

Simulations of all three crowding regimes were performed in the presence of either PMB1 or PME. Each regime was prepared under two different concentrations of KCl; with neutralising counterions alone, or, with neutralising counterions and an excess salt concentration of 150 mM. These are referred to as the “neutralised” and “concentrated” systems from here on. Triplicate replica simulations of all system compositions were prepared; a summary of all simulations performed can be found in Table 4.1.

The CGenFF protocol Vanommeslaeghe et al. (2010) was used to generate parameters for PMB1, the CHARMM-GUI ligand reader tool Kim et al. (2017) was used to generate parameters for spermidine and glycerol whilst the CHARMM-GUI glycan reader tool Park et al. (2019) was used to generate those for OPG. The crystal structure of ubiquitin (1UBQ) was obtained from the RSCB database, determined at a resolution of 1.8 Å Vijay-Kumar et al. (1987).



Crowding Regime	Polymyxin Type	Salt Conc.	Osmolytes and Ions	Length (ns)
Poly	PMB1 (x8)	Neutralised	Cl^- (38)	3 x 250
		150 mM	Cl^- (247), K^+ (209)	3 x 250
	PME (x8)	Neutralised	Cl^- (38)	3 x 250
		150 mM	Cl^- (247), K^+ (209)	3 x 250
Osmo	PMB1 (x8)	Neutralised	Cl^- (41), glycerol (17), spermidine (1), OPG (9)	3 x 250
		150 mM	Cl^- (243), K^+ (202), glycerol (17), spermidine (1), OPG (9)	3 x 250
	PME (x8)	Neutralised	Cl^- (41), glycerol (17), spermidine (1), OPG (9)	3 x 250
		150 mM	Cl^- (243), K^+ (202), glycerol (17), spermidine (1), OPG (9)	3 x 250
Ubiq	PMB1 (x8)	Neutralised	Cl^- (41), glycerol (17), spermidine (1), OPG (9), ubiquitin (11)	3 x 250
		150 mM	Cl^- (243), K^+ (202), glycerol (17), spermidine (1), OPG (9), ubiquitin (11)	3 x 250
	PME (x8)	Neutralised	Cl^- (41), glycerol (17), spermidine (1), OPG (9), ubiquitin (11)	3 x 250
		150 mM	Cl^- (243), K^+ (202), glycerol (17), spermidine (1), OPG (9), ubiquitin (11)	3 x 250

TABLE 4.1: Summary of all simulations performed for this work. Bracketed numbers give the number of each molecule present within the specific simulation regime.

4.2.3 Simulation protocols

Simulations were performed using the GROMACS 2021.2 molecular dynamics package Abraham et al. (2015), utilising the CHARMM36m force field Huang et al. (2017) and TIP3 water model Jorgensen et al. (1983). Simulations were divided into two parts: equilibration simulations in NVT and NPT ensembles lasting for 200 ps and 40 ns respectively; and production simulations in the NPT ensemble, which ran for 250 ns. A constant temperature of 310 K was maintained using the velocity rescale thermostat Bussi et al. (2007) with a time constant of 1 ps. The pressure was maintained anisotropically at 1 atm using the Parrinello-Rahman barostat Parrinello and Rahman (1981) with a time constant of 1 ps. Hydrogen bonds were constrained using the LINCS algorithm Hess et al. (1997); Hess (2008); stable treatment of these constraints required the use of a 1 fs integration time step. Long-range electrostatics were treated using the particle mesh Ewald method Darden et al. (1993). The short-range electrostatic and van der Waals cut-offs were both set to 1.2 nm.

For the replicates; new initial configurations of all polymyxin, osmolyte and ubiquitin molecules were generated, along with re-solvation and ionisation of each system before being passed through the equilibration and production simulation phases. The initial velocities of all atoms were modified between each replicate at the start of NVT equilibration to ensure an unbiased sampling of the simulation phase space.

The trend in interactions with the cell wall oxygens was calculated by considering the number of cell wall oxygen contacts (atomic separation < 4 Å) with K^+ and the mean number of coincident cell wall oxygen contacts with the DAB residues of both peptides. Linear regression models were fitted against the mean number of DAB residue contacts with cell wall oxygens, as a function of the number of coincident K^+ contacts with cell wall oxygens; this analysis was performed using the one dimensional polynomial fitting algorithm provided by the NumPy Harris et al. (2020) python package. Confidence intervals were calculated, at a confidence level of 95 %, against each complete set of DAB - cell wall oxygen contact counts corresponding to unique values of coincident K^+ - cell wall oxygen contacts, according to the standard form for the confidence interval Simundic et al. (2008).

4.3 Results

A preliminary analysis was performed of the number of hydrogen bonds formed between the cell wall and water molecules throughout each replicate simulation of the neutralised Poly and concentrated Ubiq regimes; representing, respectively, the most compositionally simple and complex systems simulated in this work. All three replicates of the concentrated Ubiq regime exhibited consistently fewer TIP3-PGN hydrogen bonds during the final 125 ns of simulation than were observed in any of the three replicates of the neutralised Poly regime during the same period (Figure 4.2).

The observed decrease in the availability of TIP3-PGN hydrogen bonding under a more complex chemical environment is indicative of the propensity for the various ions, osmolytes and crowding proteins to interact with the cell wall, coating the surface of the cell wall (Figure 4.11) and decreasing the number of available hydrogen bonding sites for the surrounding water molecules. This effect is likely to also impact the nature of interactions between the cell wall and polymyxin molecules as the presence of an abundance of other biomolecules forces the polymyxins to compete for interaction sites on the cell wall surface.

FIGURE 4.2: Number of hydrogen bonds formed between water molecules (TIP3) and the cell wall (PGN) throughout each replicate simulation of the neutralised Poly and concentrated Ubiq regimes in the presence of PMB1. Data is plotted as a moving average across 25 ns windows.

Consideration of the simple system alone is therefore unlikely to be indicative of the behaviour of the system *in vivo* and thus, in conjunction with the more detailed analysis that follows, this result begins to highlight the importance of considering the true biological complexity of a system when determining the nature of biomolecular interactions through simulation.

The subsequent analyses were split into two components; the duration of binding between the polymyxins and the cell wall, and the specific residue interactions observed at each frame during their binding. Kernel density estimates (KDE) were fitted to the observed binding durations across all replicates of each system, and are presented in Figure 4.3B. The specific residue interactions were categorised according to interaction type (*i.e.* involving either hydrophobic, DAB or Thr residues of the polymyxins) and the aggregated results for each polymyxin species across all simulation regimes are presented in Figures 4.3C and 4.3D. The complete dataset for each simulation regime can be found in Tables 4.2 & 4.3.

4.3.1 Binding Durations

4.3.1.1 PMB1

Focusing first on the analysis of the binding duration between PMB1 and the cell wall; the chemical complexity of the simulation environment was seen to have a clear effect (Figure 4.3B). Across all replicate simulations of the neutralised Poly regime, all but one PMB1 molecules were bound to the cell wall at the onset of production MD. These molecules bound to the cell wall during the first 10 ns of NPT equilibration and did not dissociate from the cell wall for the duration of the subsequent production MD simulations. The anomalous PMB1 molecule bound to the cell wall within 10 ns of production MD and thereafter also remained associated with the cell wall. A similar result was observed in the neutralised Osmo regime; all PMB1 molecules were initially bound to the cell wall and remained there for the duration of the simulation.

When these same systems were simulated under concentrated conditions, multiple PMB1 molecules in all replicate simulations of both the Poly and Osmo regimes were not bound to the cell wall at the onset of production MD. Dissociation of PMB1 molecules from the cell wall was prevalent in both regimes and resulted in an abundance of short ($\leq$ 50 ns) and intermediate (50 - 200 ns) duration PMB1-cell wall interactions.

The addition of osmolytes further increased the preference for short and intermediate duration interactions between PMB1 and the cell wall. Interactions persisting for less than 200 ns accounted for 39.4 % of the total number of PMB1-cell wall interactions in the concentrated Poly regime, compared to 69.4 % in the concentrated Osmo regime.

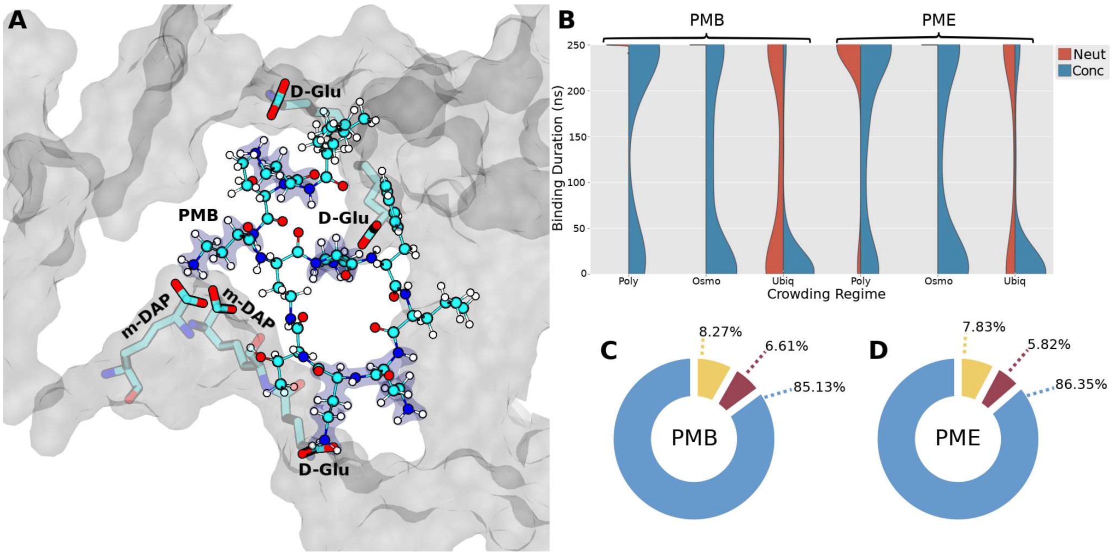


FIGURE 4.3: A) Image of PMB1 molecule inserted into a pore in the cell wall. Specific cell wall residues coordinated with PMB1 are shown in semi-transparent licorice. Cell wall is shown as a transparent surface, PMB1 in CPK. DAB residues of PMB1 are highlighted in blue bubbles. B) Kernel density estimate (KDE) curves fitted to the binding durations of all unique instances of binding between polymyxin molecules and the cell wall. C/D) Pie charts of residue interaction types between the cell wall and PMB1 (C) / PME (D), aggregated across all simulation regimes. Interactions involving the DAB / Thr / hydrophobic residues of polymyxins are coloured in blue / red / yellow respectively.

This effect can be seen explicitly in Figure 4.3B, where comparison of the KDE curves for the two regimes shows that the concentrated Osmo regime exhibited a higher relative density of short duration interactions.

Increased periplasmic crowding also impacted the interactions between PMB1 molecules and the cell wall under both neutralising and excess salt concentrations. In the neutralised Ubiq regime, multiple PMB1 molecules in each replicate simulation were not bound to the cell wall at the onset of production MD; contrary to the behaviour observed in the less crowded neutralised systems. Furthermore, of all the neutralised simulations reported in this study, dissociation of PMB1 from the cell wall was observed only in these, the most crowded, systems. The median PMB1-cell wall binding duration in the neutralised Ubiq regime was ~ 96 ns, compared to just ~ 1 ns in the concentrated Ubiq regime. This dramatic decrease is indicative of a strong preference for PMB1 to form short duration interactions with the cell wall under higher salt concentration; indeed, 87 % of PMB1-cell wall interactions were classified as short (< 50 ns) in the concentrated Ubiq regime, compared to just ~ 43 % under neutralised conditions.

Notably, this result also indicates that the concentrated Ubiq regime (*i.e.* the most compositionally complex and crowded system) exhibits the largest proportion of short and intermediate duration interactions of any simulated system. This follows the observed trend from the concentrated Poly and Osmo regimes, whereby the addition of osmolytes led to a relative increase in the number of short and intermediate duration interactions; indicating that the presence of crowding ubiquitin proteins further enhanced the effects caused by the inclusion of osmolytes.

4.3.1.2 PME

PME molecules in both the neutralised Poly and Osmo regimes exhibited a preference for long duration ($>$ 200 ns) interactions with the cell wall; with both systems exhibiting median binding durations of 250 ns. Indeed, all PME molecules in these two systems were found to be bound to the cell wall at the beginning and end of production MD. In the neutralised Osmo regime, no dissociation of PME from the cell wall was observed in any of the replicate simulations. In one replicate of the neutralised Poly regime, however, a single PME molecule was seen to dissociate from the cell wall after 1.6 ns of production MD. This molecule was situated on the IM-facing surface of the cell wall at the onset of production MD, having passed through a pore in the cell wall during NPT equilibration. After moving across the cell wall surface for approximately 26 ns, it bound to a non-crosslinked meso-DAP residue on the cell wall where it remained for the rest of the simulation.

The similarity of the KDE curve profiles for equivalent systems of PME and PMB1 in the Poly and Osmo regimes highlights that the effects of an increase in salt concentration on the binding of PME to the cell wall were similar to those discussed for PMB1 (Figure 4.3B). Multiple PME molecules in each replicate of the concentrated Poly and Osmo regimes were not bound to the cell wall at the onset of production MD, and dissociation of PME from the cell wall was observed in all simulations of these regimes. Furthermore, interactions persisting for less than 200 ns accounted for 51.4 % of PME-cell wall interactions in the concentrated Poly regime, compared to 65.1 % in the concentrated Osmo regime; highlighting that, similar to the behaviour observed in systems containing PMB1, the inclusion of osmolytes under concentrated conditions increased the preference for short and intermediate duration interactions between PME and the cell wall.

Increased periplasmic crowding further impacted the interactions between PME and the cell wall under both neutralising and excess salt concentrations. In the neutralised Ubiq regime, at least one PME molecule in each replicate simulation was not bound to the cell wall at the onset of production MD. Furthermore, dissociation of PME from the cell wall under neutralised conditions was most prevalent in these most crowded systems: with a mean $\sim$ 1.4 instances of PME-cell wall binding per polymyxin in the Ubiq regime, compared to $\sim$ 1.1 and 1.0 instances per polymyxin in the Poly and Osmo regimes, respectively. The median PME-cell wall binding duration in the neutralised Ubiq regime was $\sim$ 210 ns, compared to just $\sim$ 1 ns in the concentrated

Ubiq regime. This dramatic decrease is indicative of a strong preference for PME to form short duration interactions with the cell wall under higher salt concentration; indeed, $\sim$ 81 % of PME-cell wall interactions were classified as short ($<$ 50 ns) in the concentrated Ubiq regime, compared to just $\sim$ 38 % under neutralised conditions. These results indicate that increases in the chemical complexity and crowding of the simulation environment affect compounding effects on the behaviour of PME; disrupting PME-cell wall interactions and thus leading to a relative increase in the number of short and intermediate duration interactions, closely following the behaviour observed in simulations of PMB1.

4.3.2 Biochemical nature of polymyxin - cell wall interactions

Analysis of the biochemical nature of the interactions between the two polymyxin species and the cell wall showed that their binding was underpinned predominantly by interactions between the charged DAB residues of the polymyxins and polar residues of the cell wall. This trend was consistent across all simulated systems. Cell wall interactions involving the DAB residues of polymyxins accounted for 85.4 ± 1.1 % and 86.8 ± 2.0 % (mean and standard deviation) for PMB1 and PME respectively. The polar meso-DAP, D-Glu and D-Ala residues on the peptide stems of the cell wall were most prevalently involved in these interactions for both peptides ($\sim$ 32 %, $\sim$ 23 % and $\sim$ 19 % respectively), precise details regarding these residue interaction percentages are given in Tables 4.2 & 4.3.

4.3.3 Cationic disruption of polymyxin – cell wall binding

The observation that both PMB1 and PME interact with the cell wall predominantly *via* polar interactions between their cationic DAB residues and the meso-DAP, D-Glu and D-Ala residues of the cell wall peptide stems enabled subsequent investigation

Residues		Simulation Regime											
PGN	PMB	poly-neut		poly-conc		osmo-neut		osmo-conc		ubiq-neut		ubiq-conc	
m-DAP	PBD1	31.21	18.44	34.49	16.67	31.75	14.02	33.83	14.78	35.51	18.04	30.48	13.63
D-Glu	PBD2	21.65	19.51	22.51	17.18	21.95	20.93	22.7	19.34	21.97	19.28	22.79	21.59
D-Ala	PBD3	18.3	15.01	20.23	15.89	18.64	15.68	18.43	16.81	19.28	15.38	17.45	19.55
L-Ala	PBD4	7.51	15.24	8.57	18.78	9.08	18.76	9.88	17.22	6.56	20.39	6.75	22.46
GlcNAc	PBD5	8.36	17.68	5.84	15.86	7.35	15.06	6.53	16.28	6.66	12.87	7.52	9.88
MurNAc	PBT1	12.97	3.2	8.35	3.55	11.22	3.23	8.62	2.85	10.02	4.45	15.01	2.57
	PBT2		3.41		3.49		3.25		2.83		3.02		3.58
	PBTL		3.63		3.49		2.55		3.59		1.25		2.87
	PBDP		2.09		2.78		4.12		3.36		2.6		1.99
	PBLU		1.79		2.31		2.41		2.94		2.72		1.89
🟨	hydrophobic	7.51		8.57		9.08		9.88		6.56		6.75	
🟦	DAB	85.88		84.38		84.44		84.43		85.97		87.1	
🟥	Thr	6.61		7.04		6.48		5.68		7.47		6.15	

TABLE 4.2: PMB1-PGN residue interaction percentages.

Residues		Simulation Regime											
PGN	PME	poly-neut		poly-conc		osmo-neut		osmo-conc		ubiq-neut		ubiq-conc	
m-DAP	PED1	28.76	15.02	34.77	17.93	30.08	14.85	30.81	15.04	30.81	15.74	36.04	17.45
D-Glu	PED2	22.07	22.01	22.19	16.13	24.2	17.84	24.21	16.78	22.75	18.66	22.31	17.84
D-Ala	PED3	19.36	20.56	19.55	18.47	16.18	22.05	19.66	20.23	18.05	19.77	20.25	20.01
L-Ala	PED4	5.71	15.35	8.54	15.64	7.91	17.88	9.69	17.47	7.47	18.53	6.56	16.27
GlcNAc	PED5	10.14	16.51	5.71	16.31	9.71	13.67	5.96	15.07	8.85	14.84	4.16	16.74
MurNAc	PET1	13.96	2.93	9.24	3	11.92	2.74	9.67	2.57	12.06	2.61	10.68	2.06
	PET2		1.91		3.98		3.06		3.15		2.37		3.07
	PETL		3.1		3.64		2.63		3.78		2.14		2.53
	PEL1		1.04		1.89		2.34		2.89		3.02		1.63
	PEL2		1.57		3.01		2.95		3.02		2.31		2.39
🟨	hydrophobic	5.71		8.54		7.91		9.69		7.47		6.56	
🟦	DAB	89.45		84.47		86.29		84.59		87.54		88.31	
🟥	Thr	4.84		6.98		5.8		5.72		4.98		5.13	

TABLE 4.3: PME-PGN residue interaction percentages.

into the biochemical nature of the disruption of their binding with the cell wall upon the addition of osmolytes and/or an excess salt concentration to the system.

The meso-DAP, D-Glu and D-Ala residues of the cell wall each contain anionic carboxylate groups within their structure and it was these functional groups in particular that were responsible for the prolific binding of these residues to the polymyxins; as these carboxylate groups readily formed salt bridges with the cationic amine groups on the polymyxin DAB residues (Figure 4.3A).

Within the concentrated simulation regimes, K^+ were seen to coordinate with these same carboxylate groups on the cell wall peptide residues. In some cases this coordination occurred in close proximity to polymyxin molecules that were already bound to the cell wall, leading to direct competition between the cationic moieties to form salt bridges with the carboxylate groups (Figure 4.4A). It is noted here that K^+ are not present in the neutralised simulation regimes and therefore no such coordination was observed in any of the neutralised simulations.

Within the concentrated Poly regime, a negative linear correlation was calculated between the number of cell wall oxygen contacts with K^+ and the mean number of coincident cell wall oxygen contacts with the DAB residues of both PMB1 (Figure 4.4B) and PME (Figure 4.4C); supporting the idea that K^+ compete with polymyxin molecules for carboxylate interaction sites on the cell wall.

A time series analysis of the example PMB1/K^+ proximal binding event is presented in Figure 4.5 and provides a more detailed insight into the nature of competition between the two molecules. The DAB residues of PMB1 formed a multitude of residue interactions with the various acidic residues of the cell wall, whilst the subsequent interaction of K^+ with the cell wall comprised fewer interactions with the same residue types. Interactions involving both K^+ and PMB1 were observed with the D-Glu4 and m-DAP5 residues of the cell wall. In both cases these residue interactions corresponded to coordination of K^+ with the cell wall carboxylate group and h-bond formation between a DAB residue of PMB1 and the amide linkage of the same cell wall residue. These data further emphasize how interactions between K^+ and the

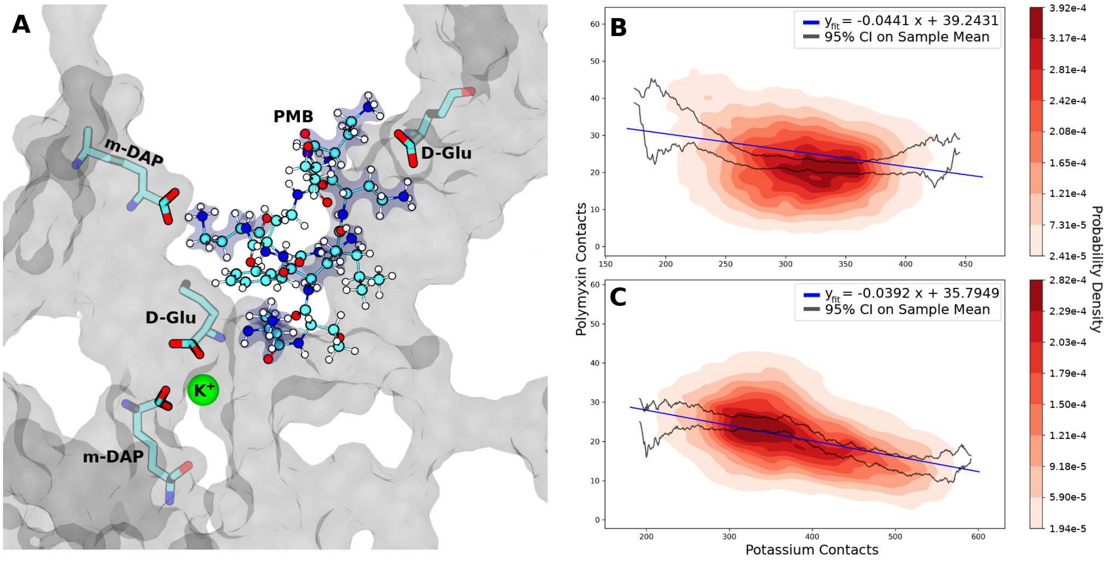


FIGURE 4.4: A) Potassium cation coordination with cell wall carboxylate groups in proximity of a PMB1 molecule. Representations are the same as in Figure 4.3. B & C) Linear regression and probability densities fitted to coincident contact counts for potassium ions and B) PMB1 or C) PME with all cell wall oxygens. Confidence intervals calculated using standard error of the mean. Data included from all three replicas of the concentrated Poly regime.

acidic cell wall residues occlude cell wall carboxylate groups from forming interactions with the nearby polymyxin molecule; leading to the direct competition between the two molecules for cell wall binding sites.

The binding of K^+ to the cell wall in proximity to PMB1, illustrated in Figure 4.4A, was seen to explicitly inhibit hydrogen bond formation between the proximal PMB1 molecule and the carboxylate groups that were coordinated with K^+ (Figure 4.6C). Prior to the binding of K^+ to the cell wall, PMB1 was bound in a compact conformation to the edge of a cell wall pore *via* 3 distinct hydrogen bonds; DAB4-DGlu16, DAB3-DGlu16 and DAB5-mDAP65 (Figure 4.6A). Approximately 3 ns after K^+ bound to the cell wall, disruption of the DAB3-DGlu16 interaction was observed (Figure 4.6B). The dissociation of this hydrogen bond resulted in increased mobility of the PMB1 molecule, highlighted by the subsequent brief dissociation of the DAB4-DGlu16 and DAB5-mDAP65 bonds. Approximately 6 ns after K^+ bound to the cell wall, the DAB4-DGlu16 and DAB5-mDAP65 bonds were permanently broken and replaced by the DAB5-DAla66 hydrogen bond (Figure 4.6D). The DAla66 residue of the cell wall was located further from the K^+ ion than the mDAP65 residue, and so the transition of the DAB5 interaction between these residues represents a movement of the PMB1 molecule away from the cell wall bound K^+. Dissociation of the DAB5-DAla66 interaction occurred 7 ns after the initial binding of K^+ with the cell wall (Figure 4.5A), with the complete dissociation of PMB1 from the cell wall occurring within a further 2 ns of simulation.

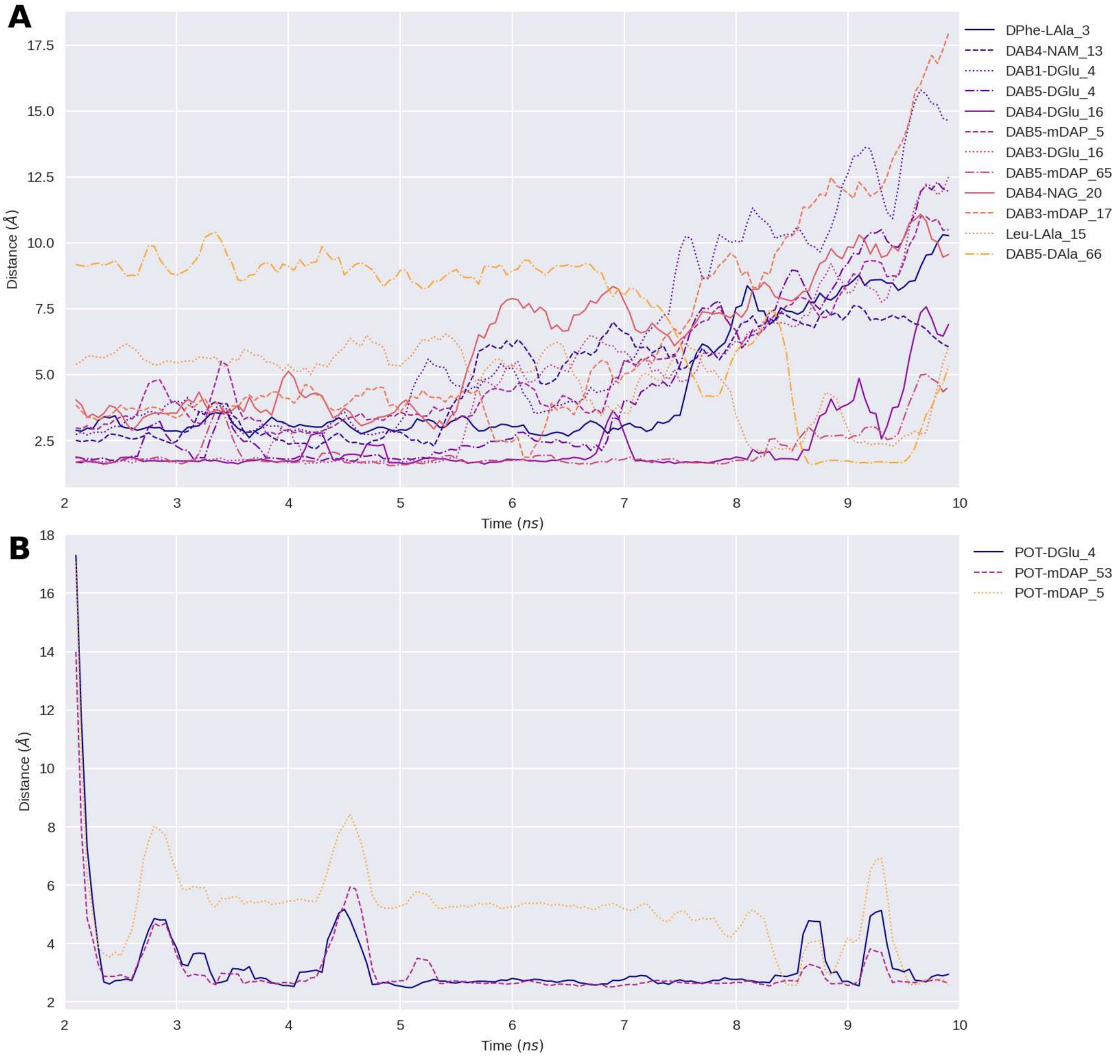


FIGURE 4.5: Minimum distances between interacting residue pairs of PGN and A) PMB1 or B) K^+ moieties observed in competition with each other for cell wall interaction sites (Figure 4.4). Data plotted for the time during which PMB1/potassium were within $\sim$ 3 nm of each other. Residue names on the left correspond to A) PMB1 / B) potassium, residue names on the right correspond to PGN.

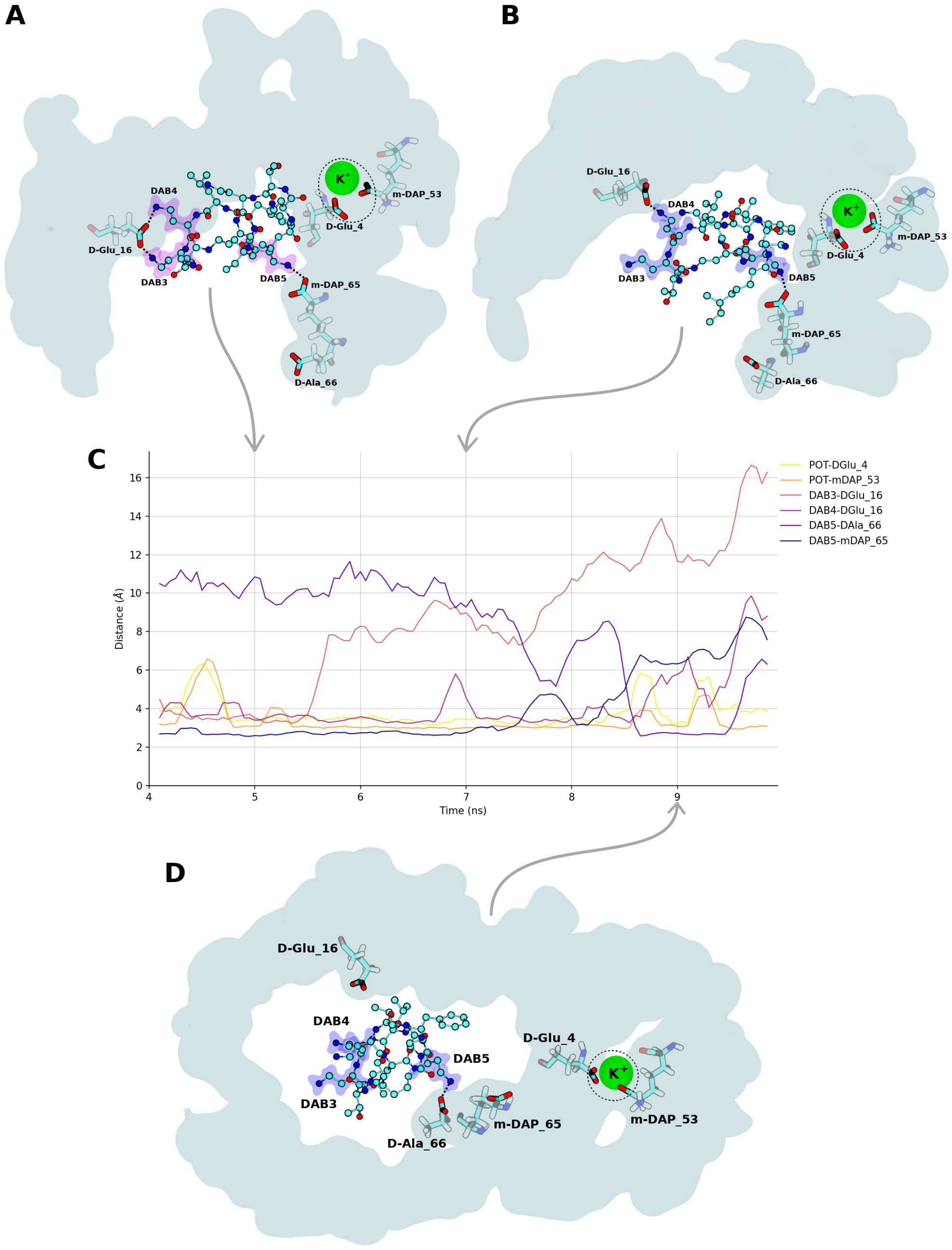


FIGURE 4.6: Time series of polymyxin-cell wall disruption resulting from proximal binding of K$^+$ ion during simulation of the concentrated Poly regime. A) Initial binding location of PMB1, 2.5 ns after initial contact between K$^+$ and the cell wall. B) First transition state of PMB1. C) Hydrogen bond distances between key residue interactions underpinning the coordination of PMB1 / K$^+$ with the cell wall. D) Final transition state of PMB1. PMB1 is represented by Goodsell CPK, K$^+$ in green vdW, interacting cell wall residues in transparent Goodsell licorice and the cell wall surface in solid cyan. Dotted lines highlight specific hydrogen bonds, dotted circles highlight K$^+$ - cell wall coordination. PMB1 hydrogens have been omitted for visual clarity.

Similar to K^+, spermidine was also observed to coordinate with the carboxylate groups on the cell wall peptide residues. This coordination was observed across all simulation regimes and occasionally occurred in close proximity to polymyxin molecules that were already bound to the cell wall (Figure 4.7A).

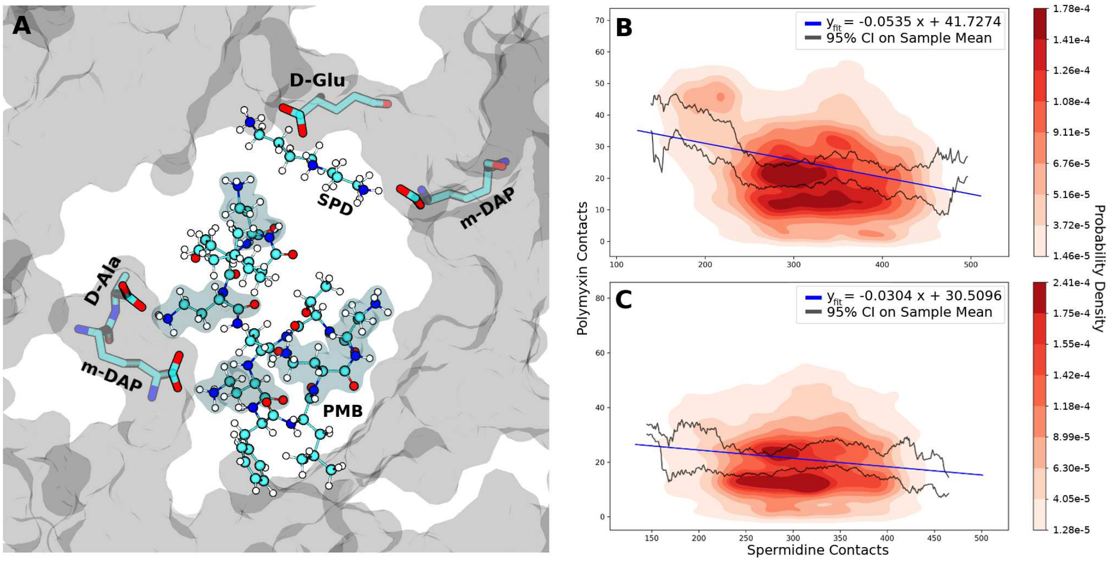


FIGURE 4.7: A) Spermidine (SPD) coordination with cell wall carboxylate groups in proximity of a PMB1 molecule. Representations are the same as in Figure 4.3. B & C) Linear regression and probability densities fitted to coincident contact counts for SPD molecules and B) PMB1 or C) PME with all cell wall oxygens. Confidence intervals calculated using standard error of the mean. Data included from all three replicas of the concentrated Osmo regime.

Given the generally linear structure of spermidine and its small size relative to the polymyxins, spermidine regularly inserted into the junction points of the cell wall pores inaccessible to polymyxin molecules (Figure 4.8). This behaviour, alongside the concentration of spermidine being much lower than other molecular components within the system, resulted in rare observations of direct competition between polymyxins and spermidine for carboxylate interaction sites. Despite this, within the concentrated Osmo regime, a negative linear correlation was calculated between the number of cell wall oxygen contacts with spermidine molecules and the mean number of coincident cell wall oxygen contacts with the DAB residues of both PMB1 (Figure 4.7B) and PME (Figure 4.7C).

The probability densities underlying this data do not provide such a clear visual indication of this trend, and the presence of density clusters perhaps indicate an unidentified variable not accounted for in the linear regression model. However, the existence of such a negative correlation in conjunction with explicit observations of spermidine binding to cell wall carboxylate groups in close proximity to polymyxin molecules (Figure 4.7A) supports the idea that, similar to K^+, spermidine competes with polymyxin molecules for carboxylate interaction sites on the cell wall.

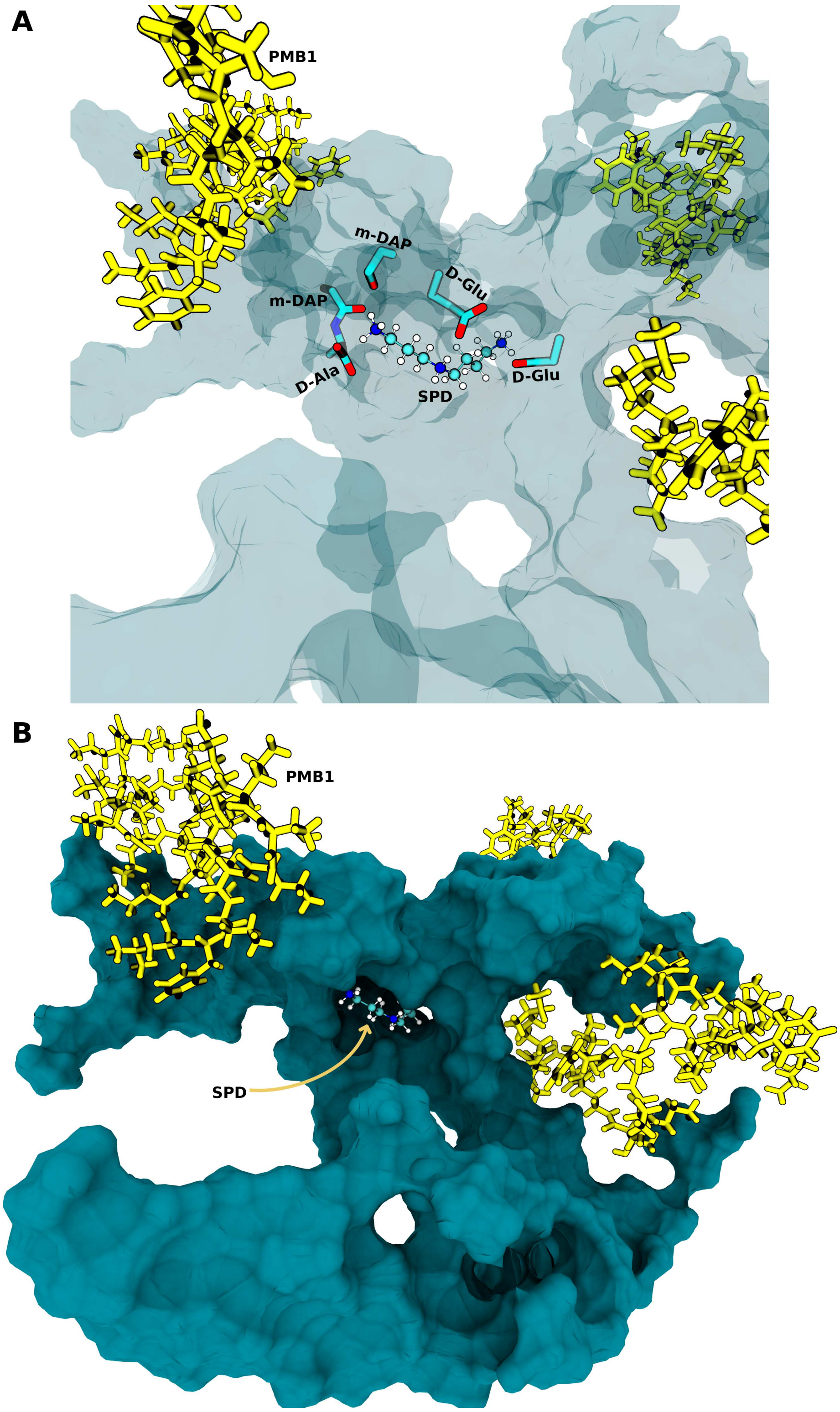


FIGURE 4.8: Insertion of spermidine (SPD) into clustered junction region of PGN cell wall, snapshots taken from the neutralised Osmo regime. A) Orthographic view with specific residues of PGN interacting with SPD highlighted in licorice. B) Perspective view highlighting the cell wall cavity into which SPD is inserted. The cell wall is represented by the transparent (A) or diffuse (B) cyan surface plot.

Time series analysis of the residue interactions between the cell wall and PMB1/spermidine (Figure 4.9) during their proximal binding highlighted that both PMB1 and spermidine formed a multitude of interactions with the various acidic residues of the cell wall.

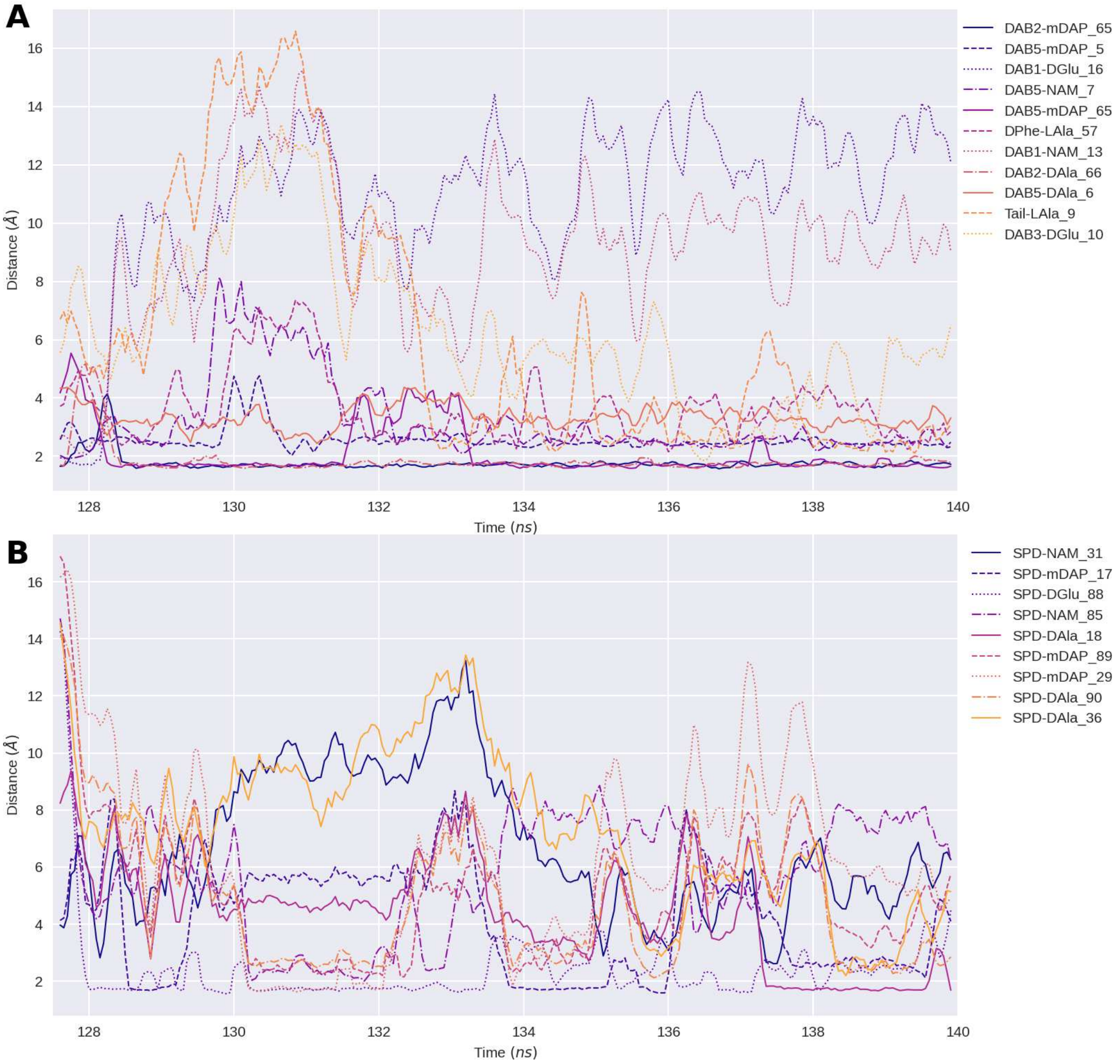


FIGURE 4.9: Minimum distances between interacting residue pairs of PGN and A) PMB1 or B) spermidine (SPD) observed in competition with each other for cell wall interaction sites (Figure 4.7). Data plotted for the time during which PMB1/SPD were within ∼ 3 nm of each other. Residue names on the left of legend labels correspond to A) PMB1 / B) SPD, residue names on the right correspond to PGN.

No specific cell wall residues were found to interact with both molecules throughout this period; indicating that PMB1 was unable to form residue interactions with cell wall residues that were concurrently engaged in interactions with spermidine and thus providing further support for the notion that PMB1 and spermidine compete for cell wall interaction sites.

In earlier analysis of the binding durations of polymyxins with the cell wall it was shown that the presence of crowding ubiquitin proteins resulted in shorter duration interactions between the cell wall and PMB1; leading to the abundant dissociation of PMB1 from the cell wall under neutralised conditions. Given these observations, and the nitrogen rich structure of ubiquitin, it was relevant to therefore ascertain whether ubiquitin may also compete with PMB1 for cell wall carboxylate interaction sites.

Similar to both potassium and spermidine, ubiquitin was observed to interact with the various carboxylate groups of the cell wall peptide residues (Figure 4.10A) predominantly *via* basic Lys and Arg residues. These interactions were observed under both neutralised and concentrated salt conditions. Due to the large size of ubiquitin and the relatively high concentration of the crowding protein within the system, the interactions regularly occurred in close proximity to polymyxin molecules.

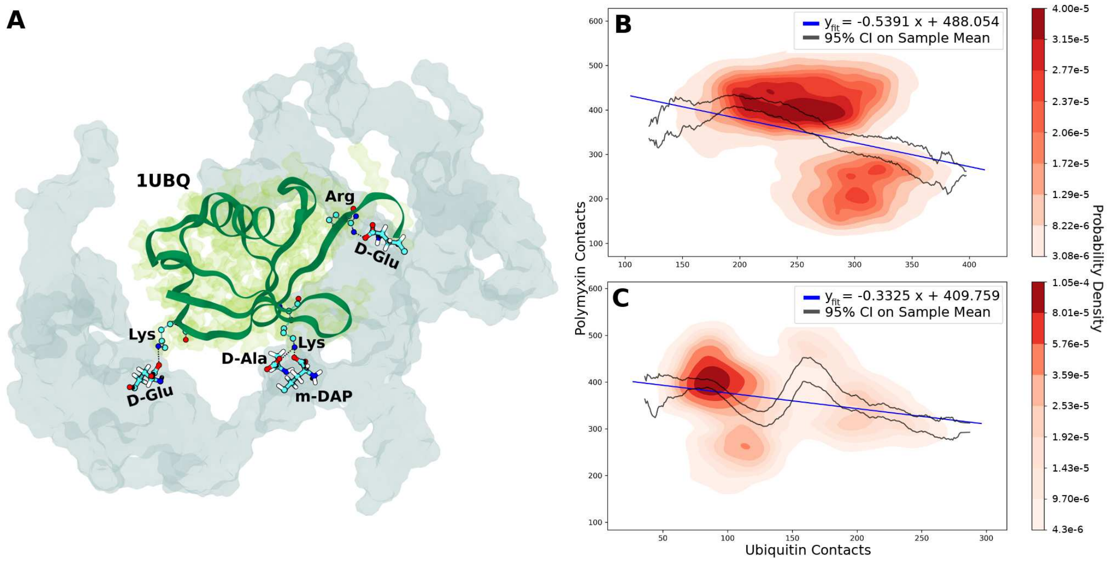


FIGURE 4.10: A) Ubiquitin coordination with cell wall carboxylate groups. Cell wall is represented in transparent cyan, ubiquitin is represented with green ribbons and a transparent green surface plot. B & C) Linear regression and probability densities fitted to coincident contact counts for ubiquitin molecules and PMB1 with all cell wall oxygens in the B) neutralised or C) 150 mM concentration regimes. Confidence intervals calculated using standard error of the mean. Data included from all three replicas of the Ubiq regime under each concentration.

Negative linear correlations were calculated between the number of cell wall oxygen contacts with ubiquitin molecules and the mean number of coincident cell wall oxygen contacts with the DAB residues of PMB1 under both neutralised (Figure 4.10B) and concentrated (Figure 4.10C) salt conditions. The probability densities underlying these two datasets show the formation of density clusters, more distinct than those observed in the spermidine data, likely indicating once more that there are unidentified variables not accounted for in the linear regression model. However, the overall negative correlation and explicit observation of ubiquitin binding to the carboxylate groups on the cell wall peptide residues likely indicate that the cationic and amine-rich residues of ubiquitin, similar to K^+ and spermidine, compete with polymyxin molecules for carboxylate interaction sites on the cell wall.

4.4 Discussion and conclusions

Here, simulations of an all-atom model of the *E. coli* cell envelope have been performed under various levels of biomolecular crowding to investigate the molecular interactions experienced by PMB1 and PME as a function of environmental complexity.

It has been shown that in the absence of a diverse chemical environment both PMB1 and PME tend to bind rapidly and irreversibly to the cell wall, predominantly *via* polar interactions between the cationic DAB residues of the polymyxins and the carboxylate groups on the meso-DAP, D-Glu and D-Ala residues of the cell wall peptide stems. Such interactions underpinned the binding of polymyxins to the cell wall across all simulations performed in this study, indicating that this behaviour was independent of environmental complexity.

The dependency of polymyxin – cell wall binding on this singular type of interaction presented an opportunity for other cationic moieties to disrupt the interactions between polymyxins and the cell wall. Indeed, it was found that the addition of physiological salt concentrations, osmolytes and increased biomolecular crowding all acted to decrease the duration of binding between the polymyxins and the cell wall. However, dissociation of polymyxins from the cell wall was only observed in simulations with physiological salt concentrations or with neutralising salt concentrations in the presence of crowding ubiquitin proteins. These observations may all be well explained through the lens of this cationic disruption.

Since Cl^- ions were sufficient to neutralise the overall charge of the model periplasm (in addition to the Ca^{2+} ions that neutralised and remained bound to LPS), it was only under excess salt concentrations that K^+ were included within the system. These freely diffusing cations were seen to coordinate with the same cell wall carboxylate groups to which the polymyxins preferentially bound; this coordination regularly occurred in close proximity to polymyxin molecules, resulting in direct competition for cell wall interaction sites. In this way, the K^+ ions disrupted the interactions responsible for polymyxin - cell wall binding, allowing for the dissociation of polymyxin molecules from the cell wall under all simulation regimes. In the most crowded simulation regime, in which polymyxin – cell wall dissociation was observed under neutralising salt concentrations, the addition of K^+ ions was again seen to decrease the duration of polymyxin – cell wall binding, providing further evidence that the presence of K^+ ions disrupts polymyxin – cell wall interactions under all environmental conditions.

These observations follow the established notion of the "salting-in" of proteins, whereby low (< 0.2 - 0.5 M) salt concentrations lead to an increase in protein solubility Arakawa and Timasheff (1982); Dumetz et al. (2007); Hassan (2005). This increase in solubility is attributed to ions 'coating' proteins in solution, screening the electrostatic interactions between neighbouring proteins and increasing the relative activity of nearby solvent molecules. With this context in mind, K^+ ions were seen to preferentially aggregate at the cell wall surface across all simulations performed in this study (Figure 4.11), thus 'coating' the cell wall. The resulting abundance of K^+ coordination with the peptide residues of the cell wall screened the electrostatic interactions of these cell wall residues with nearby polymyxin molecules. It is expected that this screening acted to decrease the strength of the electrostatic interaction between these polymyxin molecules and the cell wall, leading to a relative increase in the activity of nearby solvent molecules on the polymyxins, thus resulting in the observed increase in their solubility.

Similarly, upon the addition of osmolytes to the system, spermidine was seen to interact with the same carboxylate groups of the meso-DAP, D-Glu and D-Ala residues of the cell wall. Spermidine is a cationic polyamine and so it is of no surprise that it too would compete with the cationic DAB residues of the polymyxins for interaction sites on the cell wall. Indeed, similar to potassium, spermidine was observed in direct competition with neighbouring polymyxin molecules for cell wall interaction sites, however, the low concentration of spermidine within the periplasm inherently limited the sampling of such events. Despite this, the number of contacts between cell wall oxygen atoms and spermidine was seen to have a negative correlation with the number of contacts between cell wall oxygen atoms and the DAB residues of the polymyxins; indicating that despite limited sampling, the presence of spermidine still imposed a measurable disruption to the binding of polymyxin molecules to the cell wall.

Finally, the ubiquitin proteins present in the most crowded models of the periplasm were also seen to interact with the cell wall *via* polar interactions with the carboxylate groups on the meso-DAP, D-Glu and D-Ala residues of the cell wall peptide stems. Ubiquitin is a large, nitrogen rich molecule with multiple cationic lysine and amine-rich arginine residues on its surface and it is these residues in particular that were seen to coordinate with the carboxylate groups on the cell wall. It was only in the presence of these ubiquitin proteins that polymyxins were seen to dissociate from the cell wall under neutralising salt concentrations, and this result may also be understood in the context of cationic disruption.

Within the neutralised Poly regime, the polymyxins were the only freely diffusing cationic compounds within the periplasm and thus had no direct competition for cell wall interaction sites, concurrent with no observations of polymyxins dissociating from the cell wall under these conditions. In simulations of the neutralised Osmo regime, only a single cationic spermidine molecule was included in the system; thus competition between spermidine and polymyxins for cell wall interaction sites was inherently limited and was again concurrent with no observations of polymyxins

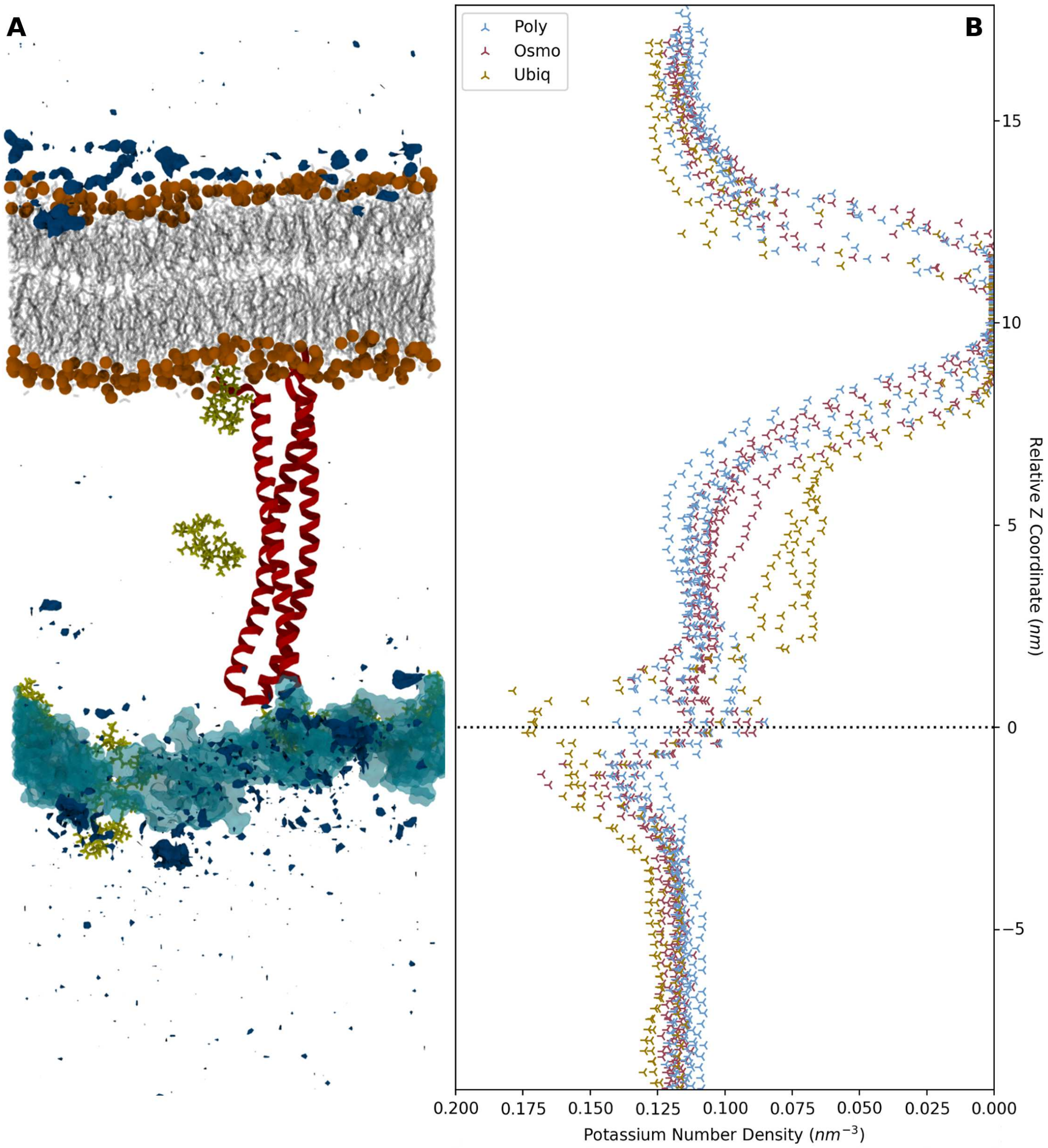


FIGURE 4.11: A) Representative average volume map of potassium ions (dark blue) calculated across one replica of the concentrated Poly regime. PMB1 represented in yellow, BLP in red ribbons, PGN in transparent cyan, OM headgroup phosphates in orange vdW. B) Partial z-densities of potassium ions across all simulations. Z coordinates are measured relative to the cell wall CoM.

dissociating from the cell wall. In contrast to this, the neutralised Ubiq regime included 11 nitrogen-rich ubiquitin proteins, each with numerous basic lysine and arginine residues on its surface. The tendency for these residues to interact with the carboxylate groups on the cell wall peptide stems, along with the large excluded volume effects resulting from multiple such proteins binding to the cell wall simultaneously, represented the greatest competition for cell wall interaction sites faced by the polymyxins in any of the neutralised simulation regimes. These behaviours thus limited the ability of polymyxins to form long-lasting interactions with the cell wall and led to the abundant dissociation of polymyxin molecules from the cell wall.

A recent experimental study Tran et al. (2023) highlighted that the diffusion of a protein, OsmY, throughout the *E. coli* periplasm was best described by a two-component random walk model, comprising one fast and one slow diffusive component. The presence of the putative slow diffusion component implied that a fraction of OsmY proteins interacted with the various supramolecular structures within the periplasm, limiting their diffusion rate. In contrast, the fast diffusion component described those proteins that were freely diffusing throughout the periplasm. The diffusion coefficient of the fast diffusion component was found to increase with periplasmic volume, implying that the free diffusion rate of proteins within the periplasm was negatively correlated with the extent of macromolecular crowding within the surrounding environment.

Our observations of sustained complex formation between the polymyxins and the cell wall support the notion that interactions with the various surfaces of the *E. coli* periplasm limit the diffusion rates of proteins within this environment. We propose, however, that the extent to which such interactions disrupt the diffusion of the polymyxins is modulated by the presence of cations, such as K^+, within the environment. We also note here that this modulation, whereby K^+ enable the free diffusion of polymyxins throughout the periplasm, likely enables the polymyxins to more readily encounter lipoprotein carriers such as LolA, which have been previously proposed to provide polymyxins with a potential passive transport mechanism within the periplasm Pedebos et al. (2021).

Overall, it is clear that the nature of polymyxin - cell wall interactions are fundamentally altered by the composition of the surrounding periplasmic environment. Whether it be a single ion, small osmolyte or large protein; the work presented in this chapter has shown that the presence of cationic moieties, of any size, disrupts the interactions of polymyxin lipopeptides with the cell wall, decreasing the duration of their binding. There are many more molecular species present in the *in vivo* periplasm than it was feasible to include within even the most complex model presented in this work, and the impacts that each may have on molecular interactions within this space are not easily quantified nor predicted prior to simulation.

Therefore, whilst this work begins to shed light on the nature of polymyxin - cell wall interaction within the periplasm, it is likely that deriving a biologically accurate understanding of these systems will require considerable further work to bring the composition of future *in silico* simulation studies closer to the ’true’ complexity of the biological environment.

Chapter 5

The nature of Polymyxin interactions with Braun's Lipoprotein

5.1 Introduction

In 1969, Braun and Rehn proved the existence of a protein that was both covalently bound to the peptidoglycan cell wall and inserted into the outer membrane *via* an integral lipid tail Braun and Rehn (1969). This discovery represented the first identified member of a large family of bacterial envelope proteins, the lipoproteins, that are attached to a membrane by a lipid moiety. In 1972, the sequence of that protein was reported, revealing the presence of a C-terminal lysine residue responsible for the covalent binding with the cell wall Braun and Bosch (1972). One year later, the protein was shown to be anchored to the OM *via* three acyl groups that were covalently attached to an N-terminal cysteine residue Hantke and Braun (1973). With the sequence known, and binding characteristics determined, this protein became known as 'the lipoprotein' and, subsequently, Braun's lipoprotein (BLP) as other members of the lipoprotein family were identified. Despite the discovery of other such lipoproteins, however, BLP is known to be the only molecule that provides a covalent attachment between the OM and cell wall.

BLP has since been found in a wide variety of Gram-negative organisms Braun et al. (1970); Halegoua et al. (1974) and is one of the most abundantly expressed proteins in the bacterial cell envelope; with *E. coli* containing an estimated 3×10^5 copies per cell Braun (1975). Such large copy numbers provide an indication of the importance of BLP to the proper function of the bacterial cell and indeed, as the only molecule to provide a covalent attachment between the OM and cell wall, BLP is vital to maintaining the structural integrity of the cell Hoekstra et al. (1976); Suzuki et al.

(1978); Schwechheimer et al. (2014). Previous work performed by the Khalid research group has further highlighted the role that BLP plays in maintaining the separation between the OM and cell wall and how, through tilting and kinking, BLP can act to modulate the distance between the two; facilitating the non-covalent interaction of other membrane bound proteins with the cell wall Samsudin et al. (2017).

As small molecules, such as antibiotics, move throughout the periplasm of Gram-negative bacteria; they inevitably pass through the immediate vicinity of one or more of the abundant BLP molecules present within this space. This thesis focuses on one particular class of antibiotics, the polymyxins, that induce cell lysis through action on the inner membrane. It is evident, therefore, that investigation of the biochemical nature of molecular interactions involving BLP are necessary to further our understanding of the transport of these molecules across the crowded periplasmic region from their point of entry into the cell, the OM, to their point of antimicrobial action, the IM.

A previous paper published by the Khalid research group, on which I am a co-author, highlighted the promiscuity of polymyxin interactions within the crowded periplasm of *E. coli* using the united-atom GROMOS54a7 forcefield Pedebos et al. (2021). In this study, PMB1 was seen to interact profusely with the various components of the periplasm, including BLP. In order to better understand how such interactions may impact the translocation of polymyxins from the OM to the IM, it is prescient to investigate the comparative interactions of PMB1 and PME with BLP within models of the periplasm crowded to varying extents.

To this end, the work presented in this chapter further analyses the simulations discussed in Chapter 4, focussing on interactions between the polymyxins and BLP. The specific residue interactions that give rise to the binding of the polymyxins to BLP are shown to differ between PMB1 and PME and, further, the balance of hydrophobic and polar residue interactions that underpin their binding is shown to be affected by variations in the chemical diversity of the simulation environment. Despite the deviations observed between the simulated regimes and polymyxin species; this chapter highlights how certain groups of residues in both PMB1 and PME are repeatedly found to play critical roles in binding with BLP. In particular, the positionally analagous Leu/D-Phe/DAB5 and Leu1/Leu2/DAB5 residue triads of PMB1 and PME, respectively, as well as the polar DAB1/Thr1/DAB2 triad of the polymyxin branched fatty acid tail are shown to be of repeated importance. The preferred binding modes of PMB1 and PME with BLP in each simulation regime are characterised and discussion is provided on how interactions with a certain type of osmolyte, osmoregulated periplasmic glucan (OPG), may contribute to the varying biochemical nature of polymyxin binding with BLP.

5.2 Methods

The construction and preparation of the systems studied in this work is discussed in detail in Section 4.2. A summary of all simulations performed is provided in Table 4.1.

5.2.1 Simulation protocols

Simulations were performed using the GROMACS 2021.2 molecular dynamics package Abraham et al. (2015), utilising the CHARMM36m force field Huang et al. (2017) and TIP3 water model Jorgensen et al. (1983). Simulations were divided into two parts: equilibration simulations in NVT and NPT ensembles lasting for 200 ps and 40 ns respectively; and production simulations in the NPT ensemble, which ran for 250 ns. A constant temperature of 310 K was maintained using the velocity rescale thermostat Bussi et al. (2007) with a time constant of 1 ps. The pressure was maintained anisotropically at 1 atm using the Parrinello-Rahman barostat Parrinello and Rahman (1981) with a time constant of 1 ps. Hydrogen bonds were constrained using the LINCS algorithm Hess et al. (1997); Hess (2008); stable treatment of these constraints required the use of a 1 fs integration time step. Long-range electrostatics were treated using the particle mesh Ewald method Darden et al. (1993). The short-range electrostatic and van der Waals cut-offs were both set to 1.2 nm.

For the replicates; new initial configurations of all polymyxin, osmolyte and ubiquitin molecules were generated, along with re-solvation and ionisation of each system before being passed through the equilibration and production simulation phases. The initial velocities of all atoms were modified between each replicate at the start of NVT equilibration to ensure an unbiased sampling of the simulation phase space.

Analyses were performed with scripts written using MDAnalysis Gowers et al. (2016); Michaud-Agrawal et al. (2011), Gromacs utilities and VMD. Kernel density estimate (KDE) curves were calculated using the Seaborn python package Waskom (2021).

5.3 Results

Similar to the analysis of polymyxin - cell wall interactions presented in Chapter 4, the analysis of polymyxin interactions with BLP was split into two components: the duration of binding between the polymyxins and BLP, and the biochemical nature of these interactions. Kernel density estimates (KDE) were fitted to the observed binding durations across all replicates of each system and are presented in Figure 5.1C. The specific residue interactions were categorised according to interaction type (*i.e.* involving either hydrophobic, DAB or Thr residues of the polymyxins) and the results

Residues		Simulation Regime									
BLP	PMB	poly-neut		poly-conc		osmo-conc		ubiq-neut		ubiq-conc	
ASP	DAB1	21.18	2.82	14.84	8.1	13.32	2.25	11.05	9.07	27.2	17.52
THR	DAB2	2.82	12.94	12.1	5.77	6.66	3.32	1.98	8.95	6.66	3.03
GLN	DAB3	18.35	11.29	21.34	13.22	6.89	1.56	22.78	4.35	8.74	1.71
SER	DAB4	1.41	5.88	10.29	5.66	3.69	1.52	8.14	0.13	3.67	9.32
ALA	DAB5	28.24	1.41	18.84	9.55	14.72	22.6	8.62	19.4	16.09	10.12
LYS	Thr1	0	4	7.31	3.68	9.91	3.22	9.55	20.12	8.29	10.02
ASN	Thr2	2.12	8.94	5.39	13.1	8.43	0.81	18.08	0.38	4.7	0.28
ARG	Acyl Tail	25.88	0	6.65	15.15	20.08	22.97	8.8	13.12	22.7	28.03
LEU	D-Phe	0	26.82	1.51	15.76	0.42	31.31	5.77	15.26	0.28	6.67
VAL	Leu	0	25.88	0.59	10.02	8.59	10.45	0.49	9.23	1.66	13.29
ILE		0		0.71		0		4.69		0	
MET		0		0.42		7.3		0.04		0	
TYR		0		0		0		0		0	
🟨	hydrophobic	52.71		40.92		64.73		37.61		48	
🟦	DAB	34.35		42.3		31.24		41.9		41.7	
🟥	Thr	12.94		16.78		4.03		20.5		10.3	

TABLE 5.1: Percentage contribution of each residue of BLP/PMB1 to the total number of observed residue interactions between PMB1 and BLP.

Residues		Simulation Regime							
BLP	PME	poly-conc		osmo-conc		ubiq-neut		ubiq-conc	
ASP	DAB1	21.07	15.74	17.57	5.68	25.92	9.25	11.01	12.65
THR	DAB2	3.34	7.35	6.59	26.28	24	0.79	27.27	6.71
GLN	DAB3	18.63	3.22	20.87	2.93	9.51	46.77	14.24	5.12
SER	DAB4	4.75	4.63	13.5	2.07	0	0	7.41	4.9
ALA	DAB5	21.03	9.41	9.19	22.87	0	0	17.37	3.59
LYS	Thr1	17.36	17.9	3.42	18.52	0	37.26	11.09	15.99
ASN	Thr2	5.8	0.99	25.2	1.59	11.61	3.75	3.51	1.5
ARG	Acyl Tail	8	25.28	2.83	5.86	28.97	2.18	7.93	30.38
LEU	Leu1	0.01	3.81	0.49	8.26	0	0	0.12	7.38
VAL	Leu2	0.01	11.66	0.07	5.96	0	0	0.02	11.78
ILE		0		0.26		0		0	
MET		0		0		0		0	
TYR		0		0.01		0		0.03	
🟨	hydrophobic	40.75		20.07		2.18		49.54	
🟦	DAB	40.36		59.82		56.81		32.97	
🟥	Thr	18.89		20.11		41.01		17.49	

TABLE 5.2: Percentage contribution of each residue of BLP/PME to the total number of observed residue interactions between PME and BLP.

for each polymyxin in the concentrated Osmo regime are presented in Figures 5.1D & 5.1E. The complete dataset for all simulation regimes are provided in Tables 5.1 and 5.2.

5.3.1 PMB1

Only one PMB1 molecule was seen to interact with BLP across all simulations of the neutralised Poly and Osmo crowding regimes. This interaction was observed in the Poly regime, where one PMB1 molecule was bound to BLP at the onset of production MD. This molecule immediately began moving along the surface of BLP, making contact with the cell wall after 9.15 ns and subsequently dissociating from BLP (Figure 5.2); the molecule then remained bound to the cell wall for the remaining duration of

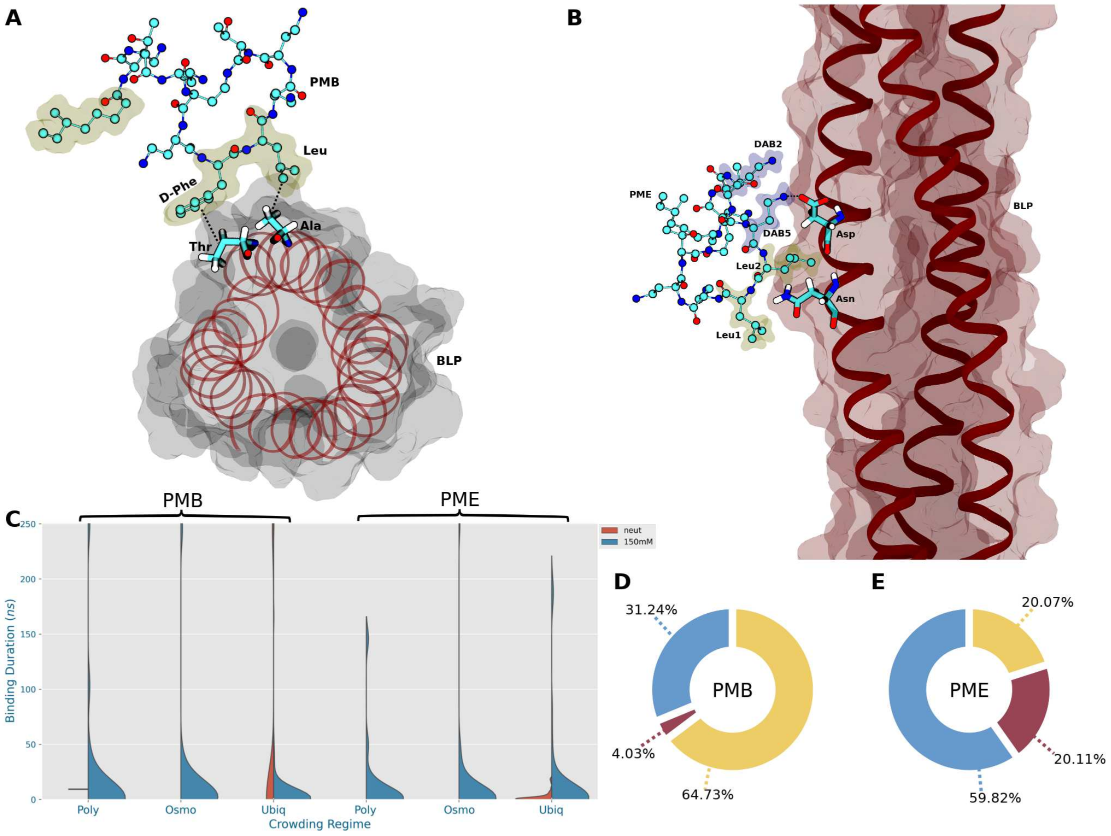


FIGURE 5.1: A) PMB1 in the hydrophobic binding mode with BLP. Hydrophobic residues of PMB1 are highlighted with yellow bubbles. B) PME in the polar binding mode with BLP. DAB residues of PME are highlighted with blue bubbles. Hydrogens have been omitted for visual clarity. C) Kernel Density Estimate (KDE) curves fitted to the binding durations of all unique instances of binding between polymyxin molecules and BLP. D/E) Pie charts of residue interaction types between BLP and PMB1 (D) / PME (E) in the concentrated Osmo regime. Interactions involving the DAB / Thr / hydrophobic residues of polymyxins are coloured in blue / red / yellow respectively.

simulation. This 'walking' behaviour of PMB1 along BLP has been reported in previous studies of this system using the GROMOS54a7 forcefield Pedebos et al. (2021).

Initial binding of the PMB1 walker to BLP was characterised by hydrophobic interactions between the acyl tail, D-Phe and Leu residues of PMB1 and Asn, Ala and Arg residues of BLP, respectively. The acyl tail of PMB1 dissociated from the Asn residue after 3.3 ns of simulation. Dissociation of the D-Phe/Ala residues occurred after a further 2.15 ns and was quickly (< 0.35 ns) followed by dissociation of the final hydrophobic Leu/Arg residue interaction. Simultaneous to the dissociation of the Leu/Arg residues, the cationic DAB3 residue of PMB1 came into contact with an anionic Asp residue of BLP; to which it remained bound for a further 2.2 ns. Upon dissociation of the DAB3/Asp residues, the DAB2 residue of PMB1 came into contact with a second Asp residue of BLP located further down the length of BLP, closer to the cell wall. This DAB2/Asp interaction remained for a further 1.25 ns, during which

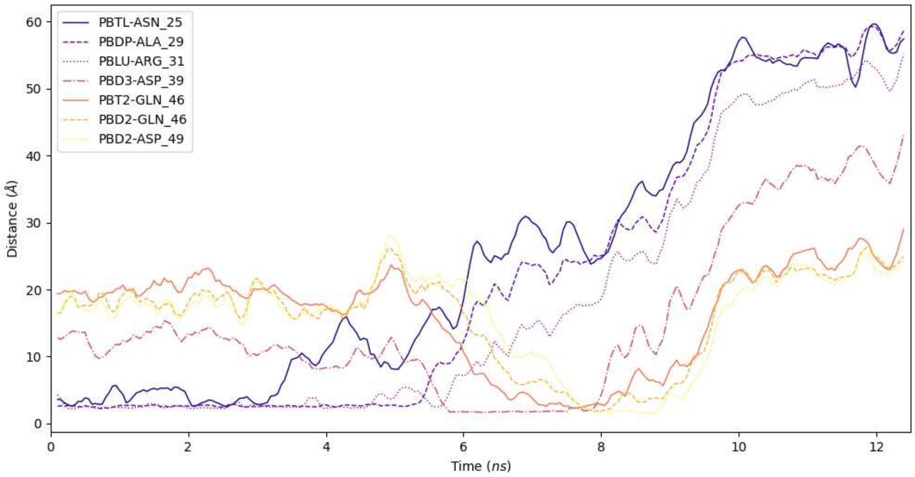


FIGURE 5.2: Minimum distances between contacting residues of PMB1 and BLP during the 'walking' of PMB1 down BLP in the neutralised Poly regime.

period the DAB2 and Thr2 residues interacted transiently with a nearby Gln residue of BLP. The PMB1 molecule made contact with the cell wall during the final 0.1 ns of the DAB2/Asp interaction. After dissociation of the final DAB2/Asp residue interaction between PMB1 and BLP, the PMB1 molecule interacted solely with the cell wall and was not observed to interact with BLP for the remaining duration of the simulation.

Interactions between PMB1 and BLP were prevalent across all concentrated crowding regimes; with 26 unique PMB1-BLP interactions observed in both the concentrated Poly and Osmo regimes and 110 in the concentrated Ubiq regime. There was a strong preference for short duration interactions in all cases, with median binding durations of 0.25 ns, 1.00 ns and 0.90 ns for the Poly, Osmo and Ubiq regimes respectively. In contrast to this trend, one replicate of each regime contained a single PMB1 molecule that remained bound to BLP for the entire duration of simulation.

One PMB1 molecule in each of the three replicate simulations of the neutralised Ubiq crowding regime was seen to remain bound to BLP for the entire duration of simulation. Indeed, other than the single PMB1 molecule in the neutralised Poly regime, it was only in these most crowded simulations that PMB1 molecules were seen to interact with BLP under neutralised conditions; with 21 unique instances of PMB1-BLP binding observed across all replicate simulations of this regime. Three of these instances correspond to the three aforementioned PMB1 molecules that remained bound to BLP for the entire duration of simulation; the remaining 19 instances, however, correspond to the repeated association/dissociation of a single PMB1 molecule to BLP throughout the duration of one replicate trajectory. It is evident therefore that, in the Ubiq regime, PMB1 exhibits a greater preference for longer duration interactions with BLP when simulated under lower salt concentrations.

It was previously reported that the Ser and acidic residues of BLP had a particular propensity to interact with PMB1 due to their ability to form hydrogen bonds and salt bridges with the cationic DAB residues of PMB1 Pedebos et al. (2021). Across all of the simulations involving PMB1, 39.80 % of all observed residue interactions between PMB1 and BLP were formed between the cationic DAB residues of PMB1 and the various polar residues of BLP. Interactions between hydrophobic residues of PMB1 and BLP accounted for 45.43 % of observed residue interactions. The remaining 14.77 % of residue interactions were formed between the Thr residues of PMB1 and the polar residues of BLP; equivalent to 27.07 % of the observed polar residue interactions. The Ser residues of BLP were seen to interact with both the DAB and Thr residues of PMB1, however, such interactions accounted for only 7.40 % of the total number of residue interactions; a lower percentage than was observed for the Gln (17.47 %), Asp (14.55 %), Ala (13.99 %), Arg (12.02 %), Asn (10.54 %) and Lys (8.75 %) residues of BLP. It is evident, therefore, that whilst interactions with the acidic residues of BLP are important to the binding of PMB1, the complete picture of their binding is a balance of both electrostatic and hydrophobic interactions, involving many different residues. These aggregated results, however, do not pick up on the nuanced effects that the complexity of each system had upon the distribution of observed residue interactions.

In the concentrated Poly regime, polar interactions involving the DAB or Thr residues of PMB1 were the dominant component in PMB1-BLP binding, making up 59.08 % of all observed residue interactions; the DAB residues were of particular importance, accounting for 42.3 %. The contribution of PMB1 hydrophobic residues to BLP binding was still significant; with the D-Phe, acyl tail and Leu residues being responsible for, respectively, 15.76 %, 15.15 % and 10.02 % of all observed residue interactions. Considering the results in context of the location of each residue in the structure of PMB1, it is found that 67.30 % of interactions involved residues residing on the heptapeptide ring of PMB1, with interactions involving the branched fatty acid tail making up the remaining 32.70 %. Of those residues that reside on the heptapeptide ring of PMB1; the hydrophobic D-Phe and Leu residues were responsible for a combined 25.78 % of residue interactions, the three DAB residues were responsible for 28.43 % and the lone Thr residue was responsible for 13.1 %. These data imply that, within this regime, PMB1 preferentially binds to BLP via its heptapeptide ring, utilising a variety of polar and hydrophobic interactions (Figure 5.3).

In the concentrated Osmo regime, hydrophobic interactions involving the acyl tail, Leu and D-Phe residues of PMB1 dominated the binding with BLP (Figure 5.4), accounting for 64.73 % of observed residue interactions; 23.81 % greater than was observed in the concentrated Poly regime. The acyl tail and Leu residues of PMB1 were responsible for, respectively, 22.97 % and 10.45 % of all observed residue interactions in the concentrated Osmo regime. The D-Phe residue was the largest single contributor, accounting for 31.31 % of all interactions. Notably, the DAB5

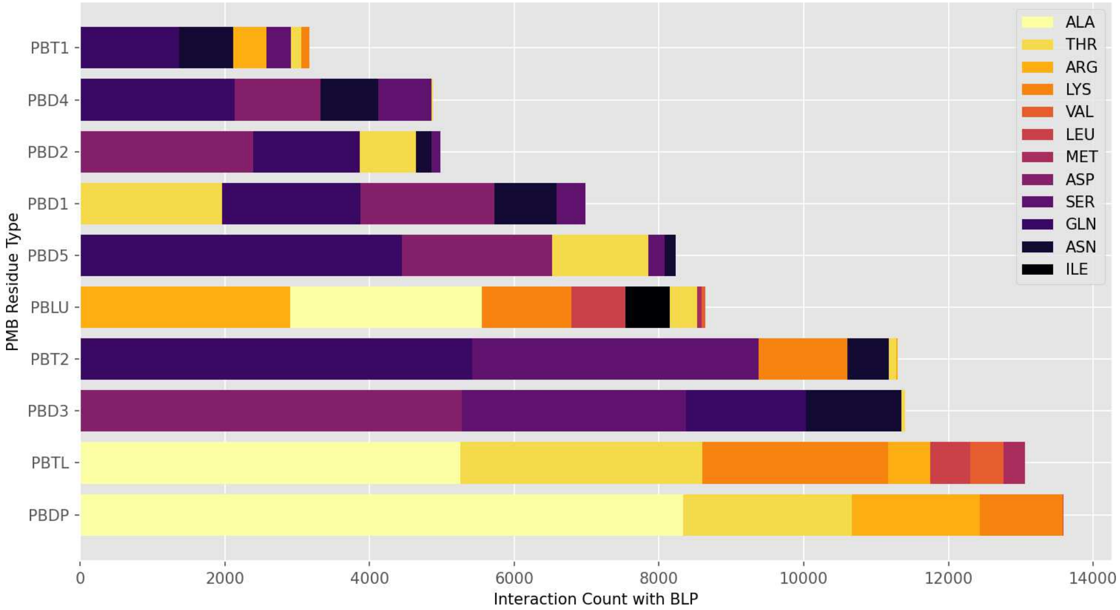


FIGURE 5.3: Residue interactions between PMB1 and BLP in the concentrated Poly regime. Amino acids in the legend refer to residues of BLP.

residue alone accounted for 22.6 % of the remaining residue interactions, with the other four DAB and two Thr residues cumulatively accounting for the final 12.68 %. This result is of particular interest given that the D-Phe residue neighbours both the Leu and DAB5 residues within the heptapeptide ring of PMB1; this triad of residues therefore represents a combined 64.36 % of all residue interactions, indicating that this specific region of PMB1 plays a crucial role in its binding to BLP. Thus it is found that the binding between BLP and PMB1 within this regime is in fact underpinned primarily (87.33 %) by interactions with the acyl tail and the Leu/D-Phe/DAB5 triad of PMB1 (Figure 5.5).

It is interesting to note here that, in the concentrated Poly regime, the broad involvement of all PMB1 residues in the binding with BLP may have masked the relative importance of the Leu/D-Phe/DAB5 triad. These residues, however, were responsible for a combined 35.33 % of residue interactions in the concentrated Poly regime. The total number of residue interactions observed in the concentrated Osmo regime was much lower (55879) than that of the concentrated Poly regime (86218); with interactions involving the polar residues of PMB1 (excluding DAB5) exhibiting the greatest relative decreases in interaction counts (Figures 5.3 & 5.4). Furthermore, in contrast to the overall decrease in the total number of residue interactions; the DAB5 and D-Phe residues of PMB1 were involved in, respectively, 4397 and 3909 more interactions in the concentrated Osmo regime compared to the concentrated Poly regime. These data indicate that whilst the Leu/D-Phe/DAB5 triad still played a considerable role in the binding of PMB1 to BLP in the concentrated Poly regime; the addition of osmolytes to the system disrupted the interaction of all other polar PMB1

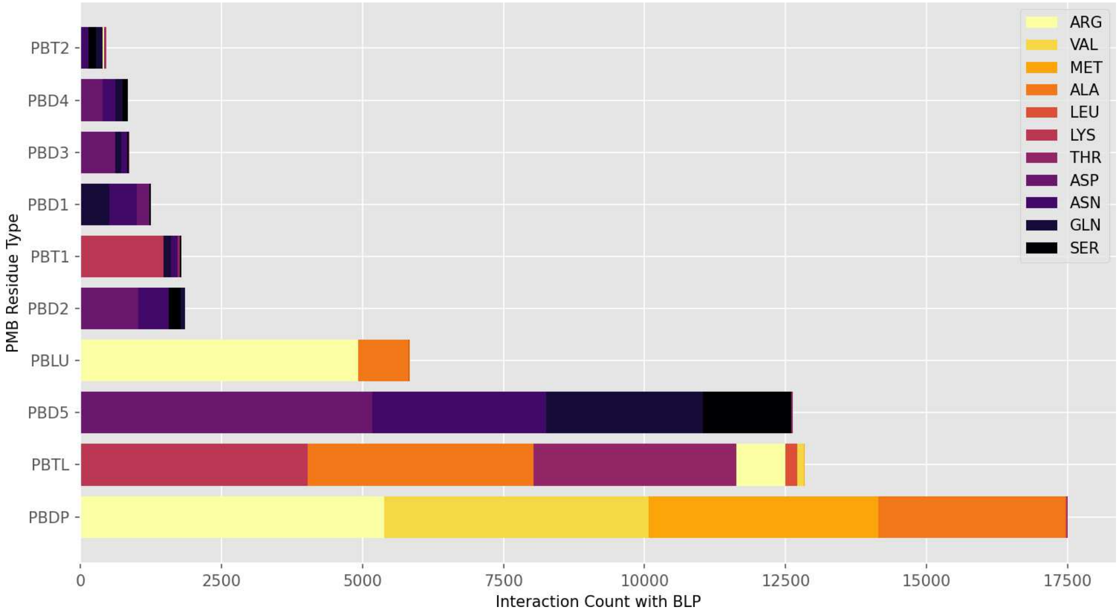


FIGURE 5.4: Residue interactions between PMB1 and BLP in the concentrated Osmo regime. Amino acids in the legend refer to residues of BLP.

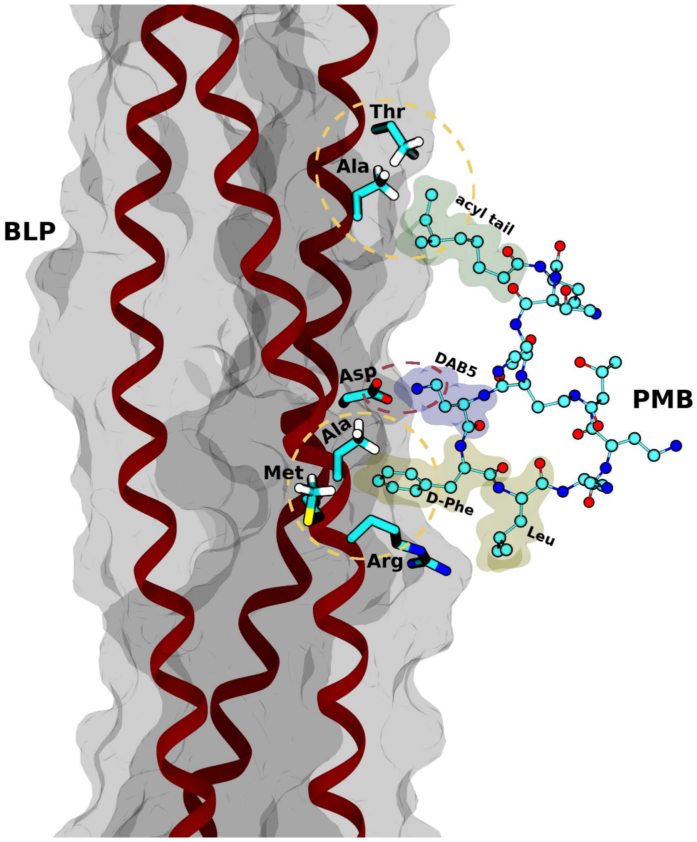


FIGURE 5.5: PMB1 bound to BLP in the concentrated Osmo regime *via* interactions involving the Leu/D-Phe/DAB5 triad and acyl tail of PMB1 with various residues of BLP. Yellow dashed circles highlight clustering of hydrophobic residues, red dashed circle highlights salt bridge formation between charged DAB5/Asp residues of PMB1/BLP.

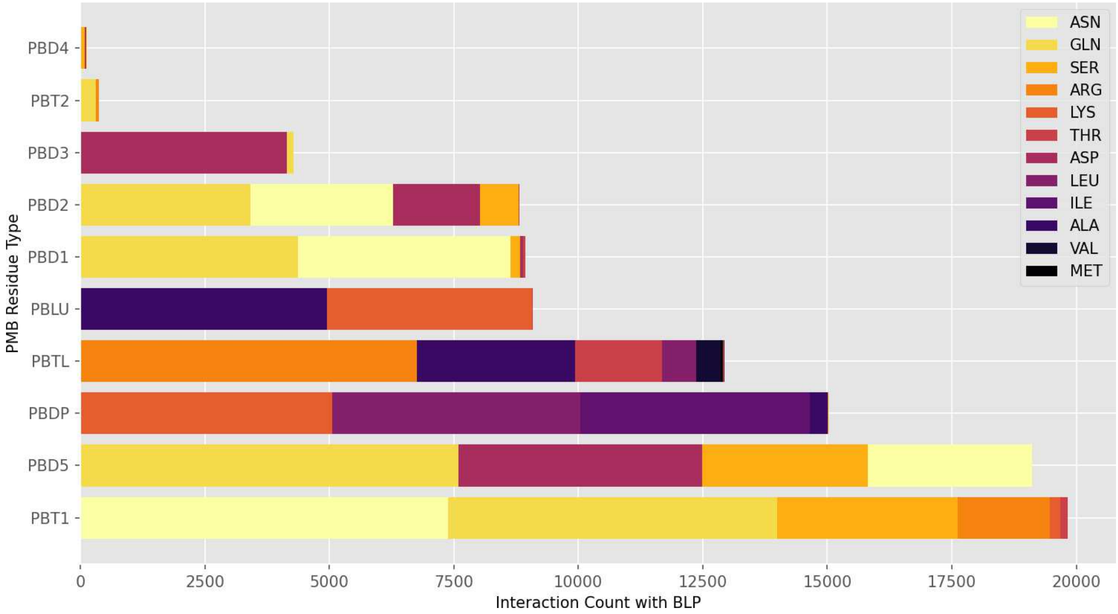


FIGURE 5.6: Residue interactions between PMB1 and BLP in the neutralised Ubiq regime. Amino acids in the legend refer to residues of BLP.

residues with BLP, exacerbating the contributions of the PMB1 Leu/D-Phe/DAB5 triad and acyl tail.

Across all replicate simulations of the neutralised Ubiq regime, hydrophobic interactions accounted for 37.61 % of observed residue interactions, with the D-Phe, acyl tail and Leu residues contributing, respectively, 15.26 %, 13.12 % and 9.23 %. The various charged and polar residues accounted for the remaining 62.39 % of observed residue interactions, with the largest contributions coming from the Thr1 (20.12 %), DAB5 (19.4 %), DAB1 (9.07 %) and DAB2 (8.95 %) residues. The residues of BLP with the greatest involvement in PMB1 binding in this regime were Gln (22.78 %), Asn (18.08 %) and Asp (11.05 %), involved in a combined 51.91 % of observed residue interactions. All three of these BLP residues were observed to interact solely with the charged/polar residues of PMB1, highlighting the importance of such electrostatic interactions in the binding of PMB1 to BLP in this regime (Figure 5.6).

Returning to the contribution of PMB1 residues; the Leu/D-Phe/DAB5 residue triad was involved in a combined 43.89 % of residue interactions, whilst the polar DAB1/Thr1/DAB2 residues of the branched fatty acid tail combined were involved in 38.14 %. These two regions were therefore responsible for 82.03 % of residue interactions between PMB1 and BLP; indicating that these regions of PMB1 play a crucial role in PMB1-BLP binding under these simulation conditions.

Across all replicate simulations of the concentrated Ubiq regime (Figure 5.7), hydrophobic interactions were responsible for 48 % of observed residue interactions. Whilst this is a lower contribution than was observed in the concentrated Osmo regime (64.73 %), it is considerably higher than the result obtained from the

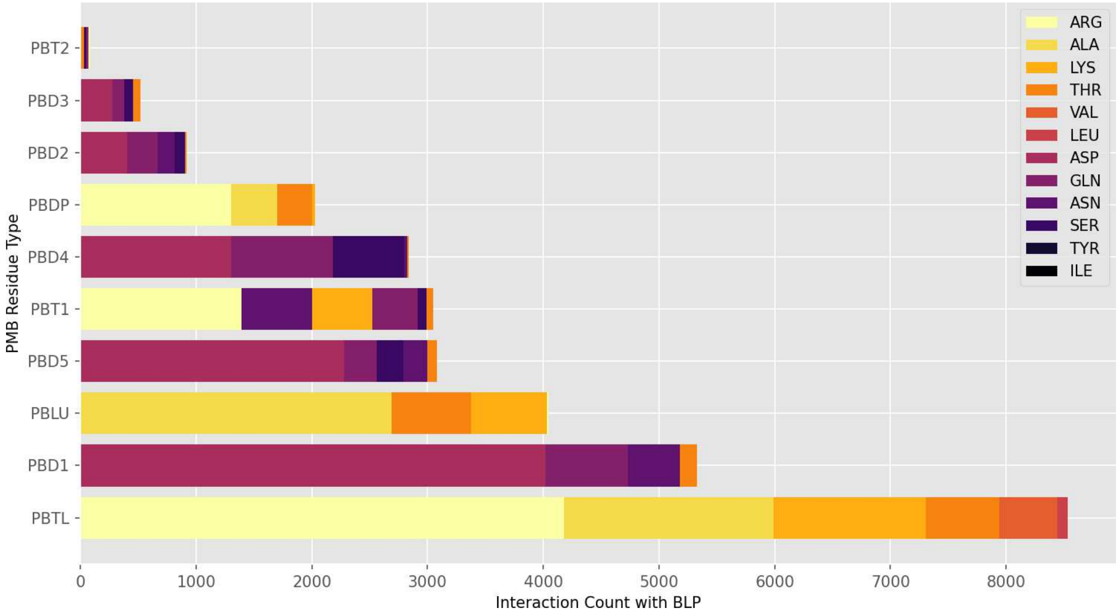


FIGURE 5.7: Residue interactions between PMB1 and BLP in the concentrated Ubiq regime. Amino acids in the legend refer to residues of BLP.

neutralised Ubiq regime (37.61 %); indicating that the addition of an excess salt concentration to this system exacerbates the involvement of the hydrophobic residues of PMB1 in binding with BLP.

The acyl tail (28.03 %), DAB1 (17.52 %), Leu (13.29 %), DAB5 (10.12 %) and Thr1 (10.02 %) residues of PMB1 were the largest contributors, accounting for a combined 78.98 % of observed residue interactions in the concentrated Ubiq regime. Notably, under these simulation conditions, the D-Phe residue of PMB1 was involved in just 6.67 % of residue interactions with BLP, corresponding to the lowest percentage contribution of this residue to the binding of PMB1 to BLP across any simulation regime; 8.59 % less than its contribution in the neutralised Ubiq regime. The Leu/D-Phe/DAB5 triad accounted for a combined 30.08 % of observed residue interactions in this regime; corresponding to the lowest percentage involvement of this triad in any of the PMB1 simulation regimes. The polar residues of the branched fatty acid tail were involved in 30.57 % of residue interactions, lower than was observed in the neutralised Ubiq regime (38.14 %), yet considerably higher than the results from the concentrated Poly (17.55 %) and Osmo (8.79 %) regimes; providing further indication that under the crowded, chemically complex conditions of the Ubiq regimes, this region plays an important role in the binding of PMB1 to BLP. Furthermore, the 28.03 % contribution from the hydrophobic acyl tail of PMB1 corresponds to the highest percentage involvement of this residue in interactions with BLP across all simulation regimes. These results therefore indicate that, under these simulation conditions, the binding of PMB1 to BLP is underpinned primarily (88.68 %) by interactions involving the amphipathic branched fatty acid tail and the Leu/D-Phe/DAB5 triad region of PMB1, albeit with a diminished contribution from the D-Phe residue itself (Figure 5.8).

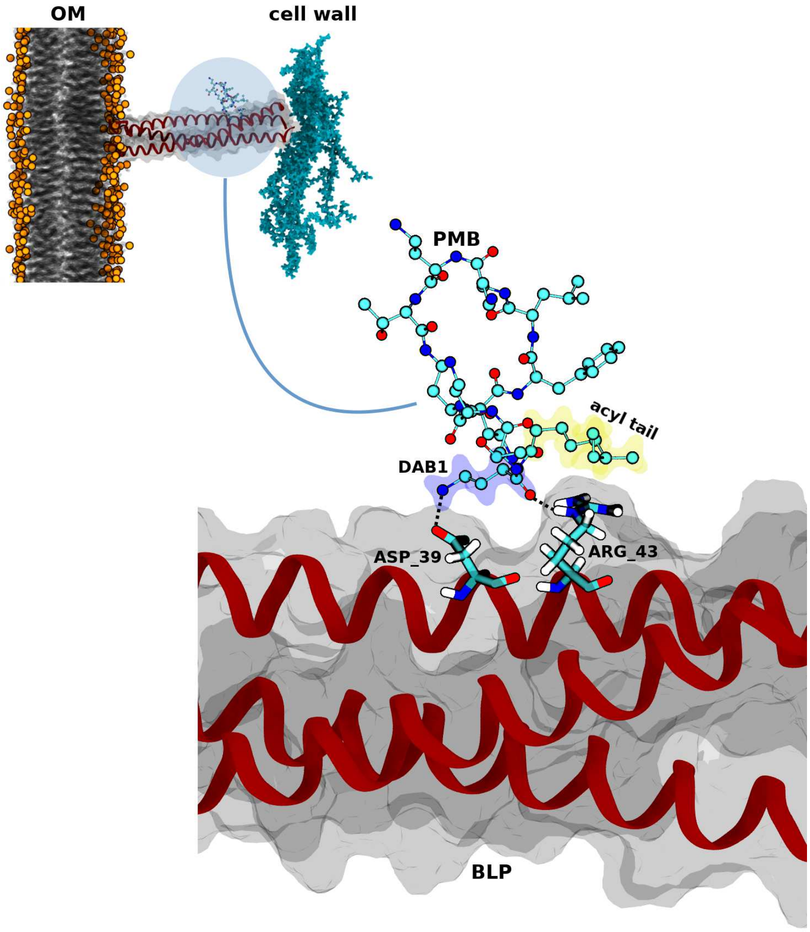


FIGURE 5.8: PMB1 bound to BLP in the concentrated Ubiq regime *via* interactions involving the acyl tail and DAB1 residues of PMB1. Hydrogen atoms of PMB1 are omitted for visual clarity. Dashed lines highlight hydrogen bond formation. The viewing perspective of the system was rotated by 90° during the rendering of this image.

Turning our attention to the specific residues of BLP involved in binding with PMB1; the Asp (27.20 %), Arg (22.70 %), Ala (16.09 %), Gln (8.74 %) and Lys (8.29 %) residues of BLP were involved in a combined 83.02 % of observed residue interactions with PMB1 in the concentrated Ubiq regime. The increased involvement of the anionic Asp (+ 16.15 %) and cationic Arg (+ 13.90 %) residues in this regime, compared to the neutralised Ubiq regime, is particularly interesting given their charged nature and the apparent decreased involvement of the charged/polar PMB1 residues (- 10.40 %) in the binding with BLP. The anionic Asp residues of BLP interacted solely with the cationic DAB residues of PMB1 and further, were involved in more residue interactions with each DAB residue than any other residue of BLP. Indeed, residue interactions between BLP and the two most prolifically interacting DAB residues, DAB1 and DAB5, comprised of 75.43 % and 73.92 %, respectively, interactions with the Asp residues of BLP (Figure 5.7).

In contrast to the Asp residues, the amphipathic nature of the cationic Arg residues of

BLP led to their involvement in a multitude of interactions with both hydrophobic and polar Thr residues of PMB1. Indeed Arg was the largest contributing residue of BLP to the observed residue interactions with the acyl tail, Thr1 and D-Phe residues of PMB1; and was responsible for, respectively, 49.02 %, 45.75 % and 64.33 % of residue interactions with these residues. Given that the acyl tail and Thr1 residues of PMB1 were involved in a combined 38.05 % of all residue interactions with BLP, and that almost half of these interactions were with the Arg residues alone; this data suggests that the amphipathic binding of Arg to these residues of PMB1 is an important contributing factor to the binding of PMB1 to BLP under these simulation conditions.

5.3.2 PME

Under neutralised conditions in the Poly and Osmo crowding regimes, no interactions were observed between PME and BLP. In contrast, in the higher salt conditions, such interactions were abundant; with 20 unique instances of PME-BLP binding observed in the Poly regime and 44 in the Osmo regime. Similar to PMB1, these interactions showed a strong preference for short durations, with median binding durations of 1.03 ns and 0.78 ns in the Poly and Osmo regimes respectively. Only one instance of a PME molecule remaining bound to BLP for the entire duration of simulation was observed across all performed simulations, occurring in the concentrated Osmo regime.

In the neutralised Ubiq regime, 64 unique instances of PME-BLP binding were observed; over double the number observed in equivalent simulations of PMB1. However, 63 of these binding events resulted from just two PME molecules that formed repeating transient interactions with BLP over the course of separate replicate simulations. Under these conditions, PME-BLP interactions were invariably brief with a median duration of 0.78 ns and, contrary to equivalent PMB1 simulations, no instances of PME molecules remaining bound to BLP for the entire duration of simulation. In the concentrated Ubiq regime, 66 unique instances of PME-BLP binding were observed, involving 8 individual PME molecules; exhibiting a median duration of 1.15 ns and no instances of PME molecules remaining bound to BLP for the entire simulation duration.

The binding of PME to BLP was largely dependent on polar interactions; accounting for $\sim$ 63.8 % of residue interactions across all regimes. Whilst these interactions predominantly involved the cationic DAB residues of PME, the Thr residues of PME accounted for $\sim$ 29.7 % of the observed polar interactions; closely matching the value calculated from the simulations of PMB1. Similar to PMB1, the Ser residues of BLP were seen to interact with both the DAB and Thr residues of PME. Such interactions, however, accounted for only $\sim$ 9.1 % of the total number of residue interactions; a lower percentage than was observed for the Gln ($\sim$ 17.6 %), Asp ($\sim$ 15.6 %), Ala ($\sim$ 14.8 %), Thr ($\sim$ 14.7 %), Asn ($\sim$ 12.3 %) and Lys ($\sim$ 9.3 %) residues of BLP. These

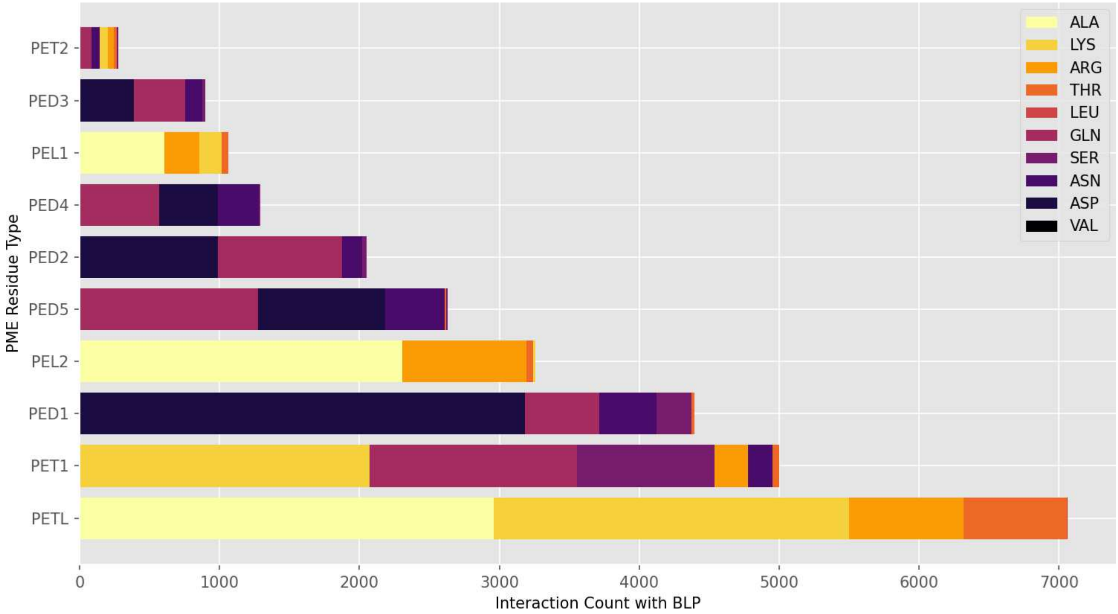


FIGURE 5.9: Residue interactions between PME and BLP in the concentrated Poly regime. Amino acids in the legend refer to residues of BLP.

results indicate that, similar to PMB1, the complete picture of PME-BLP binding is a balance of both electrostatic and hydrophobic interactions, involving many different residues.

In the concentrated Poly regime (Figure 5.9), polar interactions involving the DAB or Thr residues of PME were the dominant component in PME-BLP binding, making up 59.25 % of all observed residue interactions. The Thr1, DAB1 and DAB5 residues were of particular importance in these polar interactions; contributing to, respectively, 17.90 %, 15.74 % and 9.41 % of observed residue interactions. The combined contributions from the hydrophobic residues of PME were still significant, however; with the Leu1, Leu2 and acyl tail residues being responsible for, respectively, 3.81 %, 11.66 % and 25.28 % of observed residue interactions.

Clustering these results according to the location of each residue in the structure of PME; the polar residues of the branched fatty acid tail (DAB1, Thr1 and DAB2) accounted for 40.99 % of observed residue interactions. Furthermore, noting that PME differs from PMB1 only by the substitution of the D-Phe residue in PMB1 for the Leu2 residue in PME, the percentage contribution of the Leu/D-Phe/DAB5 triad of PMB1 may be compared to that of the analagous Leu1/Leu2/DAB5 triad of PME. In doing so, it is found that the Leu1/Leu2/DAB5 triad in PME was responsible for 24.88 % of observed residue interactions in the concentrated Poly regime, a considerably lower percentage than the 35.33 % contribution observed from the Leu/D-Phe/DAB5 triad in the equivalent PMB1 regime, providing further indication of the importance of the D-Phe residue to the function of this triad in the PMB1 simulations.

The binding of PME to BLP in this regime, however, appears to instead be dominated by interactions involving the branched fatty acid tail of PME, comprising 66.27 % of observed residue interactions; more than twice the contribution from the same residues of PMB1 (32.70 %) in equivalent simulations of the concentrated Poly regime. This result is somewhat surprising since the heptapeptide ring of PME is much bulkier than the branched fatty acid tail and comprises a greater number of polar and charged residues. Thus the apparent dominance of polar interactions on the binding of PME to BLP in this regime is not distributed equally amongst all polar residues, rather, it is concentrated within the fatty acid tail of PME. Indeed, when compared to the results obtained from the equivalent PMB1 simulation regime, the greatest decrease in involvement percentage of any polar residues are exhibited by the DAB3 (- 10.0 %) and Thr2 (- 12.11 %) residues on the heptapeptide ring of PME. Furthermore, the Leu1 and Leu2 residues of PME are involved in, respectively, 6.21 % and 4.10 % fewer of the observed residue interactions than the equivalent Leu and D-Phe residues of PMB1. These data highlight that, in the concentrated Poly regime, the variety of polar and hydrophobic interactions that the heptapeptide ring of PMB1 experiences with BLP are diminished in the binding of PME to BLP, leading to the dominant binding of PME to BLP *via* the polar and hydrophobic residues of its branched fatty acid tail (Figure 5.10).

In the concentrated Osmo regime (Figure 5.11), polar interactions involving the various DAB and Thr residues of PME dominated the binding of PME to BLP, accounting for a combined 79.93 % of observed residue interactions. The DAB2, DAB5 and Thr1 residues of PME were of particular importance, accounting for, respectively, 26.28 %, 22.87 % and 18.52 % of observed residue interactions. The hydrophobic residues of PME were responsible for 20.07 % of residue interactions; with the Leu1, Leu2 and acyl tail residues accounting for, respectively, 8.26 %, 5.96 % and 5.86 %. The polar residues of the branched fatty acid tail (DAB1, Thr1 and DAB2) accounted for a combined 50.48 % of observed residue interactions. The Leu1/Leu2/DAB5 triad in PME was responsible for a further 37.09 % of observed residue interactions; just over half of the contribution from the analagous triad in the equivalent PMB1 simulation regime (64.36 %). Whilst the importance of this triad was diminished in PME as compared to PMB1, the combination of its contribution with that of the DAB1/Thr1/DAB2 residues accounted for a combined 87.57 % of observed residue interactions with BLP. These data therefore indicate that, in this regime, the interaction of PME with BLP is underpinned by interactions involving the polar region of the branched tail of PME and the Leu1/Leu2/DAB5 triad (Figure 5.1B).

The proportion of residue interactions that involved the hydrophobic residues of PME in the concentrated Osmo regime ($\sim$ 20.1 %) was approximately half that of the concentrated Poly regime ($\sim$ 40.7 %). This decrease may largely be attributed to the decreased involvement of the acyl tail of PME; this residue was responsible for

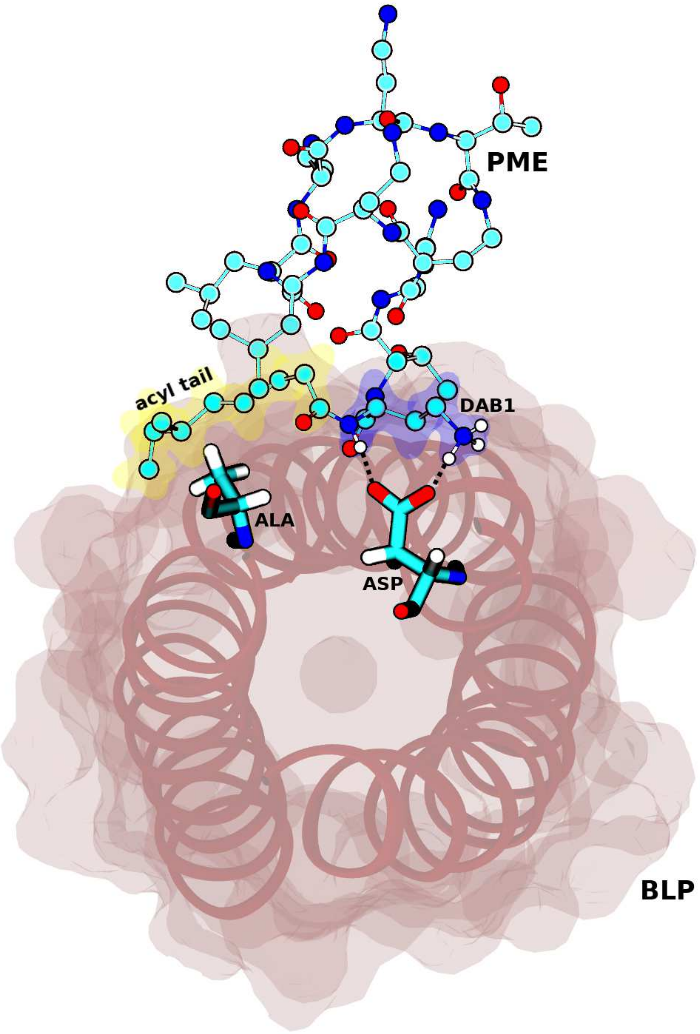


FIGURE 5.10: PME bound to BLP in the concentrated Poly regime via interactions involving the cationic DAB1 and hydrophobic acyl tail residues located on the branched fatty acid tail of PME. Dashed lines highlight hydrogen bond formation. Hydrogen atoms of PME that are not involved in hydrogen bonding with BLP are omitted for visual clarity. This image was rendered from a perspective looking down the length of BLP, from the OM towards the cell wall.

$\sim$ 25.3 % of residue interactions in the Poly regime compared to just $\sim$ 5.9 % in the Osmo regime. Particular interest may be found in this result when compared to the previously described increase in prevalence of hydrophobic interactions during the binding of PMB1 to BLP upon the inclusion of osmolytes; indicating that the presence of these osmolytes had an opposing effect on the biochemical nature of BLP binding with each type of polymyxin. However, given that the D-Phe residue of PMB1 accounted for 31.31 % of all observed residue interactions in the concentrated Osmo regime, it is of no surprise that the substitution of this moiety for a Leucine residue in PME is accompanied by a markedly different distribution of residue interactions (Figures 5.4 & 5.11).

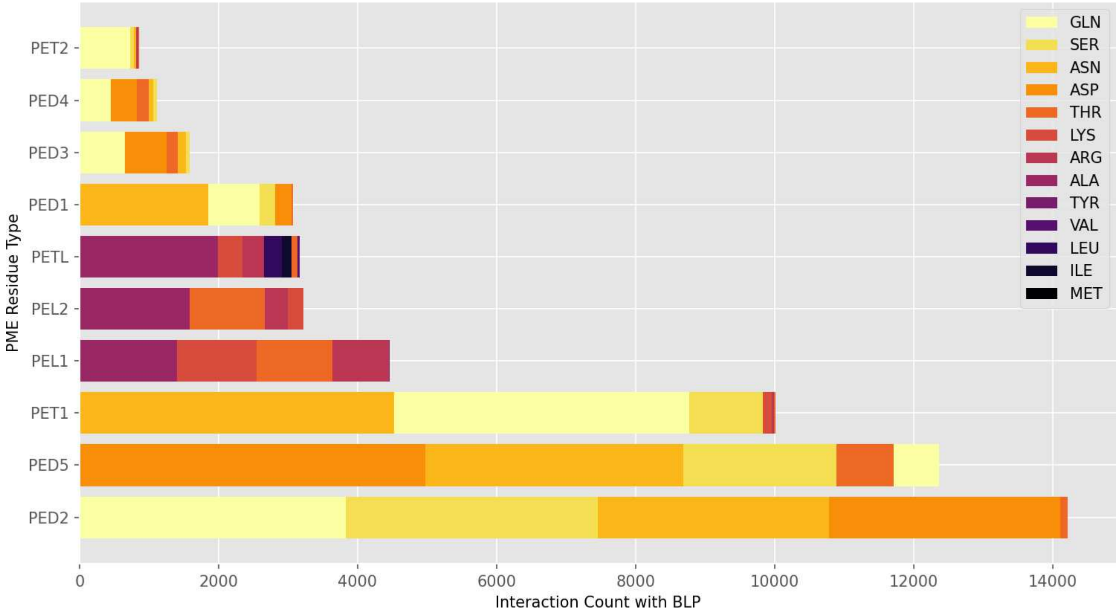


FIGURE 5.11: Residue interactions between PME and BLP in the concentrated Osmo regime. Amino acids in the legend refer to residues of BLP.

In the neutralised Ubiq regime (Figure 5.12), just two PME molecules were responsible for all but one observed instance of PME-BLP binding. Both PME molecules were bound to clusters of ubiquitin proteins that were bound to the cell wall in close proximity to BLP (Figure 5.13). Neither PME molecule was observed to dissociate from these clusters and further, the clusters did not dissociate from the cell wall. The two PME molecules thus remained in close proximity to BLP for the entire duration of simulation; forming repeated, transient interactions involving only those residues that were not already bound to the ubiquitin clusters. This behaviour gives rise to the apparent dominance of the Thr1 ($\sim$ 46.8 %) and DAB3 ($\sim$ 37.3 %) residues of PME in binding with BLP under these simulation conditions.

In the concentrated Ubiq regime (Figure 5.14), polar interactions involving the various DAB and Thr residues of PME accounted for a combined $\sim$ 50.5 % of observed residue interactions; $\sim$ 29.4 % lower than the contribution of these same residues in the concentrated Osmo regime. Of these polar residues, the Thr1 and DAB1 residues of PME were of particular importance, contributing to, respectively, $\sim$ 16.0 % and $\sim$ 12.7 % of observed residue interactions. The hydrophobic residues of PME accounted for the remaining $\sim$ 49.5 % of observed residue interactions; with the acyl tail, Leu2 and Leu1 residues accounting for, respectively, $\sim$ 30.4 %, $\sim$ 11.8 % and $\sim$ 7.4 %.

Clustering these results as before highlights that $\sim$ 65.7 % of observed residue interactions involved residues located on the branched fatty acid tail of PME; with a further $\sim$ 19.2 % arising from interactions involving the two hydrophobic Leu residues on the PME heptapeptide ring. The mean contribution of each remaining

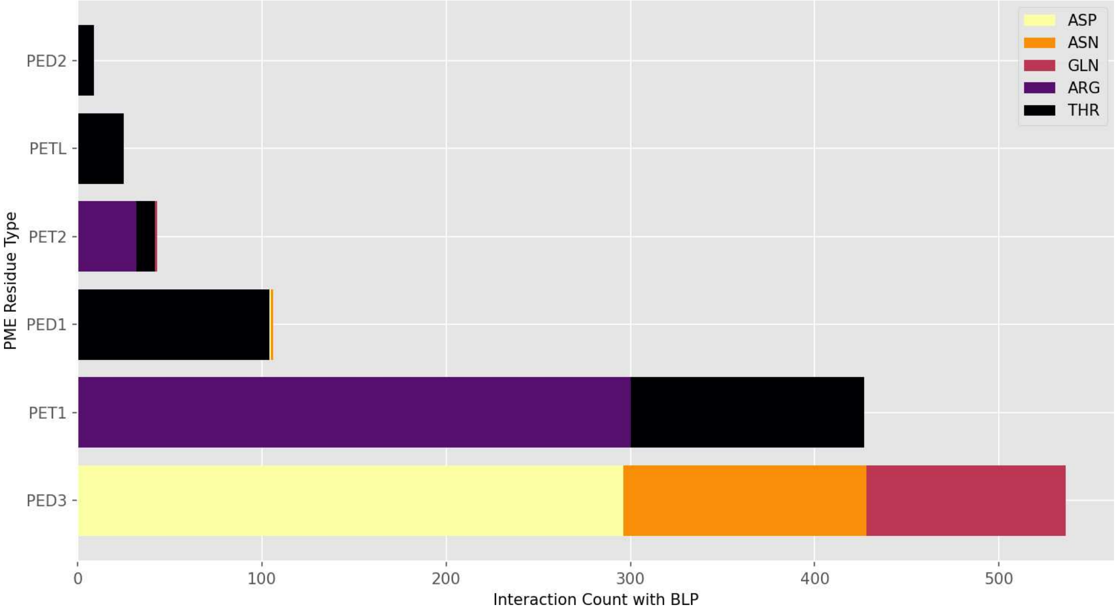


FIGURE 5.12: Residue interactions between PME and BLP in the neutralised Ubiq regime. Amino acids in the legend refer to residues of BLP.

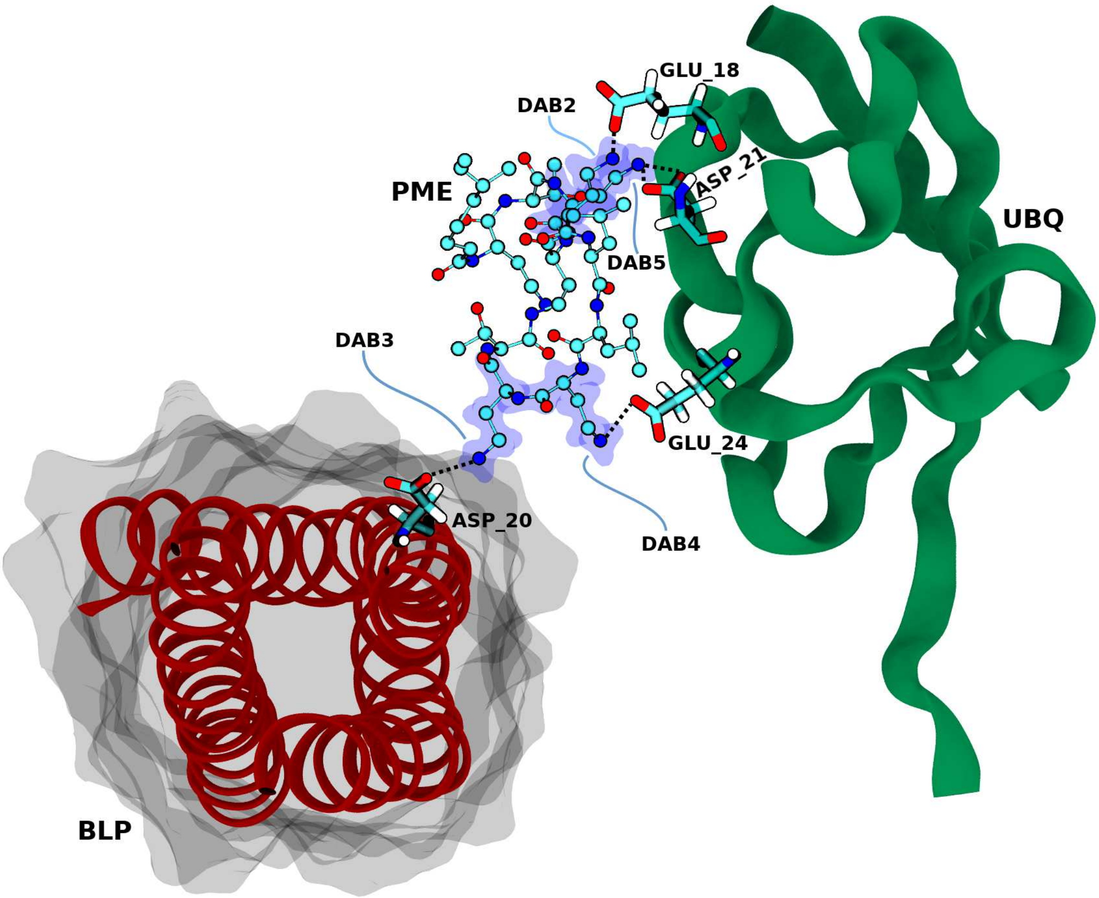


FIGURE 5.13: PME bound to BLP and Ubiq in the neutralised Ubiq regime *via* electrostatic interactions involving cationic DAB residues. Dashed lines highlight hydrogen bond formation. Hydrogen atoms of PME are omitted for visual clarity. This image was rendered from a perspective looking down the length of BLP, from the OM towards the cell wall.

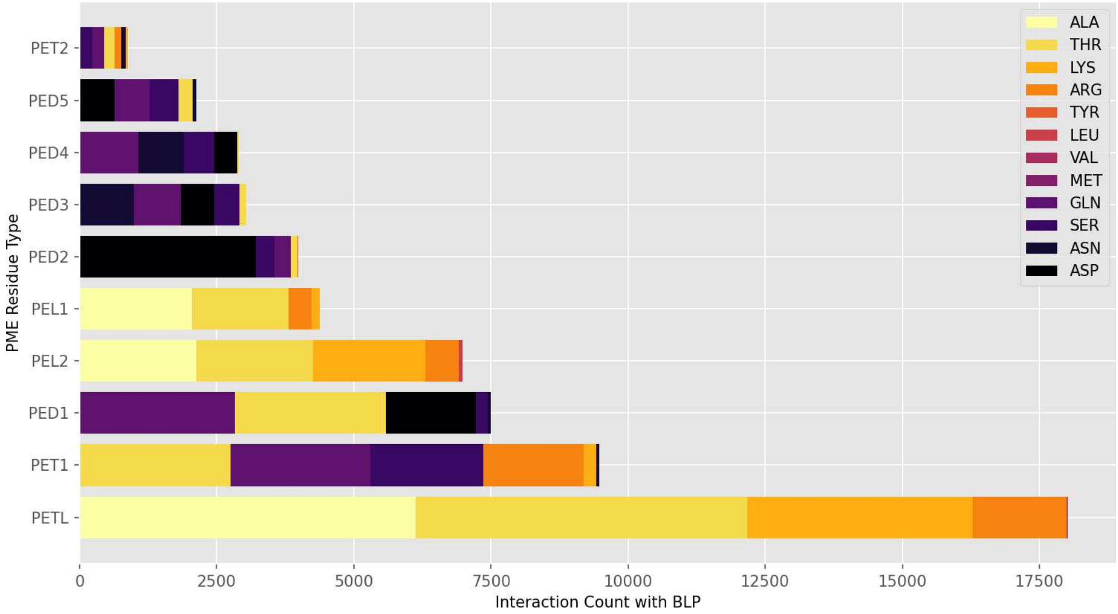


FIGURE 5.14: Residue interactions between PME and BLP in the concentrated Ubiq regime. Amino acids in the legend refer to residues of BLP.

polar residue on the heptapeptide ring was just 3.8 ± 1.7 %. The binding of PME to BLP under these simulation conditions is thus underpinned primarily (∼ 84.9 %) by interactions involving the branched fatty acid tail and hydrophobic Leu residues of PME. This result is comparable to that obtained from the equivalent simulations of PMB1 in the concentrated Ubiq regime; in which PMB1-BLP binding was dominated by interactions involving the branched fatty acid tail and Leu/D-Phe/DAB5 triad, the latter of which, excluding the DAB5 residue, is positionally analagous to the two Leu residues of PME.

5.3.3 Polymyxin - OPG Interactions

The dynamic, variational nature of residue interactions between the polymyxins and BLP in response to changes in the biochemical environment poses an immediate question regarding the mechanistic understanding of such behaviour: if the inclusion of ions, osmolytes and crowding proteins has such a profound effect on the nature of polymyxin - BLP binding; what, exactly, is the nature of interaction between these additional components and the polymyxins and further, do these interactions account for the variations observed in the interaction of the polymyxins with BLP.

To this end, the specific residue interactions that underpinned the interactions between each polymyxin species and OPG molecules were studied in each of the simulation regimes. OPG comprises 5 carbohydrate rings with numerous polar hydroxyl groups that are likely to interact with the cationic DAB residues of the polymyxins; perhaps providing one source of disruption to the formation of electrostatic interactions between the polymyxins and BLP. Furthermore, carbohydrate rings are known to form stacking interactions with aromatic amino acids Lucas et al. (2014) and thus the interplay between BLP and OPG for interactions with the D-Phe residue of PMB1 is of particular interest.

5.3.3.1 PMB1

Across all replicates of the neutralised Osmo regime, interactions involving the hydrophobic residues of PMB1 accounted for 38.59 % of all observed residue interactions with OPG; with the acyl tail, D-Phe and Leu residues responsible for, respectively, 17.47 %, 12.79 % and 8.22 %. The cationic DAB residues of PMB1 were responsible for a further 44.56 % of observed residue interactions; with the largest contributions arising from the DAB1 (14.9 %) and DAB5 (8.97 %) residues. Interactions between OPG and the two polar Thr residues of PMB1 accounted for the final 16.96 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 9.68 % and 7.28 %.

Whilst the broad involvement of all PMB1 residues in binding with OPG (Figure 5.15) in this regime indicates that there is no particular region of PMB1 that is entirely excluded from interactions with OPG; the acyl tail, DAB1 and D-Phe residues each exhibited a relatively large percentage involvement. Clustering the contributions of these residues with those of their neighbouring residues, it is found that the previously discussed Leu/D-Phe/DAB5 triad accounts for a combined 29.98 % of all observed residue interactions with OPG. The polar DAB1/Thr1/DAB2 residues located on the branched fatty acid tail of PMB1 account for a further 30.47 % of residue interactions. The contributions from these two residue triads plus the hydrophobic acyl tail are thus found to account for 77.92 % of all observed residue interactions with OPG.

These data provide an indication that all residues of PMB1, hydrophobic, charged or polar, interact with all regions of OPG under these conditions. There appears to be a slight preference for interactions involving the Leu/D-Phe/DAB5 residue triad, as well as large contributions from the DAB1, Thr1 and acyl tail residues located on the branched fatty acid tail of PMB1. These residues were found to play an important role in PMB1-BLP binding in various simulation regimes. Notably, no observations of PMB1-BLP binding were observed in any replicate of this regime and thus there was no direct competition between OPG and BLP for interactions with these regions of PMB1. Indeed, all PMB1 molecules were found to be bound to the cell wall for the entire duration of each simulation; with the DAB2, DAB4 and DAB3 residues being the largest contributing residues to PMB1 - cell wall binding. These residues account for three of the four smallest contributing residues to PMB1-OPG binding in this regime, indicating that PMB1 interactions with the cell wall limit the formation of

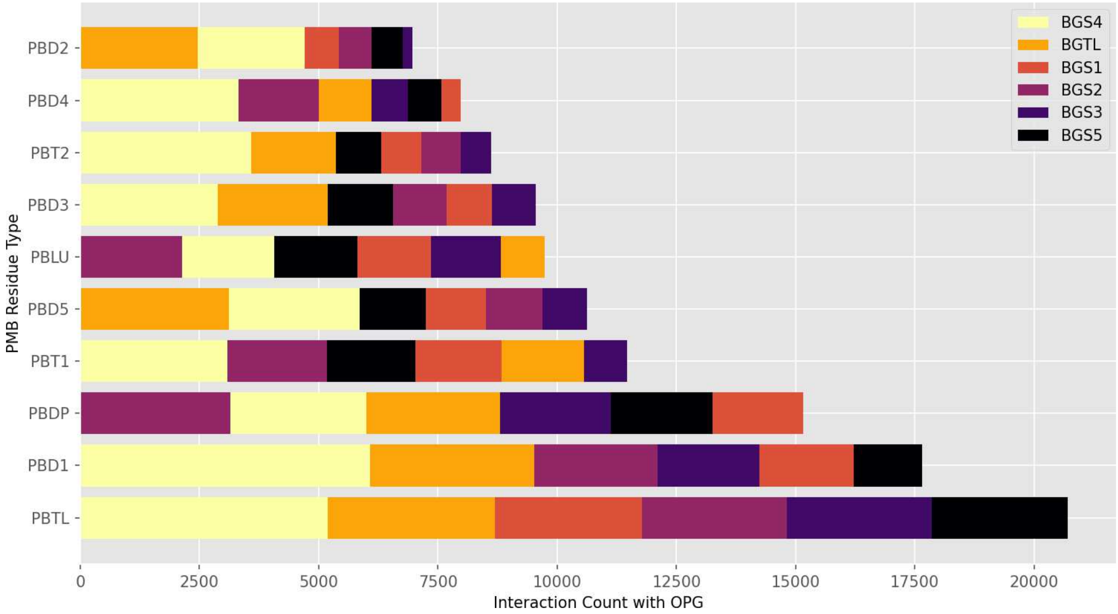


FIGURE 5.15: Residue interactions between PMB1 and OPG in the neutralised Osmo regime. Residue names in the legend refer to residues of OPG.

certain OPG-DAB residue interactions; leaving the hydrophobic and branched tail residues of PMB1 free to interact with surrounding OPG molecules (Figure 5.16).

Across all replicates of the concentrated Osmo regime (Figure 5.17), interactions involving the hydrophobic residues of PMB1 accounted for 25.91 % of all observed residue interactions with OPG; with the the acyl tail, D-Phe and Leu residues responsible for, respectively, 11.3 %, 8.66 % and 5.96 %. The cationic DAB residues of PMB1 were responsible for a further 49.12 % of observed residue interactions; with the largest contributions arising from the DAB3 (13.98 %) and DAB5 (12.38 %) residues. Interactions between OPG and the two polar Thr residues of PMB1 accounted for the final 24.97 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 8.67 % and 16.29 %.

The broad involvement of all PMB1 residues in binding with OPG under these high salt conditions indicates that, similar to the results from the neutralising salt simulations, no particular region of PMB1 was entirely excluded from interactions with OPG. The contributions of the Leu/D-Phe/DAB5 (27.0 %) and DAB1/Thr1/DAB2 (25.83 %) triads were, however, diminished in the concentrated Osmo regime compared to the neutralised Osmo regime; indicating that whilst all residues of PMB1 may still be able to interact with each region of OPG, the addition of an excess salt concentration fundamentally changes the regions of each molecule that tend to coordinate with one another. Indeed, the Thr2 and DAB3 residues were the two largest contributors to PMB1-OPG binding in these high salt conditions, compared to their position amongst the least contributing residues in the neutralised regime.

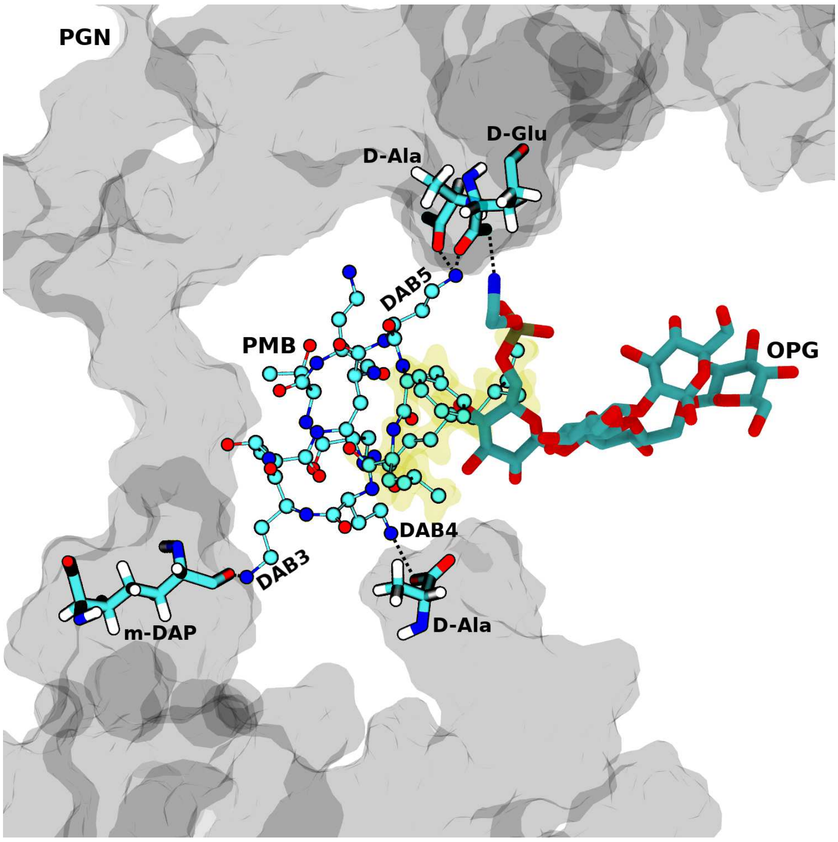


FIGURE 5.16: Cell wall bound PMB1 interacting with a nearby OPG molecule in the neutralised Osmo regime. Hydrophobic residues of PMB1 are highlighted in yellow bubbles.

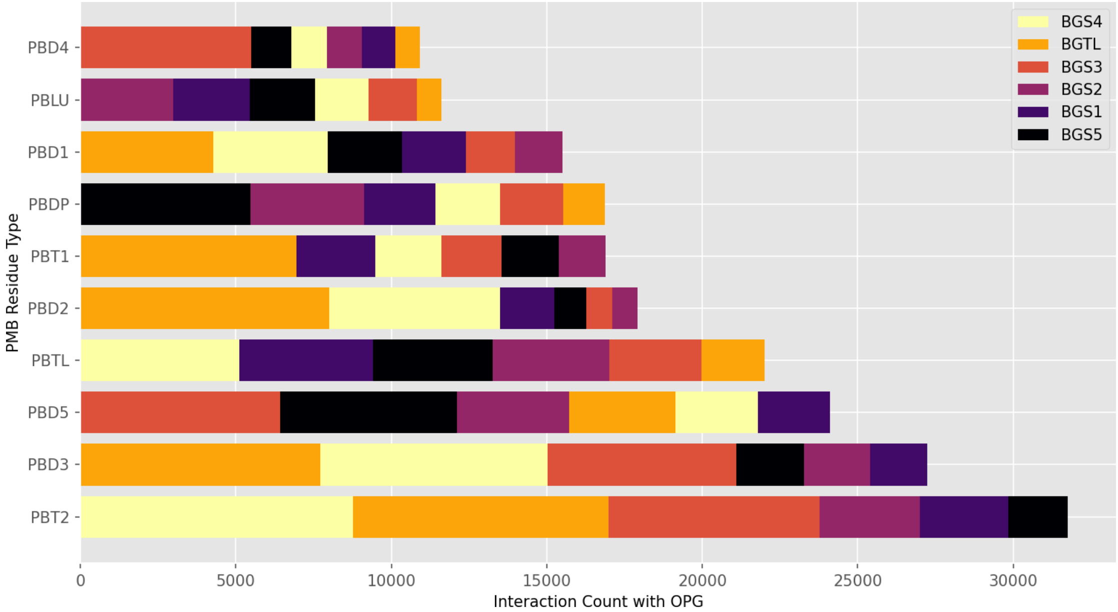


FIGURE 5.17: Residue interactions between PMB1 and OPG in the concentrated Osmo regime. Residue names in the legend refer to residues of OPG.

This change in the biochemical nature of PMB1-OPG interaction is corroborated by consideration of the percentage contribution of each residue of OPG to the observed residue interactions with PMB1; whereby it is found that the most prolifically interacting residue of OPG in the neutralised Osmo regime, BGS4, contributed to 8.02 % fewer of the total observed residue interactions in the concentrated Osmo regime. Correspondingly, the least prolifically interacting residue of OPG in the neutralised Osmo regime, BGS3, contributed to an additional 7.05 % of the total number of observed residue interactions in the concentrated Osmo regime compared to the neutralised regime.

Since the dissociation of PMB1 from the cell wall, and their subsequent interaction with BLP, was observed in the Osmo regime only under concentrated conditions; the deviation in the observed residue interactions between PMB1 and OPG in the neutralised and concentrated regimes is likely explained through the matrix of newly enabled interactions between each constituent molecule. The absence of PMB1 interactions with BLP in the neutralised Osmo regime was a result of the ubiquitous and unbroken binding of PMB1 molecules to the cell wall in all replicate simulations. Interactions of the PMB1 molecules with the cell wall were dominated by electrostatic interactions involving the cationic DAB residues; thus, it is of no surprise that the percentage contribution of each hydrophobic residue of PMB1 to OPG binding was greater in the neutralised regime than in the concentrated regime as the electrostatic interactions between PMB1 and the cell wall left the hydrophobic residues free to interact with the surrounding solutes. Conversely, in the concentrated Osmo regime, the dissociation of PMB1 molecules from the cell wall enabled the surrounding biomolecules to interact freely with the cationic DAB residues, therefore decreasing the relative contribution of PMB1 hydrophobic residues to their binding.

Indeed, the ability for PMB1 molecules to diffuse throughout the periplasm in the concentrated Osmo regime resulted in the formation of OPG-PMB1 clusters involving both BLP-bound and freely diffusing PMB1 molecules. In one case, an OPG dimer was observed to interact with a BLP-bound PMB1 molecule whilst a second, freely diffusing PMB1 bound to the opposing side of the OPG dimer (Figure 5.18). In this instance, it was evident that the binding of PMB1 to BLP *via* the Leu/D-Phe/DAB5 triad left the DAB1 residue on the branched fatty acid tail exposed and able to coordinate with the various hydroxyl groups of the incoming OPG dimer. The second, non-BLP-bound PMB1 formed interactions with the opposing side of the OPG dimer through the various polar residues of the PMB1 heptapeptide ring; leaving the branched fatty acid tail exposed to the bulk aqueous phase and thus excluded from interactions with the OPG dimer.

Across all replicates of the neutralised Ubiq regime, interactions involving the hydrophobic residues of PMB1 accounted for 33.94 % of all observed residue interactions with OPG; with the D-Phe, acyl tail and Leu residues responsible for,

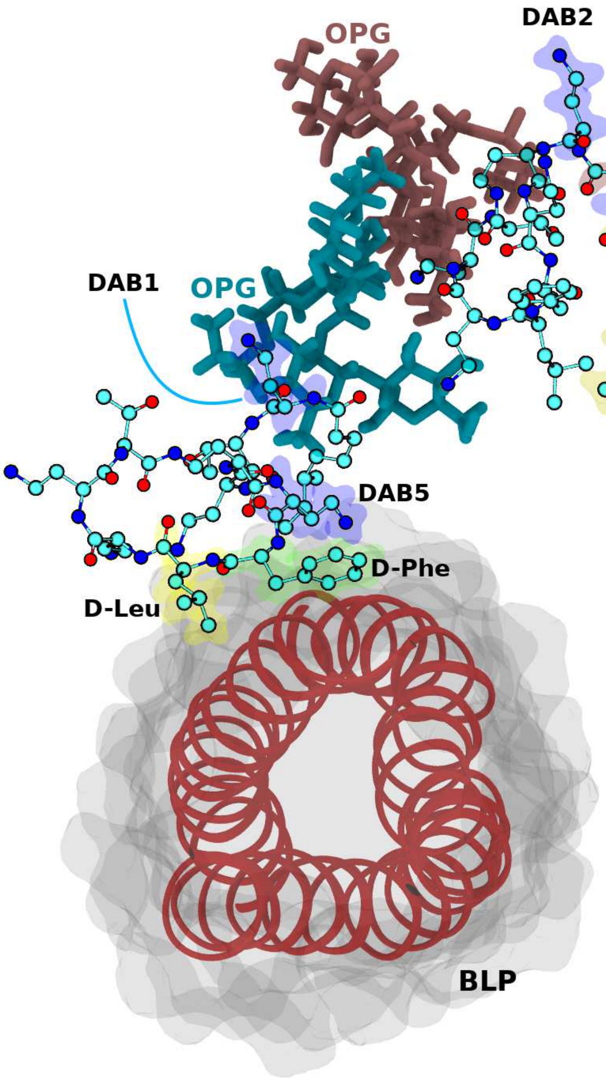


FIGURE 5.18: OPG dimer interacting with a BLP bound PMB1 in the concentrated Osmo regime; a second, freely diffusing, PMB1 molecule is also seen interacting with the OPG dimer. This image was rendered from a perspective looking down the length of BLP, from the cell wall towards the OM.

respectively, 13.0 %, 11.54 % and 9.41 %. The cationic DAB residues of PMB1 were responsible for a further 46.09 % of observed residue interactions; with the largest contributions arising from the DAB3 (10.24 %) and DAB4 (9.8 %) residues. The DAB1, DAB2 and DAB5 residues were involved in, respectively, 8.97 %, 8.39 % and 8.69 % of observed residue interactions; it is thus evident that, under these simulation conditions, there was no overwhelming preference for OPG interactions with any particular DAB residue of PMB1. Interactions between OPG and the two polar Thr residues of PMB1 accounted for the final 19.97 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 13.41 % and 6.56 %. The median binding duration of PMB1 with the cell wall was $\sim$ 210 ns in this regime, with their binding being dominated by interactions involving the DAB residues of PMB1; thus,

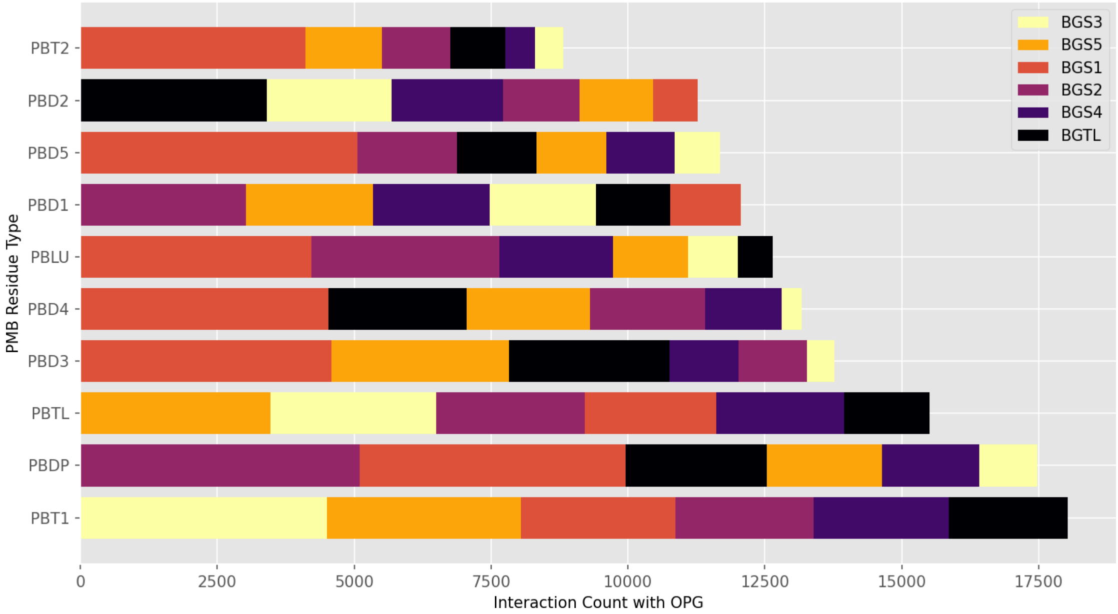


FIGURE 5.19: Residue interactions between PMB1 and OPG in the neutralised Ubiq regime. Residue names in the legend refer to residues of OPG.

for the majority of the duration of these simulations, a subset of the DAB residues of each PMB1 were unable to interact with OPG, leading to the slight preference for PMB1-OPG interactions involving the D-Phe, acyl tail and Thr1 residues of PMB1.

Across all replicates of the concentrated Ubiq regime (Figure 5.20), interactions involving the hydrophobic residues of PMB1 accounted for 39.59 % of all observed residue interactions with OPG; with the Leu, acyl tail and D-Phe residues responsible for, respectively, 14.35 %, 12.8 % and 12.44 %. The cationic DAB residues of PMB1 were responsible for a further 42.79 % of observed residue interactions; with the largest contribution arising from the DAB3 (12.86 %) residue. The DAB1, DAB2, DAB4 and DAB5 residues were involved in, respectively, 8.02 %, 5.19 %, 7.74 % and 8.99 % of observed residue interactions. Interactions between OPG and the two polar Thr residues of PMB1 accounted for the final 17.63 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 6.83 % and 10.8 %.

Under these simulation conditions, the polar DAB1/Thr1/DAB2 residues of the branched fatty acid tail of PMB1 accounted for a combined 20.04 % of observed residue interactions with OPG, whilst the Leu/D-Phe/DAB5 triad was responsible for a further 35.78 %. Notably, the complete branched tail (DAB1/Thr1/DAB2/acyl tail) of PMB1 was involved in 32.84 % of all observed residue interactions, 9.47 % fewer than in the neutralised Ubiq regime. This is a roughly equivalent decrease to that observed between the results from the neutralised and concentrated Osmo regime; indicating that the additional of an excess salt concentration to either system induces a similar decrease in the relative importance of the branched fatty acid tail of PMB1 to the binding with OPG.

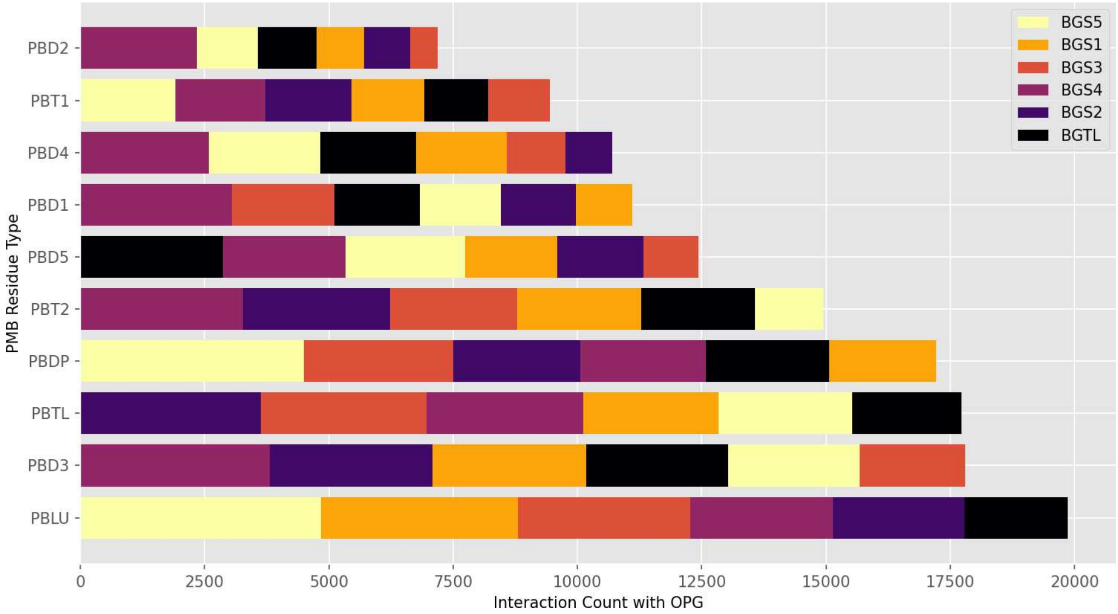


FIGURE 5.20: Residue interactions between PMB1 and OPG in the concentrated Ubiq regime. Residue names in the legend refer to residues of OPG.

5.3.3.2 PME

Across all replicates of the neutralised Osmo regime, interactions involving the hydrophobic residues of PME accounted for 33.96 % of all observed residue interactions with OPG; with the the acyl tail, Leu1 and Leu2 residues responsible for, respectively, 17.26 %, 10.33 % and 6.38 %. The cationic DAB residues of PME were responsible for a further 50.34 % of observed residue interactions; with the largest contributions arising from the DAB3 (17.44 %) and DAB4 (10.27 %) residues. Interactions between OPG and the two polar Thr residues of PME accounted for the final 15.70 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 7.69 % and 8.01 %.

Similar to the results from PMB1, the broad involvement of all PME residues in binding with OPG (Figure 5.21) in this regime indicates that there is no particular region of PME that is entirely excluded from interactions with OPG. The acyl tail and DAB3 residues, however, exhibited relatively large contributions. The previously discussed Leu1/Leu2/DAB5 triad of PME accounted for just 24.23 % of all observed residue interactions with OPG, whilst the polar DAB1/Thr1/DAB2 residues located on the branched fatty acid tail accounted for a further 22.8 % of residue interactions. These two residue triads therefore show no particular importance to the binding of PME to OPG in this regime.

Across all replicates of the concentrated Osmo regime, interactions involving the hydrophobic residues of PME accounted for 39.39 % of all observed residue interactions with OPG; with the the acyl tail, Leu1 and Leu2 residues responsible for,

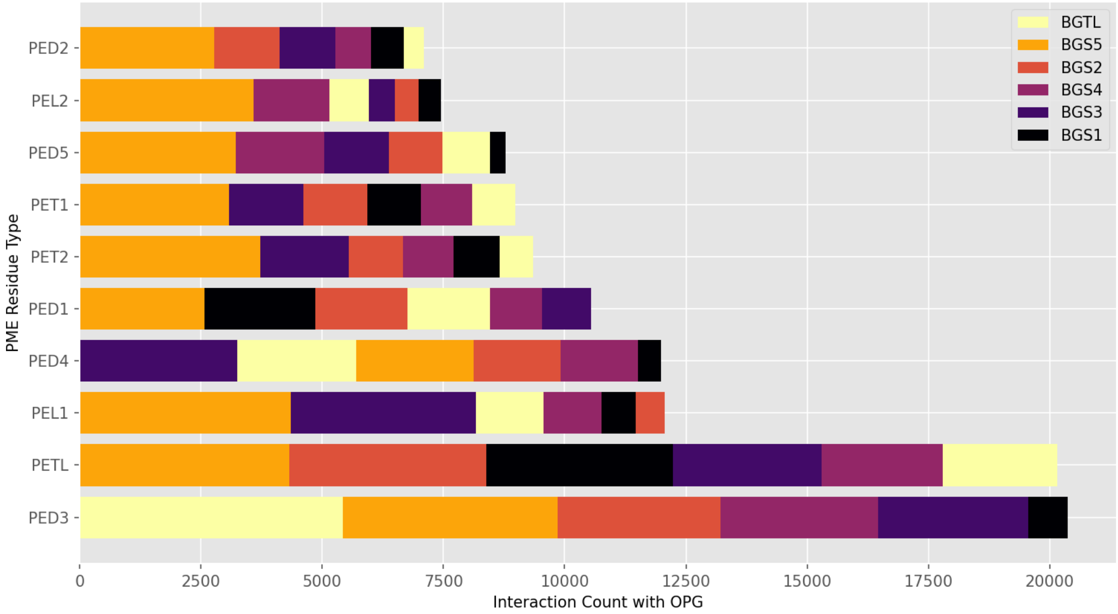


FIGURE 5.21: Residue interactions between PME and OPG in the neutralised Osmo regime. Residue names in the legend refer to residues of OPG.

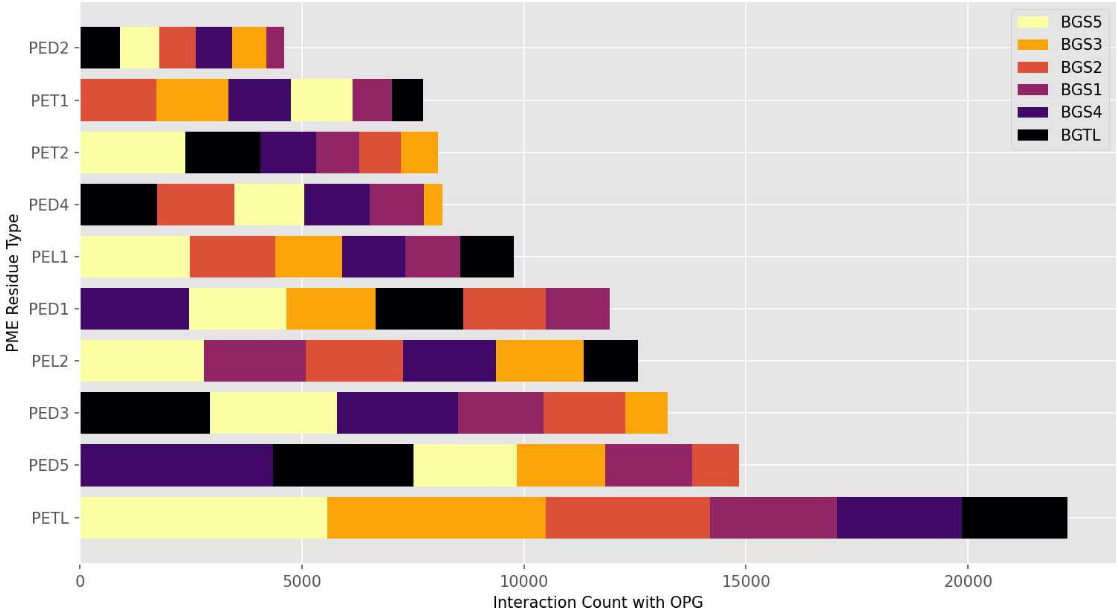


FIGURE 5.22: Residue interactions between PME and OPG in the concentrated Osmo regime. Residue names in the legend refer to residues of OPG.

respectively, 19.65 %, 8.63 % and 11.11 %. The cationic DAB residues of PME were responsible for a further 46.65 % of observed residue interactions; with the largest contributions arising from the DAB5 (13.12 %), DAB3 (11.69 %) and DAB1 (10.55 %) residues. Interactions between OPG and the two polar Thr residues of PME accounted for the final 13.96 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 6.83 % and 7.13 %.

The broad involvement of all residues of PME in binding with OPG in this regime once again indicates that their binding is underpinned by a variety of both hydrophobic and electrostatic interactions involving a variety of different residues. Notably, the smallest contributions come from the DAB2 and Thr1 residues, which are located at the junction region between the heptapeptide ring and branched fatty acid tail of PME; perhaps indicating that this region in particular is excluded from interactions with OPG. Indeed, the DAB1/Thr1/DAB2 triad contributed to just 21.45 % of observed residue interactions, closely matching the value obtained from neutralised regime; indicating that the dissociation of PME from the cell wall, observed only under the concentrated conditions, had no significant effect on the contribution of this residue triad to PME-OPG binding. Notably, the DAB2 and Thr1 residues were of particular importance to the binding of BLP to PME in the concentrated Osmo regime, contributing to, respectively, 26.28 % and 18.52 % of PME-BLP residue interactions. The meagre contributions of these two residues to binding with OPG may thus be seen to leave these residues open to form an abundance of interactions with BLP when PME-cell wall interactions are disrupted to such an extent as to allow PME to diffuse throughout the periplasm.

Furthermore, the Leu1/Leu2/DAB5 triad was involved in 32.86 % of residue interactions in this regime; an increase of 8.63 % compared to the neutralised regime. This provides some evidence that the dissociation of PME from the cell wall increases the proclivity of this residue triad to form interactions with OPG. The same triad accounted for 37.09 % of PME-BLP interactions in this regime, and thus it is likely that OPG and BLP compete for interactions involving this region of PME.

Across all replicates of the neutralised Ubiq regime (Figure 5.23), interactions involving the hydrophobic residues of PME accounted for 43.49 % of all observed residue interactions with OPG; with the the acyl tail, Leu1 and Leu2 residues responsible for, respectively, 22.83 %, 6.97 % and 13.69 %. The cationic DAB residues of PME were responsible for a further 40.57 % of observed residue interactions; with the largest contributions arising from the DAB3 (13.29 %) and DAB5 (12.36 %) residues. Interactions between OPG and the two polar Thr residues of PME accounted for the final 15.95 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 2.88 % and 13.07 %.

Similar to the results from both neutralised and concentrated Osmo regimes; the DAB2 and Thr1 residues of PME provided the smallest contributions to OPG binding in this regime. Indicating that, again, the junction region between the heptapeptide ring and fatty acid tail appears to be excluded from interactions with OPG. Indeed, the full DAB1/Thr1/DAB2 triad accounted for just 12.47 % of residue interactions with OPG; by far the lowest contribution of this triad to binding with OPG in any regime. This triad was responsible for 47.3 % of residue interactions with BLP in this regime, once again supporting the notion that the lack of OPG interactions leaves these residues free to interact profusely with BLP. As previously discussed, however, there

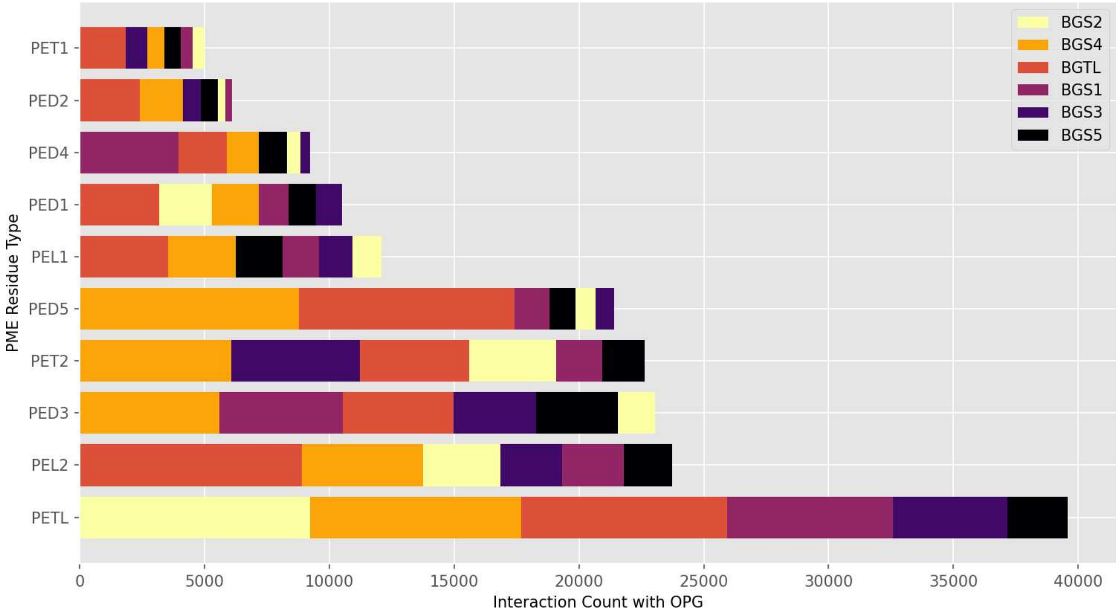


FIGURE 5.23: Residue interactions between PME and OPG in the neutralised Ubiq regime. Residue names in the legend refer to residues of OPG.

was only limited sampling of BLP-PME interactions in this regime and thus further simulation will be required to provide conclusive evidence of this correlation.

Across all replicates of the concentrated Ubiq regime (Figure 5.24), interactions involving the hydrophobic residues of PME accounted for 37.58 % of all observed residue interactions with OPG; with the the acyl tail, Leu1 and Leu2 residues responsible for, respectively, 13.19 %, 16.3 % and 8.09 %. The cationic DAB residues of PME were responsible for a further 45.14 % of observed residue interactions; with the largest contributions arising from the DAB3 (16.5 %) and DAB1 (10.32 %) residues. Interactions involving the two polar Thr residues of PME accounted for the final 17.28 % of observed residue interactions, with Thr1 and Thr2 accounting for, respectively, 6.41 % and 10.87 %.

Notably, PME-BLP binding in this regime was dominated (65.73 %) by interactions involving the branched fatty acid tail of PME, with the DAB1/Thr1/DAB2 triad contributing 35.35 % and the acyl tail contributing 30.38 %. Conversely, PME-OPG binding in this regime was dominated (66.28 %) by interactions involving the heptapeptide ring of PME, with the acyl tail contributing to a further 13.19 %. It is not clear, therefore, that BLP and OPG compete with each other for interactions involving specific residues of PME in this regime.

The DAB2 residue provided the smallest contribution of any PME residue to binding with OPG, similarly, the Thr1 residue accounted for the third smallest contribution. These results closely follow those obtained from all other simulation regimes involving PME. The junction region of PME is thus found to be largely inhibited from interacting with OPG, regardless of the simulated environment. The acyl tail of PME

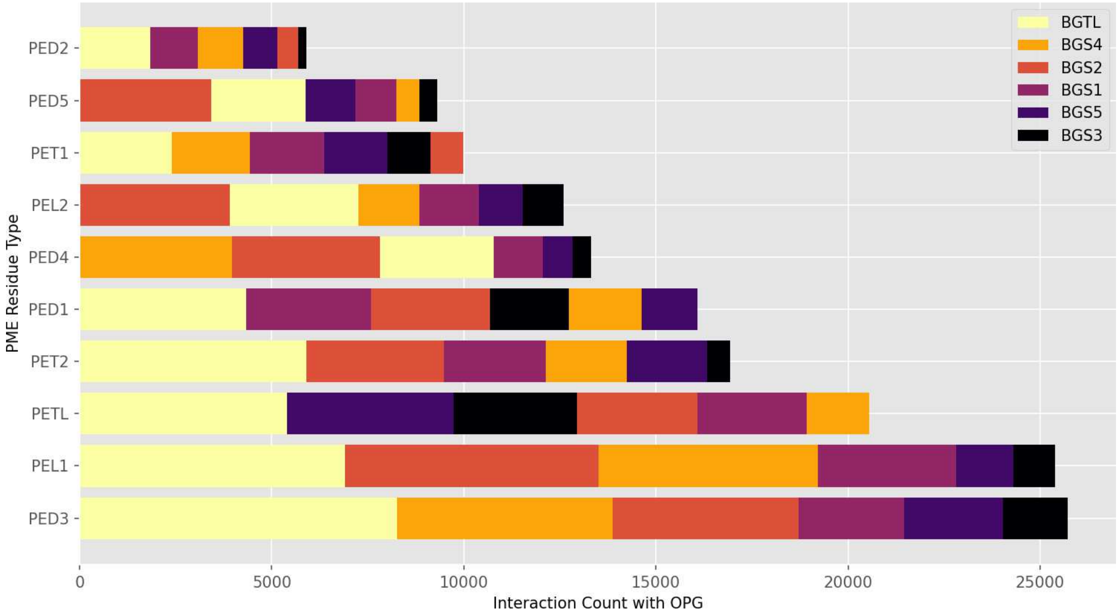


FIGURE 5.24: Residue interactions between PME and OPG in the concentrated Ubiq regime. Residue names in the legend refer to residues of OPG.

is also found to provide one of the largest contributions to OPG binding in this regime, again, closely following the data from all other simulation regimes. These results indicate that whilst there are underlying variations in the nature of PME binding with OPG, there is a conserved motif of binding via the acyl tail of PME along with exclusion of the DAB1/Thr1/DAB2 triad.

5.4 Discussion and conclusions

When compared to the analysis of polymyxin interactions with the cell wall presented in the previous chapter; the study of polymyxin interactions with BLP highlighted a different picture of how environmental complexity impacts polymyxin interactions within the periplasm of *E. coli*. Whilst the interactions between the polymyxins and BLP exhibited a preference for short durations across all simulation regimes; the residue interactions that underpinned their binding were found to be dependent both on polymyxin type and system complexity. This is in stark contrast to the effects that increased biochemical complexity had on polymyxin – cell wall binding; wherein the binding durations were observed to decrease whilst the underlying residue interactions remained consistent.

It has been previously reported that the serine and acidic residues of BLP had a particular propensity to interact with PMB1 Pedebos et al. (2021); however, the work presented in this chapter has shown that it is in fact a multitude of both polar and hydrophobic interactions that give rise to the binding of both PMB1 and PME to BLP. The balance of these polar and hydrophobic interactions was similar for both polymyxins in the Poly regime; however, deviations emerged in the Osmo and Ubiq regimes, exemplified by the concentrated Osmo regime, in which PMB1 – BLP binding was dominated by hydrophobic interactions and PME – BLP binding was dominated by electrostatic interactions. Analysis of the underlying residue interactions indicated that, under these conditions, almost a third of all PMB1 - BLP residue interactions involved the hydrophobic D-Phe residue of PMB1; since this moiety is substituted for a D-Leu residue in PME, the deviation in residue interaction distributions between the two polymyxin species is unsurprising, albeit unexplained as of yet.

Despite the deviations observed between simulation regimes and polymyxin species, this work presented in this chapter has highlighted how certain groups of residues in both PMB1 and PME are repeatedly found to play critical roles in binding with BLP under different simulation conditions. In particular, the positionally analagous Leu/D-Phe/DAB5 and Leu1/Leu2/DAB5 triads of PMB1 and PME, respectively, as well as the polar DAB1/Thr1/DAB2 triad, located on the branched fatty acid tail of the polymyxins, have been shown to be of repeated importance.

The stereochemistry of the DAB2 residue region of the polymyxins, along with the hydrophobicity of the D-Phe and Leucine residues on their heptapeptide ring has previously been reported to regulate the interactions of polymyxins with the membrane of human kidney cells Jiang et al. (2020c); it is thus interesting to note that these same residues (or their positional analogues in PME) are found to play a particularly important role in the interactions of the polymyxins with BLP within the bacterial cell envelope.

The exact mechanisms by which the structural differences between PMB1 and PME, along with their varied molecular interactions with the ions, osmolytes and proteins within each system, may give rise to the regime-dependent variations in the underlying residue interactions between each polymyxin species and BLP are not well understood as of yet: leaving open a target for future work to further our understanding of the challenges posed to the motion of the polymyxins as they traverse the complex environment of the bacterial cell envelope.

Finally, this work has begun to illustrate how interactions between the polymyxins and one particular osmolyte, OPG, may impact (and be impacted by) interactions with BLP. There are many more chemically diverse species present within the periplasm, both *in vivo* and already included within the simulation performed for this work; and thus it is likely that the true picture of polymyxin interactions within this space is even more complex than has been discussed in this chapter. However, these simulations have certainly highlighted that the modes of interaction of polymyxin antibiotics with BLP are impacted by the nature of the other species present in the surrounding environment. It is therefore imperative that any future studies of such antibiotics, and likely antibiotics in general, within the periplasm involve a more nuanced consideration of the periplasmic composition so as to ensure that any observed behaviour *in silico* may be considered to be an accurate representation of the true behaviour of the *in vivo* system.

Chapter 6

Polmyxin B1 insertion into the *E. coli* inner membrane

"We live in a world out of equilibrium - a nonequilibrium world."
(Todd and Daivis (2017))

6.1 Introduction

Whilst there is a general consensus that the polymyxins induce cell lysis *via* action on the inner membrane, the exact mechanism by which this is achieved is not well understood. Some insights have been obtained through MD studies, leading to three possible models of polymyxin action: barrel-stave insertion Hancock (1997); Shai (1999); Dupuy et al. (2018), carpet insertion Shai (1999); Deris et al. (2014); Berglund et al. (2015) and polymyxin-mediated phospholipid exchange between the outer and inner membranes Clausell et al. (2007); Cajal et al. (1996).

The reported ability for individual polymyxin molecules to penetrate the IM has been inconsistent across these studies, with some results favouring the spontaneous insertion of individual polymyxins into the membrane core Berglund et al. (2015) whilst others indicate the prevalence of polymyxin aggregation at the membrane surface Fu et al. (2020).

One of these studies, performed by the Khalid research group Berglund et al. (2015), investigated the interactions between PMB1 and both the inner and outer membranes of *E. coli*. The OM simulations were performed using two different membrane models; one that used rough LPS (ReLPS) - a form of LPS without its outer core and O-antigen regions (see Fig. 1.3), and another that used just the lipid A region - ignoring the polysaccharide region entirely.

Simulations of PMB1 in proximity to an outer membrane composed of ReLPS found that the charged DAB residues of PMB1 readily interacted with the LPS sugars, leading to the majority of PMB1 molecules binding to the membrane within the first 200 ns of simulation. Spontaneous insertion of the acyl tails of PMB1 into the hydrophobic region of the membrane was not observed. PMB1 molecules were instead seen to aggregate in the head group region of the LPS inner core forming interactions between the positively charged DAB residues and the negatively charged phosphate ions in the LPS head groups. Once within the head group region, multiple PMB1 molecules were seen to form micelle-like structures with their acyl tails facing towards the center of the aggregate. In this state, the probability of observing interactions between the acyl tails of different PMB1 molecules at any given time was found to be $\sim$ 75 %, indicating the prevalence of these hydrophobic interactions and how they played a vital role in aggregation.

The formation of similar PMB1 aggregates at the OM surface was observed in simulations using an OM model comprised entirely of the lipid A portion of LPS. In this model, however, spontaneous insertion of the acyl tails of PMB1 into the hydrophobic region of the membrane was observed for one molecule $\sim 1\ \mu$s after the formation of the PMB1 aggregate. Crucially, the charged DAB residues of the inserted molecule and other nearby PMB1 molecules were observed to interact with the negatively charged phosphate groups on the membrane surface. These interactions acted to push apart the phosphate groups of adjacent lipids, facilitating the insertion of the acyl tails of PMB1 into the membrane core. This highlights how the cooperative action of neighbouring PMB1 molecules may be a necessary component of their uptake mechanism.

Similar to interactions with the OM, PMB1 interactions with the IM were found to be initially mediated by hydrogen bond formation between the cationic DAB residues of PMB1 and the negatively charged phosphate headgroups of the IM. Subsequent to this initial contact, spontaneous insertion of the acyl tail and D-Phe residues of PMB1 into the hydrophobic core of the membrane were observed and, within 100 ns of binding to the membrane surface, the hydrophobic acyl tails of all PMB1 molecules were found to be inserted into the membrane core. Unlike the OM models, there was no obvious formation of PMB1 aggregates at the surface of the IM; with the peptides appearing to insert into the membrane as monomers and in some cases penetrating deep enough so as to contact the lipids in the opposing leaflet of the IM.

Further investigation into the nature of polymyxin aggregation is thus of significant interest for its potential to provide deeper insight into the specific uptake mechanism of these antibiotics into the bacterial cell and further, once inside the cell, determining the propensity for polymyxins to aggregate at the surface of the IM may significantly alter our view of the insertion of these antibiotics into the cytoplasmic membrane and thus the efficacy of their antimicrobial action.

6.2 Methods

6.2.1 Simulation Protocol

All simulations in this chapter were performed using the GROMACS 2021.2 molecular dynamics package Abraham et al. (2015), utilising the Martini2.2 force field de Jong et al. (2013). A constant temperature of either 310 K or 350 K, depending on the system, was maintained using the velocity rescale thermostat Bussi et al. (2007) with a time constant of 1 ps. During production MD, the pressure of each system was maintained at 1 atm using the Parrinello-Rahman barostat Parrinello and Rahman (1981) with a time constant of 1 ps. To ensure system stability during equilibration, the Berendsen thermostat Berendsen et al. (1984) was used during NPT equilibration phases with a time constant of 1 ps. Long-range electrostatics were treated using the reaction field method Tironi et al. (1995). The short-range electrostatic and van der Waals cut-offs were both set to 1.1 nm.

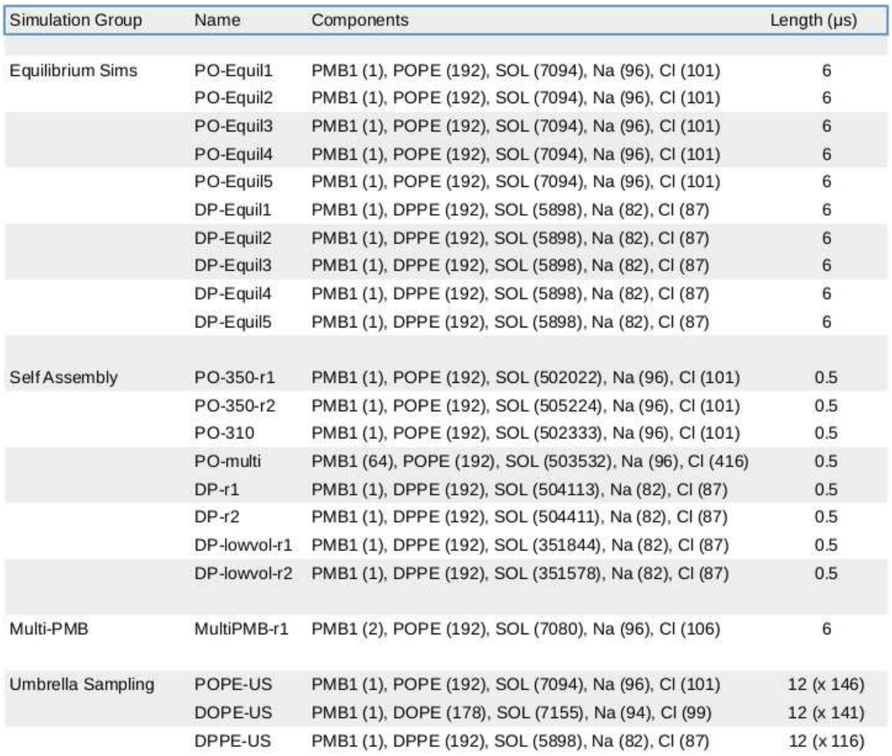
For any replicate simulations; new initial configurations of all molecules were generated, along with re-solvation and ionisation of each system before being passed through the relevant equilibration and production MD phases. The initial velocities of all atoms were modified between replicates, generated from a random seed at the start of each NVT equilibration phase so as to ensure the unbiased sampling of the simulation phase space.

A summary of all simulations performed in this work is presented in Table 6.1.

6.2.2 Umbrella Sampling

In order to determine the PMF curve profiles for polymyxin insertion into the IM, an umbrella sampling (US) procedure was performed on PMB1 insertion into three models of the IM with different lipid compositions; pure POPE, pure DOPE or pure DPPE. These lipids were chosen due to the equivalence of their phosphate headgroups and the slight variations in saturation of their lipid tails: DPPE has a carbon-carbon double bond in both of its tails, POPE in just one of its tails and, finally, DOPE is fully saturated in both of its tails.

For each membrane model, an initial steered MD (SMD) procedure was utilised to pull a single PMB1 molecule through the membrane. This was done by applying a pulling force of 1000 kJ mol^{-1} nm^{-2} at a rate of 0.01 nm s^{-1} along a collective variable ξ; in this case defined simply as the distance between the center of mass (CoM) of PMB1 and the CoM of the membrane, measured along the membrane normal, *i.e.* the z-axis. Once a pulling trajectory had been generated, system configurations were extracted at intervals of 0.1 nm along the collective variable. Additional configurations



Simulation Group	Name	Components	Length (μs)
Equilibrium Sims	PO-Equil1	PMB1 (1), POPE (192), SOL (7094), Na (96), Cl (101)	6
	PO-Equil2	PMB1 (1), POPE (192), SOL (7094), Na (96), Cl (101)	6
	PO-Equil3	PMB1 (1), POPE (192), SOL (7094), Na (96), Cl (101)	6
	PO-Equil4	PMB1 (1), POPE (192), SOL (7094), Na (96), Cl (101)	6
	PO-Equil5	PMB1 (1), POPE (192), SOL (7094), Na (96), Cl (101)	6
	DP-Equil1	PMB1 (1), DPPE (192), SOL (5898), Na (82), Cl (87)	6
	DP-Equil2	PMB1 (1), DPPE (192), SOL (5898), Na (82), Cl (87)	6
	DP-Equil3	PMB1 (1), DPPE (192), SOL (5898), Na (82), Cl (87)	6
	DP-Equil4	PMB1 (1), DPPE (192), SOL (5898), Na (82), Cl (87)	6
	DP-Equil5	PMB1 (1), DPPE (192), SOL (5898), Na (82), Cl (87)	6
Self Assembly	PO-350-r1	PMB1 (1), POPE (192), SOL (502022), Na (96), Cl (101)	0.5
	PO-350-r2	PMB1 (1), POPE (192), SOL (505224), Na (96), Cl (101)	0.5
	PO-310	PMB1 (1), POPE (192), SOL (502333), Na (96), Cl (101)	0.5
	PO-multi	PMB1 (64), POPE (192), SOL (503532), Na (96), Cl (416)	0.5
	DP-r1	PMB1 (1), DPPE (192), SOL (504113), Na (82), Cl (87)	0.5
	DP-r2	PMB1 (1), DPPE (192), SOL (504411), Na (82), Cl (87)	0.5
	DP-lowvol-r1	PMB1 (1), DPPE (192), SOL (351844), Na (82), Cl (87)	0.5
	DP-lowvol-r2	PMB1 (1), DPPE (192), SOL (351578), Na (82), Cl (87)	0.5
Multi-PMB	MultiPMB-r1	PMB1 (2), POPE (192), SOL (7080), Na (96), Cl (106)	6
Umbrella Sampling	POPE-US	PMB1 (1), POPE (192), SOL (7094), Na (96), Cl (101)	12 (x 146)
	DOPE-US	PMB1 (1), DOPE (178), SOL (7155), Na (94), Cl (99)	12 (x 141)
	DPPE-US	PMB1 (1), DPPE (192), SOL (5898), Na (82), Cl (87)	12 (x 116)

TABLE 6.1: Summary of all simulations performed for this work. Bracketed numbers in the components and length sections give the number of each component and the number of simulations performed, respectively.

in the vicinity of poorly sampled regions of the collective variable were extracted manually: these regions corresponded to positions at which the PMF was particularly large (*e.g.* near the membrane center) or at which large conformational changes were occurring within the system (*e.g.* when PMB1 was entering/exiting the interfacial-inserted binding mode). Each configuration was then passed through a brief 1 ns equilibration phase in the NPT ensemble. During this phase, the collective variable of the peptide was restrained with a force of 1000 kJ mol^{-1} nm^{-2}. This restraining force was increased to 2000 kJ mol^{-1} nm^{-2} for certain manually extracted configurations located close to the membrane center.

Each configuration then underwent a 12 μs production run during which the restraint on the collective variable of the peptide was maintained, these simulations are referred to as US windows from here on. It has been reported that compounding systematic errors in the PMF profile of small molecule permeation through lipid bilayers may be caused by the residual disturbance to the membrane structure imparted during the SMD procedure used to generate the initial structures Lee et al. (2016). Furthermore, these defects have been reported to persist on $\mathcal{O}(\mu s)$ timescales in US windows for which the inserted molecule is located close to the center of the bilayer Neale and Pomès (2016). During the simulations performed for this work, PMB1 molecules located close to the bilayer center were found to remain in contact with the phosphate beads of the leaflet through which they penetrated the membrane.

Dissociation of PMB1 from these phosphate beads occurred after up to 2 μs of simulation, with PMB1 subsequently forming interactions with the phosphate beads of the opposing leaflet. This “flipping” of PMB1 between leaflets, and the potential for long-lasting defects in the membrane structure due to the initial SMD procedure, motivated us to discard the first 6 μs of each US window during analysis in order to ensure the peptide occupied an equilibrium position within the membrane and that the membrane itself had sufficient time to equilibrate around the inserted peptide.

PMF curve profiles were then generated for each membrane model using the WHAM algorithm provided by the *gmx wham* utility. Full/half PMFs were calculated using 300/150 bins, respectively. A convergence tolerance of $1 \times 10^{-}6$ was used in all cases. Errors were estimated with 30 bootstraps and the value of the PMF was set to 0 in bulk water.

6.2.2.1 DPPE Simulation Details

Initial simulations of the DPPE membrane model at 310 K resulted in a gel-phase membrane due to the simulation temperature being below the experimentally determined melting point of DPPE, 336.15 K.

Further test simulations at 337 K were performed, resulting once again in a gel-phase membrane. It is known that the MARTINI forcefield does not completely capture the lipid structure and dynamics seen in atomistic simulations due to the projection of the system onto a coarse-grained mapping. This leads to issues of accuracy and temperature transferability between atomistic and CG systems due to the incorrect enthalpy–entropy decomposition within the potential functions of the model Jarin et al. (2021).

This effect, whereby the MARTINI forcefield overestimates enthalpy to compensate for an intrinsically low entropy due to reduced degrees of freedom, has previously been reported to affect the ability of CG membranes to sample the ripple-like membrane phase Sharma et al. (2021).

The overestimation of enthalpy is equivalent to the overestimation of the heat content of the system, and thus it is not surprising that simulation of DPPE at a temperature equal to its experimental melting temperature does not express sufficient heat content to drive the membrane into the liquid-disordered phase. Further simulations performed at a temperature of 350 K resolved this issue, resulting in the desired liquid-disordered membrane phase.

6.2.3 Equilibrium Simulations

Unrestrained, equilibrium MD simulations were performed on the output configurations of a subset of the original US windows. For each membrane model, 5 windows were chosen for which the PMB1 molecule resided: in bulk water below the lower leaflet; in bulk water above the upper leaflet; inserted into the lower leaflet headgroup; inserted into the upper leaflet headgroup or fully inserted in the hydrophobic region of the membrane center. Once specific umbrella windows corresponding to these 5 configurations had been chosen, the final structure of each window underwent a further 6 μs of unrestrained production MD. Here, the lower leaflet refers to the leaflet through which PMB1 initially penetrated into the membrane during the SMD phase of the US simulation protocol. Conversely, the upper leaflet refers to the membrane leaflet through which PMB1 exited the membrane center during the SMD phase of the US protocol.

6.2.4 Multi PMB1 Simulations

In order to construct a system with multiple PMB1 molecules within the interior of the IM, the structures generated during the US simulations, described in Section 6.2.2, were utilised. From the US simulations of the POPE membrane, one window was selected in which the PMB1 molecule was situated close to the membrane center. The coordinates of this system were extracted from the final frame of the US window, thus ensuring that the membrane had been allowed to equilibrate around the inserted PMB1 molecule for 12 μs.

Once this initial configuration had been obtained, a second PMB1 molecule was placed in the bulk water region directly beneath the original PMB1 using the standard *gmx insert-molecules* utility. Any water molecules overlapping with the newly included PMB1 were removed, and 5 additional chloride ions were added to account for the additional +5 e charge conferred to the system by the additional PMB1 molecule.

The system was passed through a brief energy minimisation protocol to remove any steric clashes resulting from the PMB1 insertion procedure. During this step, a 1000 kJ mol^{-1} nm^{-2} restraining potential was applied to the x, y and z coordinates of all atoms in the membrane and both PMB1 molecules. This restraint was required so as to hold the original PMB1 molecule at the desired position within the hydrophobic region of the membrane center without inducing unwanted deformations to the membrane structure, whilst ensuring that the two PMB1 molecules remained aligned in the z-axis.

After energy minimisation, the system was passed through a 1 ns NVT equilibration phase to impart velocities to the atoms and raise the system temperature to 310 K.

During this step, the coordinate restraints used in the EM phase were removed. Umbrella potentials were applied to each PMB1 molecule, individually, to restrain the distance between their CoM z-component and that of the membrane; *i.e.* to maintain their position along the membrane normal.

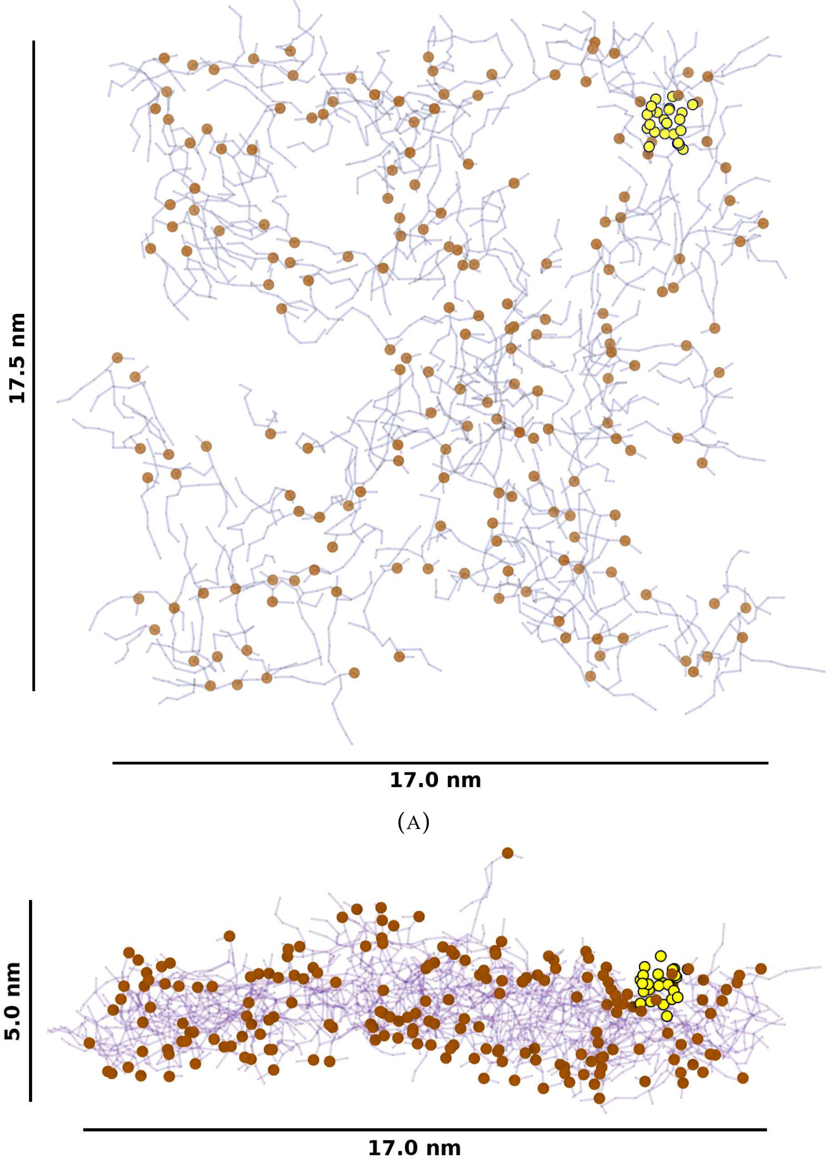
A steered MD procedure was then applied to the resulting system to pull the new PMB1 molecule through the membrane. The umbrella potential on the original, already inserted, PMB1 molecule was maintained during this step to hold it in the desired position within the center of the membrane. A pulling force of 1000 kJ mol^{-1} nm^{-2} was applied to the new PMB1 molecule along the membrane normal with a pulling rate of 0.01 nm s^{-1} for 1.1 ns; this duration was sufficient to ensure that the new PMB1 molecule had been pulled through both leaflets of the membrane.

A structure file was then extracted from one frame of the pulling trajectory in which both PMB1 molecules were situated within the hydrophobic region at the center of the membrane. This structure was then simulated for 6 μs, with the inclusion of umbrella potentials restraining the position of each PMB1 molecule along the membrane normal, with respect to the CoM of the membrane. This step was necessary to ensure that the membrane lipids were correctly packed around the inserted PMB1 molecules after any deformation induced during the insertion of the second PMB1 molecule. After this, the system underwent a further 6 μs of unrestrained production MD.

6.2.5 Self Assembly Simulations

Initial configurations were chosen for the POPE and DPPE membrane models that contained a PMB1 molecule situated in the bulk water phase outside of the membrane; this criteria was chosen so as to mitigate any potential hysteresis effects resulting from initial lipid conformations favouring interactions with PMB1. These configurations were selected from the final structures of the US window simulations discussed in Section 6.2.2. A multi-step alchemical procedure was performed, using the GROMACS free energy calculation functionality, to dissolve the lipid bilayers and generate random lipid distributions and configurations for each replicate simulation.

Each system was first ”exploded” by gradually turning off all inter- and intra-molecular electrostatic and van der Waals forces over the course of 10 ns. This initial phase resulted in the expansion of the periodic box, increasing the separation between all constituent molecules and leading to extensive pore formation throughout the membrane (Figure 6.1a). The overall structure of the membrane was retained, however, with the phospholipids still occupying a central region of the simulation domain (Figure 6.1b) and thus a further ”diffusion” phase was required in order to generate sufficiently random initial lipid configurations from which to reassemble the system.

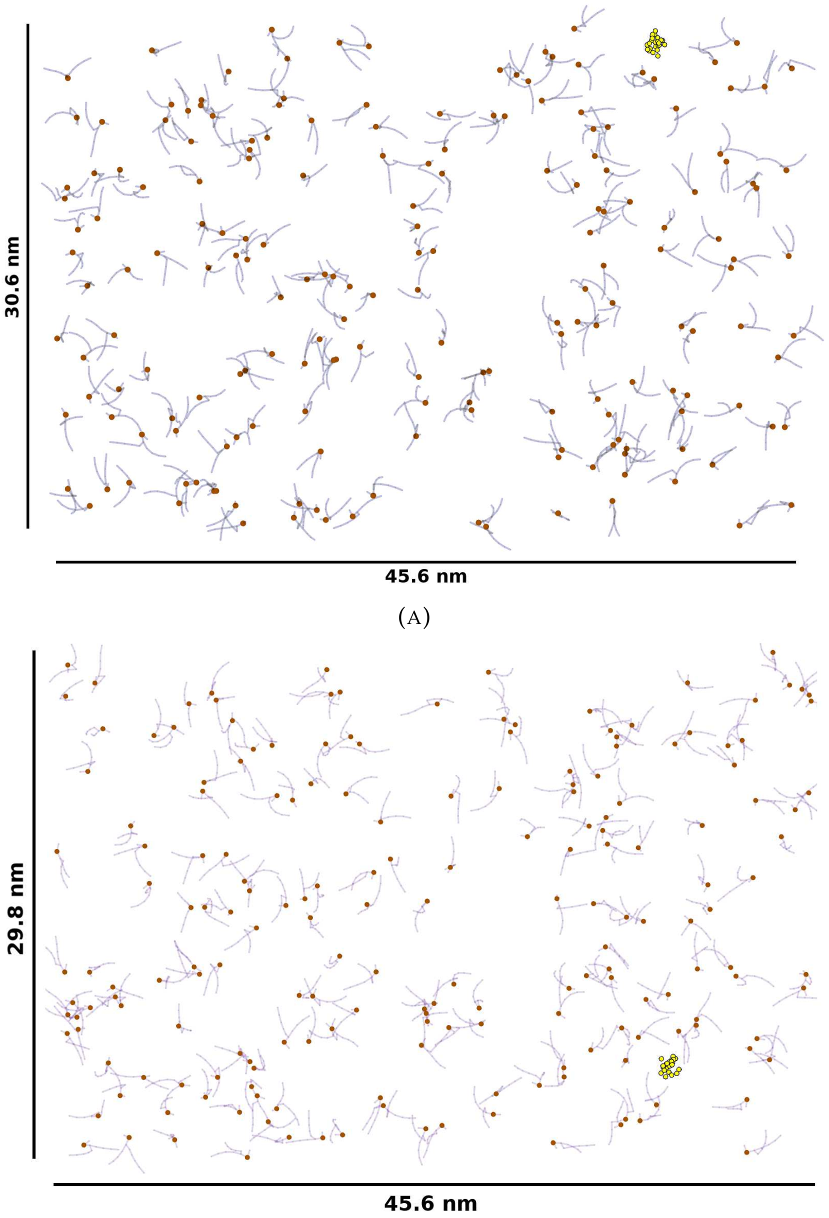


(A)

(B)

FIGURE 6.1: Top-down (A) and side-on (B) views of system configuration extracted from the initial explosion phase of the DP-lowvol-r2 simulation. PMB1 is represented by yellow vdW spheres, headgroup phosphate beads by orange vdW spheres and DPPE lipids as light blue lines.

To perform the diffusion phase, the system was subjected to a further 10 ns of unrestrained MD whilst all forces remained absent. This allowed sufficient time for the molecules to diffuse around the periodic box, resulting in a random distribution of lipids throughout the simulation domain with non-correlated conformations (Figure 6.2). To ensure that each simulation regime had comparable lipid concentrations during self-assembly, structures were extracted from the diffusion simulations at which each system had a volume of $\sim 6 \times 10^4$ nm^3 with the exception of two simulations of the DPPE model which were performed at a lower volume of $\sim 4 \times 10^4$ nm^3.



(A)

(B)

FIGURE 6.2: Top-down (A) and side-on (B) views of the energy minimised system configuration extracted from the diffusion phase of the DP-lowvol-r2 simulation. PMB1 is represented by yellow vdW spheres, headgroup phosphate beads by orange vdW spheres and DPPE lipids as light blue lines.

The removal of all non-bonded forces during the diffusion phase inherently caused the separation of beads within each molecule to increase. For any pairs of beads that had been parameterised with the inclusion of non-bonded pair interactions, this led to a potential failure condition when non-bonded forces were reapplied to the system. If such a pair of beads had drifted too far from their equilibrium separation, then the interaction distance exceeded the MD algorithm pair list cutoff and the interaction was therefore not accounted for by the integrator; leading to fatal errors at the onset of simulation. In order to correct for these improper conformations, the output structure from the diffusion phase was subjected to an energy minimisation procedure with only van der Waals forces enabled. This allowed each molecule in the system to relax into a more energetically favourable conformation, reducing the exaggerated separation between beads.

The initial protocol used to reconstruct the diffuse systems involved a 100 ns alchemical procedure, during which the forces were gradually turned back on, followed by a 100 ns equilibrium MD phase to allow for reassembly. A damped harmonic oscillation in the volume of the system was observed during the alchemical phase, which reached a temporary steady state before the simulation crashed after 90 ns (Figure 6.3).

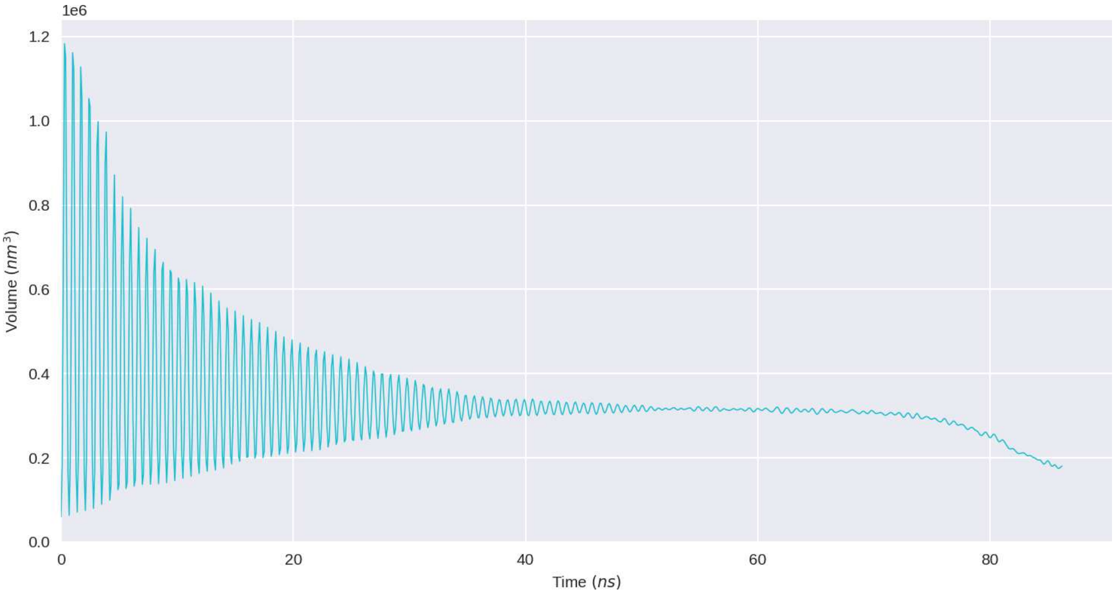


FIGURE 6.3: System volume during failed 100 ns alchemical assembly procedure.

In order to avoid the late stage failure of the simulation during the alchemical procedure, a subsequent attempt to reassemble the system was performed using a 1 ns alchemical phase during which the forces were reapplied to the system much more rapidly. A harmonic oscillation in the system volume was observed once again, however it did not reach the temporary steady state observed during the previous 100 ns alchemical protocol. The 1 ns alchemical phase ran to completion, and the resulting system configuration was passed into the equilibrium MD phase. The harmonic volume oscillation continued for 7.5 ns during the equilibrium MD phase, after which the box volume decreased linearly for a further 4 ns before reaching a temporary steady state (Figure 6.4). Approximately 25 ns later the system rapidly expanded in one dimension, resulting in the simulation crashing. Aggregates formed during this phase comprised of lipids, PMB1 and water molecules; the inclusion of water molecules in these aggregates pointed to one potential cause of failure in the simulation. As these aggregates grew larger they attracted ever more water molecules, decreasing the density in the surrounding solvent and eventually resulting in a

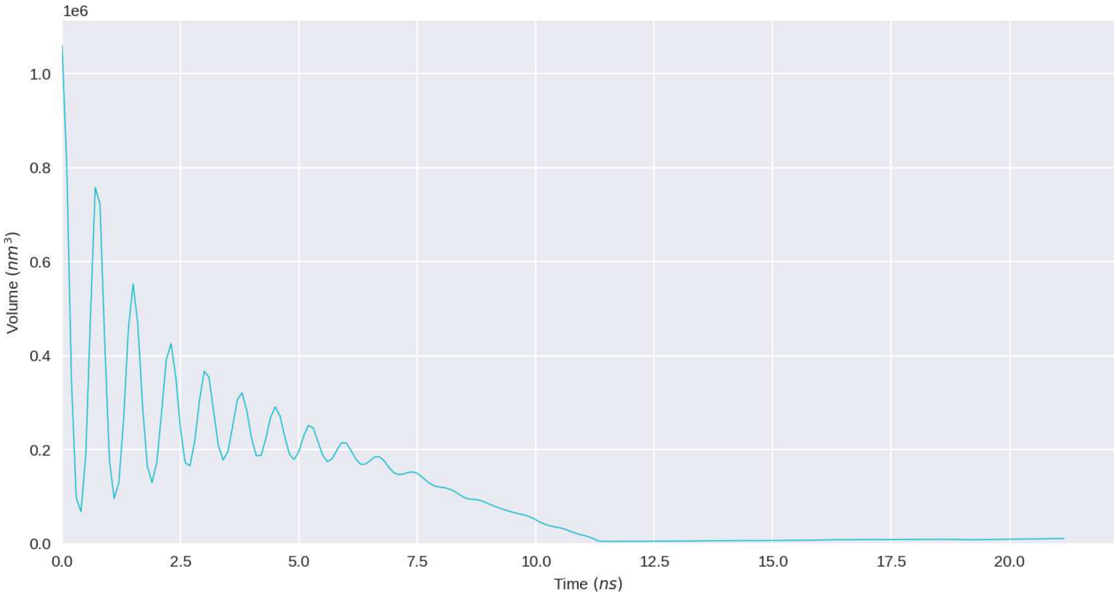


FIGURE 6.4: System volume during failed equilibrium MD reassembly phase from the 1 ns alchemical protocol.

vacuum-like simulation environment with near-infinite compressibility and thus no restoring forces against the growth of system size in one coordinate.

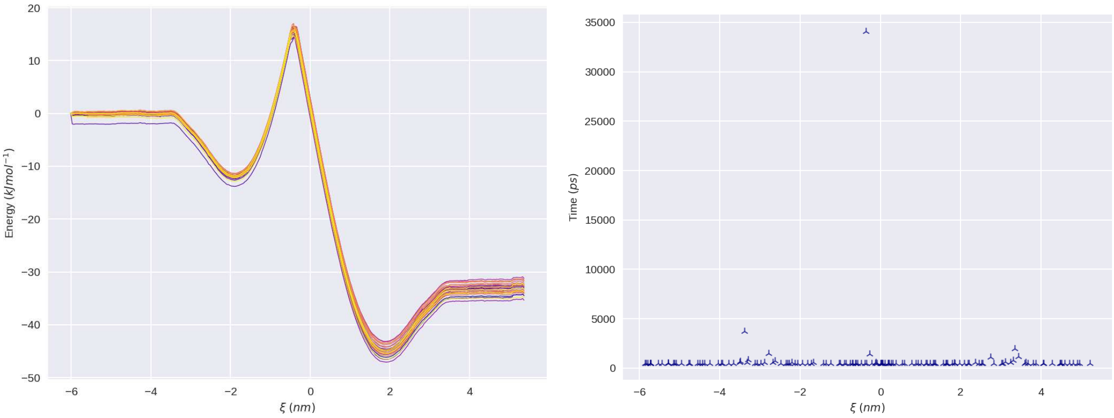
This motivated a different approach to the assembly phase; instead of attempting to reverse the alchemical procedure, it was found that resolvation and ionisation of the diffuse system structure was a more robust option to avoid the formation of vacuum-like conditions during self-assembly.

To this end, water molecules were added to the energy minimised output structures from each diffusion phase. The resolvated systems were then subjected to a 10 ns NVT equilibration phase with all intra- and inter-molecular forces enabled. This was followed by a 50 ns NPT equilibration phase using the Berendsen barostat and isotropic pressure coupling. The equilibrated systems were then simulated for 500 ns unrestrained production MD using the Parinello-Rahman barostat, maintaining the isotropic pressure coupling regime. Isotropic pressure coupling was required in all self-assembly simulations to maintain a roughly cubic periodic domain, avoiding the rapid single-dimension expansion of the systems that had previously caused simulation failure.

6.3 Results and Discussion

6.3.1 Umbrella Sampling

Initial attempts to construct PMF profiles for the insertion of PMB1 into the DOPE membrane resulted in bootstrapping profiles that appeared well converged with respect to to one another, but that did not exhibit the expected zero values of free energy when PMB1 was situated in the bulk water phase on both sides of the membrane (Figure 6.5a). This asymmetry highlighted that the analysis had not converged onto a physically reasonable PMF profile and instead the low standard deviation between bootstrapping profiles was an indication that each set of bootstrapped histograms represented a set of states that existed in the region of phase space local to that already explored by the respective US windows.



(A) Bootstrapped PMF profiles.

(B) US window integrated autocorrelation times.

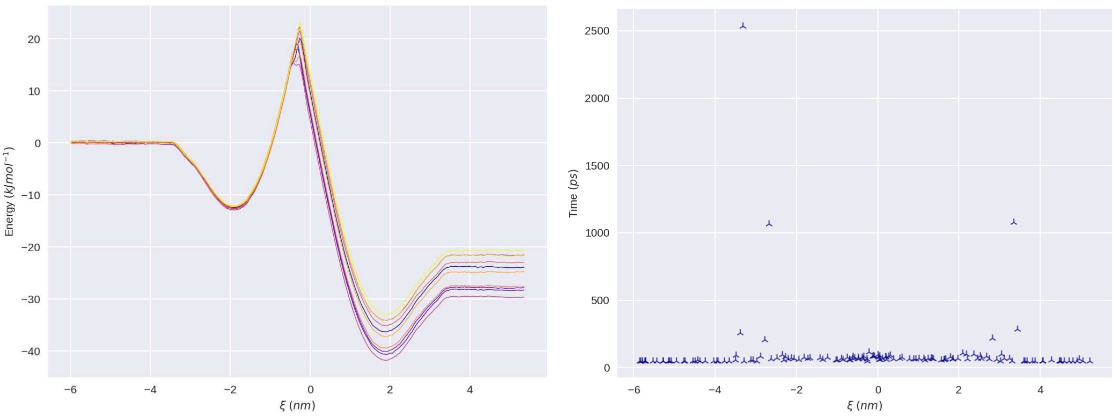
FIGURE 6.5: WHAM analysis of PMB1 insertion into DOPE membrane. Generated using early protocol that did not include the additional 6 μs equilibration phase.

This poor sampling was also evidenced by high IACT values in certain umbrella windows (Figure 6.5b). Such long autocorrelation times ($\sim$ 35 ns) indicated that these windows were not exploring an extensive region of the system's allowed phase space and were instead trapped in a local minimum; sampling similar phase states repeatedly, resulting in autocorrelation over long timescales.

Subsequent protocols allowed for much longer equilibration at the start of production MD. This was necessary to allow time for the perturbations arising from the initial steered MD phase to settle out.

As the PMB1 molecule was pulled through the membrane, phosphate beads from the membrane headgroups were dragged with it; caused by the strong polar interactions between the peptide ring of PMB1 and the headgroup phosphate beads. This membrane invagination was present in the initial configurations of all US windows for which the PMB1 molecule was situated within the center of the membrane.

The first 6 μs of each US window was subsequently excluded in an effort to account for any long timescale relaxation that the membrane required to equilibrate around the newly inserted PMB1 molecule, however under these conditions, convergence was still not achieved (Figure 6.6a). The IACT values obtained using this protocol were, however, orders of magnitude lower than those obtained without the 6 μs



(A) Bootstrapped PMF profiles.

(B) US window integrated autocorrelation times.

FIGURE 6.6: WHAM analysis of PMB1 insertion into DOPE membrane. Generated using 6 μs equilibration protocol

equilibration phase (Figure 6.6b). Further to this, the standard deviation between bootstrapped profiles was seen to increase at ξ values corresponding to PMB1 positions beyond the membrane center; indicating that the US histograms used to generate the bootstrapped ξ distributions were sampling a broader region of the system's phase space, and hence the reconstructed data exhibited a wider range of possible PMF pathways. Thus the sampling in this protocol may be considered more indicative of the system's true dynamics than without the additional equilibration phase, however it was still insufficient to obtain a properly converged PMF profile.

The lack of overall convergence was likely due to the limitations of the implementation of the US protocol itself. Whilst the position of the PMB1 molecule was correctly restrained along the membrane normal, with respect to the CoM position of the membrane; the membrane itself was not restrained during the initial SMD phase that generated the initial system configurations. As such, any US window initialised from a SMD configuration in which the inner leaflet headgroups had been dragged into the membrane center with the PMB1 molecule may not be considered to be correctly sampling the permeation of PMB1 through the hydrophobic core of the membrane, but are instead sampling the PMF pathway for the progressive membrane curvature induced by the forced passage of PMB1 through the inner leaflet.

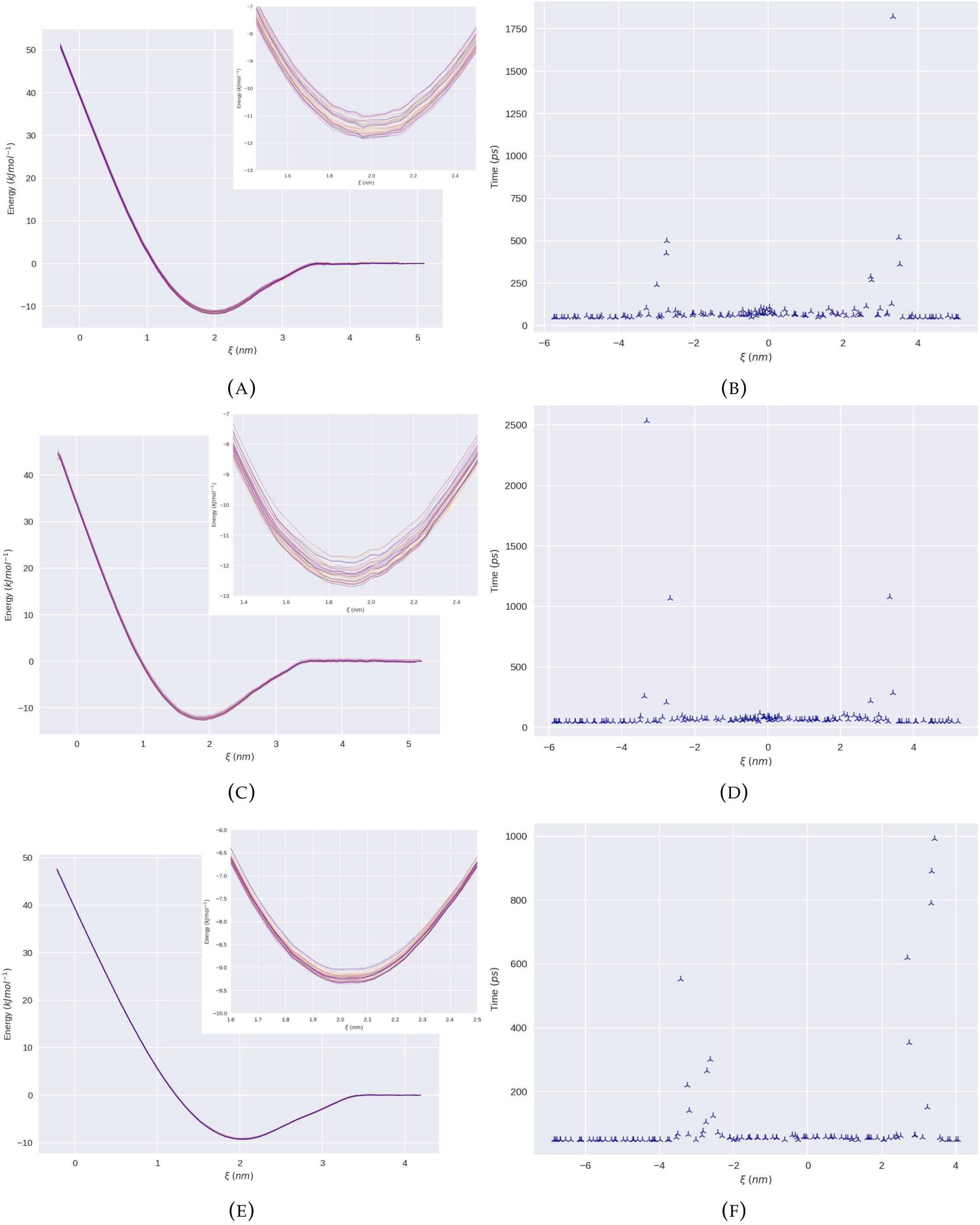
Furthermore, but perhaps of less importance than the aforementioned concerns; the use of a single SMD trajectory to generate the initial configurations of all US windows inherently limited the sampling to this single insertion pathway. Whilst the US protocol allows for each window to explore the phase space surrounding its initial configuration, it may be the case that energetic barriers orthogonal to the chosen collective variable restrict the available phase space of each window, hindering the sampling efficiency. A better approach would be to use multiple SMD trajectories to generate a set of initial configurations for each value of ξ, broadening the scope of phase space exploration at each position and reducing the risk of single configurations becoming restricted to some unseen energetic minimum.

The poor sampling of PMB1 in the membrane center thus motivated us to restrict the analysis of the PMF profiles to the region of the collective variable domain that corresponded to PMB1 positions before the peak of the central energy barrier. A well converged profile across all bootstrap iterations was obtained with low IACT values in all US windows (Figure 6.7). This result indicated that the initial energy minimum was well sampled and that the lack of an initial energetic barrier to PMB1 as it enters this state is reasonably certain.

This follows the so-called Type 3 insertion described in literature Neale and Pomès (2016) and exemplified by the insertion of n-propylguanidinium into DOPC, POPC and DPPC bilayers: n-propylguanidinium is an amphipathic molecule with a cationic guanidino moiety and a short hydrophobic acyl tail. It was reported Neale et al. (2011) that as n-propylguanidinium approaches the bilayer from the aqueous phase, favorable charge–charge interactions between the cationic guanidino group and the anionic phosphate headgroups of the membrane act to orient the n-propylguanidinium to face the bilayer. Its orientation is then reversed as ξ decreases toward the global free energy minimum and the solute becomes embedded as a mini-detergent in the bilayer, which invaginates slightly around it. As is discussed in the following sections; this behaviour closely matches the insertion of PMB1 into the membrane *via* the interfacial-inserted binding mode observed in the equilibrium MD and self-assembly simulations performed for this chapter. The similarity between these insertion mechanisms provides support for the validity of the general profiles of the PMF curves presented in Figure 6.7.

The *gmx cluster* utility was used to perform a structural clustering procedure on PMB1 conformations observed during representative US windows of each membrane model that were located at the position of the initial energy minimum as PMB1 inserted into the membrane. The gromos clustering algorithm Daura et al. (1999) was used with an RMSD cutoff of 0.20 nm; this cutoff was determined to provide an adequate trade off between minimization of the number of structural clusters and minimization of the cutoff value itself.

The primary, most populated, structural clusters obtained from this analysis corresponded to the PMB1 molecule occupying the so-called membrane interfacial-inserted state; in which the PMB1 molecule adopted a folded amphipathic conformation at the interface between the surface of the membrane and the bulk water phase. In this position, the hydrophobic residues of PMB1 were inserted into the core of the membrane whilst its polar and charged residues were exposed to the external aqueous environment (Figure 6.8). Indeed, visual inspection of the US window trajectories found PMB1 to invariably occupy this folded amphipathic,



(A) (B)

(C) (D)

(E) (F)

FIGURE 6.7: Bootstrapped half PMF profiles for A) POPE, C) DOPE and E) DPPE membrane models. A magnified region around the energy minimum of each PMF profile is also shown. Integrated autocorrelation times for all US windows of the B) POPE, D) DOPE and F) DPPE membrane models are also shown.

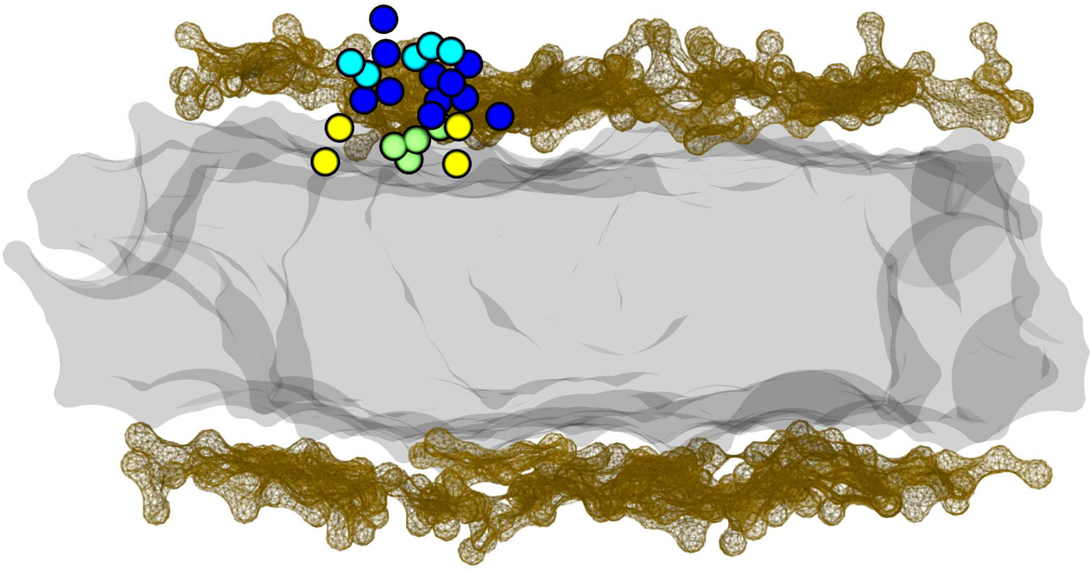


FIGURE 6.8: Representative structure of the primary cluster of PMB1 conformations when bound to the DOPE membrane at the location of the initial energy minimum as PMB1 penetrates into the membrane. PMB1 is found to adopt a folded-amphipathic conformation at the interface between the membrane surface and bulk water phase. The membrane headgroup region is represented by the orange wireframe surface, the hydrophobic membrane core by the transparent grey surface. The PMB1 molecule is represented by Goodsell spheres with: non-cyclised DAB residues in blue, cyclised DAB and Thr residues in cyan, acyl tail and Leu residues in yellow, D-Phe residue in lime green.

interfacial-inserted state throughout the entire duration of US windows corresponding to the location of the initial energy minimum.

Regardless of the issues with exact convergence, the exceedingly large free energy values obtained at the peak of the central barrier indicates that it is unlikely that any individual PMB1 molecule is able to penetrate the membrane without some secondary energetic contribution. This provides some initial evidence to support the theory reported in recent literature that the true mechanism for PMB1 disruption of the IM is not as an individual molecule, but instead requires the aggregation of multiple PMB1 molecules Fu et al. (2020).

6.3.2 Equilibrium Simulations

In the PO-Equil1 simulation, the PMB1 molecule was initially situated in bulk water close to the surface of the inner leaflet of the IM (< 1 nm). After 17 ns, the cationic DAB residues of the PMB1 molecule came into contact (separation < 5.3 Å) with the phosphate beads of the membrane. 4 ns after this initial contact, the hydrophobic tail of the PMB1 molecule inserted into the hydrophobic region of the membrane. The PMB1 molecule remained in this interfacial-inserted state (Figure 6.9) for the rest of the 6 μs simulation.

Similarly, in the DP-Equil1 simulation, the PMB1 molecule was initially situated in bulk water close to the surface of the inner leaflet of the IM. Within 1 ns, the DAB residues of PMB1 contacted the membrane headgroup region, however this interaction was transient, with the PMB1 molecule dissociating from the membrane after a further 3 ns of simulation. The PMB1 molecule freely diffused throughout the bulk water phase for a further 30 ns before once more coming into contact with the headgroup region of the inner leaflet. This contact was initially mediated by the polar residues of PMB1, however the molecule rapidly (< 1 ns) adopted the folded amphipathic interfacial-inserted state where it remained for the remaining duration of the 6 μs simulation.

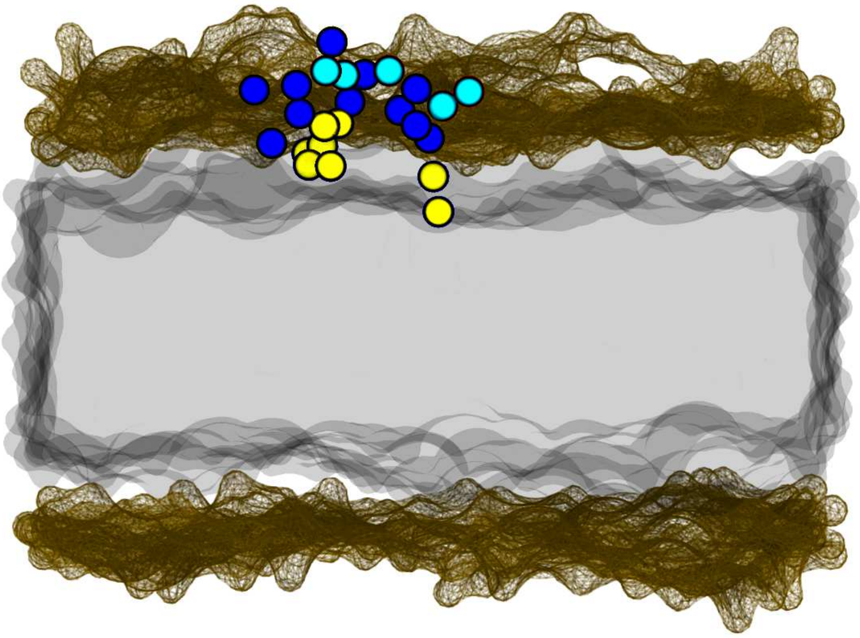


FIGURE 6.9: PMB1 molecule in the folded amphipathic conformation, situated in the interfacial-inserted binding mode with the DPPE membrane. Snapshot taken from the Equil3 simulation. PMB1 is respresented by Goodsell spheres with DAB residues in blue, hydrophobic Leu and tail residues in yellow, hydrophobic D-Phe residue in lime green and polar noncharged residues in cyan. The membrane headgroup region is represented by the orange wireframe surface, the hydrophobic membrane core by the transparent grey surface. The PMB1 molecule is represented by Goodsell spheres with: non-cyclised DAB residues in blue, cyclised DAB and Thr residues in cyan, hydrophobic Lys and D-Phe residues in yellow.

In the PO-Equil2 and DP-Equil2 simulations, the PMB1 molecule was initially situated in the interfacial-inserted state observed in both Equil1 simulations; with the peptide ring associated with the phosphate beads of the membrane headgroup region and the hydrophobic tail inserted into the hydrophobic membrane core. Similar to the Equil1 simulations, PMB1 remained in this interfacial-inserted state for the duration of the 6 μs simulation.

In both the DP-Equil3 and PO-Equil3 simulations, the PMB1 molecule was initially situated at the membrane center. The inner leaflet of the membrane was deformed towards the membrane center allowing for the interaction of the headgroup phosphate beads with the PMB1 molecule. This initial deformation was remnant from the initial SMD phase used to pull the PMB1 molecule through the membrane and was conserved through the US phase due to the applied restraint on the CoM distance between the PMB1 molecule and the membrane. Within 1 ns of equilibrium MD the deformed leaflet returned to a typical planar configuration, parallel to the opposing leaflet. As this deformation settled out, the interaction between the phosphate beads and inserted PMB1 molecule pulled the PMB1 molecule out of the membrane center and into the interfacial-inserted state observed in the Equil1 and Equil2 simulations. The PMB1 molecule remained in this position for the remaining duration of each 6 μs simulation.

In the DP-Equil4 and PO-Equil4 simulation, the PMB1 molecule was situated in the membrane interfacial-inserted state on the outer leaflet (opposing to the Equil2 simulation). Notably, the PMB1 molecule had therefore been pulled through the center of the membrane during the SMD phase prior to the extraction of each configuration from the US simulations. Nevertheless, the PMB1 molecule in each simulation was once again observed to remain in the interfacial-inserted state for the entire duration of both 6 μs simulations.

In the DP-Equil5 and PO-Equil5 simulations, the PMB1 molecule was initially situated in the bulk water region close ($<$ 1 nm) to the outer leaflet of the membrane. After 5 ns / 4 ns of the DP-Equil5 / PO-Equil5 simulations, respectively, the polar residues of PMB1 came into contact with the phosphate beads of the outer leaflet; in both simulations, the PMB1 molecule immediately ($<$ 1 ns) adopted a folded amphipathic conformation, followed by the rotation and insertion of the hydrophobic tail, Leu and D-Phe residues into the hydrophobic core of the membrane.

The adoption of a folded amphipathic conformation prior to the insertion of AMPs into lipid membranes is already noted in the literature Fernández-Vidal et al. (2007); Jiang et al. (2020a,b). This behaviour arises from the hydrophobic moment caused by the polar/non-polar asymmetry of PMB1 and results in its folding upon binding to the membrane surface, followed by a subsequent rotation of the PMB1 molecule so as to expose the polar and charged residues to the aqueous bulk water phase and the hydrophobic residues to the more energetically favourable hydrophobic environment of the membrane core. It is this process that was observed to drive the insertion of the PMB1 molecules into the membrane interfacial-inserted state in the Equil1 and Equil5 simulations of both the POPE and DPPE membrane models (Figure 6.10).

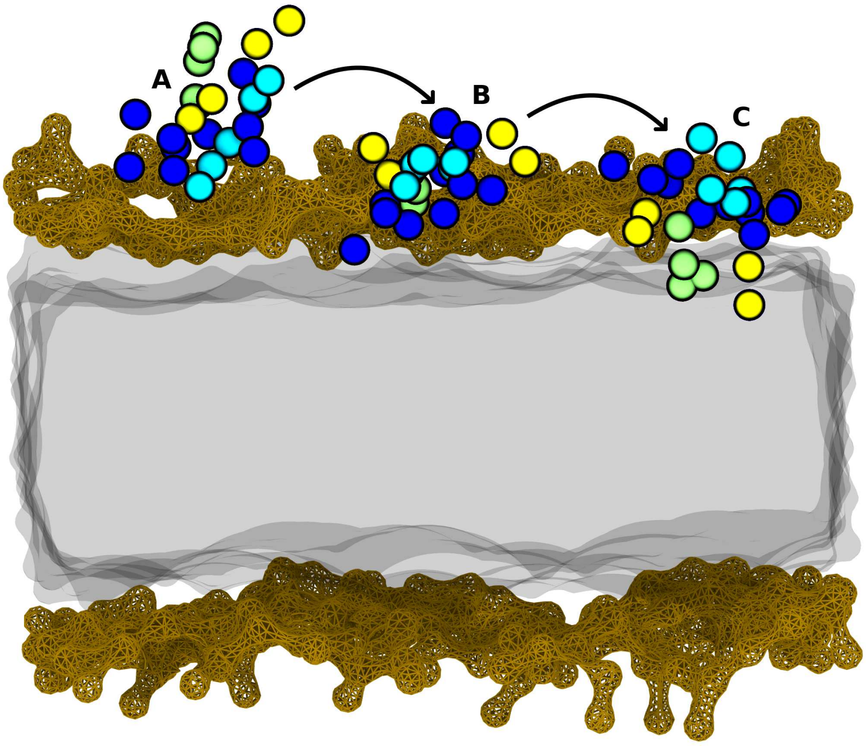


FIGURE 6.10: Rotation of PMB1 molecule into the folded-amphipathic interfacial-inserted state during the PO-Equil1 simulation. A) Initial contact of PMB1 polar residues with membrane headgroup phosphate beads. B) PMB1 molecule maximises contacts between cationic DAB residues and headgroup phosphates. C) PMB1 adopts the folded amphipathic interfacial-inserted binding mode upon insertion of hydrophobic residues into the membrane core. System components have the same representations as in Figure 6.8.

6.3.2.1 PMB1 Conformational Analysis

To further investigate the adoption of the folded amphipathic conformation upon membrane binding, the angle formed between the peptide ring and hydrophobic tail of PMB1 was calculated at each frame of the equilibrium MD simulations of both POPE and DPPE membranes. An equivalent analysis was then performed over a selection of reference US windows in which the PMB1 molecule was restrained within the bulk water phase outside of the membrane. The results of these analyses are presented in Figures 6.11 & 6.12.

The mean PMB1 ring-tail angle calculated across all equilibrium POPE simulations was $89.50 \pm 0.30\ ^\circ$, where the error denotes a 95 % confidence interval on the mean. A standard deviation of $26.22\ ^\circ$ was calculated across these data. In comparison, the results across all of the selected US windows returned a mean of $98.94 \pm 0.44\ ^\circ$ with a standard deviation of $38.57\ ^\circ$.

Similar results were obtained across the DPPE simulations; for which equilibrium simulations returned a mean ring-tail angle of $88.94 \pm 0.31\ ^\circ$ with a standard deviation

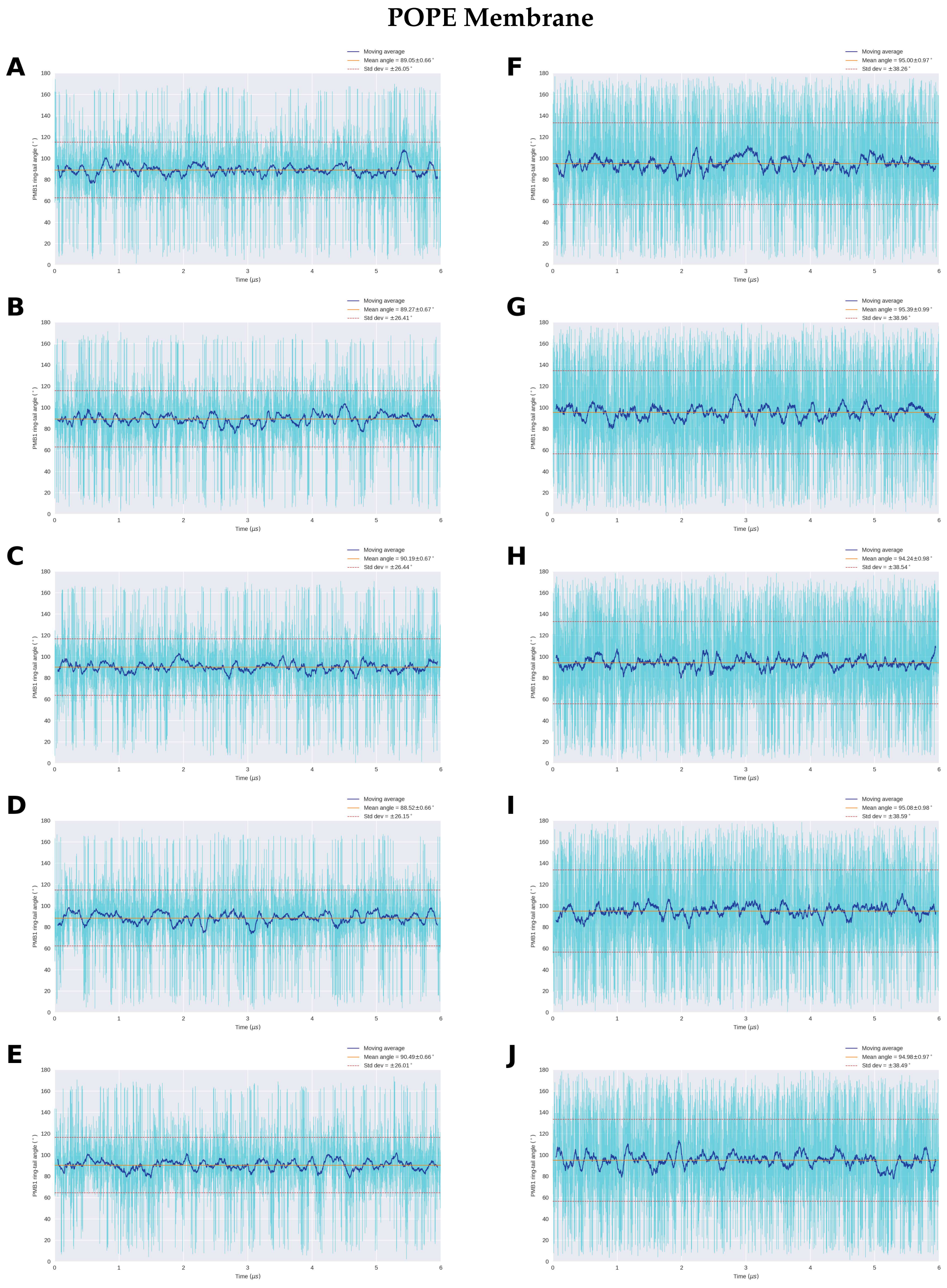


FIGURE 6.11: Angle formed between the peptide ring and hydrophobic tail of PMB1 in A-E) POPE equil1 - equil5 simulations respectively, and F-J) US windows in which PMB1 was restrained in the bulk water phase outside the POPE membrane. Transparent lines show raw data.

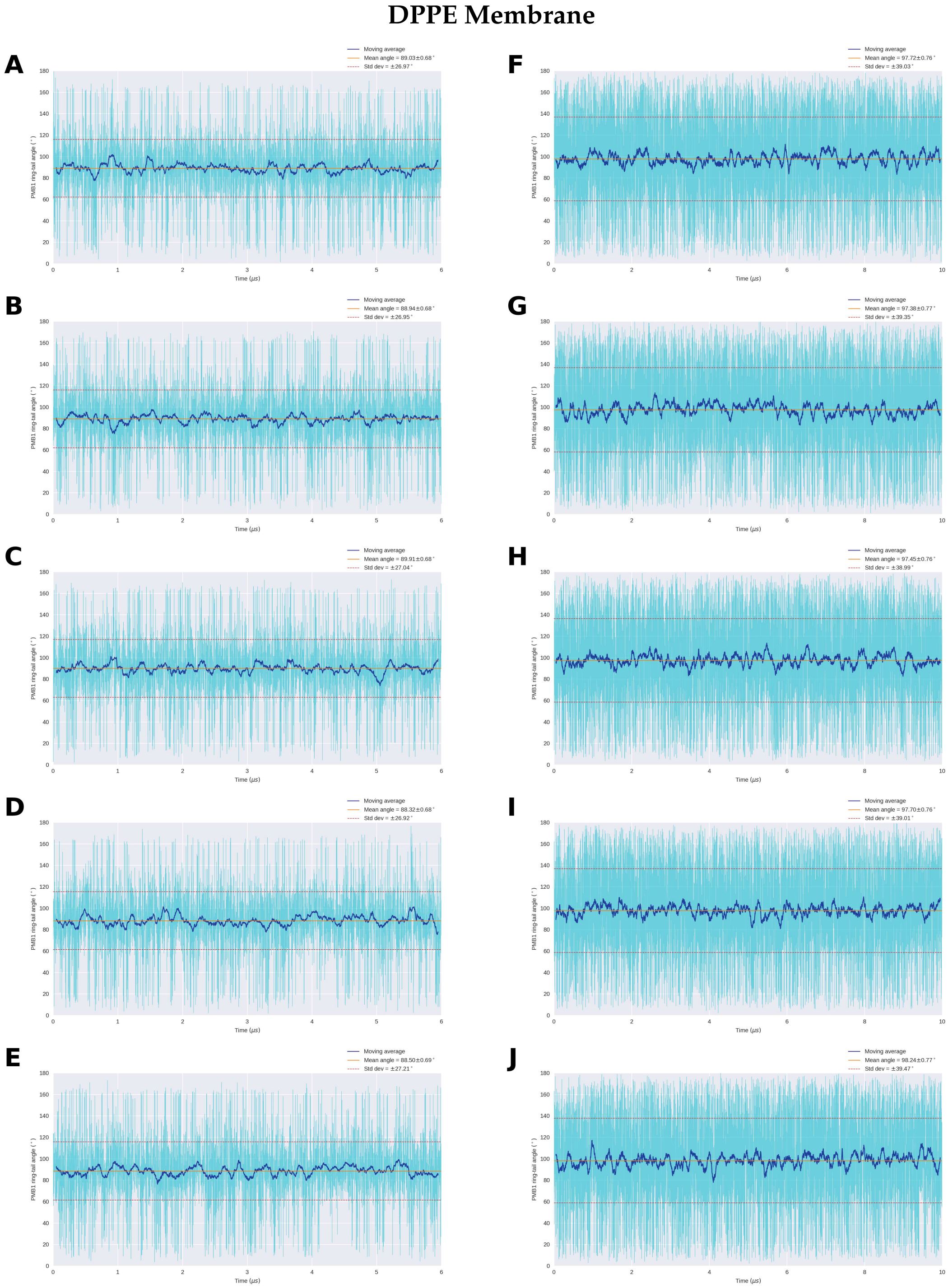


FIGURE 6.12: Angle formed between the peptide ring and hydrophobic tail of PMB1 in A-E) DPPE equil1 - equil5 simulations respectively, and F-J) US windows in which PMB1 was restrained in the bulk water phase outside the DPPE membrane. Transparent lines show raw data.

of 27.02 °. Analysis of the selected US windows returned a mean of 97.70 ± 0.34 ° with a standard deviation of 39.17 °.

The first behaviour to note is that the mean ring-tail angle obtained when PMB1 is in the interfacial-inserted state (Equil simulations) is approximately 9 ° smaller than when PMB1 is held in solution (US simulations) for both POPE and DPPE membrane models. Furthermore, in both membrane models, the mean ring-tail angle in the interfacial-inserted state was less than 90 °. These results in tandem provide evidence of the transition of PMB1 into the folded amphipathic conformation (Figure 6.9) upon binding to the membrane surface. The similarity between the results for both membrane models is also to be expected given that POPE and DPPE have identical headgroups; thus one may reasonably suspect that the behaviour of PMB1 would be comparable as it interacts with the surface of either membrane.

Now considering the results for the standard deviation of the data across each simulated environment, another notable insight may be obtained. Alongside the decrease in mean ring-tail angle, the standard deviation in the data is seen to decrease by roughly 12 ° when PMB1 is situated in the interfacial-inserted state with both membrane models, compared to the simulations when PMB1 is held in bulk water. This result indicates that PMB1 exhibits a greater degree of flexibility when in bulk water; highlighting the restraint on PMB1 conformation when in coordination with the membrane headgroup region. This restraint naturally arises from the various hydrophobic and polar interactions between PMB1 and the membrane that act to maintain the preferred folded amphipathic binding mode.

6.3.2.2 The Impact of PMB1 on Membrane Order

Lipid order parameters were calculated for each equilibrium simulation of the POPE and DPPE membranes. The analysis was performed on chain A (fully saturated) of each lipid, and the results for each molecule were calculated individually. These "per molecule" order parameters were then compared to the results obtained from applying an equivalent analysis on the same set of reference US windows used in the previous PMB1 ring-tail angle analysis; in which PMB1 was situated in bulk water outside of the membrane. The comparison to this reference data allowed us to ascertain what impact, if any, the interaction of PMB1 molecules with the surface of the membrane had on membrane order.

For simulations involving POPE (Figure 6.13), the reference data showed little deviation in order parameters between all molecules in both the inner and outer leaflets. This was closely matched by the results of the PO-Equil1, PO-Equil2 and PO-Equil3 simulations; indicating that the presence of PMB1 induced a negligible change to membrane order in these systems. Contrary to this, results from the

PO-Equil4 and PO-Equil5 simulations exhibited a disparity between lipid order parameters between the inner and outer leaflets. In these systems, the leaflet interacting with PMB1 (lipid residues 1 to 96) exhibited larger order parameters than the equivalent leaflet in the reference data, whilst the leaflet opposite to the PMB1 molecule (lipid residues 97 to 192) showed a corresponding decrease in per molecule order parameters as compared to the reference data.

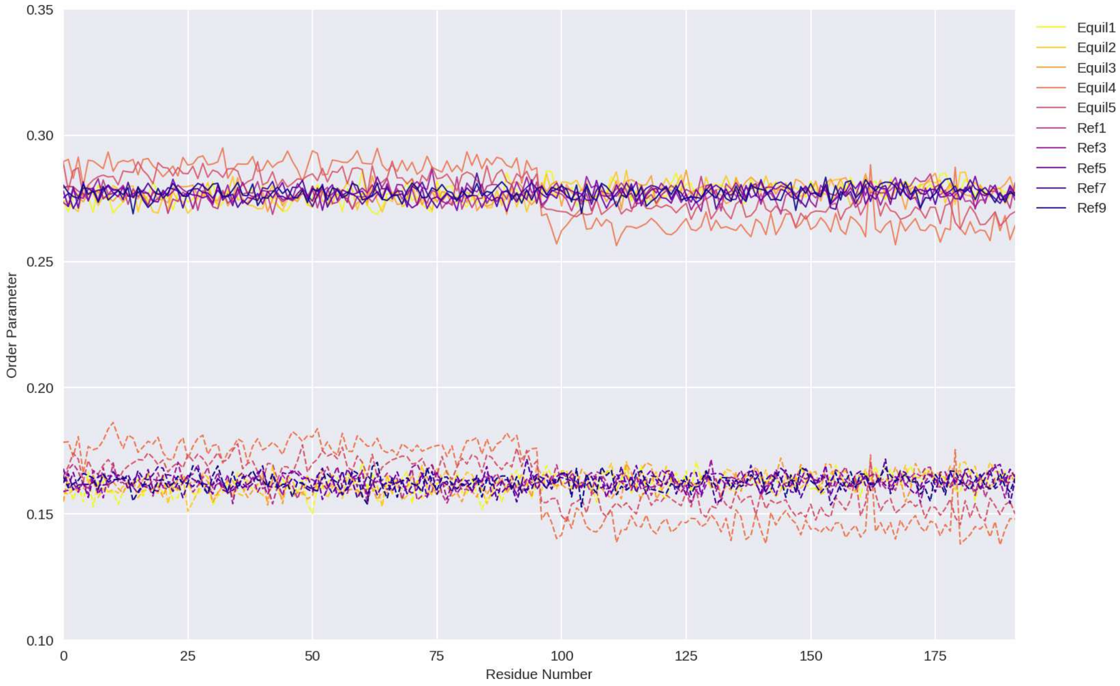


FIGURE 6.13: Comparison of per molecule order parameters for lipids in the POPE membrane when PMB1 is situated in the inter-facial inserted binding mode (Equil1-5) and when PMB1 is restrained in the bulk water phase external to the membrane (Ref1-9).

Similarly, for simulations involving DPPE (Figure 6.14), the reference data exhibited little difference between the order parameters of all lipids in both the inner and outer leaflets, with these results matched by those from the DP-Equil1, DP-Equil2 and DP-Equil3 simulations. Similar to the results from the POPE simulations, the DP-Equil4 simulation expressed a large deviation between the order parameters of the two membrane leaflets, with an increase in order in the leaflet interacting with PMB1 and a decrease in the opposing leaflet. Whilst the DP-Equil5 simulation also shows some evidence of this trend, a smaller inter-leaflet deviation was observed in this system.

The Equil4 and Equil5 simulations of both POPE and DPPE membrane systems were initialised from the final structure of US windows in which the PMB1 molecule was being held at or near to the outer leaflet surface, having previously passed through the center of the membrane during the SMD phase. In contrast, the Equil1 and Equil2 simulations were initialised from US windows in which PMB1 was being held at or near to the inner leaflet surface, and the PMB1 molecules had therefore not previously

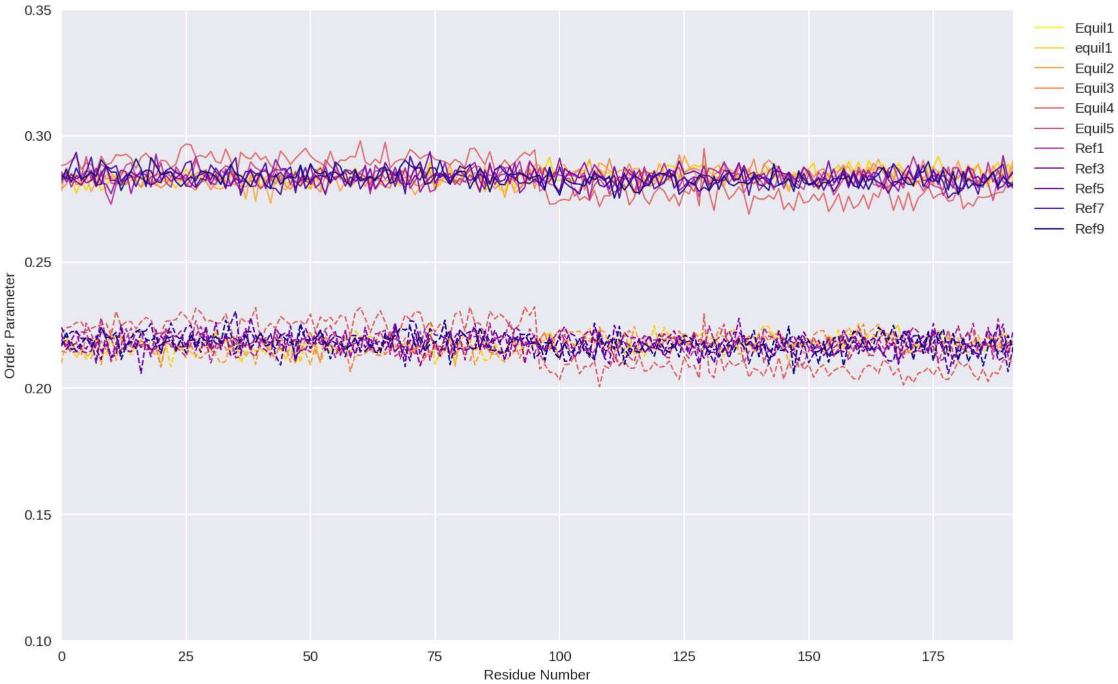


FIGURE 6.14: Comparison of per molecule order parameters for lipids in the DPPE membrane when PMB1 is situated in the inter-facial inserted binding mode (Equil1-5) and when PMB1 is restrained in the bulk water phase external to the membrane (Ref1-9).

passed through the membrane center. The presence of a deviation between leaflet order parameters in the Equil4 and Equil5 simulations and the lack of such a deviation in the Equil1 and Equil2 simulations indicates that perhaps there is a long lasting (> 6 μs) impact on membrane order resulting from the translocation of PMB1 molecules between the two leaflets, across the membrane center.

6.3.3 Multi-PMB Simulations

At the end of the initial 6 μs US equilibration phase, the two PMB1 molecules were found in complex with each other close to the membrane center. The inner leaflet headgroups were deformed towards the center of the membrane so as to remain in contact with the aggregate (Figure 6.15). This deformation was a remnant from the SMD simulation in which the first PMB1 was dragged into the membrane, as discussed previously in Section 6.3.1. This deformation once again remained for the duration of the US simulation; indicating that the most stable configuration of the aggregated PMB1 molecules at this location was to remain in coordination with the headgroup phosphate beads. This maintained coordination is unsurprising given the polarity of the PMB1 peptide rings which would, in the absence of such membrane deformation, be wholly surrounded by the hydrophobic environment of the membrane center.

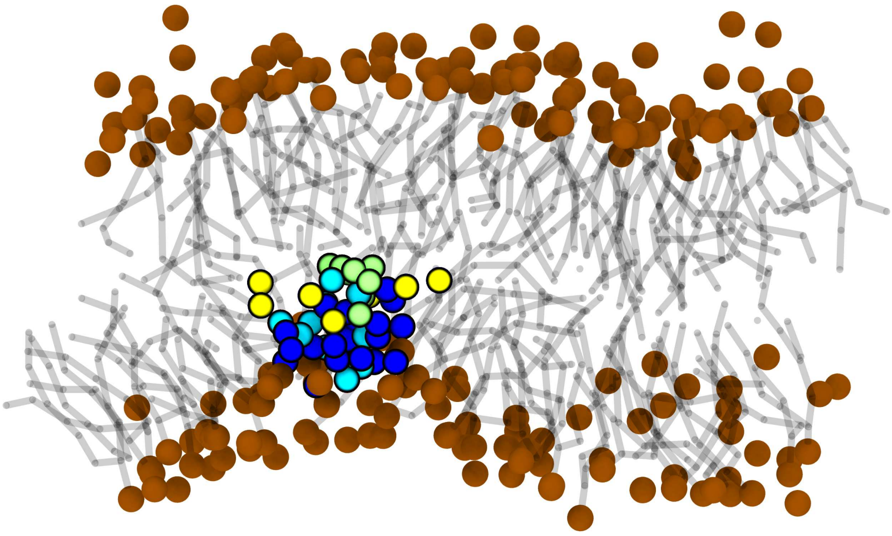


FIGURE 6.15: PMB1 molecules in complex with one another at the center of POPE membrane. For the PMB1 molecules: cationic DAB residues are represented by blue vdW spheres, hydrophobic residues are represented by yellow vdW spheres, polar noncharged residues are represented by cyan vdW spheres. White vdW spheres represent the anionic phosphate beads of the membrane headgroups.

During equilibrium MD simulation (multiPMB-r1 in Table 6.1), the aforementioned membrane deformation rapidly disappeared, with the membrane returning to a typical parallel planar bilayer configuration within 1 ns of simulation. The PMB1 aggregate was concomitantly ejected from the membrane center, remaining coordinated with the headgroup phosphate beads as they returned to the plane of their host leaflet. The aggregate remained intact at the membrane surface for roughly 20 ns, after which the two PMB1 molecules diffused away from each other whilst remaining individually bound to the membrane in the interfacial-inserted state. The PMB1 molecules were observed to diffuse freely across the surface of the membrane, occasionally coming into contact with one another and forming transient complexes before once more diffusing away from each other.

These observations provide us with further insight into the stability of the interfacial-inserted state, corroborating the existence of the energetic minimum obtained from the PMF profile analysis for individual PMB1 molecules. Furthermore, the dissolution of the PMB1 aggregate at the surface of the membrane and the subsequent transient interactions between the PMB1 molecules, indicates that any potential aggregation of polymyxins at the membrane surface likely requires some threshold concentration of PMB1 greater than was present in the system.

6.3.4 Self-Assembly Simulations

In order to further investigate the nature of PMB1 binding with the surface of the IM, self-assembly simulations were performed of either POPE or DPPE lipids in the presence of PMB1. An alchemical procedure was utilised to disassemble the systems used in the US simulations in a manner that ensured random lipid distributions and conformations. The simulation of these ″randomized″ systems could then be used to ascertain the propensity for PMB1 to interact with POPE and DPPE micelles in the absence of any hysteresis effects resulting from the presence of a fully formed membrane or any imposed choice of initial PMB1 location.

6.3.4.1 DPPE

Four self-assembly simulations were performed on the DPPE membrane model in the presence of a single PMB1 molecule. These simulations were all performed at a temperature of 350 K. Two of these simulations were performed at the standard volume of 6×10^4 nm^3 used in all POPE simulations and are referred to as the DP-r1 and DP-r2 simulations. The other two simulations were performed at a reduced volume of 4×10^4 nm^3 to enable investigation of how an increase in lipid concentration may affect micelle formation, these simulations are referred to as the DP-lowvol-r1 and DP-lowvol-r2 simulations.

In all four simulations, the PMB1 molecule was located in the interfacial-inserted state with a DPPE micelle at the onset of the self-assembly production MD phase; with the formation of the micelles and their interaction with the PMB1 molecule both occurring during the previous NPT equilibration phases of each system. In each case, the PMB1 molecules initially exhibited the folded amphipathic configuration observed in previous simulations; with the hydrophobic acyl tail and D-Phe residues both inserted into the hydrophobic core of the micelle, whilst the polar peptide region was exposed to the bulk aqueous phase. The PMB1 molecules remained bound to the micelle in this interfacial-inserted state for the entire duration of each 500 ns simulations.

Notably, in the DP-lowvol-r1 simulation, the hydrophobic tail of PMB1 was briefly ($\sim$ 10 ns) expelled from the hydrophobic core of the micelle after 375 ns of simulation (Figure 6.16). During this time, the hydrophobic D-Phe and D-Leu residues remained inserted into the core of the micelle and so whilst this event represents a rare deviation from the consistent residence of PMB1 in the standard interfacial-inserted state; it still exhibits a comparable binding mode whereby hydrophobic residues are inserted into the micelle core accompanied by an array of polar interactions with the micelle headgroups, anchoring the PMB1 molecule at the interfacial region of the micelle (Figure 6.17).

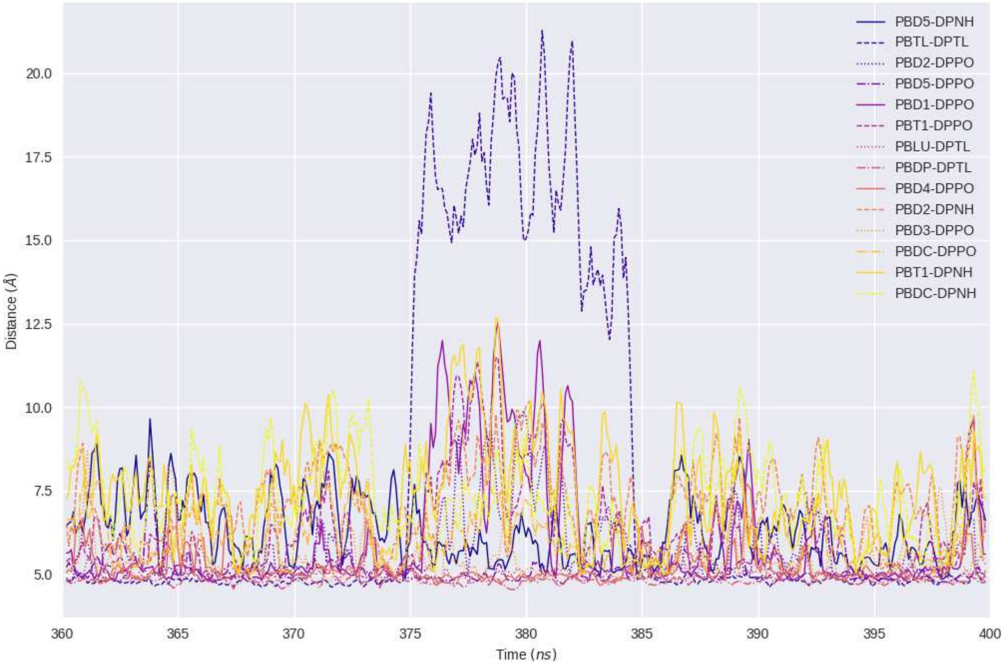


FIGURE 6.16: Residue interactions between PMB1 and DPPE micelle in the DP-r1 simulation during the period of tail expulsion from the micelle core.

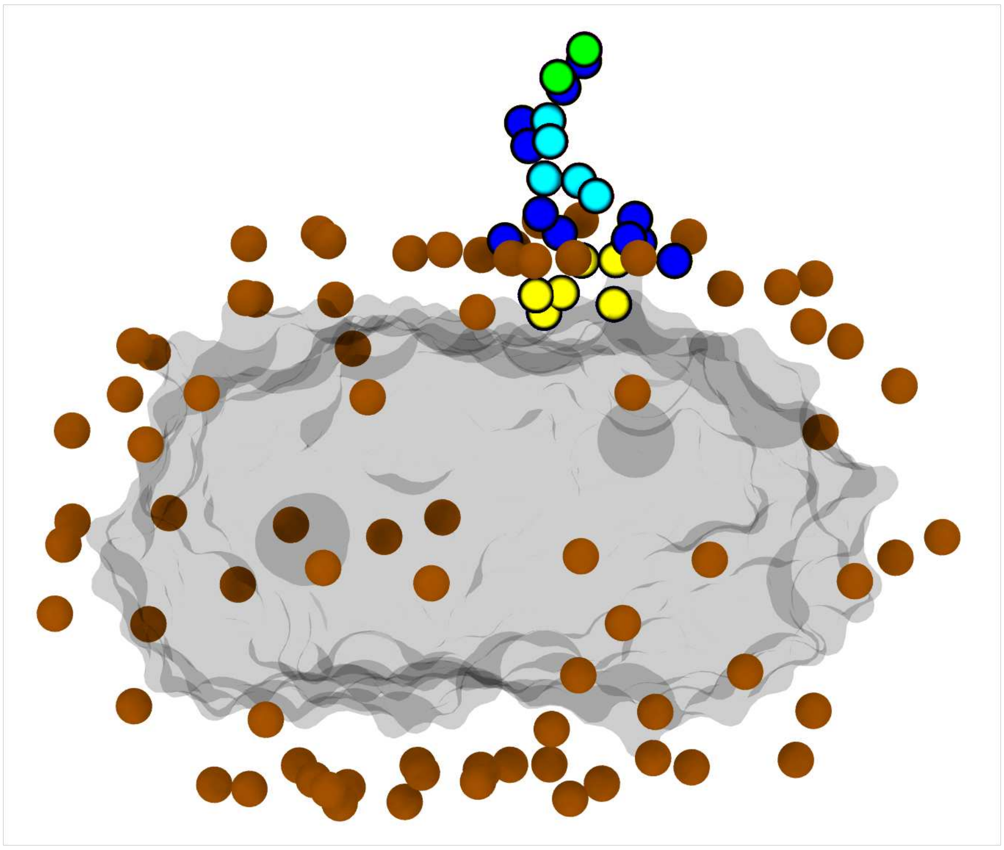


FIGURE 6.17: PMB1 molecule bound to a DPPE micelle in the DP-r1 simulation. The hydrophobic tail of PMB1, shown in green, is ejected from the micelle core whilst the D-Phe and D-Leu residues, shown in yellow, remain inserted.

6.3.4.2 POPE

Three self-assembly simulations were performed using the POPE membrane model in the presence of a single PMB1 molecule. Two simulations were performed at a temperature of 350 K to enable comparison with the DP-r1 and DP-r2 simulations whilst one other simulation was performed at a temperature of 310 K, closer to the lipid melting temperature for POPE; these three POPE simulations will be referred to respectively as the PO-350-r1, PO-350-r2 and PO-310 simulations from here on. An additional POPE self-assembly simulation was performed in the presence of 64 PMB1 molecules, equivalent to a 3:1 ratio of lipids to PMB1. This simulation was performed at a temperature of 310 K and will be referred to as the PO-multi simulation from here on.

In the PO-310 simulation, the PMB1 molecule was not associated with a micelle at the onset of production MD. Indeed, it remained freely diffusing throughout the bulk water phase for $\sim$ 487.5 ns before coming into contact with a micelle (Figure 6.19).

Contact between PMB1 and the POPE micelle was initially mediated by the interaction of the hydrophobic tail of PMB1 with an exposed region of lipid tails on the micelle surface, quickly followed ($<$ 0.5 ns) by a polar interaction between one of the cationic DAB residues of PMB1 and an anionic phosphate bead in the membrane headgroup region (Figure 6.20). Both of these residue interactions remained unperturbed for the remaining duration of the simulation, anchoring the PMB1 molecule onto the micelle surface.

Several further residue interactions between the remaining DAB residues of PMB1 and other phosphate beads in the membrane headgroup region were recorded, as well as the insertion of the hydrophobic D-Phe and D-Leu residues of PMB1 into the hydrophobic core of the micelle $\sim$ 8.75 ns after the initial contact was formed. These results once more support the notion that PMB1 adopts a folded amphipathic conformation upon binding to lipid micelles; whereby the formation of numerous polar interactions anchor the PMB1 molecule to the lipid headgroup region whilst the hydrophobic residues that remain exposed to the bulk water phase are subsequently rotated into the hydrophobic core of the lipid aggregate. Thus PMB1 was found to ultimately occupy the same interfacial-inserted state observed in previous simulations containing POPE bilayers.

In both the PO-350-r1 and PO-350-r2 simulations, PMB1 was bound to a micelle at the onset of production MD in the standard interfacial-inserted state and remained so for the entire duration of both 500 ns simulations. The deviation in this behaviour from the PO-310 simulation is likely a result of the increase in atomic velocities, corresponding to the increased system temperature, leading to the more rapid exploration of the system's available phase space.

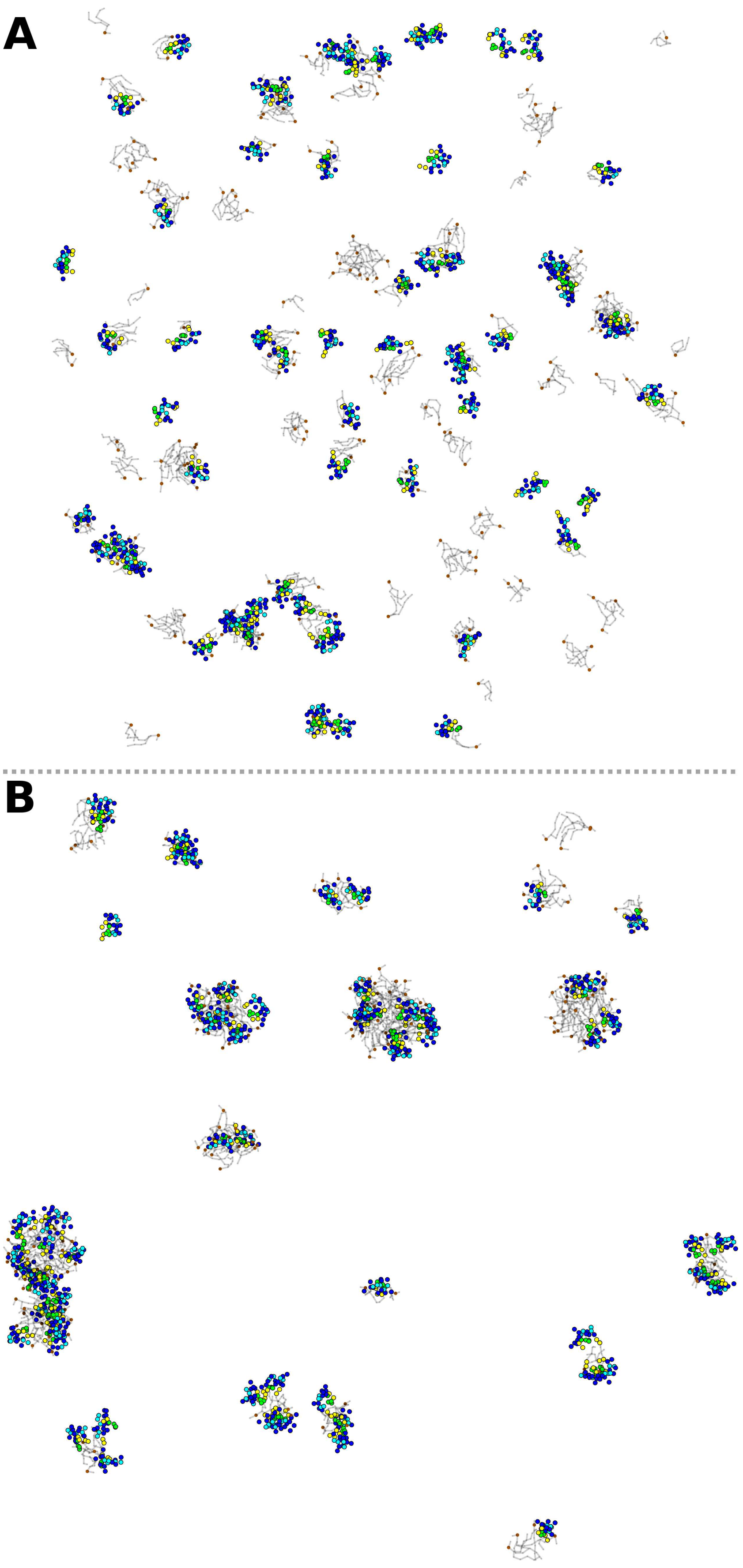


FIGURE 6.18: Snapshots of the PO-multi self-assembly simulation after A) 0 ns and B) 500 ns of production MD.

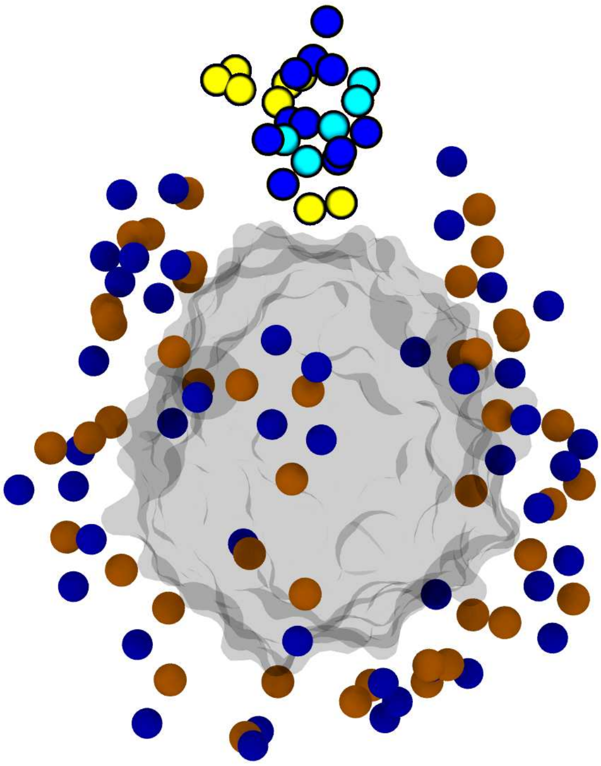


FIGURE 6.19: PMB1 molecule contacting POPE micelle in the PO-310 simulation. Diffuse orange and blue spheres represent phosphate and amine beads of the membrane headgroups respectively. Transparent grey surface represents hydrophobic core of the micelle. PMB1 shown in Goodsell representation with hydrophobic residues in yellow, cationic DAB residues in blue and polar non-charged residues in cyan.

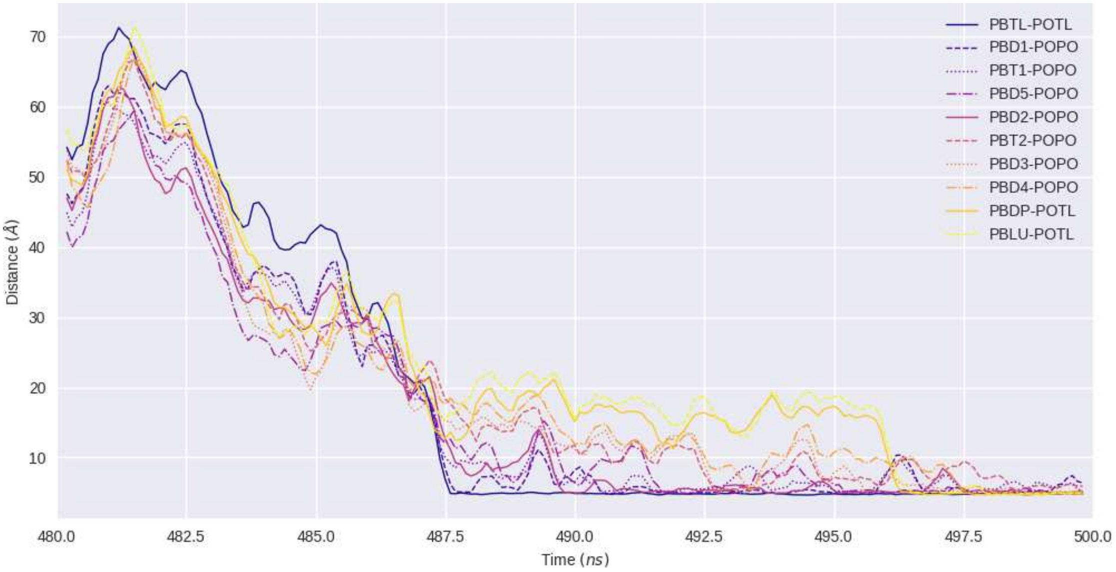


FIGURE 6.20: Residue interactions between PMB1 and POPE micelle in the PO-310 simulation.

6.3.4.3 Micelle Size

In order to investigate the effects of PMB1 on lipid aggregation; the number of micelles and the average number of lipids per micelle were calculated, as a function of time across all of the self-assembly simulations (Figures 6.21 & 6.22). In all cases, the micelle counts calculated by the clustering algorithm at the initial and final frames of simulation were validated through visual inspection of the trajectories.

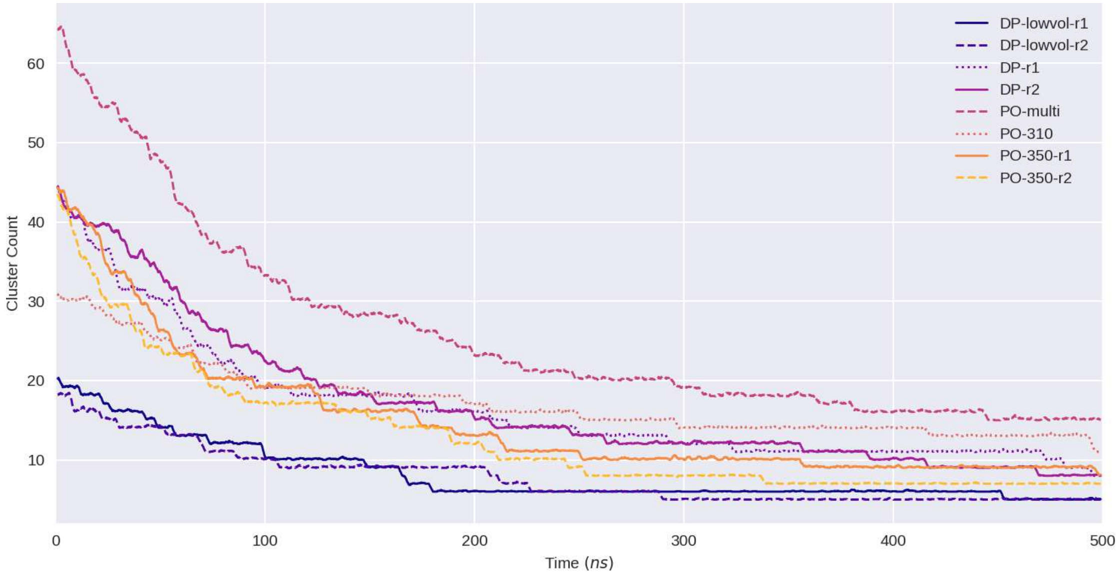


FIGURE 6.21: Number of lipid clusters in each self-assembly simulation.

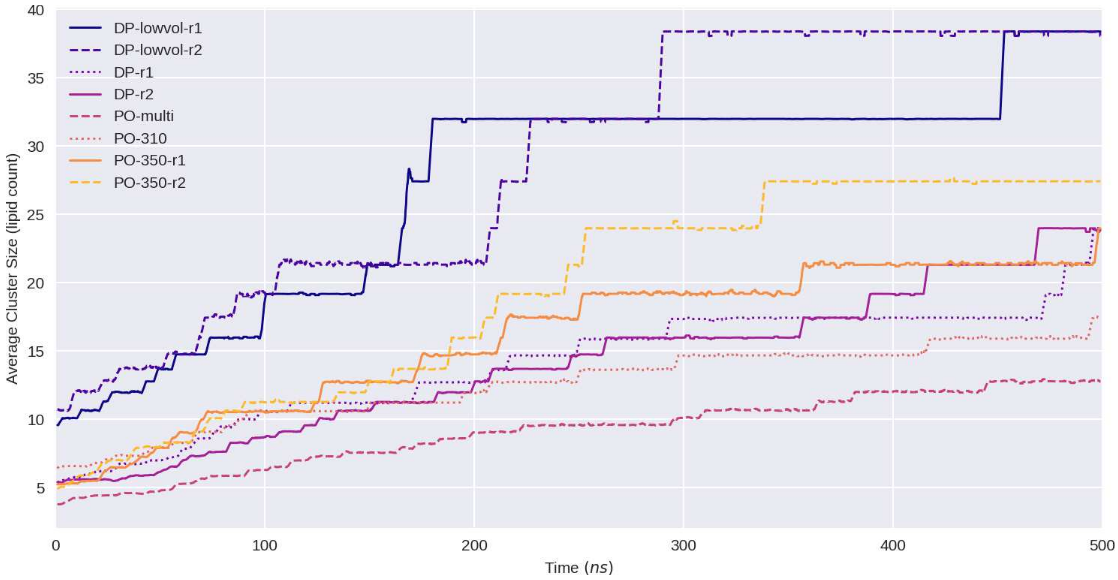


FIGURE 6.22: Average number of lipids per cluster in each self-assembly simulation.

Comparable results were obtained from the DP-lowvol-r1 and DP-lowvol-r2 trajectories; with initial micelle counts of 20 and 19 respectively, and a final micelle count of 5 at the end of both simulations. The average number of lipids per micelle (LpM) was also comparable between these two systems, with initial values of 9.6 /

11.2 and final values of 38.4 / 38.4, respectively, in the DP-lowvol-r1 and DP-lowvol-r2 simulations.

The DP-r1 and DP-r2 simulations both had initial micelle counts of 44, and final micelle counts of 8. The average LpM for these systems had initial values of 5.2 / 5.4, and final values of 24.0 / 24.0, respectively, in the DP-r1 and DP-r2 simulations. These data indicate that the decrease in lipid concentration in the DP-r1 and DP-r2 simulations resulted in a larger number of smaller micelles as compared to the DP-lowvol-r1 and DP-lowvol-r2 simulations. This result is to be expected given that the lipids in the DP-r1 and DP-r2 simulations were diffusing throughout a larger simulation domain and would thus contact one another less frequently than in the lower volume simulations, leading to a decreased rate of micelle fusion.

The PO-350-r1 and PO-350-r2 simulations had initial micelle counts of 44 / 44 and final micelle counts of 8 / 7, respectively. The average LpM for these systems had initial values of 5.2 / 4.8 and final values of 24.0 / 27.4, respectively, in the PO-350-r1 and PO-350-r2 simulations. Whilst these values are similar to those observed for the DP-r1 and DP-r2 simulations, consideration of Figures 6.21 & 6.22 illustrates that the PO-350 simulations exhibited smaller micelle counts and larger average LpM values throughout the majority of the simulation duration.

The PO-310 simulation had initial / final micelle counts of 32 / 11, and initial / final LpM values of 6.2 / 17.5, respectively. Comparison of these results to those obtained from the PO-350 systems highlights how the rate of micelle growth was limited by the lower temperature of the PO-310 system; once more drawing attention to the effect that the rate of phase space exploration has on the rate of micelle formation.

The PO-multi simulation exhibited the largest micelle counts and smallest average LpM values of any of the self-assembly simulations at all times throughout the 500 ns simulation (Figures 6.21 & 6.22). The system had initial / final micelle counts of 64 / 16, and initial / final LpM values of 3.8 / 12.8, respectively. Comparison of these data to those from the PO-310 simulation highlight that in the absence of any change in system temperature or volume, the presence of an excess of PMB1 molecules leads to a doubling in the initial micelle count, and a slower rate of micelle fusion throughout the simulation duration (Figure 6.22).

6.3.4.4 PMB1 Aggregation

The propensity for PMB1 molecules to form aggregates at the IM surface, and the role that such aggregates may play in the permeation of PMB1 through the IM, are still relatively unknown factors in the translocation of PMB1 across the cell envelope. In order to investigate this behaviour further, the nature of PMB1-PMB1 interactions in the PO-multi simulation were analysed.

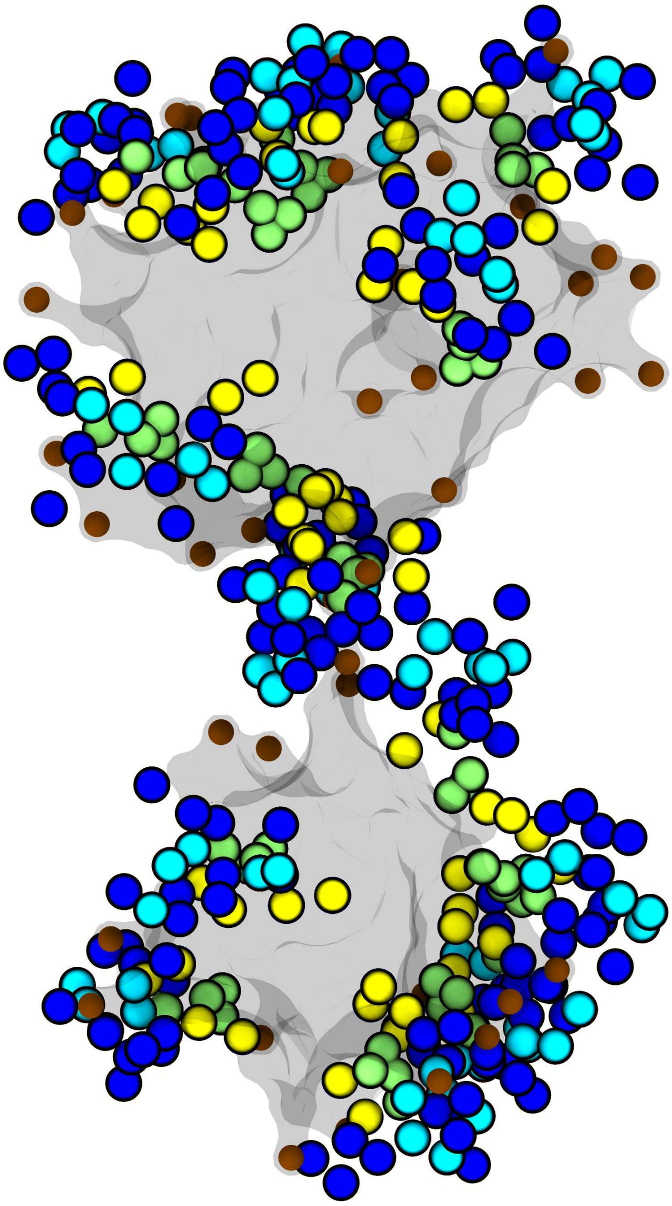


FIGURE 6.23: POPE micelle fusion occurring after 500 ns in the PO-multi simulation. Transparent grey surface represents hydrophobic core of the micelle. PMB1 shown in Goodsell representation with hydrophobic residues in yellow, cationic DAB residues in blue and polar non-charged residues in cyan.

First, the radial distribution function (RDF) between the terminal hydrophobic beads from the tails of all PMB1 molecules during separate 100 ns segments of the trajectory was calculated (Figure 6.24). The number of micelles with interfacial-inserted PMB1 molecules and the mean number of PMB1 molecules per micelle (Figure 6.25) were then approximated using the *gmx clustsize* utility with a radial distance cutoff of 3.8 nm, corresponding to the hydrophobic width of the largest micelle in the trajectory. This was followed by a more detailed analysis of the residue types that were involved in all residue-residue contacts between different PMB1 molecules (Figure 6.26).

During the first 100 ns of simulation, PMB1 molecules began to interact with and

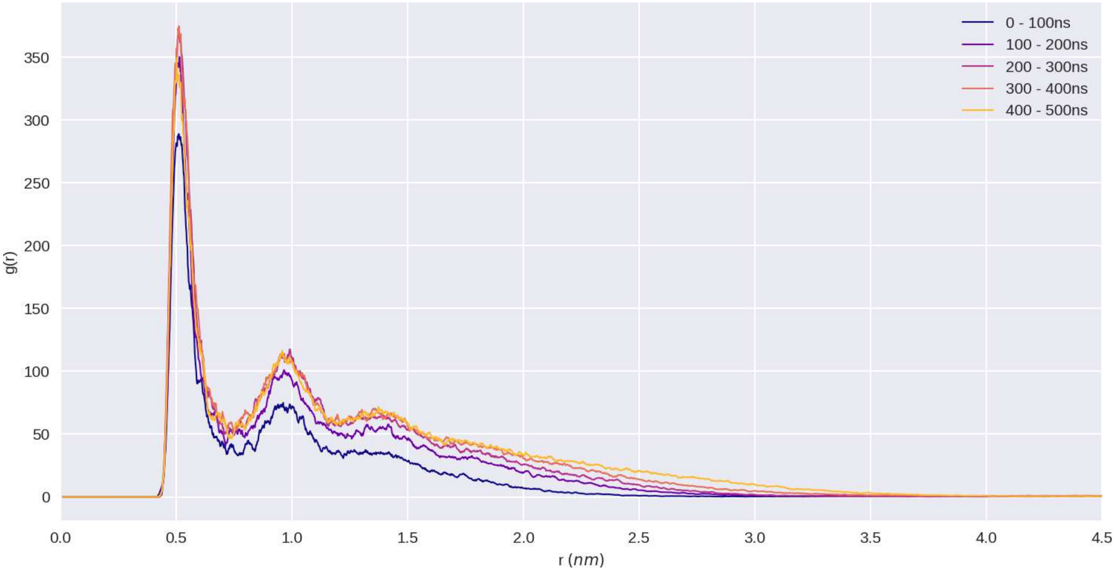


FIGURE 6.24: PMB1-PMB1 radial distribution function calculated over 100 ns chunks of the PO-multi simulation. Distances are measured between the terminal beads of the hydrophobic tail of each PMB1. All values of $g(r)$ continue to asymptotically approach zero as r increases beyond the plotted domain.

insert into the small lipid micelles (radii = $r < 3$ nm) that had formed during this time. The binding of PMB1 molecules onto these small micelles is reflected by the non-zero values of the RDF at $r < 2.5$ nm for the 0-100 ns simulation segment (Figure 6.24).

Subsequent simulation segments show this non-zero region of the RDF expanding to include larger radial distance values, with larger values of $g(r)$ across the domain; culminating in the 400-500 ns simulation segment which exhibits the greatest range of radial values for which the RDF is non-zero. This gradual increase in the amplitude of the RDF throughout the simulation reflects the gradual increase in the micelle size (Figure 6.22) and mean number of PMB1 molecules associated with each micelle (approximated by the PMB1 cluster size in Figure 6.25).

Noteably, the RDFs of all simulation segments exhibit three conserved peaks of decreasing magnitude at values of r = 0.53, 1.03 and 1.45 nm. These values roughly correspond to integer multiples of the standard Martini bead diameter (0.53 nm). These peaks therefore indicate that the terminal hydrophobic beads of the PMB1 tails are preferentially clustering in contact with one another, or separated by only one or two other beads. Furthermore, the increasing amplitude of these peaks in successive simulation segments, combined with the results from previous analyses, implies that this behaviour is exacerbated by the increasing size of lipid micelles and the corresponding increase in PMB1 molecules per micelle throughout the simulation; providing evidence that PMB1 molecules aggregate together when bound to the surface of a micelle, rather than distributing themselves uniformly across its surface.

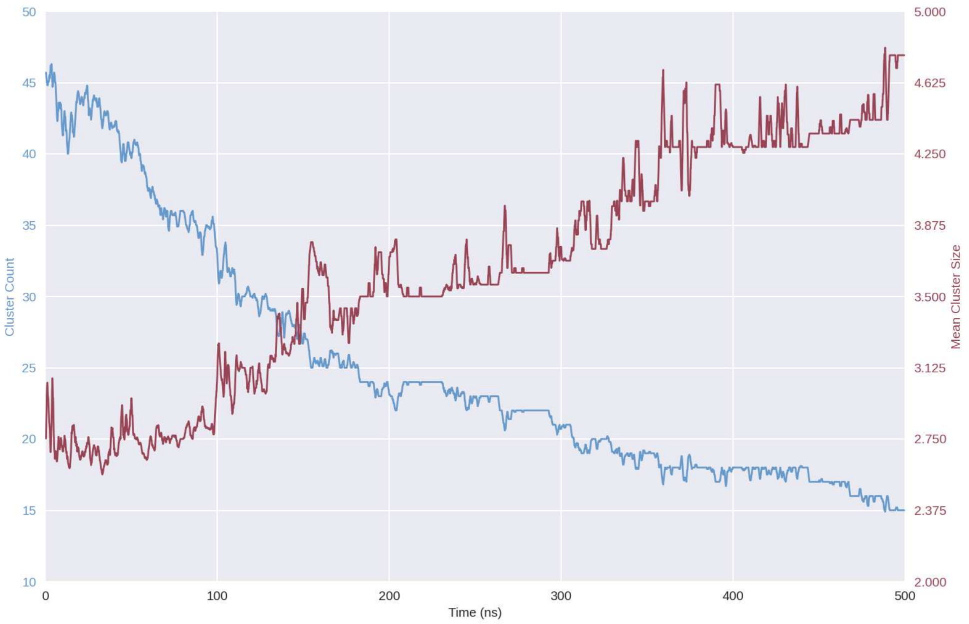


FIGURE 6.25: Number of PMB1 clusters (blue) and the mean cluster size (red) calculated at each frame of the PO-multi simulation. Cluster size is calculated as the number of PMB1 molecules per cluster.

Visual inspection of PMB1 interactions with various micelles throughout the simulation provided further evidence towards their preference for aggregation at the micelle surface. This is exemplified by the fusion of the two largest micelles at the end of the 500 ns PO-multi trajectory, illustrated in Figure 6.23: numerous PMB1 molecules were situated on the surfaces of each micelle, however distinct regions of each micelle surface were left unoccupied by any PMB1 molecule. Further, these unoccupied regions contained several exposed (anionic) phosphate beads that could provide suitable interaction sites for the cationic DAB residues of the PMB1 molecules. The distribution of PMB1 molecules across the micelle surface therefore did not appear to be linked solely to the distribution of the anionic binding sites, rather, it was also dependent upon the preferential aggregation of the PMB1 molecules.

This behaviour was also evident upon inspection of several smaller micelles throughout the simulation; with which PMB1 molecules were observed to interact as monomers, dimers and trimers whilst once more leaving regions of the micelle surface unoccupied and with several exposed phosphate beads (Figure 6.27).

Analysis of the specific residue interactions between PMB1 molecules highlighted the importance of the hydrophobic D-Phe, D-Leu and acyl tail residues to PMB1 - PMB1 interactions; with these residues accounting for a combined 53 % of all residue contacts (Figure 6.26A). Furthermore, at every frame of the simulation, the D-Phe

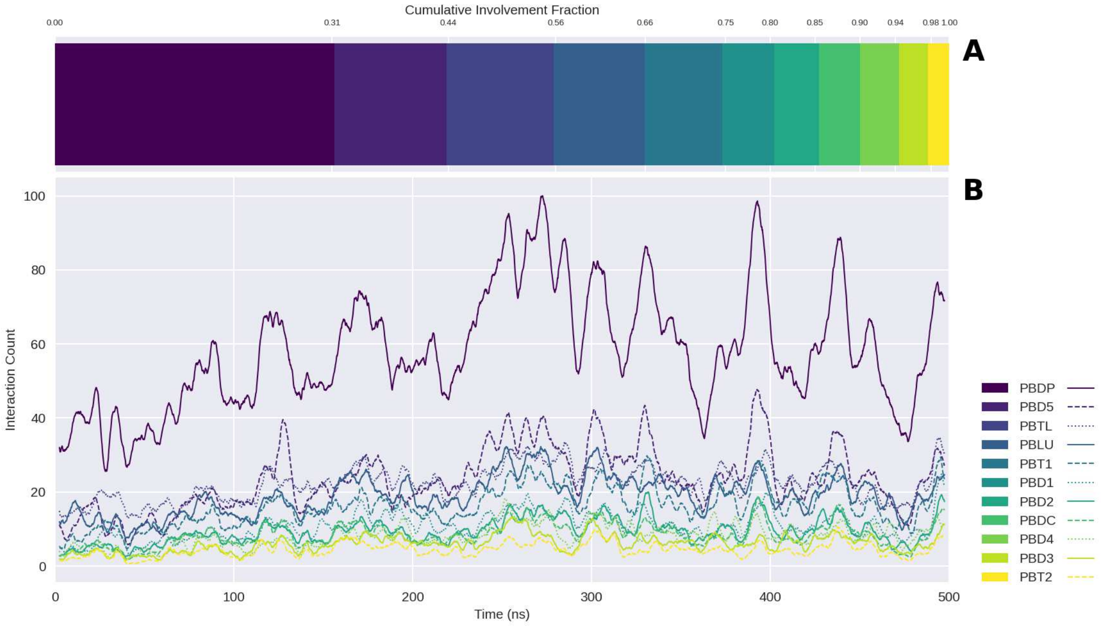


FIGURE 6.26: Residue interactions between PMB1 molecules in the PO-multi simulation. A) Proportion of residue contacts that involve each specific residue type, plotted cumulatively. B) Number of residue contacts involving each specific residue type in each frame of the 500 ns trajectory.

residues experienced more combined residue contacts than any other residue of PMB1 (Figure 6.26B); indicating that the D-Phe residues were of particular importance to the mediation of PMB1-PMB1 interactions.

Indeed, visual inspection of several PMB1 aggregates throughout the simulation highlighted the presence of a cluster of D-Phe residues at the center of each (Figure 6.27), corroborating the importance of these residues in the aggregation of PMB1 molecules at the micelle surface. The clustering of these hydrophobic D-Phe residues may therefore act as the nucleation point for the formation of such PMB1 aggregates, providing an energetic justification for why PMB1 molecules do not simply distribute themselves evenly across the various anionic phosphate beads of the micelle surface.

6.4 Conclusions

The simulations presented in this chapter have begun to illustrate a picture of how PMB1 molecules interact with and insert into lipid bilayers and micelles composed of lipids representative of those in the *E. coli* IM. Through the use of umbrella sampling techniques, evidence has been provided that there is no initial energetic barrier to the insertion of the hydrophobic tail of PMB1 into the inner membrane and further equilibrium MD simulations have highlighted how this insertion is likely driven by the adoption of a folded amphipathic conformation by the PMB1 molecule at the point of initial contact with the membrane surface. This process of interfacial folding and

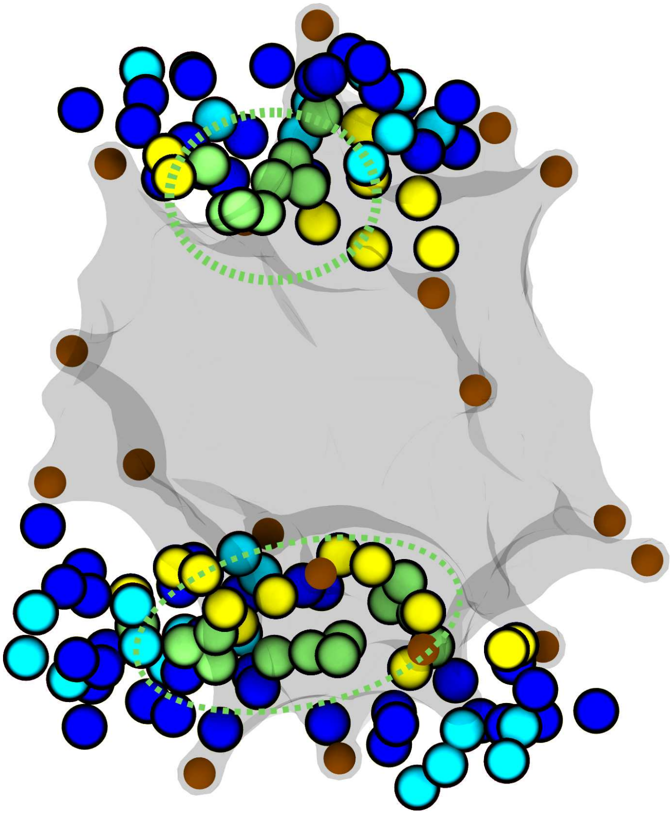


FIGURE 6.27: Two PMB1 aggregates bound to the surface of a POPE micelle 250 ns into the PO-multi simulation. Dotted green circles highlight clustered PMB1 D-Phe residues. PMB1 is shown in Goodsell representation with D-Phe in lime green, D-Leu and tail residues in yellow, cationic DAB residues in blue and polar non-charged residues in cyan. Diffuse orange spheres represent POPE headgroup phosphates, surface plot of the entire micelle is shown in transparent grey.

spontaneous membrane insertion is not uncommon among AMPs and designed peptides Im and Brooks III (2005); Jiang et al. (2020c); Fernández-Vidal et al. (2007), and thus it is not surprising that PMB1 is seen to behave in this manner.

This work has demonstrated the stability of the so-called interfacial-inserted binding mode through extensive equilibrium MD and self-assembly simulations, in which PMB1 molecules were found to invariably occupy this state by the end of each simulation. Furthermore, the lack of any observed instance of PMB1 dissociation from the IM when occupying the interfacial-inserted state corroborates the observation of the energetic minimum in the PMB1-IM insertion PMF curve when PMB1 was situated in this configuration.

Whilst the large energetic barrier to PMB1 permeation across the center of the membrane, obtained from the umbrella sampling analysis, contains spurious contributions from the improper membrane deformation resulting from a naïve treatment of the initial configuration generation procedure; the existence of such a barrier, albeit of some unknown magnitude, is likely supported by the extensive sampling across all other simulation regimes.

If the trajectories of each PMB1 molecule in all of the equilibrium MD simulations within this chapter are concatenated; this work may be considered to have sampled a combined 107 μs of PMB1 dynamics in the presence of various lipid environments. Throughout this extensive sampling period, no PMB1 molecule was seen to insert deeper into its host bilayer/micelle than the interfacial-inserted state. Thus it is reasonably certain that such an event would require a significant external impetus not present in any of these simulations; consistent with the existence of a large energetic barrier to the permeation of PMB1 across the center of the membrane hydrophobic core.

It is therefore likely that PMB1 does not permeate into the inner membrane as a monomer and instead relies upon some coordination between numerous PMB1 molecules in order to perform its antimicrobial action on the IM. Through simulation of the self-assembly of POPE lipids in the presence of an excess of PMB1 molecules, this work has provided evidence that the initial stages of this coordination takes the form of PMB1 aggregate formation at the membrane/micelle surface; mediated by the clustering of hydrophobic D-Phe residues.

Chapter 7

Conclusions

7.1 Summary

The work presented in this thesis has sought to provide detail on the behaviour of polymyxin antibiotics within the *E. coli* cell envelope. Through the use of either all atom or coarse grain modelling techniques, the nature of polymyxin interactions with the cell wall, Braun's lipoprotein and the inner membrane have been studied in depth, highlighting the myriad ways in which molecular crowding and complex formation affect the functional dynamics of polymyxin lipopeptides. Where possible, discussion of comparable experimental and computational literature has been provided so as to place this work within its proper context.

Chapter 3 outlined the protocol used to construct the all atom model of the *E. coli* periplasm presented in this thesis. The pitfalls encountered during this process were discussed and the solutions implemented to avoid them were described and motivated. Particular focus was given to the construction of the cell wall, as well as the necessity for the independent equilibration of individual system components prior to the construction of the combined model. The compositions of the three periplasmic fluid models, or ″crowding regimes″, studied in this thesis were also discussed, with the choice of molecules and their concentrations motivated through reference to existing experimental and computational literature. The details provided in this chapter highlighted the complexities that arise from constructing biophysical models involving both the cell wall and bacterial membranes; thus, this discussion will likely be of use for the future construction of more complex models of the Gram-negative cell envelope.

Chapter 4 analysed the nature of polymyxin interactions with the cell wall as a function of periplasmic composition using the aforementioned all atom model of the *E. coli* periplasm. In the absence of a diverse chemical environment both PMB1 and

PME were shown to bind rapidly and irreversibly to the cell wall. The presence of ions, osmolytes and crowding ubiquitin proteins were each shown to disrupt these interactions to various extents, leading to shorter duration interactions between the polymyxins and the cell wall. The effect of K^+ ions was shown to be of particular importance, with the regular dissociation of polymyxins from the cell wall, in the absence of crowding ubiquitin proteins, observed only in those simulations involving K^+. As mentioned, the presence of ubiquitin proteins in the most compositionally complex simulations was also shown to induce the dissociation of polymyxins from the cell wall, even in the absence of K^+. The primary result of this chapter was the proposed mechanism underlying the disruption of polymyxin - cell wall binding by such cationic moieties.

The binding of both polymyxins to the cell wall was shown to be underpinned predominantly by polar interactions between the cationic DAB residues of the polymyxins and the various carboxylate groups on the peptide residues of the cell wall. The dependence of polymyxin - cell wall binding on this single mode of interaction was shown to enable other cationic moieties within the system to disrupt the otherwise stable complex formation between the polymyxins and the cell wall. Explicit observations of K^+, spermidine and ubiquitin binding to the carboxylate groups of the cell wall in the immediate vicinity of polymyxins indicated that these moieties were in direct competition for cell wall binding sites. The small size and relatively high concentration of K^+ ions, as compared to spermidine and ubiquitin, allowed for K^+ to coat the cell wall. This chapter therefore proposed that the dissociation of polymyxins from the cell wall, induced by the presence of K^+, follows the experimentally established notion of the salting-in of proteins, whereby low salt concentrations lead to an increase in protein solubility due to the screening of electrostatic interactions Arakawa and Timasheff (1982); Dumetz et al. (2007); Hassan (2005). Furthermore, the induced dissociation of polymyxins from the cell wall would likely enable the polymyxins to more readily encounter lipoprotein carriers such as LolA within the surrounding aqueous phase of the periplasm. LolA has been previously proposed to provide polymyxins with a potential passive transport mechanism within the periplasm Pedebos et al. (2021); thus, the cationic disruption of polymyxin - cell wall binding may be a requirement to the overall transport mechanism of polymyxins from their point of entry into the cell, the OM, to the target of their antimicrobial action, the IM.

Chapter 5 utilised the simulations presented in the previous chapter to provide another perspective on how the complexity of the periplasm can affect the interactions of the polymyxins with different components of the Gram-negative cell envelope. To this end, the analysis in this chapter focused on polymyxin interactions with BLP. The specific residue interactions that underpinned the binding of the polymyxins to BLP were shown to vary both between polymyxin species and as a function of periplasmic composition, in stark contrast to the analysis of polymyxin interactions with the cell wall.

An in depth analysis of residue interaction percentages in each crowding regime highlighted that despite these deviations, certain groups of residues within PMB1 and PME were repeatedly found to play important roles in the binding with BLP under different simulation conditions. Particular attention was drawn to the importance of the positionally analagous Leu/D-Phe/DAB5 and Leu1/Leu2/DAB5 residue triads of PMB1 and PME, respectively, as well as the polar DAB1/Thr1/DAB2 triad located on the branched fatty acid tail of both polymyxins.

The origin of the variation in polymyxin - BLP binding between simulation regimes was investigated through the analysis of polymyxin interactions with one particular osmolyte, OPG. It was shown that OPG interacted with both polymyxins *via* a range of both hydrophobic and electrostatic interactions. The balance between these interactions was shown to vary as a function of polymyxin species and crowding regime; and evidence was provided that these variations could, at least in part, be correlated with variations in the residues involved in polymyxin binding with either BLP or the cell wall. These results serve to highlight the highly coupled nature of molecular interactions within complex biological systems.

As has been exemplified by the binding of the polymyxins with BLP, it is not reasonable to derive a confident mechanistic understanding of the interactive behaviour between pairs of molecules without consideration of the multitude of interactions with the other constituents of the surrounding environment. As was discussed in Section 1.5, it is still commonplace in many fields of computational chemistry to consider the functional dynamics of specific molecules in highly idealised environments. As the global threat posed by multi-drug resistant bacteria continues to develop O'Neill (2016); Antimicrobial Resistance Collaborators (2022), it is imperative that the field of computational antibiotic research ensures the efficacy and validity of its results so as to accurately inform the future development of novel therapeutics. It is in this context that this chapter emphasised the necessity for a nuanced, biologically motivated consideration of the periplasmic composition used in any future studies of antibiotics within the bacterial cell envelope.

Chapter 6 presented coarse grain simulations of PMB1 in the vicinity of lipid bilayers representative of the *E. coli* inner membrane. Three membrane models were constructed, comprised entirely of either POPE, DOPE or DPPE phospholipids.

Umbrella sampling was used to determine the PMF profiles for PMB1 insertion into each membrane. Details of the umbrella sampling protocol were discussed and attention was drawn to the source of common systematic errors present when performing umbrella sampling procedures using configurations generated using steered MD techniques. Whilst the full PMF profiles of PMB1 permeation through both leaflets of each membrane did not converge, likely due to these systematic errors, notable details were able to be extracted from the convergent half profiles that corresponded to polymyxin permeation through the proximal leaflet of each membrane. It was shown that there was no initial energetic barrier to PMB1 insertion into the headgroup region of each membrane, consistent with the Type 3 insertion described in literature Neale and Pomès (2016) that has been reported for other amphipathic molecules Neale et al. (2011).

Through structural cluster analysis, it was shown that PMB1 adopted a folded-amphipathic conformation when bound to each membrane at the position of the initial energetic minimum in the respective PMF profiles. This conformation was characterised by the insertion of PMB1 hydrophobic residues into the membrane core, whilst the polar and charged PMB1 residues remained bound to the polar headgroup region of the membrane, exposed to the external aqueous environment.

During equilibrium MD simulations of PMB1 in the vicinity of the POPE and DPPE membranes, PMB1 was shown to spontaneously insert into this so-called membrane interfacial-inserted folded-amphipathic binding mode, confirming the lack of an initial energetic barrier to PMB1 insertion observed in the half PMF profiles. Spontaneous insertion of PMB1 into the interfacial-inserted state was mediated by initial contact between the charged DAB residues of PMB1 and the polar headgroup region of the membrane, followed by the rotation of the hydrophobic residues of PMB1 into the membrane core. This process was consistent with previously reported mechanisms for the insertion of amphipathic AMPs into lipid membranes Fernández-Vidal et al. (2007); Jiang et al. (2020a,b).

Simulations of POPE and DPPE lipid self assembly in the presence of a single PMB1 molecule were performed under different conditions of temperature and lipid concentration. A single simulation using a 3:1 ratio of POPE to PMB1 was also performed. Here, PMB1 molecules were shown to invariably occupy the interfacial-inserted binding mode by the end of each simulation. Analysis of the number and size of micelles in each simulation indicated that the presence of an abundance of PMB1 molecules was correlated with a greater number of smaller micelles, as well as a slower rate of micelle fusion throughout the simulation.

Of particular interest were the results from the analysis of PMB1 aggregation during self assembly. It was shown that PMB1 molecules formed aggregates at the micelle surface, mediated by the clustering of hydrophobic D-Phe residues. This result was in contrast to a previous study of a similar system that reported no obvious aggregation of PMB1 at the surface of the *E. coli* IM Berglund et al. (2015).

Overall, the work presented in this thesis has provided an in depth view into the transport of polymyxin lipopeptides throughout each layer of the *E. coli* periplasm. The impacts of crowding and aggregation have been demonstrated at each point.

Whether through the observation of cationic competition for cell wall binding sites, protein-peptide complex formation disrupting interactions with BLP, or the formation of polymyxin aggregates on the surface of the IM: it is clear that the complex composition of the bacterial cell envelope plays a pivotal role in the nature of polymyxin interactions within this space.

7.2 Future Work

The work presented in Chapters 4 and 5 of this thesis utilised all atom simulations of the *E. coli* periplasm to investigate the effects of periplasmic composition and biomolecular crowding on the interactions of polymyxin antibiotics with the various structural components of the Gram-negative cell envelope. Whilst some notable conclusions have been drawn, there are a number of obvious paths for future work.

Perhaps the simplest of these options would involve further analysis of the existing trajectories generated during the course of this thesis. Explicit calculation of molecular diffusion rates would provide a valuable metric to compare against recently published experimental data Tran et al. (2023). This would provide an additional quantitative measure of the relative impacts of environmental complexity and crowding on the transport behaviour of the polymyxins throughout the *E. coli* cell envelope.

Another avenue of interest would be a more rigorous analysis of the correlations between the residue interaction percentages for each pair of molecular constituents within the periplasm. Whilst this would involve a considerably more sophisticated approach than has been implemented for this thesis, it would enable a truly quantitative perspective on how each molecular component affects the interconnected network of interactions involving all other molecules within the system.

Further investigation into the energetics of the proposed cationic disruption of polymyxin - cell wall binding would also be of use. The analysis of PMF profiles for polymyxin dissociation from the cell wall in the presence or absence of proximally cell wall bound K^+ ions would provide quantitative support for the proposed disruption mechanism.

Aside from further consideration of the existing trajectories; there are clear improvements that may be made to bring the composition of the model closer to the true *in vivo* environment. The addition of an inner membrane would allow for the inclusion of trans-envelope protein complexes, such as the Tol-Pal complex that connects the OM to the IM *via* a tether to the cell wall Szczepaniak et al. (2020). How the presence of such supramolecular complexes may impact the molecular interactions and diffusion rates of the polymyxins as they traverse the Gram-negative cell envelope would be of key future interest. Furthermore, substitution of the crowding ubiquitin proteins for a more biologically relevant selection of *E. coli* periplasmic proteins Lalgudi and Elcock (2016) would further close the gap between the *in silico* model and the true *in vivo* environment. Another key avenue of interest would be the construction of a larger envelope model containing multiple BLP molecules: with only one BLP present, any kinking or bending induces significant changes to the dimensions of the periplasm. The addition of multiple BLP molecules would therefore provide a more robust and biologically accurate scaffold between the OM and cell wall.

The work presented in Chapter 6 of this thesis utilised coarse grain simulations of the *E. coli* inner membrane to investigate the energetics and conformational dynamics of polymyxin insertion into the IM, as well as polymyxin aggregation at the IM surface. There are two main avenues for future work that are of particular interest.

Due to the clinical toxicity and antimicrobial potency of the polymyxins, there is an active effort within the scientific community to derive novel therapeutics *via* the modification of the chemical structure of the polymyxins Kaguchi et al. (2023b). The use of the coarse grain Martini2.2 force field enables the modification of the polarity and charge of each constituent particle within a molecule with trivial difficulty. It would therefore be of use to construct a number of new polymyxin models with modifications to certain residues in order to investigate to what extent particular regions of the polymyxin structure impact their ability to disrupt the bacterial IM. Modifications to the DAB2, D-Phe and D-Leu residues would be of particular interest due to reports of these residues regulating the interactions of polymyxins with the cell membrane of human kidney tubular cells Jiang et al. (2020c). Furthermore, the results presented in Chapter 4 of this thesis highlighted how the cell wall sequestered polymyxins *via* interactions with the cationic DAB residues, thus reducing the overall charge of the polymyxins, through modification of the charge on the DAB2 residue, may ease the passage of these antimicrobial agents across the cell wall as they move through the crowded periplasm towards the target of their antimicrobial function, the IM.

Finally, it would be interesting to repeat the analysis of the PMF profiles for polymyxin insertion into models of the IM with the inclusion of varying concentrations of LPS. Reports of polymyxins targeting LPS within the IM Sabnis et al. (2021), and the importance of LPS targeting to the initial uptake mechanism of polymyxins through the OM Khadka et al. (2018); Ledger et al. (2022) motivate a more nuanced analysis of the energetics of polymyxin insertion that would provide insight into the effects of LPS concentration on the efficacy of the bactericidal mechanism of the polymyxins.

References

Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2: 19–25, September 2015. ISSN 2352-7110. . URL `https://www.sciencedirect.com/science/article/pii/S2352711015000059`.

M. P Allen and D. J Tildesley. *Computer simulation of liquids*. Oxford University Press, 1 edition, 1987.

Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet (London, England)*, 399(10325): 629–655, February 2022. ISSN 1474-547X. .

Tsutomu Arakawa and Serge N. Timasheff. Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry*, 21(25):6545–6552, December 1982. ISSN 0006-2960. . URL `https://doi.org/10.1021/bi00268a034`. Publisher: American Chemical Society.

Sean N. Avedissian, Jiajun Liu, Nathaniel J. Rhodes, Andrew Lee, Gwendolyn M. Pais, Alan R. Hauser, and Marc H. Scheetz. A Review of the Clinical Pharmacokinetics of Polymyxin B. *Antibiotics (Basel, Switzerland)*, 8(1):31, March 2019. ISSN 2079-6382. .

Fernando Baquero and Bruce R Levin. Proximate and ultimate causes of the bactericidal action of antibiotics. *Nature Reviews Microbiology*, 19(2):123–132, 2021.

Herman JC Berendsen, JPM van Postma, Wilfred F Van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

Nils A. Berglund, Thomas J. Piggot, Damien Jefferies, Richard B. Sessions, Peter J. Bond, and Syma Khalid. Interaction of the Antimicrobial Peptide Polymyxin B1 with Both Membranes of E. coli: A Molecular Dynamics Study. *PLOS Computational Biology*, 11(4):e1004180, April 2015. ISSN 1553-7358. . URL `https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004180`. Publisher: Public Library of Science.

Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.

Silvie Bernatová, Ota Samek, Zdeněk Pilát, Mojmír Šerỳ, Jan Ježek, Petr Jákl, Martin Šiler, Vladislav Krzyžánek, Pavel Zemánek, Veronika Holá, et al. Following the mechanisms of bacteriostatic versus bactericidal action using raman spectroscopy. *Molecules*, 18(11):13188–13199, 2013.

Blake Bertani and Natividad Ruiz. Function and biogenesis of lipopolysaccharides. *EcoSal Plus*, 8(1):10–1128, 2018.

Matthew J Betts and Michael JE Sternberg. An analysis of conformational changes on protein–protein association: implications for predictive docking. *Protein Engineering*, 12(4):271–283, 1999.

Jessica MA Blair, Grace E Richmond, and Laura JV Piddock. Multidrug efflux pumps in gram-negative bacteria and their role in antibiotic resistance. *Future microbiology*, 9(10):1165–1177, 2014.

Alister T. Boags, Firdaus Samsudin, and Syma Khalid. Binding from Both Sides: TolR and Full-Length OmpA Bind and Maintain the Local Structure of the E. coli Cell Wall. *Structure (London, England: 1993)*, 27(4):713–724.e2, April 2019. ISSN 1878-4186. .

Mikhail Bogdanov, Kyrylo Pyrshev, Semen Yesylevskyy, Sergey Ryabichko, Vitalii Boiko, Pavlo Ivanchenko, Ramziya Kiyamova, Ziqiang Guan, Christophe Ramseyer, and William Dowhan. Phospholipid distribution in the cytoplasmic membrane of gram-negative bacteria is highly asymmetric, dynamic, and cell shape-dependent. *Science Advances*, 6(23):eaaz6333, 2020.

Sébastien Bontemps-Gallo, Jean-Pierre Bohin, and Jean-Marie Lacroix. Osmoregulated Periplasmic Glucans. *EcoSal Plus*, 7(2), June 2017. ISSN 2324-6200. .

Emilio Bouza and Roger Finch. Infections caused by gram-positive bacteria: situation and challenges of treatment. *Clinical Microbiology and Infection*, 7:iii, 2001.

V. Braun. Covalent lipoprotein from the outer membrane of escherichia coli. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 415(3):335–377, October 1975. ISSN 0304-4157. . URL `https://www.sciencedirect.com/science/article/pii/0304415775900131`.

V Braun and K Rehn. Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the e. coli cell wall: the specific effect of trypsin on the membrane structure. *European Journal of Biochemistry*, 10(3):426–438, 1969.

Volkmar Braun and Valerie Bosch. Sequence of the murein· lipoprotein and the attachment site of the lipid. *European Journal of Biochemistry*, 28(1):51–69, 1972.

Volkmar Braun, Kurt Rehn, and Helga Wolff. Supramolecular structure of the rigid layer of the cell wall of Salmonella, Serratia, Proteus, and Pseudomonas fluorescens. Number of lipoprotein molecules in a membrane layer. *Biochemistry*, 9(26): 5041–5049, December 1970. ISSN 0006-2960. . URL `https://doi.org/10.1021/bi00828a001`. Publisher: American Chemical Society.

Robin Bruyndonckx, Niels Adriaenssens, Ann Versporten, Niel Hens, Dominique L Monnet, Geert Molenberghs, Herman Goossens, Klaus Weist, and Samuel Coenen. Consumption of antibiotics in the community, european union/european economic area, 1997–2017. *Journal of Antimicrobial Chemotherapy*, 76(Supplement_2):ii7–ii13, 2021.

Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101, 2007. .

Yolanda Cajal, Joseph Rogers, Otto G Berg, and Mahendra K Jain. Intermembrane molecular contacts by polymyxin b mediate exchange of phospholipids. *Biochemistry*, 35(1):299–308, 1996.

David Cameron. *Prime Minister warns of global threat of antibiotic resistance [Press Release]*. July 2014. URL `www.gov.uk/government/news/prime-minister-warns-of-global-threat-of-antibiotic-resistance`.

Martine Caroff and Doris Karibian. Structure of bacterial lipopolysaccharides. *Carbohydrate research*, 338(23):2431–2447, 2003.

D. S. Cayley, H. J. Guttman, and M. T. Record. Biophysical characterization of changes in amounts and activity of Escherichia coli cell and compartment water and turgor pressure in response to osmotic stress. *Biophysical Journal*, 78(4):1748–1764, April 2000. ISSN 0006-3495. .

Center for Disease Control, Prevention, et al. Outpatient antibiotic prescriptions—united states, 2014. *Availble at: https://www. cdc. gov/antibiotic-use/community/programs-measurement/state-local-activities/outpatient-antibiotic-prescriptions-US-2018. html (Accessed February 10, 2022)*, 2020.

Aparajita Chakraborty, Elisey Kobzev, Jonathan Chan, Gayan Heruka de Zoysa, Vijayalekshmi Sarojini, Thomas J Piggot, and Jane R Allison. Molecular dynamics simulation of the interaction of two linear battacin analogs with model gram-positive and gram-negative bacterial cell membranes. *ACS omega*, 6(1): 388–400, 2020.

Adrià Clausell, Maria Garcia-Subirats, Montserrat Pujol, M Antonia Busquets, Francesc Rabanal, and Yolanda Cajal. Gram-negative outer and inner membrane models: insertion of cyclic cationic lipopeptides. *The journal of physical chemistry B*, 111(3):551–563, 2007.

Eli J Cohen, Josie L Ferreira, Mark S Ladinsky, Morgan Beeby, and Kelly T Hughes. Nanoscale-length control of the flagellar driveshaft requires hitting the tethered outer membrane. *Science*, 356(6334):197–200, 2017.

Seymour S Cohen. *A Guide to the Polyamines*. Oxford Univ Press, New York, NY USA, 1997.

Robin A Corey, Wanling Song, Anna L Duncan, T Bertie Ansell, Mark SP Sansom, and Phillip J Stansfeld. Identification and assessment of cardiolipin interactions with e. coli inner membrane proteins. *Science advances*, 7(34):eabh2217, 2021.

Felsis Angelene Daison, Nitheeshkumar Kumar, Siranjeevi Balakrishnan, Kavyashree Venugopal, Sangamithra Elango, and Pandian Sokkar. Molecular dynamics studies on the bacterial membrane pore formation by small molecule antimicrobial agents. *Journal of Chemical Information and Modeling*, 62(1):40–48, 2021.

Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12): 10089–10092, June 1993. ISSN 0021-9606. . URL `https://doi.org/10.1063/1.464397`.

K Heran Darwin. Prokaryotic ubiquitin-like protein (pup), proteasomes and pathogenesis. *Nature Reviews Microbiology*, 7(7):485–491, 2009.

Xavier Daura, Karl Gademann, Bernhard Jaun, Dieter Seebach, Wilfred F Van Gunsteren, and Alan E Mark. Peptide folding: when simulation meets experiment. *Angewandte Chemie International Edition*, 38(1-2):236–240, 1999.

Djurre H de Jong, Gurpreet Singh, WF Drew Bennett, Clement Arnarez, Tsjerk A Wassenaar, Lars V Schafer, Xavier Periole, D Peter Tieleman, and Siewert J Marrink. Improved parameters for the martini coarse-grained protein force field. *Journal of chemical theory and computation*, 9(1):687–697, 2013.

Zakuan Z. Deris, James D. Swarbrick, Kade D. Roberts, Mohammad A. K. Azad, Jesmin Akter, Andrew S. Horne, Roger L. Nation, Kelly L. Rogers, Phillip E. Thompson, Tony Velkov, and Jian Li. Probing the penetration of antimicrobial polymyxin lipopeptides into gram-negative bacteria. *Bioconjugate Chemistry*, 25(4): 750–760, April 2014. ISSN 1520-4812. .

Shweta Devi, Minal Chaturvedi, Siraj Fatima, and Smriti Priya. Environmental factors modulating protein conformations and their role in protein aggregation diseases. *Toxicology*, 465:153049, 2022.

Thibaud Dieudonné, Felix Kümmerer, Michelle Juknaviciute Laursen, Charlott Stock, Rasmus Kock Flygaard, Syma Khalid, Guillaume Lenoir, Joseph A. Lyons, Kresten Lindorff-Larsen, and Poul Nissen. Activation and substrate specificity of the human P4-ATPase ATP8B1. *Nature Communications*, 14(1):7492, November 2023. ISSN 2041-1723. . URL `https://doi.org/10.1038/s41467-023-42828-9`.

Natallia V. Dubashynskaya and Yury A. Skorik. Polymyxin Delivery Systems: Recent Advances and Challenges. *Pharmaceuticals*, 13(5):83, April 2020. ISSN 1424-8247. . URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7281078/`.

André C. Dumetz, Ann M. Snellinger-O'Brien, Eric W. Kaler, and Abraham M. Lenhoff. Patterns of protein–protein interactions in salt solutions and implications for protein crystallization. *Protein Science : A Publication of the Protein Society*, 16(9): 1867–1877, September 2007. ISSN 0961-8368. . URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2206983/`.

Fernando G. Dupuy, Isabella Pagano, Kathryn Andenoro, Maria F. Peralta, Yasmene Elhady, Frank Heinrich, and Stephanie Tristram-Nagle. Selective Interaction of Colistin with Lipid Model Membranes. *Biophysical Journal*, 114(4):919–928, February 2018. ISSN 1542-0086. .

Jacob D Durrant and J Andrew McCammon. Molecular dynamics simulations and drug discovery. *BMC biology*, 9(1):1–9, 2011.

Frank Eisenhaber, Philip Lijnzaad, Patrick Argos, Chris Sander, and Michael Scharf. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *Journal of computational chemistry*, 16(3):273–284, 1995.

Khaled M Elokely and Robert J Doerksen. Docking challenge: protein sampling and molecular docking performance. *Journal of chemical information and modeling*, 53(8): 1934–1945, 2013.

Robert Ernst, Christer S Ejsing, and Bruno Antonny. Homeoviscous adaptation and the regulation of membrane lipids. *Journal of molecular biology*, 428(24):4776–4791, 2016.

Martin Exner, Sanjay Bhattacharya, Bärbel Christiansen, Jürgen Gebel, Peter Goroncy-Bermes, Philippe Hartemann, Peter Heeg, Carola Ilschner, Axel Kramer, Elaine Larson, et al. Antibiotic resistance: What is so special about multidrug-resistant gram-negative bacteria? *GMS hygiene and infection control*, 12, 2017.

Matthew E. Falagas and Sofia K. Kasiakou. Toxicity of polymyxins: a systematic review of the evidence from old and recent studies. *Critical Care (London, England)*, 10(1):R27, February 2006. ISSN 1466-609X. .

Matthew E. Falagas, Margarita Kyriakidou, Georgios L. Voulgaris, Filippos Vokos, Sevasti Politi, and Konstantinos S. Kechagias. Clinical use of intravenous polymyxin B for the treatment of patients with multidrug-resistant Gram-negative bacterial infections: An evaluation of the current evidence. *Journal of Global Antimicrobial Resistance*, 24:342–359, March 2021. ISSN 2213-7173. .

K. Anton Feenstra, Berk Hess, and Herman J. C. Berendsen. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *Journal of Computational Chemistry*, 20(8):786–798, 1999. .

Mónica Fernández-Vidal, Sajith Jayasinghe, Alexey S Ladokhin, and Stephen H White. Folding amphipathic helices into membranes: amphiphilicity trumps hydrophobicity. *Journal of molecular biology*, 370(3):459–470, 2007.

Daan Frenkel and Berend Smit. page 87–88. Academic Press, 2 edition, 2002.

Lei Fu, Mingwei Wan, Shan Zhang, Lianghui Gao, and Weihai Fang. Polymyxin b loosens lipopolysaccharide bilayer but stiffens phospholipid bilayer. *Biophysical journal*, 118(1):138–150, 2020.

Y Funahara and H Nikaido. Asymmetric localization of lipopolysaccharides on the outer membrane of salmonella typhimurium. *Journal of bacteriology*, 141(3): 1463–1465, 1980.

Ana C Gales, Ronald N Jones, and Hélio S Sader. Contemporary activity of colistin and polymyxin b against a worldwide collection of gram-negative pathogens: results from the sentry antimicrobial surveillance program (2006–09). *Journal of Antimicrobial Chemotherapy*, 66(9):2070–2074, 2011.

Anne Gershenson, Shachi Gosavi, Pietro Faccioli, and Patrick L Wintrode. Successes and challenges in simulating the folding of large proteins. *Journal of Biological Chemistry*, 295(1):15–33, 2020.

Bernd Glauner, JV Höltje, and Uli Schwarz. The composition of the murein of escherichia coli. *Journal of Biological Chemistry*, 263(21):10088–10095, 1988.

Camille Goemans, Katleen Denoncin, and Jean-François Collet. Folding mechanisms of periplasmic proteins. *Biochimica Et Biophysica Acta*, 1843(8):1517–1528, August 2014. ISSN 0006-3002. .

Richard J. Gowers, Max Linke, Jonathan Barnoud, Tyler J. E. Reddy, Manuel N. Melo, Sean L. Seyler, Jan Domański, David L. Dotson, Sébastien Buchoux, Ian M. Kenney, and Oliver Beckstein. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proceedings of the 15th Python in Science Conference*, pages 98–105, 2016. . URL `https://conference.scipy.org/proceedings/scipy2016/oliver_beckstein.html`. Conference Name: Proceedings of the 15th Python in Science Conference.

Christian Gram. Ueber die isolirte farbung der schizomyceten in schnitt-und trockenpraparaten. *Fortschritte der Medicin*, 2:185–189, 1884.

Erin R Green and Joan Mecsas. Bacterial secretion systems: an overview. *Virulence mechanisms of bacterial pathogens*, pages 213–239, 2016.

Gernot Guigas and Matthias Weiss. Effects of protein crowding on membrane systems. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(10):2441 – 2450, 2016. ISSN 0005-2736. .

James C Gumbart, Morgan Beeby, Grant J Jensen, and Benoît Roux. Escherichia coli peptidoglycan structure and mechanics as predicted by atomic-scale simulations. *PLoS computational biology*, 10(2):e1003475, 2014.

Simon Halegoua, Akikazu Hirashima, and Masayori Inouye. Existence of a Free Form of a Specific Membrane Lipoprotein in Gram-Negative Bacteria. *Journal of Bacteriology*, 120(3):1204–1208, December 1974. . URL `https://journals.asm.org/doi/10.1128/jb.120.3.1204-1208.1974`. Publisher: American Society for Microbiology.

Robert EW Hancock. Peptide antibiotics. *The Lancet*, 349(9049):418–422, February 1997. ISSN 0140-6736. . URL `https://www.sciencedirect.com/science/article/pii/S0140673697800517`.

Klaus Hantke and Volkmar Braun. Covalent binding of lipid to protein: Diglyceride and amide-linked fatty acid at the n-terminal end of the murein-lipoprotein of the escherichia coli outer membrane. *European Journal of Biochemistry*, 34(2):284–296, 1973.

Marcus D. Hanwell, Donald E. Curtis, David C. Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R. Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4(1):17, August 2012. ISSN 1758-2946. . URL `https://doi.org/10.1186/1758-2946-4-17`.

Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020. . URL `https://doi.org/10.1038/s41586-020-2649-2`.

Sergio A. Hassan. Amino Acid Side Chain Interactions in the Presence of Salts. *The Journal of Physical Chemistry B*, 109(46):21989–21996, November 2005. ISSN

1520-6106. . URL `https://doi.org/10.1021/jp054042r`. Publisher: American Chemical Society.

Pablo Herrera-Nieto, Adrià Pérez, and Gianni De Fabritiis. Binding-and-folding recognition of an intrinsically disordered protein using online learning molecular dynamics. *Journal of Chemical Theory and Computation*, 2023.

Berk Hess. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, January 2008. ISSN 1549-9618. . URL `https://doi.org/10.1021/ct700200b`. Publisher: American Chemical Society.

Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1997. ISSN 1096-987X. .

Dick Hoekstra, Jan Willem van der Laan, Loe de Leij, and Bernard Witholt. Release of outer membrane fragments from normally growing Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 455(3):889–899, December 1976. ISSN 0005-2736. . URL `https://www.sciencedirect.com/science/article/pii/0005273676900584`.

Chad W. Hopkins, Scott Le Grand, Ross C. Walker, and Adrian E. Roitberg. Long-time-step molecular dynamics through hydrogen mass repartitioning. *Journal of Chemical Theory and Computation*, 11(4):1864–1874, 2015. .

Adam Hospital, Josep Ramon Goñi, Modesto Orozco, and Josep L Gelpí. Molecular dynamics simulations: advances and applications. *Advances and Applications in Bioinformatics and Chemistry*, pages 37–47, 2015.

Jing Huang, Sarah Rauscher, Grzegorz Nawrocki, Ting Ran, Michael Feig, Bert L. de Groot, Helmut Grubmüller, and Alexander D. MacKerell. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, 14(1):71–73, January 2017. ISSN 1548-7105. . URL `https://www.nature.com/articles/nmeth.4067`. Number: 1 Publisher: Nature Publishing Group.

Kerwyn Casey Huang, Ranjan Mukhopadhyay, Bingni Wen, Zemer Gitai, and Ned S Wingreen. Cell shape and cell-wall organization in gram-negative bacteria. *Proceedings of the National Academy of Sciences*, 105(49):19282–19287, 2008.

W. Humphrey, A. Dalke, and K. Schulten. VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 27–28, February 1996. ISSN 0263-7855. .

J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. .

Kevin S Ikuta, Lucien R Swetschinski, Gisela Robles Aguilar, Fablina Sharara, Tomislav Mestrovic, Authia P Gray, Nicole Davis Weaver, Eve E Wool, Chieh Han, Anna Gershberg Hayoon, et al. Global mortality associated with 33 bacterial pathogens in 2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 400(10369):2221–2248, 2022.

Wonpil Im and Charles L Brooks III. Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proceedings of the National Academy of Sciences*, 102(19):6771–6776, 2005.

Arash Izadpanah and Richard L. Gallo. Antimicrobial peptides. *Journal of the American Academy of Dermatology*, 52(3):381–390, March 2005. ISSN 0190-9622. . URL `https://www.sciencedirect.com/science/article/pii/S0190962204022066`.

Zack Jarin, James Newhouse, and Gregory A Voth. Coarse-grained force fields from the perspective of statistical mechanics: Better understanding of the origins of a martini hangover. *Journal of Chemical Theory and Computation*, 17(2):1170–1180, 2021.

Dušan Jasovský, Jasper Littmann, Anna Zorzet, and Otto Cars. Antimicrobial resistance-a threat to the world's sustainable development. *Upsala Journal of Medical Sciences*, 121(3):159–164, August 2016. ISSN 2000-1967. .

Damien Jefferies, Jonathan Shearer, and Syma Khalid. Role of o-antigen in response to mechanical stress of the e. coli outer membrane: Insights from coarse-grained md simulations. *The Journal of Physical Chemistry B*, 123(17):3567–3575, 2019. . PMID: 30971088.

Xukai Jiang, Kai Yang, Bing Yuan, Bin Gong, Lin Wan, Nitin A Patil, James D Swarbrick, Kade D Roberts, Falk Schreiber, Lushan Wang, et al. Simulations of octapeptin–outer membrane interactions reveal conformational flexibility is linked to antimicrobial potency. *Journal of Biological Chemistry*, 295(47):15902–15912, 2020a.

Xukai Jiang, Kai Yang, Bing Yuan, Meiling Han, Yan Zhu, Kade D Roberts, Nitin A Patil, Jingliang Li, Bin Gong, Robert EW Hancock, et al. Molecular dynamics simulations informed by membrane lipidomics reveal the structure–interaction relationship of polymyxins with the lipid a-based outer membrane of acinetobacter baumannii. *Journal of Antimicrobial Chemotherapy*, 75(12):3534–3543, 2020b.

Xukai Jiang, Shuo Zhang, Mohammad A. K. Azad, Kade D. Roberts, Lin Wan, Bin Gong, Kai Yang, Bing Yuan, Hemayet Uddin, Jingliang Li, Philip E. Thompson, Tony Velkov, Jing Fu, Lushan Wang, and Jian Li. Structure-Interaction Relationship of Polymyxins with the Membrane of Human Kidney Proximal Tubular Cells. *ACS infectious diseases*, 6(8):2110–2119, August 2020c. ISSN 2373-8227. . URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7485602/`.

Sunhwan Jo, Taehoon Kim, Vidyashankara G Iyer, and Wonpil Im. Charmm-gui: a web-based graphical user interface for charmm. *Journal of computational chemistry*, 29 (11):1859–1865, 2008.

William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, July 1983. ISSN 0021-9606. . URL `https://doi.org/10.1063/1.445869`.

Rintaro Kaguchi, Akira Katsuyama, Toyotaka Sato, Satoshi Takahashi, Motohiro Horiuchi, Shin-ichi Yokota, and Satoshi Ichikawa. Discovery of biologically optimized polymyxin derivatives facilitated by peptide scanning and in situ screening chemistry. *Journal of the American Chemical Society*, 145(6):3665–3681, 2023a.

Rintaro Kaguchi, Akira Katsuyama, Toyotaka Sato, Satoshi Takahashi, Motohiro Horiuchi, Shin-ichi Yokota, and Satoshi Ichikawa. Discovery of Biologically Optimized Polymyxin Derivatives Facilitated by Peptide Scanning and In Situ Screening Chemistry. *Journal of the American Chemical Society*, 145(6):3665–3681, February 2023b. ISSN 0002-7863. . URL `https://doi.org/10.1021/jacs.2c12971`. Publisher: American Chemical Society.

Peter M. Kekenes-Huskey, Caitlin E. Scott, and Selcuk Atalay. Quantifying the influence of the crowded cytoplasm on small molecule diffusion. *The Journal of Physical Chemistry B*, 120(33):8696–8706, 2016. .

Nawal K Khadka, Chinta M Aryal, and Jianjun Pan. Lipopolysaccharide-dependent membrane permeation and lipid clustering caused by cyclic lipopeptide colistin. *ACS omega*, 3(12):17828–17834, 2018.

Syma Khalid, Thomas J. Piggot, and Firdaus Samsudin. Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned? *Accounts of Chemical Research*, 52(1):180–188, January 2019. ISSN 1520-4898. .

Seonghoon Kim, Jumin Lee, Sunhwan Jo, Charles L. Brooks, Hui Sun Lee, and Wonpil Im. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *Journal of Computational Chemistry*, 38(21):1879–1886, June 2017. ISSN 1096-987X. .

Y. Kimura, H. Matsunaga, and M. Vaara. Polymyxin B octapeptide and polymyxin B heptapeptide are potent outer membrane permeability-increasing agents. *The Journal of Antibiotics*, 45(5):742–749, May 1992. ISSN 0021-8820. .

Ralf Koebnik, Kaspar P Locher, and Patrick Van Gelder. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Molecular microbiology*, 37 (2):239–253, 2000.

Bjoern F. Kraemer, Robert A. Campbell, Hansjörg Schwertz, Mark J. Cody, Zechariah Franks, Neal D. Tolley, Walter H. A. Kahr, Stephan Lindemann, Peter Seizer, Christian C. Yost, Guy A. Zimmerman, and Andrew S. Weyrich. Novel anti-bacterial activities of -defensin 1 in human platelets: suppression of pathogen growth and signaling of neutrophil extracellular trap formation. *PLoS pathogens*, 7 (11):e1002355, November 2011. ISSN 1553-7374. .

Irina M. Kuznetsova, Konstantin K. Turoverov, and Vladimir N. Uversky. What Macromolecular Crowding Can Do to a Protein. *International Journal of Molecular Sciences*, 15(12):23090–23140, December 2014. ISSN 1422-0067. . URL `https://www.mdpi.com/1422-0067/15/12/23090`. Number: 12 Publisher: Multidisciplinary Digital Publishing Institute.

P Lalgudi and AH Elcock. Molecular dynamics simulations of periplasmic proteins interacting with the peptidoglycan layer of escherichia coli. *Journal of Emerging Investigators*, 2016.

Nick Lane. *Life ascending: the ten great inventions of evolution*. Profile books, 2010.

Elizabeth VK Ledger, Akshay Sabnis, and Andrew M Edwards. Polymyxin and lipopeptide antibiotics: membrane-targeting drugs of last resort. *Microbiology*, 168 (2), 2022.

Christopher T Lee, Jeffrey Comer, Conner Herndon, Nelson Leung, Anna Pavlova, Robert V Swift, Chris Tung, Christopher N Rowley, Rommie E Amaro, Christophe Chipot, et al. Simulation-based approaches for determining membrane permeability of small compounds. *Journal of chemical information and modeling*, 56(4):721–733, 2016.

Jumin Lee, Dhilon S. Patel, Jonas Ståhle, Sang-Jun Park, Nathan R. Kern, Seonghoon Kim, Joonseong Lee, Xi Cheng, Miguel A. Valvano, Otto Holst, Yuriy A. Knirel, Yifei Qi, Sunhwan Jo, Jeffery B. Klauda, Göran Widmalm, and Wonpil Im. CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *Journal of Chemical Theory and Computation*, 15(1): 775–786, January 2019. ISSN 1549-9626. .

Jun Lei, Lichun Sun, Siyu Huang, Chenhong Zhu, Ping Li, Jun He, Vienna Mackey, David H. Coy, and Quanyong He. The antimicrobial peptides and their potential clinical applications. *American Journal of Translational Research*, 11(7):3919–3931, 2019. ISSN 1943-8141.

Ming Li, Yang Li, Yue Lu, Jianhui Li, Xuhang Lu, Yue Ren, Tianlei Wen, Yaojie Wang, Shenghai Chang, Xing Zhang, et al. Molecular basis of mg2+ permeation through the human mitochondrial mrs2 channel. *Nature communications*, 14(1):4713, 2023.

Zekun Li, Yuping Cao, Lingxian Yi, Jian-Hua Liu, and Qiwen Yang. Emergent polymyxin resistance: end of an era? In *Open forum infectious diseases*, volume 6, page ofz368. Oxford University Press US, 2019.

M. Yu. Lobanov, N. S. Bogatyreva, and O. V. Galzitskaya. Radius of gyration as an indicator of protein structure compactness. *Molecular Biology*, 42(4):623–628, Aug 2008. ISSN 1608-3245. .

John Loree and Sarah L Lappin. Bacteriostatic antibiotics. 2019. URL `https://pubmed.ncbi.nlm.nih.gov/31613458/`.

Ricardo Lucas, Pablo Peñalver, Irene Gómez-Pinto, Empar Vengut-Climent, Lewis Mtashobya, Jonathan Cousin, Olivia S. Maldonado, Violaine Perez, Virginie Reynes, Anna Aviñó, Ramón Eritja, Carlos González, Bruno Linclau, and Juan C. Morales. Effects of Sugar Functional Groups, Hydrophobicity, and Fluorination on Carbohydrate–DNA Stacking Interactions in Water. *The Journal of Organic Chemistry*, 79(6):2419–2429, March 2014. ISSN 0022-3263. . URL `https://doi.org/10.1021/jo402700y`. Publisher: American Chemical Society.

Ben Lugtenberg and Loek Van Alphen. Molecular architecture and functioning of the outer membrane of escherichia coli and other gram-negative bacteria. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 737(1):51–115, 1983.

Dmitry Lyumkis. Challenges and opportunities in cryo-em single-particle analysis. *Journal of Biological Chemistry*, 294(13):5181–5197, 2019.

J Randy Macdonald and W Curtis Johnson JR. Environmental features are important in determining protein secondary structure. *Protein Science*, 10(6):1172–1177, 2001.

Alexander D MacKerell Jr, Michael Feig, and Charles L Brooks. Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society*, 126(3):698–699, 2004.

Valério RF Matias, Ashraf Al-Amoudi, Jacques Dubochet, and Terry J Beveridge. Cryo-transmission electron microscopy of frozen-hydrated sections of escherichia coli and pseudomonas aeruginosa. *Journal of bacteriology*, 185(20):6112–6118, 2003.

Melanie A McDowell, Michael Heimes, Giray Enkavi, Ákos Farkas, Daniel Saar, Klemens Wild, Blanche Schwappach, Ilpo Vattulainen, and Irmgard Sinning. The get insertase exhibits conformational plasticity and induces membrane thinning. *Nature Communications*, 14(1):7355, 2023.

Xuan-Yu Meng, Hong-Xing Zhang, Mihaly Mezei, and Meng Cui. Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2):146–157, 2011.

Naveen Michaud-Agrawal, Elizabeth J. Denning, Thomas B. Woolf, and Oliver Beckstein. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, 32(10):2319–2327, 2011. ISSN 1096-987X. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21787. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.21787.

Samuel I Miller. Antibiotic resistance and regulation of the gram-negative bacterial outer membrane barrier by host innate immune molecules. *MBio*, 7(5):10–1128, 2016.

Grzegorz Nawrocki, Wonpil Im, Yuji Sugita, and Michael Feig. Clustering and dynamics of crowded proteins near membranes and their influence on membrane bending. *Proceedings of the National Academy of Sciences*, 116(49):24562–24567, 2019. ISSN 0027-8424. .

Chris Neale and Régis Pomès. Sampling errors in free energy simulations of small molecules in lipid bilayers. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1858 (10):2539–2548, 2016.

Chris Neale, WF Drew Bennett, D Peter Tieleman, and Régis Pomes. Statistical convergence of equilibrium properties in simulations of molecular solutes embedded in lipid bilayers. *Journal of Chemical Theory and Computation*, 7(12): 4175–4188, 2011.

Junio Oliveira and Wanda C. Reygaert. Gram-negative bacteria, 2019. URL https://www.ncbi.nlm.nih.gov/books/NBK538213/.

J O'Neill. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. Technical report, Wellcome Collection, May 2016. URL https://wellcomecollection.org/works/thvwsuba.

Maite L. Ortiz-Suarez, Firdaus Samsudin, Thomas J. Piggot, Peter J. Bond, and Syma Khalid. Full-Length OmpA: Structure, Function, and Membrane Interactions Predicted by Molecular Dynamics Simulations. *Biophysical Journal*, 111(8):1692–1702, October 2016. ISSN 0006-3495. . URL https://www.sciencedirect.com/science/article/pii/S0006349516308025.

Natalia Ostrowska, Michael Feig, and Joanna Trylska. Modeling crowded environment in molecular simulations. *Frontiers in Molecular Biosciences*, 6:86, 2019.

Malvina Papanastasiou, Georgia Orfanoudaki, Marina Koukaki, Nikos Kountourakis, Marios Frantzeskos Sardis, Michalis Aivaliotis, Spyridoula Karamanou, and Anastassios Economou. The escherichia coli peripheral inner membrane proteome. *Molecular & Cellular Proteomics*, 12(3):599–610, 2013.

Jae-Hyun Park, Kouki Kawakami, Naito Ishimoto, Tatsuya Ikuta, Mio Ohki, Toru Ekimoto, Mitsunori Ikeguchi, Dong-Sun Lee, Young-Ho Lee, Jeremy RH Tame, et al. Structural basis for ligand recognition and signaling of hydroxy-carboxylic acid receptor 2. *Nature Communications*, 14(1):7150, 2023.

Sang-Jun Park, Jumin Lee, Yifei Qi, Nathan R Kern, Hui Sun Lee, Sunhwan Jo, InSuk Joung, Keehyung Joo, Jooyoung Lee, and Wonpil Im. CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology*, 29(4):320–331, April 2019. ISSN 1460-2423. . URL `https://doi.org/10.1093/glycob/cwz003`.

M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, December 1981. ISSN 0021-8979. . URL `https://doi.org/10.1063/1.328693`.

Conrado Pedebos, Iain Peter Shand Smith, Alister Boags, and Syma Khalid. The hitchhiker's guide to the periplasm: Unexpected molecular interactions of polymyxin B1 in E. coli. *Structure*, 29(5):444–456, May 2021. ISSN 1878-4186. .

Thomas J. Piggot, Daniel A. Holdbrook, and Syma Khalid. Electroporation of the E. coli and S. Aureus membranes: molecular dynamics simulations of complex bacterial membranes. *The Journal of Physical Chemistry. B*, 115(45):13381–13388, November 2011. ISSN 1520-5207. .

Luca Pinzi and Giulio Rastelli. Molecular docking: shifting paradigms in drug discovery. *International journal of molecular sciences*, 20(18):4331, 2019.

Benjamin Plackett. Why big pharma has abandoned antibiotics. *Nature*, 586(7830): S50–S52, October 2020. . URL `https://www.nature.com/articles/d41586-020-02884-3`. Bandiera_abtest: a Cg_type: Outlook Number: 7830 Publisher: Nature Publishing Group Subject_term: Economics, Drug discovery, Policy, Health care.

Mithila Rajagopal and Suzanne Walker. Envelope structures of gram-positive bacteria. *Protein and sugar export and assembly in Gram-positive bacteria*, pages 1–44, 2017.

Chet Ramey et al. Bash reference manual. *Network Theory Limited*, 15, 1998.

Darnelle Ramirez and Mariana Giron. Enterobacter infections, 2023. URL `http://europepmc.org/books/NBK559296`.

Peter Reeves. Role of o-antigen variation in the immune response. *Trends in microbiology*, 3(10):381–386, 1995.

Ernst T Rietschel, Teruo Kirikae, F Ulrich Schade, Uwe Mamat, Günter Schmidt, Harald Loppnow, Artur J Ulmer, Ulrich Zähringer, Ulrich Seydel, Franco Di Padova, et al. Bacterial endotoxin: molecular relationships of structure to activity and function. *The FASEB Journal*, 8(2):217–225, 1994.

Leif Rilfors and Göran Lindblom. Regulation of lipid composition in biological membranes—biophysical studies of lipids and lipid synthesizing enzymes. *Colloids and Surfaces B: Biointerfaces*, 26(1-2):112–124, 2002.

Paul Robustelli, Stefano Piana, and David E. Shaw. Mechanism of coupled folding-upon-binding of an intrinsically disordered protein. *Journal of the American Chemical Society*, 142(25):11092–11101, 2020. .

Veronica W Rowlett, Venkata KPS Mallampalli, Anja Karlstaedt, William Dowhan, Heinrich Taegtmeyer, William Margolin, and Heidi Vitrac. Impact of membrane phospholipid alterations in escherichia coli on cellular function and bacterial stress adaptation. *Journal of bacteriology*, 199(13):10–1128, 2017.

Akshay Sabnis, Katheryn LH Hagart, Anna Klöckner, Michele Becce, Lindsay E Evans, R Christopher D Furniss, Despoina AI Mavridou, Ronan Murphy, Molly M Stevens, Jane C Davies, et al. Colistin kills bacteria by targeting lipopolysaccharide in the cytoplasmic membrane. *elife*, 10:e65836, 2021.

Milton RJ Salton, Kwang-Shin Kim, and S Baron. Medical microbiology. *University of Texas Medical Branch at Galveston: Galveston, TX*, 1996.

Firdaus Samsudin, Maite L. Ortiz-Suarez, Thomas J. Piggot, Peter J. Bond, and Syma Khalid. OmpA: A Flexible Clamp for Bacterial Cell Wall Attachment. *Structure (London, England: 1993)*, 24(12):2227–2235, December 2016. ISSN 1878-4186. .

Firdaus Samsudin, Alister Boags, Thomas J. Piggot, and Syma Khalid. Braun's Lipoprotein Facilitates OmpA Interaction with the Escherichia coli Cell Wall. *Biophysical Journal*, 113(7):1496–1504, October 2017. ISSN 0006-3495. . URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5627309/`.

Paramita Sarkar, Venkateswarlu Yarlagadda, Chandradhish Ghosh, and Jayanta Haldar. A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics. *Medchemcomm*, 8(3):516–533, 2017.

KARL HEINZ Schleifer and Otto Kandler. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological reviews*, 36(4):407–477, 1972.

Carmen Schwechheimer, Adam Kulp, and Meta J. Kuehn. Modulation of bacterial outer membrane vesicle production by envelope structure and content. *BMC Microbiology*, 14(1):324, December 2014. ISSN 1471-2180. . URL `https://doi.org/10.1186/s12866-014-0324-1`.

Pratik Shah and Edwin Swiatlo. A multifaceted role for polyamines in bacterial pathogens. *Molecular Microbiology*, 68(1):4–16, April 2008. ISSN 1365-2958. .

Y. Shai. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochimica Et Biophysica Acta*, 1462(1-2):55–70, December 1999. ISSN 0006-3002. .

Pradyumn Sharma and K Ganapathy Ayappa. A molecular dynamics study of antimicrobial peptide interactions with the lipopolysaccharides of the outer bacterial membrane. *The Journal of Membrane Biology*, 255(6):665–675, 2022.

Pradyumn Sharma, Rajat Desikan, and K Ganapathy Ayappa. Evaluating coarse-grained martini force-fields for capturing the ripple phase of lipid membranes. *The Journal of Physical Chemistry B*, 125(24):6587–6599, 2021.

Jonathan Shearer, Damien F Jefferies, and Syma Khalid. The outer membrane proteins ompa, fhua, ompf, esta, btub and ompx have unique lipopolysaccharide fingerprints. *Biophysical Journal*, 116(3):374a, 2019.

Yigong Shi. A glimpse of structural biology through x-ray crystallography. *Cell*, 159 (5):995–1014, 2014.

Utsab R Shrestha, Jeremy C Smith, and Loukas Petridis. Full structural ensembles of intrinsically disordered proteins from unbiased molecular dynamics simulations. *Communications biology*, 4(1):243, 2021.

W. Shu, J. Liu, H. Ji, and M. Lu. Core structure of the outer membrane lipoprotein from Escherichia coli at 1.9 A resolution. *Journal of Molecular Biology*, 299(4): 1101–1112, June 2000. ISSN 0022-2836. .

Thomas J Silhavy, Daniel Kahne, and Suzanne Walker. The bacterial cell envelope. *Cold Spring Harbor perspectives in biology*, 2(5):a000414, 2010.

Ana-Maria Simundic et al. Confidence interval. *Biochemia Medica*, 18(2):154–161, 2008.

Michael Sinensky. Temperature control of phospholipid biosynthesis in escherichia coli. *Journal of bacteriology*, 106(2):449–455, 1971.

Iain Peter Shand Smith, Conrado Pedebos, and Syma Khalid. Molecular crowding alters the interactions of polymyxin lipopeptides within the periplasm of e. coli: insights from molecular dynamics. *The Journal of Physical Chemistry B*, 128(11): 2717–2733, 2024.

Sara M Soto. Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. *Virulence*, 4(3):223–229, 2013.

D. R. Storm, K. S. Rosenthal, and P. E. Swanson. Polymyxin and related peptide antibiotics. *Annual Review of Biochemistry*, 46:723–763, 1977. ISSN 0066-4154. .

Hideho Suzuki, Yukinobu Nishimura, Seiichi Yasuda, Akiko Nishimura, Masao Yamada, and Yukinori Hirota. Murein-lipoprotein of Escherichia coli: A protein involved in the stabilization of bacterial cell envelope. *Molecular and General Genetics MGG*, 167(1):1–9, January 1978. ISSN 1432-1874. . URL `https://doi.org/10.1007/BF00270315`.

Joanna Szczepaniak, Cara Press, and Colin Kleanthous. The multifarious roles of tol-pal in gram-negative bacteria. *FEMS Microbiology Reviews*, 44(4):490–506, 2020.

Ilario G Tironi, René Sperb, Paul E Smith, and Wilfred F van Gunsteren. A generalized reaction field method for molecular dynamics simulations. *The Journal of chemical physics*, 102(13):5451–5459, 1995.

B.D. Todd and P.J. Daivis. *Nonequilibrium Molecular Dynamics: Theory, Algorithms and Applications*, page 1–4. Cambridge University Press, 2017. .

Buu Minh Tran, Christiaan Michiel Punter, Dmitrii Linnik, Aditya Iyer, and Bert Poolman. Single-protein diffusion in the periplasm of escherichia coli. *Journal of Molecular Biology*, page 168420, 2023.

Nishant Tripathi and Amit Sapra. Gram staining, 2023. URL `http://europepmc.org/books/NBK562156`.

Martti Vaara. Polymyxins and Their Potential Next Generation as Therapeutic Antibiotics. *Frontiers in Microbiology*, 10:1689, July 2019. ISSN 1664-302X. . URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6671869/`.

Rakesh Vaiwala, Pradyumn Sharma, and K Ganapathy Ayappa. Differentiating interactions of antimicrobials with gram-negative and gram-positive bacterial cell walls using molecular dynamics simulations. *Biointerphases*, 17(6), 2022.

Guido Van Rossum, Fred L Drake, et al. *Python reference manual*, volume 111. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

Muriel CF van Teeseling, Miguel A de Pedro, and Felipe Cava. Determinants of bacterial morphology: from fundamentals to possibilities for antimicrobial targeting. *Frontiers in microbiology*, 8:1264, 2017.

K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell Jr. 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*, 31(4):671–690, 2010. ISSN 1096-987X. . URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21367`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.21367.

Tony Velkov, Philip E. Thompson, Roger L. Nation, and Jian Li. Structure—Activity Relationships of Polymyxin Antibiotics. *Journal of medicinal chemistry*, 53(5):1898, March 2010. . URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2907661/`. Publisher: NIH Public Access.

S. Vijay-Kumar, C. E. Bugg, and W. J. Cook. Structure of ubiquitin refined at 1.8 A resolution. *Journal of Molecular Biology*, 194(3):531–544, April 1987. ISSN 0022-2836. .

Waldemar Vollmer, Didier Blanot, and Miguel A De Pedro. Peptidoglycan structure and architecture. *FEMS microbiology reviews*, 32(2):149–167, 2008.

Sören von Bülow, Marc Siggel, Max Linke, and Gerhard Hummer. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proceedings of the National Academy of Sciences*, 116(20):9843–9852, 2019. ISSN 0027-8424. .

Callum Waller, Jan K Marzinek, Eilish McBurnie, Peter J Bond, Philip TF Williamson, and Syma Khalid. Impact on s. aureus and e. coli membranes of treatment with chlorhexidine and alcohol solutions: insights from molecular simulations and nuclear magnetic resonance. *Journal of Molecular Biology*, 435(11):167953, 2023.

Dongdong Wang, Jingwei Weng, and Wenning Wang. Glycerol transport through the aquaglyceroporin GlpF: bridging dynamics and kinetics with atomic simulation. *Chemical Science*, 10(29):6957–6965, July 2019. ISSN 2041-6539. . URL `https://pubs.rsc.org/en/content/articlelanding/2019/sc/c9sc01690b`. Publisher: The Royal Society of Chemistry.

Michael L. Waskom. Seaborn: Statistical Data Visualization. *Journal of Open Source Software*, 6(60):3021, April 2021. ISSN 2475-9066. . URL `https://joss.theoj.org/papers/10.21105/joss.03021`.

Melissa N Webby, Abraham O Oluwole, Conrado Pedebos, Patrick G Inns, Anna Olerinyova, Dheeraj Prakaash, Nicholas G Housden, Georgina Benn, Dawei Sun, Bart W Hoogenboom, et al. Lipids mediate supramolecular outer membrane protein assembly in bacteria. *Science Advances*, 8(44):eadc9566, 2022.

Joel H. Weiner and Liang Li. Proteome of the Escherichia coli envelope and technological challenges in membrane proteome analysis. *Biochimica Et Biophysica Acta*, 1778(9):1698–1713, September 2008. ISSN 0006-3002. .

Brent B Welch, Ken Jones, and Jeffrey Hobbs. *Practical programming in Tcl and Tk*. Prentice Hall Professional, 2003.

Emilia L Wu, Olof Engström, Sunhwan Jo, Danielle Stuhlsatz, Min Sun Yeom, Jeffery B Klauda, Göran Widmalm, and Wonpil Im. Molecular dynamics and nmr spectroscopy studies of e. coli lipopolysaccharide structure and dynamics. *Biophysical journal*, 105(6):1444–1455, 2013.

Xingjian Xu, Igor Dikiy, Matthew R Evans, Leandro P Marcelino, and Kevin H Gardner. Fragile protein folds: sequence and environmental factors affecting the equilibrium of two interconverting, stably folded protein conformations. *Magnetic Resonance*, 2(1):63–76, 2021.

X Yao, Manfred Jericho, D Pink, and T Beveridge. Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy. *Journal of bacteriology*, 181(22):6865–6875, 1999.

S B Zimmerman and A P Minton. Macromolecular crowding: Biochemical, biophysical, and physiological consequences. *Annual Review of Biophysics and Biomolecular Structure*, 22(1):27–65, 1993. .