MULTISCALE MODELS OF ANTIMICROBIAL PEPTIDES

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA BY

DAN BOLINTINEANU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

YIANNIS N. KAZNESSIS, Adviser

December, 2010

To my family

Abstract

Antimicrobial peptides (AMPs) are small proteins that constitute a first line of defense against invading pathogens in the innate immune systems of countless plant and animal species. Their mechanism of action relies to a large extent on selectively disrupting the cell membranes of bacteria, which makes them promising therapeutic agents in the fight against infectious pathogens, including antibiotic-resistant bacterial strains. However, AMPs also exhibit toxicity towards mammalian cells, which presents a significant bottleneck in the development of AMPs for antibiotic treatment in humans. Thus, before the full potential of AMPs as therapeutic agents can be unlocked, their fundamental mechanism of action must be understood in order to identify targets for mutation that can improve activity and ameliorate toxic effects. To this end, we carry out computer simulations and modeling studies in order to understand the interactions of protegrin, a particularly potent and well-studied AMP, with lipid bilayers that mimic the compositions of bacterial and mammalian cell membranes. In particular, we attempt to connect molecular-level thermodynamic information to biologically relevant membrane association equilibria; we elucidate the ion transport characteristics of protegrin pores, and show how the atomistic structure leads to experimentally observed conductance behavior; we provide multi-scale models that can quantitatively link the structure of protegrin and protegrin pores to the leakage of potassium ions from bacterial cells, which appears to be an essential element in the bactericidal mechanism of action of protegrin; finally, we investigate the structure of protegrin pores using molecular dynamics simulations with atomistic resolution. Our work reveals this link for the first time, and provides a united, quantitative connection between molecular-level structural and thermodynamic information and mesoscopic, measurable quantities. Our modeling tools can be easily extended to other AMPs, and protein-membrane systems in general.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 The growing problem of antibiotic resistance

The discovery of antibiotics in the first half of the twentieth century has arguably been the most life-saving invention in human history. As a result, millions of otherwise fatal infections are now routinely treated with antibiotics. However, within only four years of the start of mass production of penicillin, a strain of the bacterium *Staphylococcus aureus* emerged which possessed high levels of resistance to this early antibiotic. Modifications were subsequently made to penicillin, and several new classes of antibiotics were developed to combat the growing number of resistant strains. Unfortunately, with each pharmacological breakthrough, bacterial strains have consistently emerged that exhibit resistance to novel antibiotics, particularly in hospital intensive care wards. This has resulted in a continuous competition between drug design efforts aimed at combating emerging multi-drug resistant 'superbugs' on the one hand, and evolutionary forces in bacteria leading to novel resistance mechanisms on the other. Despite the growing problem of antibiotic-resistant bacteria, research and development of novel antibiotic agents has yielded relatively few novel therapeutics recently: in the past thirty years, only two new classes of antibiotics have been developed, neither of which are effective against Gram-negative infections (Gram-positive and Gram-negative bacteria are discussed in section 1.4), as compared to more than a dozen classes introduced between the 1930s and 1960s [1]. To make matters worse, the development of antibiotics is eschewed by major drug companies, as it does not entail significant financial incentives: a small number of doses typically resolve an infection, whereas other afflictions require life-long drug dependencies, resulting in much higher profit margins. Meanwhile, infectious diseases remain the third leading cause of death in the United States [2], with more people dying from methicillin-resistant *Staphylococcus aureus* (MRSA) infections in US hospitals than of HIV/AIDS and tuberculosis combined [3-4]. Furthermore, Gram-negative pathogens such as *Acinetobacter baumannii*,

Pseudonoma aeruginosa, and *Escherichia coli* are becoming increasingly resistant to commonly used antibiotics and emerging as major threats to human health [4-5]. According to reports compiled by the Infectious Disease Society of America (IDSA), nearly 2 million patients are afflicted with bacterial infections in U.S. hospitals each year, of which 90,000 die as a result, of which seventy percent are related to infections caused by antibiotic-resistant bacteria [6]. It was initially hoped that most strains of drug-resistant bacteria would disappear when treatment with the respective drug was halted for a sufficiently long period of time, due to the loss of any competitive advantage associated with the resistance-conferring mutation. However, this prediction has proven to be largely inaccurate, since many mutations that confer drug resistance have little or no associated evolutionary disadvantages [7]. The rapid emergence of antibiotic-resistant bacteria and the associated potential medical crisis show no signs of abatement in the near future.

1.2 Mechanisms of antibiotic resistance development

Traditional antibiotic design strategies are based on identifying key molecular targets in bacteria, and developing drug molecules that bind to and inhibit the functions of these targets. For instance, penicillin and methicillin both work by binding to and disabling the transpeptidase enzyme used by Gram-positive bacteria to construct their cell walls, which results in bacterial death upon replication. These approaches are advantageous in that they typically result in highly selective drugs that are only mildly toxic to human cells, but suffer from major drawbacks with regards to the potential for bacterial resistance - minor mutations in a target molecule can often render it impervious to antibiotic inhibition, while still retaining its natural function. The exposure of bacteria to nonlethal doses of antibiotics provides a strong evolutionary stress that leads to the emergence and proliferation of antibiotic-resistant bacterial strains. Given the extremely short reproduction times of bacteria (doubling times on the order of 30 minutes to several hours, even in suboptimal conditions), such evolutionary effects are highly amplified and

quickly apparent. The transfer of resistance genes among organisms under certain conditions further exacerbates the problem, leading to the rapid proliferation of antibiotic resistance genes among bacterial populations.

Although antibiotic design based on specific intracellular targeting by small molecules remains a viable strategy, it is highly unlikely that any such drug design efforts will provide a long-lasting solution to antibiotic-resistant bacteria. Furthermore, similarities among small molecule drugs and among their targets often translate to multiple drug resistance upon mutations, wherein a bacterial strain that develops resistance to a particular antibiotic thereby gains resistance to a host of other similar therapeutics or therapeutics that share similar targets. The need for a radically novel approach to antibiotic design is increasingly apparent.

1.3 The promise of antimicrobial peptides

In the search for new classes of antibiotic agents, researchers have recently turned to nature for clues. Since potentially infectious microbes are ubiquitous in most environments where living organisms thrive, virtually all organisms must possess some form of natural microbial defenses. In vertebrates, this consists of a complex adaptive immune system, which involves a sophisticated interplay among various immune system components, such as B-cells, T-cells and antibodies. Simpler organisms, including bacteria, possess much more primitive defense mechanisms, which rely on the secretion of various proteins that can directly attack hostile viruses and other microbes. These small peptides can have highly specific target pathogens, or they can exhibit a broad spectrum of activity against numerous viruses, fungi and bacteria. Such molecules are loosely classified as antimicrobial peptides (AMPs). As it turns out, AMPs are also present in the innate immune system of higher organisms, including humans. Although they are likely to be evolutionary vestiges of a more primitive immune system, AMPs still form an effective first line of defense to many infections and play an integral role alongside the adaptive immune system.

Their mechanism of action, although in many ways still a mystery, appears to be based primarily on non-specific interactions with the bacterial cell membrane, resulting in rapid membrane depolarization and bacterial death. The cationic nature of most AMPs accounts for their preference for bacterial membranes, which contain a high fraction of anionic lipids such as phosphatidylglycerols, as opposed to zwitterionic lipids, such as phosphatidylcholines and cholesterol found in eukaryotic cell membranes (see section 1.4). In part because of this non-specific targeting, very little bacterial resistance has been observed to AMPs, as evidenced by their widespread presence and continued effectiveness in nature. Furthermore, AMPs have a broad spectrum of activity, with many peptides showing comparably strong activity against Gram-positive and Gram-negative bacteria, as well as fungi and enveloped viruses [8]. Unfortunately, this broad spectrum of activity also means that most AMPs are toxic to human cells at the concentration levels required to treat systemic infections [9]. Nonetheless, peptide analogues have been engineered with reduced cytotoxicity and enhanced antimicrobial activity [10]. AMPs therefore appear to hold the potential for a lasting solution to the current crisis in antibiotic resistance.

Despite the ever-growing number of known AMPs – almost 1,000 known examples since their discovery in the early 1980s[11] – as well as the high level of research activity in the area, very few AMP-based drugs have been commercialized to date. A review by Hancock and Sahl [12] identified fifteen AMP-based therapeutics in various stages of development. To date, only four AMPs have advanced as far as phase 3 clinical efficacy trials [12], two of which have already failed due to insufficient antimicrobial effectiveness. One of these, Iseganan, originally developed by Intrabiotics Pharmaceuticals Inc., was an analogue of protegrin-type peptides, the type of peptide studies extensively herein. Its intended purpose was the treatment of oral mucositis in patients receiving chemotherapy, but it failed to demonstrate adequate efficacy in phase 3 trials [13-14]. A more successful story is that of Omiganan, an indolicidin analogue developed for catheter-associated infections, which achieved success in reducing

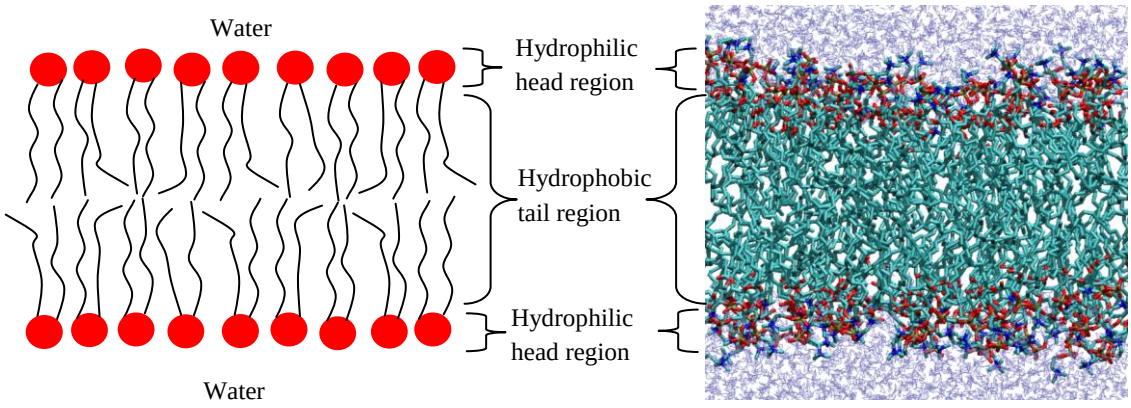
infections in phase 3a trials and is currently undergoing phase 3b trials [12]. For a more thorough review of the history and current status of AMP clinical development, the reader is referred to [12, 15-16]. The picture that emerges is that of a nascent technology with strong potential to provide a lasting solution to the problem of antibiotic resistance. However, before this potential can be fully unlocked for clinical use, a better understanding of the mechanism of action of AMPs is required to guide rational drug design efforts aimed at ameliorating the toxicity of AMPs and enhancing their activity against various pathogens.

We next provide a brief discussion of some of the biophysical characteristics of cellular membranes that are important for the action of AMPs. We then discuss protegrins, the model peptides that we have selected for all of the work presented herein.

1.4 Cell membranes – the targets of AMPs

The cellular membrane, sometimes called the plasma membrane, is a thin membrane that envelops all living cells. It consists primarily of amphipathic lipid molecules, which possess a hydrophobic tail and a hydrophilic head. Lipids form a membrane by arranging themselves in a lamellar bilayer structure, with the hydrophobic tails confined to the interior of the membrane, and the hydrophilic heads on the exterior.

Figure 1-1: The lipid bilayer. A schematic representation (left) and a snapshot from a molecular dynamics simulation of a POPC bilayer (right).



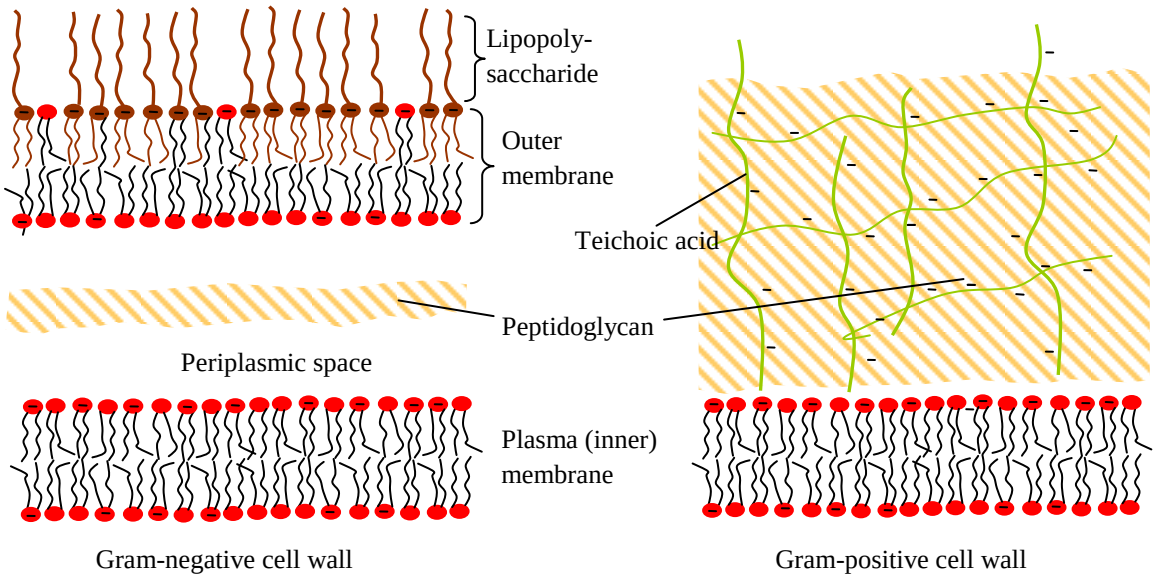
The plasma membrane forms a highly effective barrier between the cell exterior and the cytoplasm. Pure lipid bilayer models represent simplifications of real cellular membranes, which also contain a plethora of membrane-anchored proteins, sugars and other chemicals. However, since the action of most AMPs appears to rely on non-specific interactions with lipid bilayers, the subsequent discussion will focus primarily on the lipid composition of plasma membranes.

1.4.1 Bacterial cell membranes

Different organisms, as well as different cells within the same organism, often exhibit significant variations in membrane architecture and composition. However, common features of bacterial cells appear to be sufficiently conserved across species, such that AMPs possess a broad spectrum of activity against different bacterial species, while retaining significant selectivity over host (mammalian) cells. The outer periphery of bacterial cells consists of the lipid bilayer plasma membrane, surrounded by a cell wall, which is absent in mammalian membranes. Common bacterial species can be classified into two large categories, denoted as Gram-positive and Gram-negative. The structure of the cell wall of both bacterial classes is shown in Figure 1-2 below. The

envelope of Gram-positive bacteria consists of a phospholipid bilayer, surrounded by a thick layer of peptidoglycan (20-80 nm). The peptidoglycan layer is a relatively permeable polymer matrix, consisting of sugars and amino acids, and teichoic acid in the case of Gram-positive bacteria. In Gram-negative bacteria, the plasma membrane is surrounded by an aqueous compartment known as the periplasmic space, which includes a peptidoglycan layer only a few nanometers thick. In addition, Gram-negative bacteria possess an outer phospholipid membrane, which has a high concentration of lipopolysaccharide (LPS). It is worth mentioning that in addition to the negative charges in the plasma membrane, both the teichoic acid in the case of Gram-positive bacteria and LPS in the case of Gram-negative bacteria are negatively charged at physiological pH, which may contribute to the selectivity of AMPs.

Figure 1-2: A comparison of the Gram-negative and Gram-positive cell walls

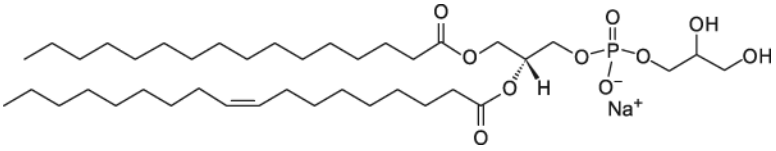


Despite this highly complex structure, it is believed that the primary target of AMPs is the plasma (inner) membrane of bacteria [15, 17-19]. The role of the bacterial cell wall in the action of AMPs is not clear, but its relatively high permeability suggests

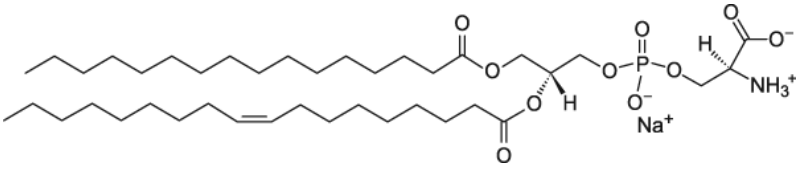
that peptides can penetrate it with ease and interact with the plasma membrane. Furthermore, it has been experimentally determined that for many classes of AMPs, their ability to cause leakage from single-layered lipid bilayer vesicles correlates well with their activity against real bacteria [20-24], corroborating the hypothesis that the plasma membrane is the primary target of AMPs. As such, the nature of the cell wall and its interactions with AMPs will not be further discussed; the interested reader is referred to [18, 25-26]. All subsequent discussion of the bacterial membrane will refer to the inner membrane in the case of Gram-negative bacteria.

As already mentioned, the defining feature of bacterial cell membranes is the much higher abundance of anionic (acidic) lipids. The structures of several such anionic as lipids are shown in Figure 1-3 below.

Figure 1-3: The structures of two anionic lipids commonly found in bacterial cell membranes



(a) 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-[Phospho-rac-(1-glycerol)] (POPG)



(b) 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-[Phospho-L-Serine] (POPS)

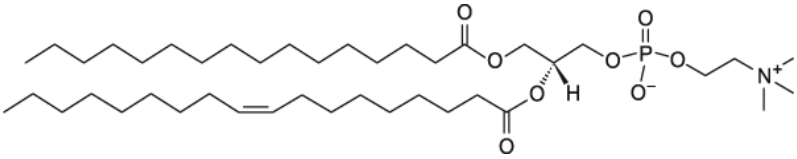
The examples shown above belong to the larger class of glycerophospholipids, one of the most common classes of lipids found in living organisms; they consist of two hydrophobic tails, connected to two sites on the glycerol backbone, and a hydrophilic

phosphate-based head group connected to the third site of the glycerol backbone. The nature of the head group in phospholipids is determined by the group attached to the phosphate moiety. Despite variations in membrane composition across different species, the membranes of common bacteria such as *S. aureus* and *E. coli* consist primarily of phosphatidylglycerols (e.g. POPG, shown in Figure 1-3, cardiolipin (effectively a dimer of two phosphatidylglycerols), phosphatidylserine (PS), as well as zwitterionic phosphatidylcholine (PC) and phosphoethanolamine (PE) head groups (Figure 1-4 below).

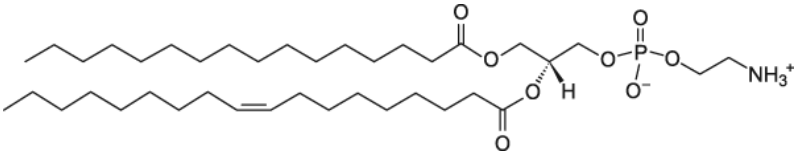
1.4.2 Mammalian cell membranes

In contrast, mammalian cell membranes consist almost entirely of lipids with zwitterionic head groups, such as the PC and PE groups already mentioned (depicted in Figure 1-4). Additionally, mammalian membranes contain a high percentage of sterols such as cholesterol, which also possesses no net charge.

Figure 1-4: Examples of zwitterionic lipids, commonly found in mammalian cell membranes



(a) 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine (POPC)



(b) 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphoethanolamine (POPE)

The heterogeneous nature of both mammalian and bacterial cells creates substantial difficulties for the development of effective membrane mimics, both in experiments and simulations. In real cells, membrane proteins (e.g. flippases) and ion gradients are often

employed to control local composition, as well as to induce compositional variations between the inner and outer leaflets of the membrane. In mammalian cells, this effect is less pronounced, and results in transmembrane potentials on the order of -90 to -110 mV; in contrast, bacterial membranes possess electrochemical gradients ranging from -130 to -150 mV [17]. This additional difference between mammalian and bacterial cells has been postulated to aid in the selectivity of AMPs towards the latter [17]. Furthermore, the transmembrane potential is an essential feature to a variety of membrane functions, and is maintained in both bacterial and mammalian cells by the active transport of ions against their chemical potential gradients. As we show in the present work, an important feature in the mechanism of action of pore-forming AMPs is their ability to permeate the membrane and facilitate unrestricted ion transport across the cell membrane. This results in the loss of the transmembrane potential, which is essential to membrane function, and may lead to cell death.

1.5 Protegrins

Fundamental studies of antimicrobial peptides to date have focused on several model peptides, all of which show promising therapeutic potential. The most studied peptides include alamethicin, magainin, melittin and ovispirin from the α-helical structural class, defensins and protegrins from the β-sheet class, and indolicidin from the unstructured class. For broad reviews of AMPs across all three major structural classes, the interested reader is referred to [8-9, 17, 19, 27-28]. In our work, we have focused on protegrin [29], an excellent model peptide due to its high efficacy against a broad spectrum of bacterial pathogens, its small, rigid structure, which makes it particularly convenient for computer simulation work, and the wealth of biological and biophysical data available in the literature for this class of peptides. The common theme among experimental and computational biophysical studies of protegrins (and indeed of all of the peptides mentioned above) is the interaction of these peptides with lipid bilayers and other membrane-mimicking environments. Although some studies suggest that AMPs

may in fact act at an intracellular level [30-32], the bulk of the evidence points to the bacterial membrane as the primary target [9, 11, 17, 19, 33-34]. This is especially true of protegrin, for which it has been shown that membrane disruption correlates with antimicrobial activity [35-36]; that substitution of D-amino acids causes no loss of activity, ruling out specific intracellular targets [37]; and that exposure to protegrins causes uncontrolled potassium leakage consistent with membrane permeabilization prior to cell death [38]. As such, protegrins have emerged as one of the leading model peptides for both experiments and simulations on account of their antibiotic efficacy, the wealth of biophysical data available, and their amenability to computer simulations due to their small size and rigid structure.

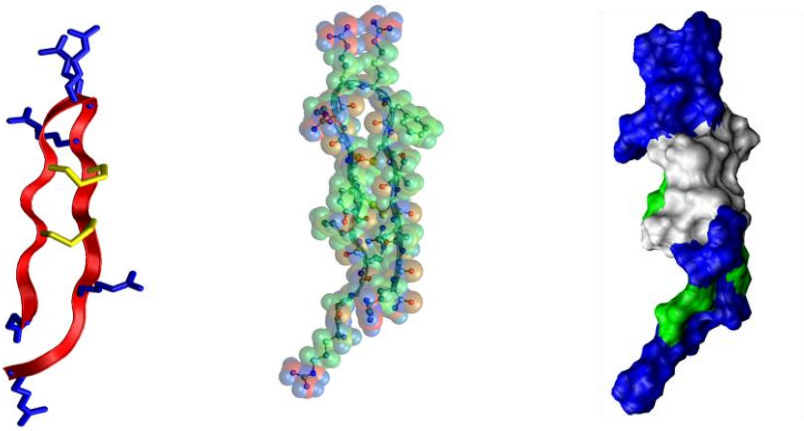
1.5.1 Biology and structure of protegrins

Protegrins are a family of peptides originally isolated from porcine leukocytes (white blood cells of pigs) [29]. Currently, there are five known naturally occurring members of this family, and dozens of mutants have been synthesized [39]. Protegrins have shown a broad spectrum of activity against both Gram-positive and Gram-negative bacteria (see section 1.4.1), including *E. coli*, *P. aeruginosa*, and *N. gonnorhoeae* [29, 40], as well as the fungus *C. albicans* and the HIV-1 virus. The minimum inhibitory concentration (MIC) is a measure of antibacterial potency, corresponding to the minimum peptide concentration at which bacterial growth ceases. Protegrins exhibit a particularly promising activity profile, with MICs on the order of 1 μg/mL against several bacterial strains [41]. Typical MIC values for AMPs are in the range of 2-50 μg/mL, making protegrin one of the most effective known AMPs [42]. However, the potent naturally occurring peptides also exhibit high levels of toxicity against human epithelial cells (lining of tissues) and red blood cells. For instance, PG-1, the most common naturally occurring protegrin variant, kills 50% of epithelial cells at concentrations in the range of 25-50 μg/mL, while a concentration of 80 μg/mL of PG-1 results in the death of over 50% of red blood cells. Despite MICs in the range of 1 μg/mL, these toxicity levels are

unacceptably high for therapeutic use. A larger data set presented in Table 1 of reference [41] includes more detailed activity/toxicity measurements for numerous protegrin mutants.

The naturally-occurring protegrins are between 16 and 18 amino acids long, and are rich in positively charged arginine residues. Their defining feature is an anti-parallel β-hairpin structure, stabilized by two cysteine-cysteine disulfide bonds between the two strands of the hairpin (see Figure 1-5 (a)). The solution structure of the most common naturally occurring protegrin variant, denoted PG-1 (amino acid sequence RGGRLCYCRRRFCVCVGR-NH2), is shown in Figure 1-5. This structure is based on the NMR study conducted by Fahrner et al [39], and is available in the protein data bank [43] (PDB code 1PG1). The disulfide bonds spanning the β-hairpin impose rigid constraints on the peptide structure. This has the effect of confining the positively charged, highly hydrophilic arginine residues in distinct spatial regions – the two termini, and the β-hairpin turn – while the remaining structure consists largely of hydrophobic residues (Figure 1-5(c)). Its rigid structure thus gives PG-1 its amphipathic nature, common to many other AMPs.

Figure 1-5: Three graphical representations of protegrin-1 (PG-1)



(a) Cartoon representation showing the backbone (red ribbon), the constraining cysteine-cysteine bonds (yellow tubes), arginine residues (blue) and amino acid names and positions.

(b) Colouring by charge:

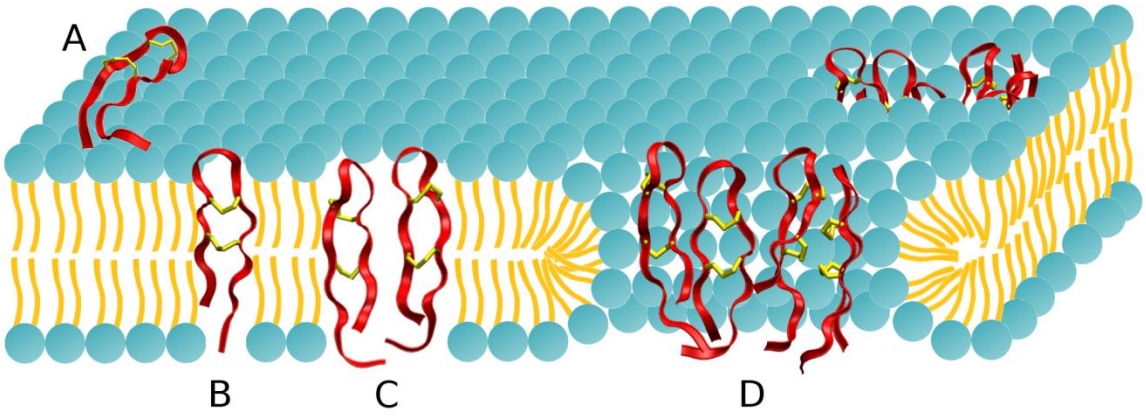
(c) Accessible surface area, with colouring based on residue types. Polar but weakly hydrophilic residues in green, hydrophobic residues in white, hydrophilic residues in blue

1.5.2 Available experimental data

In addition to the extensive activity and toxicity data described, numerous biophysical experimental investigations of protegrins have been reported. The emerging model of protegrin action against bacteria involves the electrostatic attraction and adsorption of cationic protegrin peptides to anionic bacterial membranes [44-47], followed by dimerization of the peptides [48-50], insertion into the hydrophobic membrane core [51-53], formation of higher aggregates (4-5 dimers) [54-57], and eventually formation of transmembrane pores that cause an unrestricted, lethal flux of ions and other molecules from the cytoplasm [38, 58-60] (see Figure 1-6). The precise order of these steps, the molecular interactions responsible for them, the thermodynamic driving forces and the kinetics involved in all of these steps are largely unknown. We briefly review currently available experimental data pertaining to these phenomena in greater detail below.

Figure 1-6: The proposed mechanism of action of protegrin

A. Peptides adsorb to the membrane surface, likely driven by electrostatic attraction to negatively charged bacterial membranes. **B.** Peptides insert into the hydrophobic core in a transmembrane orientation. **C**. Peptide monomers form dimers (this likely also occurs on the membrane surface and in the bulk solution). **D.** Peptide dimers oligomerize, likely as four or five pairs, to form water-filled, transmembrane pores. These pores result in uncontrolled leakage of ions and small molecules, leading to cell death.



Numerous studies have investigated the interactions of protegrins with various membrane mimics. Lee and coworkers have conducted a series of experiments using lipid monolayers of assorted compositions as bacterial and mammalian membrane mimics. Their findings confirm that PG-1 readily inserts into anionic films and films composed of lipid A, a primary component in the bacterial outer membrane, but does not readily interact with zwitterionic lipids [44]. They have also been able to distinguish a higher affinity of PG-1 for phosphoethanolamine (PE) monolayers over phosphocholine (PC) monolayers, both of which are zwitterionic, but the former is more abundant in ruminants than in humans [47]. These results emphasize the importance of membrane lipid composition in the action of protegrin (a brief discussion of membranes and the relevant constituent lipid molecules was given in section 1.4). Work from the same group using atomic force microscopy on bilayers has provided direct evidence of increased bilayer

disruption by PG-1 with increased peptide concentration [61]. The selectivity of PG-1 towards anionic phosphoglycerol lipids over zwitterionic phosphocholine lipids was confirmed through NMR investigations by Mani et al [45]. From their NMR studies of protegrin interacting with lipid bilayers, Hong and coworkers have established that PG-1 causes disruption of the order of longer tailed, unsaturated lipids (1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine (POPC) with tails 16 and 18 carbon atoms long), but does not disrupt thinner bilayers (1,2-Dilauroyl-sn-Glycero-3-Phosphocholine (DLPC), with tails containing 12 carbon atoms) [51]. This suggests that geometrical considerations are also vital for the action of PG-1, presumably in connection with hydrophobic matching between the peptide and the hydrophobic bilayer core. The preservation of order in the thinner bilayers allowed Yamaguchi et al [51] to carry out further NMR experiments to determine that PG-1 was fully inserted in a DLPC bilayer, where it adopts an orientation in which its long axis is tilted approximately 55º from the bilayer normal. This was confirmed by molecular dynamics simulations of PG-1 in a DLPC bilayer in our group [62].

Several groups have investigated the ability of protegrin to aggregate into dimers and higher structures. The first such investigation was a series of NMR experiments carried out by Roumestand et al [63] with PG-1 in the presence of dodecylphosphocholine micelles (DPC, a zwitterionic surfactant). The nuclear Overhauser effect (NOE) distance constraint data obtained from their experiments was most readily explained by a peptide dimer structure, in which the two protegrin monomers aligned in an anti-parallel fashion, with the C-terminal strands interacting through backbone hydrogen bonds [63] . In contrast, a solid-state NMR study conducted by Tang et al found that PG-1 forms aggregates in a phosphate buffer saline solution, with ^{13}C and ^{1}H spin diffusion data indicating a parallel orientation of the monomers [64] (see Figure 1-6 C). Further NMR investigations by Hong and coworkers have shown that PG-1 likewise forms parallel dimers in POPC bilayers, with the C-terminal strands forming the dimer interface (Figure 1-6 C) [48, 50]. Since this appears to be the

dominant dimerization motif in lipid bilayers, we do not concern ourselves with anti-parallel dimers any further in this work.

A solid-state NMR study by the Mani et al [54] provides strong evidence for the formation of higher aggregates, in which four protegrin dimers aggregate to form a trans-membrane pore structure in a bacterial membrane mimic (3:1 molar ratio of zwitterionic POPE to anionic POPG; see sections 1.4.1 and 1.4.2). Interestingly, the same study found that PG-1 dimers did not form pore-like aggregates, but rather remained surface-bound in a mammalian membrane mimic (1.2:1 molar ratio of zwitterionic POPC to neutral cholesterol; see Figure 1-4). This significant difference was hypothesized to be primarily a result of the high cholesterol content, which exerts a negative curvature strain on the bilayer and prevents peptide insertion [54]. This finding has particularly encouraging therapeutic implications, as it suggests a different mechanism of action for antimicrobial activity as opposed to toxicity, potentially permitting the amelioration of the latter without sacrificing the former.

More recently, several investigations have probed the nature of protegrin pores. Early evidence of pore formation by protegrins came from electrophysiological studies, which showed that PG-1 forms anion channels in *Xenopus laevis* oocytes [59], and that PG-1 and PG-3 (a naturally-occurring analogue of PG-1) cause an increase in the conductivity of lipid bilayers consistent with the formation of ion channels [58]. Based on neutron diffraction data, Yang and coworkers were able crystallize membrane pore structures for protegrins and magainins [60]. More recently, Mani et al reported solid state NMR (ssNMR) data suggesting that PG-1 forms transmembrane, NCCN multimers in 3:1 POPE:POPG lipid bilayers, where the repeat unit is an NCCN parallel dimer [54]. Combining this with the fact that membranes treated with PG-1 allowed the permeation of polyethyleneglycol beads with a hydrodynamic radius up to 10.5Å, the authors concluded that PG-1 forms transmembrane pores composed of four or five dimer repeat units, creating a water-filled pore with an inner diameter around 21Å [54]. In contrast, they found that PG-1 dimers adsorb to the surface of mammalian-mimicking 1.2:1 POPC:cholesterol lipid bilayers, where they likely form tetrameric β-sheet clusters.

Subsequent work from the same group using ssNMR measured distances between several ^{13}C atoms of PG-1 and ^{31}P atoms in anionic POPE/POPG lipid bilayers, from which they concluded that lipids surrounding the PG-1 pore most likely adopt a toroidal conformation (see Figure 5-1)[56]. Further ssNMR experiments revealed that Arg-11 is more mobile than Arg-4 and Leu-5 in the transmembrane pore state [55], that dimethylation of the PG-1 arginine groups results in a peptide that does not aggregate or insert in a transmembrane orientation [57] and that the arginine side chains of PG-1 bind both water and lipid head groups [65]. Additionally, Wi and Kim have reported ^{31}P and ^{2}H ssNMR experiments that yield anisotropic line shapes consistent with toroidal pore formation and bilayer thinning in the binding of oriented lipids with PG-1 [66].

The experimental investigations just described provide critical information as well as motivation for a number of simulation studies of protegrin. Given the extremely fast dynamics, small scales and complex solution-state nature of all the processes described, computer simulations can provide an excellent complement to experimental investigations and greatly enhance the level of detail attainable experimentally. In particular, several significant phenomena emerge, all of which are well suited for computer simulation studies: the interplay between the monomer and dimer forms in different states of membrane association (e.g. transmembrane inserted or surface-adsorbed); the nature of protegrin pores, in particular their ability to conduct ions; and the effects of multiple protegrin pores on entire cells. All of these phenomena can be captured and analyzed in detail using a variety of computer modeling and simulation techniques, thereby providing useful insight into the mechanism of action of this potent AMP, and ultimately aiding in the design of optimal therapeutic peptides.

2 Connecting molecular thermodynamic information to membrane association equilibria

2.1 Introduction

The work presented in this chapter addresses the physical phenomena that precede the formation of pores, namely the adsorption of peptides to the membrane surface, the formation of dimers, and the insertion of peptides into the membrane core. Previous work in our group [67-69] has provided molecular-level thermodynamic data (specifically, potentials of mean force) that describe these steps separately for individual peptides. The goal of the work presented herein is to connect and unify these data into a model that can describe the complex equilibria between the important states of protegrin that are precursors to pore formation (solvated, membrane surface-bound and membrane-inserted monomers and dimers), while accounting for non-ideal effects that arise due to peptide-peptide interactions. This type of modeling can in principle provide a quantitative link between molecular thermodynamic properties and experimentally measurable quantities, such as adsorption isotherms or monolayer surface pressure measurements. It will also allow us to understand experimentally measurable peptide-membrane association affinities in greater detail, and identify the relative importance of different physical phenomena.

The standard free energy of binding ΔG and the associated equilibrium constant $K = \exp(-\Delta G^0/RT)$ are typically the most convenient and most complete thermodynamic measures of biomolecular association events, such as the binding of a peptide to a lipid bilayer or the association of peptide monomers to form dimers. Experimentally, free energies of binding can be determined by a variety of methods, including isothermal calorimetry, surface plasmon resonance, or any method that can accurately measure the ratios of the relevant bound and unbound states. Simulations can reach beyond simply measuring the free energy differences between different states, and

compute the free energy of a system as a continuous function of one or more reaction coordinates that separate the two states. This quantity is known as a potential of mean force, and represents complete information about the thermodynamic equilibrium of a process, as well as clues about its kinetic aspects. For a system that can be described by a general set of coordinates $\{\boldsymbol{r}_1,..,\boldsymbol{r}_N,z\}$, the PMF along the reaction coordinate z is defined as:

$$W(z) = -k_B T\, ln \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\big(-U(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N,z)\big)\, d\boldsymbol{r}_1 \dots d\boldsymbol{r}_N \qquad (2.1)$$

Note that the integral on the right-hand side is carried out over all coordinates except z, the reaction coordinate of interest. For instance, in the case of a peptide binding to a membrane, the relevant reaction coordinate is the separation distance between the peptide's center of mass and the membrane surface. The potential of mean force (PMF) then represents the free energy of the system as a function of this separation distance. For a microscopic, solvated lipid bilayer patch with a single peptide, the number of relevant coordinates is on the order of 10^5, and the system potential energy U is a complex function of these coordinates (for instance, the CHARMM potential function, equation A.12), which makes the integral in equation (2.1) intractable. Instead, computational approaches are used to estimate the derivative of the potential of mean force, which is equivalent to the negative of the mean force F_z acting along the reaction coordinate z:

$$\frac{\partial \mathrm{W(z)}}{\partial \mathrm{z}} = -\langle \mathrm{F_z} \rangle \qquad (2.2)$$

In order to measure this force, constraints can be applied to the reaction coordinate at different values, known as windows, and the remaining degrees of freedom are sampled with MD simulation. This enhances sampling of high energy regions of the PMF that are otherwise poorly sampled, and monitoring the constraint force yields the desired force along the reaction coordinate, F_z. The PMF, $W(z)$, is then calculated by integrating the average constraint forces from a large separation to the location of interest:

$$\langle F_{cons} \rangle_z = -\langle F_{peptide} \rangle = \frac{dW}{dz} \; ; \quad W(z^*) = \int_{\infty}^{z^*} \langle F_{cons} \rangle_z dz \tag{2.3}$$

Previous work in our group has used this constraint force integration method to calculate the PMFs that describe PG-1 monomer and dimer adsorption to 3:1 POPE:POPG membranes [67], as well as dimerization in bulk water, on the membrane surface, and inside the membrane core [69]. Preliminary results from Dr. Vivcharuk's work using the adaptive biasing force (ABF) method [70] have also yielded the PMF corresponding to the insertion of a PG-1 monomer into a 3:1 POPE:POPG membrane [68].

Although potentials of mean force and ΔG values for these different processes yield useful information about the interactions of peptides with lipid bilayers, it is not clear how these processes behave when they occur in unison, as they do in biological situations, or how they are affected by non-idealities arising from peptide crowding in the membrane, as we will show is the case for biologically relevant concentrations. The work presented in this chapter provides a modeling framework that connects molecular-level thermodynamic information in the form of PMFs to mesoscopic quantities, such as adsorption isotherms. We augment the ideal equilibrium relations that follow from the PMFs with additional statistical mechanical considerations of entropic contributions, as well as simple models that account for area exclusion and crowding effects.

2.2 Methods

2.2.1 Model description

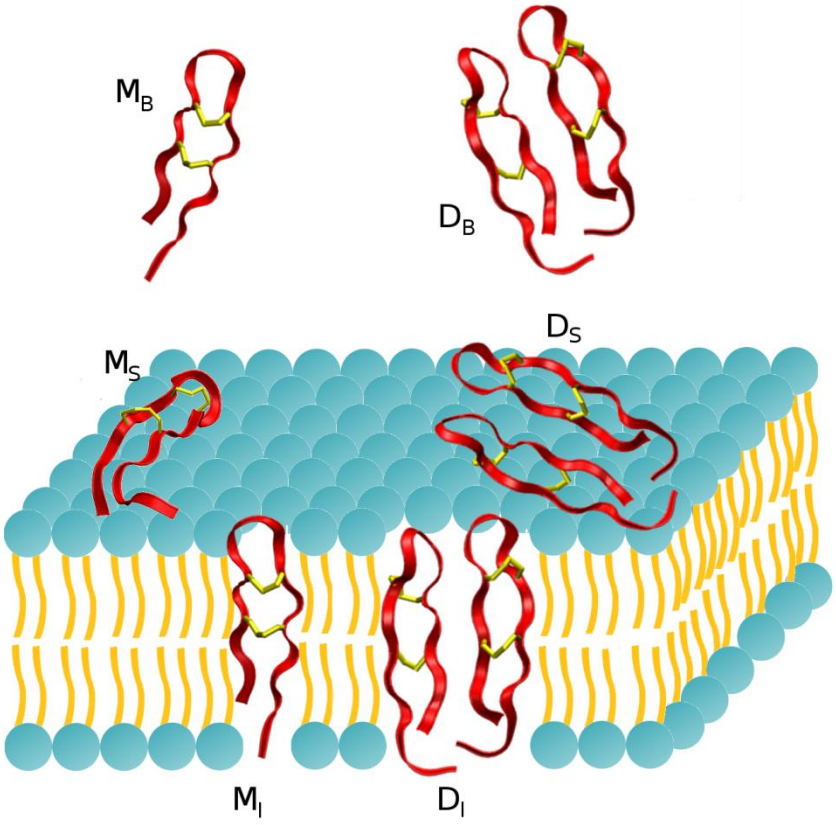
The overall model aims to capture the equilibrium behavior of peptide adsorption onto the membrane surface, the dimerization of peptides - in a parallel, NCCN dimer configuration, the predominant form observed in lipid bilayers [48] - and the insertion of peptides into the membrane core. We discuss the application to both lipid monolayers and bilayers, but restrict our calculations to lipid bilayers, as they are more relevant to biological applications. The discussion that follows is applicable to both cases, with the

only difference being in the interpretation of the surface density values, as we will show later.

Peptide monomers (M) and dimers (D) are assumed to exist in one of three states, or phases: in bulk solution (B), adsorbed to the membrane surface (S), or inserted in the membrane core (I) (see Figure 2-1).

Figure 2-1: The six states of protegrin accounted for in the present model.

M_B: monomer in the bulk phase; D_B: dimer in the bulk phase; M_S: monomer adsorbed to the membrane surface; D_S: dimer adsorbed to the membrane surface; M_I: monomer inserted in the membrane core; D_I: dimer inserted in the membrane core.



A particular species and its chemical potential are denoted accordingly; for example, M_s represent a monomer (M) in the surface-adsorbed state (S), while $\mu_{M,S}$ represents its chemical potential. We write the relevant processes as reversible reactions,

and the equality of chemical potentials provides the required equations. These are summarized in Table 2-1, along with the final model equations, which are derived in the next few sections.

Table 2-1: Relevant reactions and corresponding equilibrium relations.

Description	Reaction	Statistical mechanical equilibrium relation
Dimerization in the bulk solution	$M_B + M_B \rightleftharpoons D_B$ (1)	$c_{D,B} = \frac{1}{16\pi^2} c_{M,B}^2 \int\limits_{\substack{dimer \\ state}} e^{-\beta W_{dim,B}} r^2 dr$
Monomer adsorption	$M_B \rightleftharpoons M_S$ (2)	$\rho_{M,S} = \frac{c_{M,B}}{4\pi} e^{-\beta \mu_{M,S}^R} \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{M,ads}(z)} dz$
Dimer adsorption	$D_B \rightleftharpoons D_S$ (3)	$\rho_{D,S} = \frac{c_{D,B}}{4\pi} e^{-\beta \mu_{D,S}^R} \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{D,ads}(z)} dz$
Monomer insertion	$M_S \rightleftharpoons M_I$ (4)	$\frac{\rho_{M,S}}{\rho_{M,I}} = \frac{\int_{ads} e^{-\beta W_{M,ins}(z)} dz}{\int_{ins} e^{-\beta W_{M,ins}(z)} dz} e^{-\beta(\mu_{M,S}^R - \mu_{M,I}^R)}$
Dimerization in the membrane	$M_I + M_I \rightleftharpoons D_I$ (5)	$\frac{\rho_{D,I}}{\rho_{M,I}^2} = \frac{1}{4\pi} \int\limits_{\substack{dimer \\ state}} e^{-\beta W_{dim,ins}(r)} r dr \, e^{-\beta(\mu_{D,I}^R - 2\mu_{M,I}^R)}$

Note that other reactions can be written, but they are simply linear combinations of the existing reactions (e.g. dimerization on the surface is a combination of reactions 1, 2 and 3; the resulting equilibrium relation, $2\mu_{M,S} = \mu_{D,S}$, can be obtained by substituting relation 2 for the left hand side of relation 1, and relation 3 for the right hand side of relation 1).

2.2.2 Ideal chemical potential

A detailed derivation is presented for the case of monomer adsorption onto the membrane. The remaining cases are treated more briefly, with reference to the derivation that follows. We first split up the chemical potential of the adsorbed species into the chemical potential for ideal adsorption, μ_M^*, and the residual chemical potential, μ_M^R ,

which accounts for non-ideality in adsorption in the form of interactions with other adsorbed species:

$$\mu_{M,S} = \mu^*_{M,S} + \mu^R_{M,S} \tag{2.4}$$

In calculating $\mu^*_{M,S}$, we assume that an adsorbed peptide is mobile on the surface of the membrane, with its lateral motion in the plane of the membrane completely unhindered. We denote the total membrane area available to an adsorbed peptide as A. The chemical potential of N peptides adsorbed on a membrane can be written as a function of the single-peptide partition function, $q_{M,S}$, as follows:

$$\mu^*_{M,S} = kTlnN - kT \ln q_{M,S} - \frac{NkT}{q}\frac{dq_{M,S}}{dN} \tag{2.5}$$

For peptides that do not interact with one another (corresponding to the present ideal case, since we are now dealing only with $\mu^*_{M,S}$), the last term disappears. The single-particle partition function can further be split into a purely configurational partition function $z_{M,S}$ and a partition function that describes all momentum contributions, q_P; the configurational partition function $z_{M,S}$ is further split into rotational, translational and internal partition functions [71]:

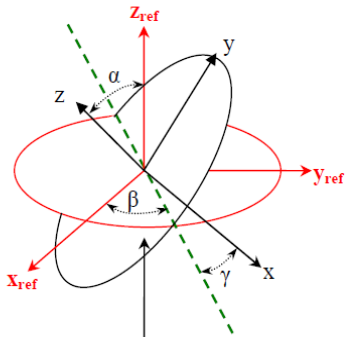
$$q_{M,S} = q_P z_{M,S} = q_P z_{rot} z_{trans} z_{int} \tag{2.6}$$

The rotational degrees of freedom are described by three Euler angles, α, β, γ, discussed below. The translational degrees of freedom are taken to be the Cartesian x, y, z coordinates of the peptide with respect to some point on the membrane. The internal degrees of freedom are lumped into a set of internal coordinates $\boldsymbol{r}_{int}$. The configurational partition function for the adsorbed state can then be written as:

$$z_{M,S} = \int_{adsorbed\ state} \dots \int e^{-U(x,y,z,\boldsymbol{r}_{int},\alpha,\beta,\gamma)/k_BT} \sin\alpha \, d\alpha d\gamma d\beta \, dxdydz \, d\boldsymbol{r}_{int} \quad (2.7)$$

Here, the integration is carried over the phase space regions that define the adsorbed state, in this case a particular value of z, the peptide-membrane separation. An adsorbed peptide is considered free to translate in the plane of the membrane (the x and y directions). It is also free to rotate around an axis perpendicular to the membrane (this is an orientational degree of freedom, the Euler angle β, described in the figure below).

Figure 2-2: Describing the orientation of the peptide



Membrane surface normal

The Euler angles α, β, γ define the orientation of the peptide (**x,y,z**) with respect to some reference frame (**xref,yref,zref**). The line of nodes is shown as a green dashed line; it is the intersection of the x-y plane with the **xref-yref** plane. The angle α is the angle between **z** and **zref**, β is the angle between the line of nodes and **xref**, and γ is the angle between **x** and the line of nodes. **Note that if the reference z-axis (zref) is chosen to align with the membrane surface normal, the coordinate β fully describes rotation about this surface normal (and α and γ have no components that do).**

Since x, y and β are unhindered degrees of freedom - in the sense that configurations with the peptide translated along the membrane surface or rotated around the membrane normal are indistinguishable due to the membrane symmetry - the potential of interaction between the peptide and the membrane will have no functional dependence on them, and we can perform the integration over these degrees of freedom immediately to yield a factor of A (from integrating over the x and y directions) and a factor of 2π (from integrating over β).

$$z_{M,S} = 2\pi A \int_{adsorbed\ state} \ldots \int e^{-\beta U(z,\boldsymbol{r}_{int},\alpha,\gamma)} \sin\alpha \, dz d\boldsymbol{r}_{int} d\gamma d\alpha \qquad (2.8)$$

This implicitly defines *A* as the total area of the membrane available for peptide translation; thus, it is equal to the area of a monolayer, but is twice the area of a lipid bilayer (since both the top and bottom surfaces are available for peptide adsorption and completely equivalent in an equilibrium context). Also note that from this point forward, $\beta = 1/k_B T$, rather than referring to the Euler angle β. The integral over the remaining degrees of freedom can be shown to be identical to the integral over the potential of mean force (PMF) with respect to *z*, the degree of freedom that is constrained in the PMF. This is equivalent to setting $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ in equation (2.1), which are completely general coordinates, to (α, γ, $\boldsymbol{r}_{int}$). The implicit assumption here is that the PMF calculations adequately sample both the orientations of the peptide with respect to the membrane (α and γ) and the internal coordinates of the peptide ($\boldsymbol{r}_{int}$). Thus, we have:

$$z_{M,S} = 2\pi A \int_{\substack{z\ \epsilon\ adsorbed \\ state}} e^{-\beta W_{M,ads}(z)} \, dz \qquad (2.9)$$

Where $W_{M,ads}$ is the potential of mean force for the adsorption of the peptide onto the membrane. Note that in sampling this potential of mean force, the separation between the peptide center of mass and the membrane surface was constrained (*z*-direction), while the remaining internal and orientational degrees of freedom ($\boldsymbol{r}_{int}, \alpha, \gamma$) were sampled by molecular dynamics simulations. Finally, substituting everything in equation (2.5), we have, for the ideal part of the chemical potential of a surface-bound monomer:

$$\mu^*_{M,S} = kT \ln N_{M,S} - kT \ln \left(q_P 2\pi A \int_{\substack{z\ \epsilon\ adsorbed \\ state}} e^{-\beta W_{M,ads}(z)} \, dz \right) \qquad (2.10)$$

$$\mu^*_{M,S} = kT \ln \rho_{M,S} - kT \ln \left(q_P 2\pi \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{M,ads}(z)}\, dz \right) \qquad (2.11)$$

Where $\rho_{M,S} = N_{M,S}/A$ is the surface density of surface-bound monomers. In this case, since the area A represents total area available for lateral translation, it will include the top and bottom surfaces of the bilayer, so it is equal to twice the lateral area of the bilayer. In the case of inserted species, this area A is equal to the area of the bilayer. In a lipid monolayer, the area A is equal to the monolayer area for both adsorbed and inserted species. **Thus, the only difference between the treatment of monolayers and bilayers is in the interpretation of the resulting surface density values.** This is discussed again at the end of section 2.2.3 in the context of the residual chemical potential, and dealt with explicitly in section 2.2.6.

For the case of the monomer in the bulk solution, which is free to translate in an entire volume V and has full orientational freedom, the corresponding expression is:

$$\mu^*_{M,B} = kT \ln N_{M,B} - kT \ln \left(q_P 8\pi^2 \int\limits_{\substack{bulk \\ state}} e^{-\beta W_{M,ads}(z)} dx dy\, dz \right) \qquad (2.12)$$

Since the PMF $W_{M,ads}$ is defined so that it vanishes as $z \rightarrow \infty$ (i.e. in the bulk solution), the integral term above evaluates to V, the volume of the bulk solution. This allows us to write:

$$\mu^*_{M,B} = kT \ln c_{M,B} - kT \ln(q_P 8\pi^2) \qquad (2.13)$$

where $c_{M,B} = N_{M,B}/V$ is the concentration of the monomer in the bulk solution. Finally, setting $\mu_{M,B} = \mu_{M,S}$ (the condition for equilibrium), and assuming that the residual chemical potential in the bulk vanishes, we have:

$$kT \ln \rho_{M,S} - kT \ln \left(q_P 2\pi \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{M,ads}(z)}\, dz \right) + \mu^R_{M,S} = kT \ln c_{M,B} - kT \ln(q_P 8\pi^2) \tag{2.14}$$

Note that the momentum partition functions cancel, which implies that the same momentum phase space is available to the peptide atoms in both the adsorbed and bulk solution state. This is generally true for 'soft' binding, as is the case here, but may not be valid when rigid, covalent binding events take place. Rearranging and solving for $\rho_{M,S}$ yields:

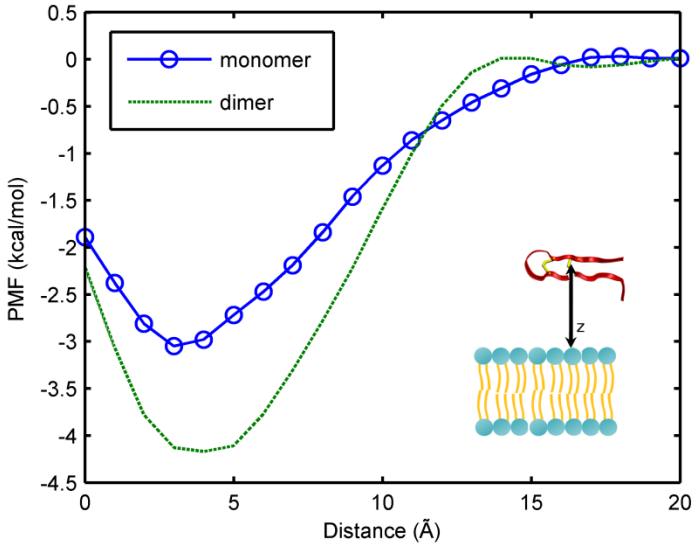
$$\rho_{M,S} = \frac{c_{M,B}}{4\pi}\, e^{-\beta \mu^R_{MS}} \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{M,ads}(z)}\, dz \tag{2.15}$$

With an identical development, we obtain the equivalent relation for dimer adsorption:

$$\rho_{D,S} = \frac{c_{D,B}}{4\pi}\, e^{-\beta \mu^R_{D,S}} \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{D,ads}(z)}\, dz \tag{2.16}$$

Here, $W_{D,ads}(z)$ is the potential of mean force for dimer adsorption, reported by Vivcharuk and Kaznessis in [67] and reproduced in Figure 2-3.

Figure 2-3: PMFs of adsorption for the PG-1 monomer and dimer. Data from [67]



2.2.3 The residual chemical potential

We now turn to the residual chemical potential, $\mu_{M,S}^{R}$. This is a term that accounts for the interactions between peptides, assumed here to be primarily the result of hard-core repulsions (i.e. area exclusion effects in the case of membrane-associated peptides). As already mentioned, we assume that this effect is negligible in the bulk solution, an approximation which is certainly valid in the case of solutions with the low peptide concentrations that are biologically relevant. However, as we will show later, in a two-dimensional membrane environment, the crowding effect is extremely important, even at low bulk solution concentrations. While our assumption of neglecting long-range peptide-peptide interactions in all cases may seem suspect for such highly charged species, recall that all interactions take place in an environment with high ionic strength due to the presence of relatively high concentrations of ions, as well as the presence of charged lipid head groups. In a 0.15 M 1:1 salt solution (e.g. KCl or NaCl), which likely represents a lower bound for the ionic strength near biological membranes, the Debye-Hückel length is on the order of 7Å, smaller than the dimensions of a single PG-1

peptide. Furthermore, as discussed later, the potentials of mean force of dimerization show **no** repulsive long-range interactions in any of the relevant environments, which further validates this assumption.

In order to treat area exclusion effects, we turn to an approach known as scaled particle theory (SPT), originally developed by Reiss et al for the study of hard-sphere fluids. The derivation is quite involved, and is not included here. Briefly, the derivation is based on first determining the equation of state of a system of rigid spherical particles of radius a based on the radial distribution function, $g(r)$, which essentially describes the probability density of other particles from the center of a particle of reference. In scaled particle theory [72], the reference particle is first scaled with a parameter λ, such that $\lambda = 0$ corresponds to the absence of that particle from the system, while $\lambda = a$ corresponds to the full-sized particle. The chemical potential of this system can be shown to have the following form [72-73]:

$$\mu = kT \ln \rho\Lambda^3 + kT\rho \int_0^a 4\pi\lambda^2 g(\lambda,\lambda) d\lambda \qquad (2.17)$$

Here, $\Lambda = h^3/(2\pi mkT)$ denotes the thermal DeBrogile wavelength, which arises by integrating over all possible particle momenta, while $g(\lambda,\lambda)$ is the notation of Reiss et al for the radial distribution function of regular particles from the scaled particle, evaluated at the point of contact (i.e. at the radius of the scaled particle) [72]. The core contribution of scaled particle theory is to approximate $g(\lambda,\lambda)$ based on a number of physical and thermodynamic properties that dictate its general form and its value at various limits, and use this approximation to evaluate the chemical potential (and hence the equation of state, and any other thermodynamic properties) using (2.17). The theory was eventually extended to treat non-spherical, convex particles in two dimensions by Boublik [74], and multicomponent systems by Talbot[75]. In two dimensions, the

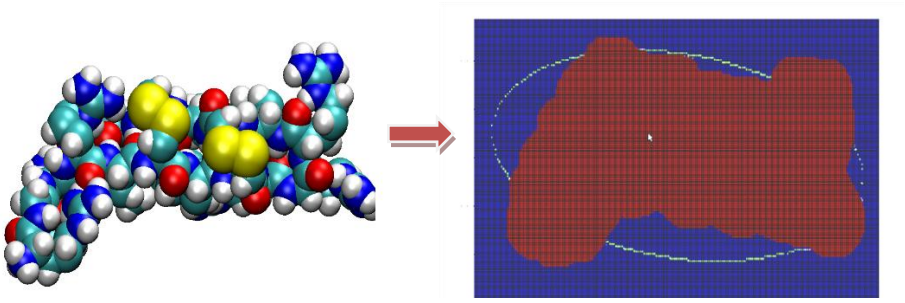
residual chemical potential of a species i in a mixture simplifies to the following expression:

$$-\frac{\mu_i^R}{kT} = \ln\left(1 - \sum_j \rho_j A_j\right) - \frac{A_i \sum_j \rho_j + L_i \sum_j \rho_j L_j / 2\pi}{1 - \sum_j \rho_j A_j} - \frac{A_i (\sum_j \rho_j A_j)^2}{4\pi\left(1 - \rho_j A_j\right)^2} \qquad (2.18)$$

Here, the summations are carried out over all species in the mixture, with ρ_j denoting the surface number density of species j, while A_j and L_j denote its projected molecular area and perimeter, respectively.

This equation is valid for any smooth convex shapes; that is, any shape such that if any two points on the periphery of the shape are connected by a straight line segment, all points on that segment lie inside the shape. For our purposes, we model the membrane-associated species as elliptical particles, with the different peptide states approximated as ellipses with a projected area equal to the projected area of the peptide in its respective surface-bound conformation. This corresponds to a total of four states: surface bound monomers (M_S), surface-bound dimers (D_S), membrane-inserted monomers (M_I) and membrane-inserted dimers (D_I). The major and minor axes of the ellipse are selected based on the ratio of the principal moments of area of the projected species surface. The conformations selected for fitting were based on the native PG-1 monomer and dimer structures in the appropriate orientations. An example of this ellipse-fitting is depicted in the figure below:

Figure 2-4: Fitting the footprints of different peptide states with ellipses



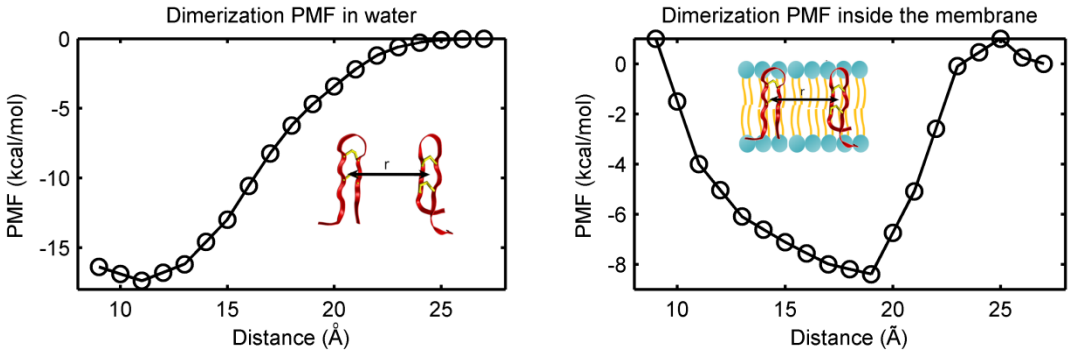
Note that given the definitions of the surface densities already discussed - for surface-bound species, the total number divided by twice the area of the bilayer; for inserted species, the total number divided by the bilayer area - the expression for the residual chemical potential needs no alterations to account for bilayers as opposed to monolayers. One can imagine that in the context of a lipid bilayer, the system is identical to a plane with twice the area of the lipid bilayer, where surface-bound peptides are represented by single particles with equivalent footprints, whereas inserted peptides would be represented by two particles, one for each leaflet. The net densities of both species are then equal to their definitions above (see section 2.2.2), and no further changes need to be made to equation (2.18). In the simpler case of the lipid monolayer, the surface density definitions change accordingly, yielding the same result.

2.2.4 Equilibrium relationships for dimerization

The derivation of the equilibrium relationship for dimerization in the bulk solution follows the same general scheme as above. In this case, both the monomer and dimer can be treated as components of an ideal solution with no residual chemical potential, since they are both very dilute for any reasonable range of biologically relevant experimental concentrations. The relevant potential of mean force in this case is along a reaction coordinate separating the centres of mass of two monomers, which was reported

by Vivcharuk and Kaznessis in previous work in different environments [69]. The relevant data are reproduced below (courtesy of Dr. Victor Vivcharuk):

Figure 2-5: Dimerization PMFs in two different environments. Data from [69]



The contribution of the relevant reaction coordinate to the dimer partition function arises as an internal degree of freedom of the dimer, in which one constituent monomer is located at some distance r with respect to the other monomer. We consider the potential of mean force to be spherically symmetrical in this case, provided the orientation of the two monomers is favorable, which yields a transformation Jacobian of r^2 in the bulk solution, and r in the membrane-inserted (two-dimensional) case. This implicitly assumes that contributions to the dimerization PMF from configurations in which the monomers are not oriented with respect to one another in the same way as they are in the parallel dimer are negligible in comparison to those where the orientations correspond to dimer formation. As before, the potential of mean force vanishes as $r \to \infty$. With these considerations, we obtain the following expression for the chemical potentials of the monomer and dimer in the bulk solution phase:

$$\mu_{M,B} = kT \ln c_{M,B} - kT \ln 8\pi^2 q_P \qquad (2.19)$$

$$\mu_{D,B} = kT \ln c_{D,B} - kT \ln \left(\frac{8\pi^2}{2} q_P^2 \int\limits_{\substack{dimer \\ state}} e^{-\beta W_{dim,B}} r^2 dr \right) \tag{2.20}$$

The limits of the integration ("$dimer\ state$") indicate that the integration should only be carried out on that part of the PMF that corresponds to the existence of a dimer; the factor of $8\pi^2/2$ arises from the configurational part of the rotational partition function in the dimer, which has a symmetry number [71] of 2 in the case of the dominant parallel dimer conformation (i.e. there are two indistinguishable orientations of the parallel dimer), and the momentum contribution is squared because the dimer has exactly twice as many atoms as the monomer. Substituting these in the equilibrium relation for dimerization, $2\mu_{M,B} = \mu_{D,B}$ and simplifying, we obtain the following equilibrium relationship:

$$c_{D,B} = \frac{1}{16\pi^2} c_{M,B}^2 \int\limits_{\substack{dimer \\ state}} e^{-\beta W_{dim,B}(r)} r^2 dr \tag{2.21}$$

Similarly, for dimerization in the membrane-inserted state, we obtain the following relation for the chemical potentials of the membrane-inserted monomers and dimers:

$$\mu_{M,I} = kT \ln \rho_{M,I} - kT \ln 2\pi q_P + \mu_R^{M,I} \tag{2.22}$$

$$\mu_{D,I} = kT \ln \rho_{D,I} - kT \ln \left(\frac{2\pi}{2} q_P^2 \int\limits_{\substack{dimer \\ state}} e^{-\beta W_{dim,I}(r)} r dr \right) + \mu_R^{D,I} \tag{2.23}$$

The rotational partition function is only 2π in this case (rotation around the membrane normal), the symmetry number of 2 remains applicable for the dimer (rotating the dimer 180º yields a conformation indistinguishable from the original), and the Jacobian of

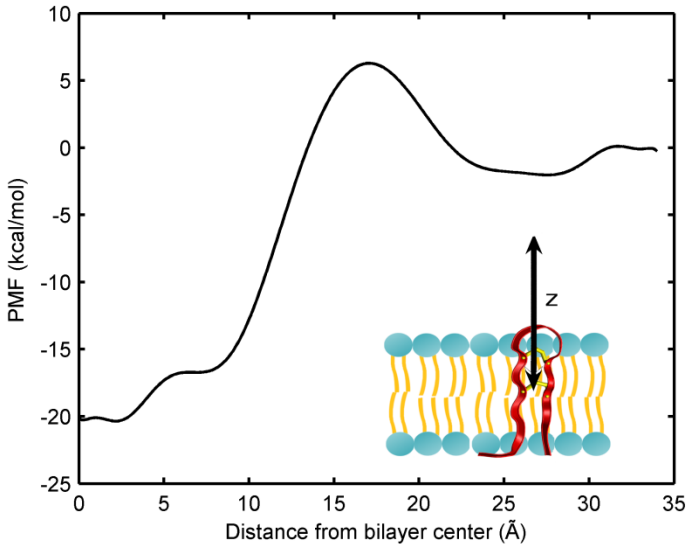
transformation is r in the integral over the spherically-symmetrical PMF. Setting $2\mu_{M,I} = \mu_{D,I}$ we obtain:

$$\rho_{D,I} = \rho_{M,I}^2 \frac{1}{4\pi} \int_{\substack{dimer \\ state}} e^{-\beta W_{dim,ins}(r)} r dr \, e^{-\beta(\mu_{D,I}^R - 2\mu_{M,I}^R)} \tag{2.24}$$

2.2.5 Equilibrium relationships for membrane insertion

The process of membrane insertion consists of an adsorbed monomer or an adsorbed dimer penetrating the membrane so that its center of mass lies near the center of the lipid bilayer. The potential of mean force for this process was recently computed using similar techniques as described for the adsorption and dimerization PMFs by Dr. Victor Vivcharuk. We present these preliminary data in Figure 2-6, again courtesy of Dr. Vivcharuk.

Figure 2-6: Preliminary data for the PMF of monomer insertion



Following the same procedure outlined above for peptide adsorption yields the following relation for the equilibrium between inserted and adsorbed monomers:

$$\frac{\rho_{M,S}}{\rho_{M,I}} = \frac{\int_{ads} e^{-\beta W_{M,ins}(z)} dz}{\int_{ins} e^{-\beta W_{M,ins}(z)} dz} e^{-\beta(\mu_{R,M}^{ads} - \mu_{R,M}^{ins})} \tag{2.25}$$

In this case, all orientational degrees of freedom are the same in both states – even though the peptide reorients upon insertion, rotation in the plane of the membrane remains unhindered in both the inserted and membrane-adsorbed states, while the remaining rotational degrees of freedom are constrained (and sampled as part of the PMF) in both states. The integrals are carried out over the corresponding regions of the PMFs, which in this case are clearly defined by the two wells at $-10<z<10$ (symmetry of the PMF is assumed due to symmetry of the bilayer) and $20<z<30$. As before, the precise bounds that define the two states are not particularly relevant, as most of the contribution to the integral arises from the regions near the minima.

2.2.6 Membrane expansion as a result of inserted species

In the derivations above, the surface densities of various membrane-associated species are implicitly defined as the total number of a particular species divided by the total in-plane area available to each particular species, assuming no interactions with other peptides. However, the insertion of peptides in either monolayers or bilayers causes them to expand, which increases the total available area (an exception to this are monolayers and bilayers constrained at constant area, in which case insertion causes an increase in surface tension; the model can easily be modified to account for this). The nonideal area exclusion effect is accounted for in the residual chemical potential, which means that the surface density definitions above are based on this total, expanded area. In other words, a system with a given number of inserted species that result in a total expanded area A_T is completely equivalent to an unexpanded lipid bilayer with area A_T that has the same number of species with the same footprints as the inserted species

adsorbed to its surface. In some cases, it may be more convenient to define the densities based on the area occupied by lipids. Such an approach is outlined below.

We define the total area of the lipid bilayer as A_T (not including both top and bottom surfaces, but only the lateral dimensions). As already discussed, the definitions of the surface densities are implicitly based on the total area available for translation, which may include expansion due to the insertion of peptides. In the case of membrane-adsorbed peptides, the surface density ρ_i is therefore $\rho_i = N_i/2A_T$, while in the case of membrane inserted species $\rho_i = N_i/A_T$. We therefore write:

$$\rho_i = \frac{N_i}{bA_T} \ , \tag{2.26}$$

where b is 2 for surface-bound species and 1 for membrane-inserted species; in the case of lipid monolayers, b is 1 for both states. We next define a new surface density ρ_i^L based only on the area occupied by lipid molecules, A_L, such that

$$\rho_i^L \equiv N_i/bA_L \tag{2.27}$$

The total area A_T can be written as the sum of the area occupied by lipids, A_L, and the area occupied by the inserted species, k:

$$A_T = A_L + \sum_{k\,\epsilon\,ins} N_k \bar{A}_k \tag{2.28}$$

The summation above runs over all inserted species, and $\bar{A}_k$ denotes the area increase per molecule of inserted species k, assumed here to be the same as the specific area of the molecule. We can re-arrange equation (2.27) as $N_k = \rho_k^L A_L$ in the case of inserted species, and insert this in (2.28) to obtain:

$$A_T = A_L + \sum_{k\,\epsilon\,ins} \rho_k^L A_L \bar{A}_k \tag{2.29}$$

We abbreviate the summation term $S_k \equiv \sum_k \rho_k^L A_L \bar{A}_k$ for convenience. With the above relationship for A_T, we can rewrite the density of each species (2.26) as follows:

$$\rho_i = \frac{N_i}{b(A_L + S_k)} = \frac{N_i}{bA_L} - \frac{N_i S_k}{bA_L(A_L + S_k)} = \rho_i^L \left(1 - \frac{S_k}{A_L + S_k}\right)$$
$$= \rho_i^L \left(1 - \frac{\sum_k \rho_k^L A_L \bar{A}_k}{A_L + \sum_k \rho_k^L A_L \bar{A}_k}\right) = \rho_i^L \left(\frac{1}{1 + \sum_k \rho_k^L \bar{A}_k}\right) \quad (2.30)$$

Recall that the summation here is only over the inserted species. This final expression allows us to express all the densities ρ_i discussed in sections 2.2.2-2.2.5 in terms of densities defined based on a lipid area that is constant (ρ_i^L). Once again, since *b* is included in the definition of ρ_i^L, these densities have different interpretations for surface-adsorbed and membrane-inserted species. To obtain the surface densities of adsorbed species as the number of molecules per lipid bilayer area, we re-arrange equation (2.27), and need only multiply ρ_i^L by *b*, i.e. by 2 in the case of bilayers, and by 1 in the case of monolayers:

$$\rho_i^{L*} = b\rho_i^L = b\rho_i \left(1 + \sum_k \rho_k^L \bar{A}_k\right) \quad (2.31)$$

Here, ρ_i^{L*} is the density of adsorbed species *i* expressed as the total number of molecules per lipid bilayer area (not including the expansion due to inserted species, and not including both the top and bottom surfaces). This last equation allows us to convert between the three different ways of expressing the surface densities of the different species.

To summarize, the model consists of the following set of equations (identical to those in Table 2-1, except equation (2.31) was used to express all surface densities based on the area covered by lipids):

$$c_{D,B} = \frac{1}{16\pi^2} c_{M,B}^2 \int_{\substack{dimer \\ state}} e^{-\beta W_{dim,B}(r)} r^2 dr \quad (2.32)$$

$$\rho_{M,S}^{L} = \left(1 + \sum_{k} \rho_{k}^{L} \bar{A}_{k}\right) \frac{c_{M,B}}{4\pi} \, e^{-\beta \mu_{M,S}^{R}} \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{M,ads}(z)} \, dz \tag{2.33}$$

$$\rho_{D,S}^{L} = \left(1 + \sum_{k} \rho_{k}^{L} \bar{A}_{k}\right) \frac{c_{D,B}}{4\pi} \, e^{-\beta \mu_{D,S}^{R}} \int\limits_{\substack{adsorbed \\ state}} e^{-\beta W_{D,ads}(z)} \, dz \tag{2.34}$$

$$\rho_{M,S}^{L} = \rho_{M,I}^{L} \frac{\int_{z \, \epsilon \, ads} e^{-\beta W_{M,ins}(z)} dz}{\int_{z \, \epsilon \, ins} e^{-\beta W_{M,ins}(z)} dz} e^{-\beta (\mu_{M,S}^{R} - \mu_{M,I}^{R})} \tag{2.35}$$

$$\rho_{D,I}^{L} = \frac{{\rho_{M,I}^{L}}^{2}}{4\pi (1 + \sum_{k} \rho_{k}^{L} \bar{A}_{k})} e^{-\beta (\mu_{D,I}^{R} - 2\mu_{M,I}^{R})} \int\limits_{\substack{dimer \\ state}} e^{-\beta W_{dim,I}(r)} r dr \tag{2.36}$$

These five equations contain a total of six unknowns: $c_{M,B}, c_{D,B}, \rho_{M,S}^{L}, \rho_{D,S}^{L}, \rho_{MI}^{L}$ and ρ_{DI}^{L}. The closure relation comes from specifying the total bulk concentration, as discussed next.

2.3 Results and Discussion

The data presented in this section are the results of solving equations (2.32) through (2.36) for a given total bulk peptide concentration. In typical experimental systems as well as in biological contexts, the total number of peptides in the bulk solution far exceeds the total number of peptides adsorbed or inserted, even at complete coverage. Thus, we can safely assume that the total bulk concentration in the bulk (of either monomer or dimer) will not be significantly affected by the number of peptides that are adsorbed or inserted. We can therefore write:

$$c_{M,B} + 2c_{D,B} \cong c_{o} \tag{2.37}$$

Together with equation (2.32), we can thus easily solve for the monomer and dimer bulk concentrations, and proceed to solve the remaining equations (2.33) through (2.36). Note that this simplification is not strictly required for a system in which the ratio of the total bulk solution volume to the total lipid bilayer area is known. This would not generally be a difficult quantity to estimate for experimental or biological systems: for instance, in

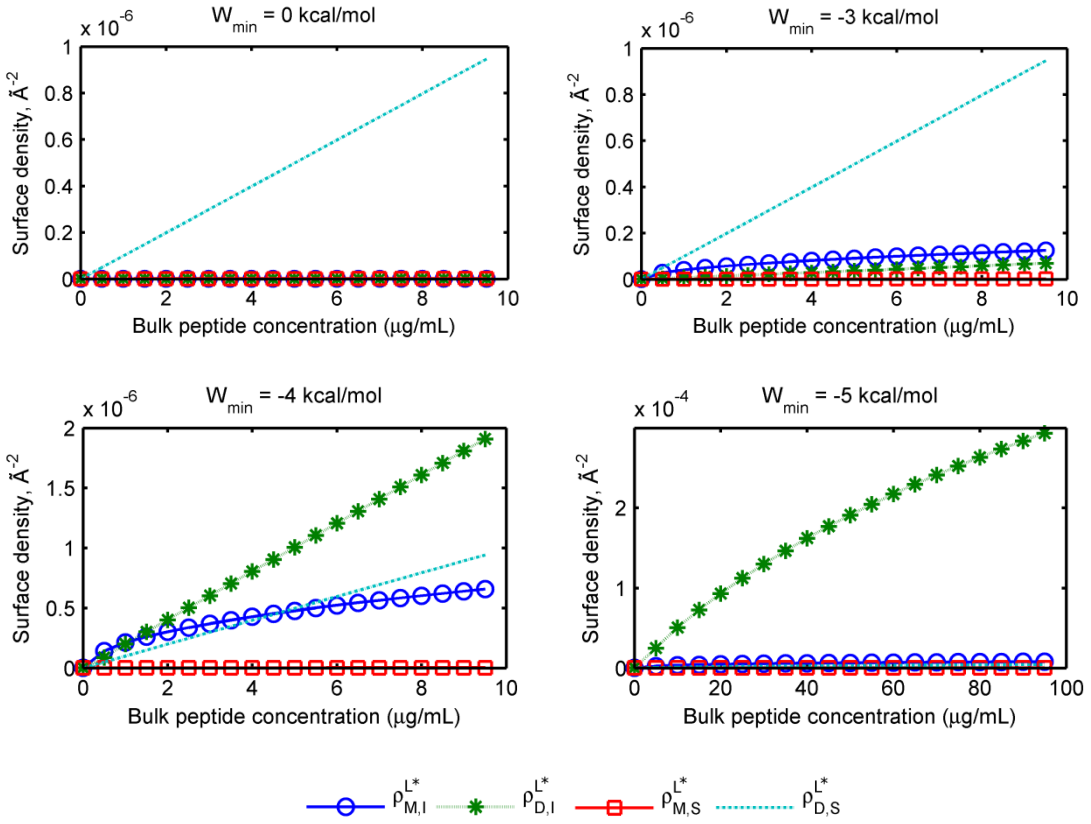
experiments with vesicles, knowledge of the lipid concentration is sufficient, while knowledge of the concentration and average size of bacteria are sufficient in a biological context. However, since we are not comparing these results to any particular experimental system, we retain the approximation of equation (2.37).

2.3.1 Results for different insertion free energies

When we began the work discussed in this chapter, the data shown in Figure 2-6 for the PMF of insertion of the PG-1 monomer were not available. Due to inherent difficulties in molecular dynamics simulations of lipid bilayer systems with a changing area, the same methods used in computing adsorption [67] and dimerization [69] PMFs were not immediately applicable in the case of insertion, and the preliminary data in Figure 2-6 may still be problematic [68]. We therefore allowed the insertion free energy to be a variable parameter, where a parabolic well shape was assumed, and the difference in the depth of the inserted monomer state and the adsorbed monomer was varied. The difference between the monomer insertion free energy minimum and the monomer adsorption free energy minimum is referred to as W_{min}, such that a negative value indicates favorable insertion. Figure 2-7 below shows the results for the surface densities of the different forms of protegrin for several values of this free energy difference, W_{min}.

Figure 2-7: Surface densities of different states of protegrin for different W_{min} values

Note the different scales of the ordinate axes, particularly for the case of W_{min}=−5 kcal/mol, as well as the different scale for the abscissa in this case.



Several interesting behaviors can be seen in these figures. At $W_{min} = 0$, which corresponds to a potential of mean force in which the energy of an inserted monomer is equal to that of an adsorbed monomer, insertion is not favored, and the dominant species are surface-adsorbed dimers (D_S). This is not due to non-idealities or due to the entropic cost associated with insertion (recall that only half the lateral area is available to inserted species for translation). Rather, it is a simply a result of the low concentration of monomers in the bulk – in equation (2.32), the integral evaluates to very large numbers due to the deep dimerization potential well (see Figure 2-5). As a result, inserted and

adsorbed monomer densities are likewise extremely low, which in turn results in low inserted dimer densities (viz. equation (2.36)). As the free energy of insertion is increased to W_{min}=−3 kcal/mol, monomer insertion starts to take place, but surface-adsorbed dimers remain favorable. The reason for the very low density of surface-adsorbed monomers ($\rho_{M,S}^{L}$) despite a favorable adsorption free energy is the same as before, namely the low concentration of monomers in the bulk phase. At W_{min}=−4 kcal/mol, the inserted dimers begin to dominate. At very low concentrations, there is a small regime where inserted monomers are the most prominent species; however, as their density increases, dimerization in the membrane becomes increasingly favorable (again, this is due to the squared dependence of $\rho_{D,I}$ on $\rho_{M,I}$ in equation (2.36)). Interestingly, this effect eventually favors the presence of surface-bound dimers over inserted monomers, despite the relatively high free energy of insertion. As the free energy of insertion is increased to W_{min} = −5 kcal/mol, the inserted species are by far the most prevalent due to the highly favorable insertion free energy, with dimers favored for the same reasons discussed above.

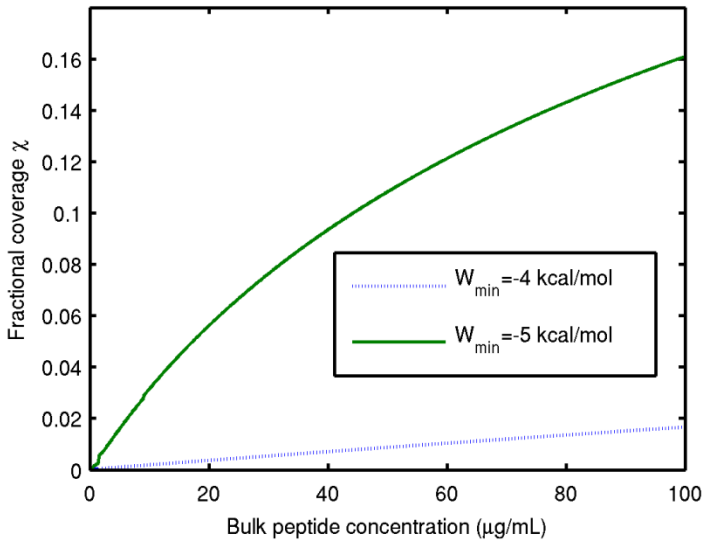
We have solved the model to higher values of the peptide bulk concentration for this last case (note the different scale of the abscissa for W_{min} = −5 kcal/mol) in order to illustrate the effects of non-idealities, which become important at this value of W_{min} and higher peptide concentrations – the residual chemical potential (equation (2.18)) is the only reason why the $\rho_{D,I}^{L}$ curve begins to level off for this case. Finally, we note that the large variations in behavior resulting from seemingly small variations in W_{min} arise due to the exponential dependence of the equilibrium on W_{min} in equation (2.35).

We also analyze the total fraction of the lipid bilayer covered by peptides, defined as follows:

$$\chi = \frac{\sum_i N_i \bar{A}_i}{A_L + \sum_k N_k \bar{A}_k} = \sum_i \rho_i^L \bar{A}_i \tag{2.38}$$

Here, χ denotes the fraction of the lipid area – consisting of both surfaces of the lipid bilayer, and accounting for expansion due to inserted species – that is covered by peptides. The summation over *k* denotes summation over inserted species only, while summation over *i* denotes summation over all species. Further details are provided in section 2.2.6. Figure 2-8 below shows plots of χ, the total fractional surface coverage, for two values of W_{min}:

Figure 2-8: Total surface coverage (χ) isotherms for two different W_{min} values



Less negative values of W_{min} predictably result in lower coverage, with a linear dependence on the bulk peptide concentration. The coverage for W_{min}= -5kcal/mol is primarily a result of inserted dimers, which are the dominant species at these conditions. The coverage is relatively large (approaching 0.2, or 20%), but by no means saturated. Nevertheless, we can see that non-idealities begin to play a role at surface coverage values less than 0.1, or 10%.

It is instructive to quantify the effects of the non-idealities separately for each species by analyzing the different residual chemical potentials, μ_i^R. For ease of

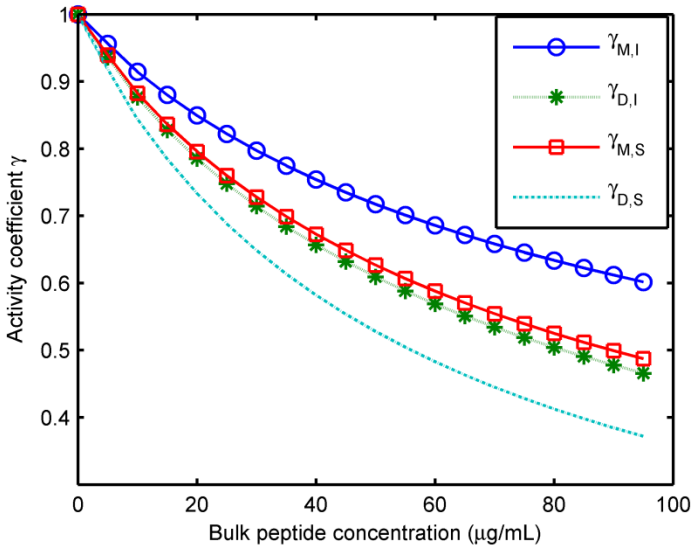
comparison, we present these as the activity coefficients γ_i of each species, which are related to μ_i^R as follows:

$$\gamma_i = \exp(-\mu_i^R / kT) \tag{2.39}$$

In the present case, activity coefficients greater than unity are not possible, as this would require negative area and perimeter values in equation (2.18) to make μ_i^R negative (physically, this would imply that crowding makes the binding of additional peptides more favorable, which clearly cannot occur with any hard-core repulsion model). Thus, γ_i will always be between 0 and 1, with 1 representing the ideal case, in which crowding has no effect. The activity coefficients for all four membrane-associated species are plotted below for the case of W_{min} = −5 kcal/mol (in the other cases discussed above, all values of $\gamma_i$were consistently above 0.95).

Figure 2-9: Acitivity coefficients γ_i of all species for W_{min}=−5 kcal/mol

Activity coefficients below unity indicate departure from ideality. Activity coefficients are shown for the species described in the text: inserted monomers ($\gamma_{M,I}$), inserted dimers ($\gamma_{D,I}$), surface-adsorbed monomers ($\gamma_{M,S}$) and surface-adsorbed dimers ($\gamma_{D,S}$).



As shown in Figure 2-9, all activity coefficients deviate significantly from unity. In general, non-ideal effects favor insertion, due to the smaller footprint of inserted species compared to surface-adsorbed species. Monomers also appear to be favored, but a closer look at equation (2.36) reveals that the right-most exponential term is equivalent to γ_D/γ_M^2, so that the net effect of non-idealities for these values of the activity coefficients is actually to favor dimerization.

2.3.2 Results based on the insertion PMF computed from MD simulations

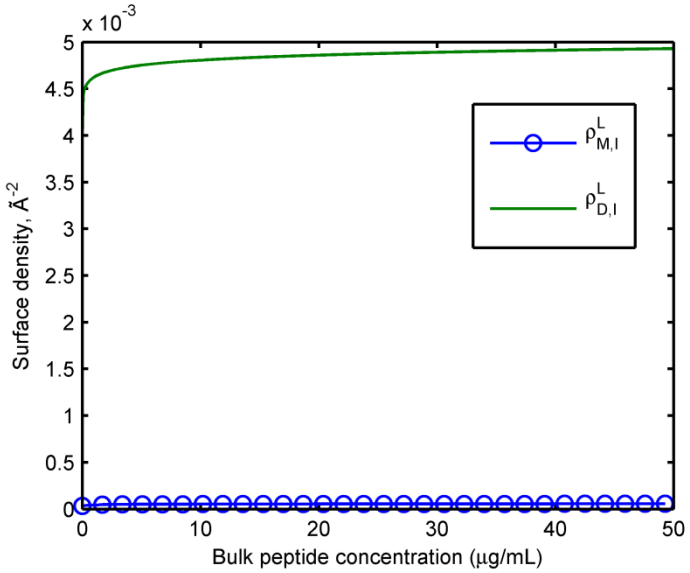
We now turn our attention to results based on the PMF extracted from the MD simulations of Vivcharuk and Kaznessis [68], shown in Figure 2-6. This PMF has a surprisingly deep potential well at insertion, corresponding approximately to $W_{min} = -20$ kcal/mol in the context of the previous section. Considering the results just discussed, this clearly indicates that inserted species are the only relevant species, so we can safely simplify our model to exclude the surface-bound species (in fact, retaining these makes the numerical solution of equations (2.32)-(2.36) extremely difficult due to variations of nearly ten orders of magnitude in the coupled variables). The reduced model consists of equation (2.32) for dimerization in the bulk, equation (2.36) for dimerization in the membrane, and an equation that treats the equilibrium of inserted monomers with monomers in the bulk phase:

$$\rho_{M,I}^{L} = \left(1 + \sum_{k} \rho_k^L \bar{A}_k\right) \frac{c_{M,B}}{4\pi}\, e^{-\beta \mu_{M,S}^{R}} \int_{\substack{inserted \\ state}} e^{-\beta W_{M,ads}(z)}\, dz \qquad (2.40)$$

Solving these equations together with (2.37) gives the adsorption isotherms for inserted dimers and monomers, shown in Figure 2-10 below.

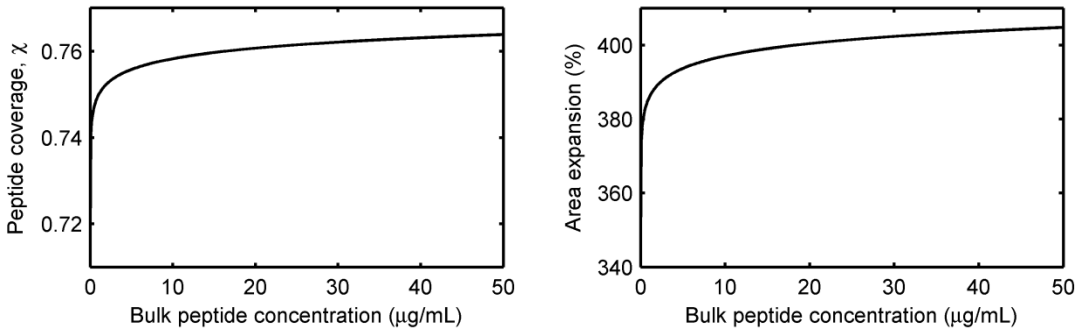
Figure 2-10: Adsorption isotherms based on the PMF computed from MD simulations [68]

The dominant species in this case are inserted dimers, with a surface density $\rho_{D,I}^{L}$. There are also some surface-inserted monomers, but at a much lower density ($\rho_{M,I}^{L}$)



Clearly, the dimer form is again dominant (viz. equation (2.36)), and both species are practically saturated due to the extremely high insertion affinity. While the surface densities shown on the ordinate axis appear reasonable, they are in fact extremely high, as discussed next. In Figure 2-11, we plot the isotherm expressed as the total surface coverage χ as a function of bulk peptide concentration. We also plot the total area expansion caused by the insertion of monomers and dimers (see section 2.2.6).

Figure 2-11: Total surface coverage (left) and membrane area expansion (right) for the PMF obtained from MD simulations



Not surprisingly, these properties follow exactly the same trends as one another, as they are driven almost exclusively by the insertion of dimers. The peptide coverage of ~0.75 refers to the portion of the total expanded area that is occupied by peptides (in this case, inserted dimers). This in turn is equivalent to a four-fold expansion of the lipid bilayer, as indicated in the right panel of Figure 2-11. We can gain further insight into the implications of this by looking at the definition of χ:

$$\chi = \frac{\sum_i N_i \bar{A}_i}{A_L + \sum_i N_i \bar{A}_i} \approx 0.75 \rightarrow \sum_i N_i \bar{A}_i \approx 3A_L$$

In other words, in the expanded bilayer, the area occupied by inserted peptides is three times the area occupied by lipids. With a typical lipid area of ~63Å^2/lipid, and the area of an inserted dimer ~600Å^2, this translates to approximately one dimer for every three lipid molecules. While this is not completely impossible, it seems highly unlikely.

There are several possible reasons for the unrealistically high surface densities just presented. First and foremost, the potential of mean force from MD simulations may be overestimating the insertion free energies. This could be a result of inadequate

sampling during the simulations, as well as inherent inaccuracies in the empirical force fields employed (CHARMM27 in this case) and the underlying assumptions of classical mechanical force fields (in particular, their failure to account for polarization and changes in ionization states). Recall from section 2.3.1 that due to the Boltzmann factor, relatively small differences in the PMF can give rise to significant differences in isotherm behavior. Second, our model only takes into account non-idealities in the context of hard-core repulsions between ellipsoidal particles corresponding to the footprints of various states (section 2.2.3). At the high surface coverage values seen here, this approximation likely begins to break down, and more specific interactions likely start to become important. Third, we have not accounted for the mechanics of the lipid bilayer itself beyond the information contained in the PMF. In reality, a lipid bilayer with a high density of embedded peptides would experience significant mechanical stresses, leading to rupture and possibly solubilization of the membrane. Given the high peptide to lipid ratios (1:3) observed herein, one would expect the hydrophobic clustering that is responsible for the integrity of lipid bilayers to largely disappear. Furthermore, in a biological context, particularly in bacteria, lipid bilayer membranes cannot expand arbitrarily, as they are confined by peptidoglycan layers (see section 1.4). Thus, significant expansion of the lipid bilayer would lead to extrusion through the peptidoglycan layer/cell wall, and result in rupture. We note that this effect is precisely what has been observed in live *E. coli* treated with sufficiently high concentrations of protegrin (see Chapter 4, particularly sections 4.3.3 and 4.3.4). Clearly, we cannot conclude that there is a direct causal connection between a high degree of peptide insertion and this extrusion/rupture at high protegrin concentrations, but it is nonetheless noteworthy that our data are compatible with this hypothesis.

2.4 Summary

In this chapter, we have presented a modeling framework that allows us to unify molecular simulation work previously carried out in our group [67-69], and connect it to macroscopic and biologically relevant quantities like adsorption isotherms. We have developed a theoretical framework that uses molecular thermodynamic information in the form of potentials of mean force as an input, and provides adsorption isotherms for the important states of protegrins – monomers and dimers bound to the surface of lipid bilayers, and inserted in the hydrophobic membrane core. Our approach accounts for a number of additional entropic contributions that are otherwise absent from the PMFs, as well as non-idealities that arise due to area exclusion effects in the crowded membrane environment. We first present data for variable insertion free energies, which shows that at low values of the insertion free energy (close to 0), the adsorbed dimer form is favored, and relatively low surface coverage is observed. As the insertion free energy decreases (i.e. becomes more negative, meaning insertion becomes more favorable), surface-inserted species become increasingly abundant, and completely dominate at values of $W_{min} < -4$ kcal/mol. We show that non-idealities become important when $W_{min} \approx -5$ kcal/mol, corresponding to total surface coverages between 10 and 20%.

If we then use the PMF of insertion provided by the work of Dr. Victor Vivcharuk using all-atom MD simulations [68], it becomes clear that inserted species are the only ones present at relevant levels (the MD-based PMF corresponds approximately to $W_{min} \approx -20$ kcal/mol). The resulting adsorption isotherms further show that inserted dimers are highly preferred over inserted monomers. A closer look also reveals that this leads to a total surface coverage of approximately 75%, and that the total membrane area must expand by a factor of 4 as a result of peptide insertion. This improbable scenario is likely a result of a combination of inaccuracies in the MD-derived PMF, as well as other shortcomings of the present model. Such high insertion affinities would probably result in membrane disintegration or possibly extrusion and rupture in the case of biological membranes. It is noteworthy that both membrane disintegration in the case of lipid

vesicles and extrusion and rupture in the case of biological membranes (see chapter 4) have been experimentally observed at sufficiently high protegrin concentrations.

The work in this chapter represents a useful starting point in creating a complete statistical thermodynamic framework for the description of the equilibrium behavior of protegrin interacting with lipid bilayer membranes. Important enhancements in the future could include the addition of a term to account for non-idealities other than area exclusion terms, such as long-range electrostatic effects, as well as including the formation of pores (octamers or higher oligomers), which are believed to be the most important structural unit for bactericidal action (chapters 3-5). Finally, this work is an important starting point for the development of kinetic models that can describe the coupled dynamics of different modes of membrane association and pore formation, and eventually couple these to ion transport dynamics (see chapters 3-4).

3 Steady-state ion diffusion through a protegrin transmembrane pore

3.1 Introduction

In the remaining three chapters, we turn our attention to protegrin pores, the structural units most relevant to bactericidal activity. The most compelling and detailed evidence of the existence of such structures comes from the NMR study reported by Mani et al [54], which suggests that PG-1 aggregates into octameric or decameric transmembrane pores in 3:1 palmitoyloleoyl phosphatidylethanolamine (POPE): palmitoyloleoyl phosphatidylglycerol (POPG) lipid bilayers. These pores are the result of the aggregation of four dimers, and NMR distance data suggests that the dimer pairs remain relatively intact in the pore structure. Related work was already briefly reviewed in section 1.5.2. Molecular dynamics simulations of these structures were subsequently carried out in our group by Langham et al [76]. These simulations provided a dynamic picture of the pore, and yielded information about these structures with an unprecedented level of detail. The authors describe the structural details of the peptide pore, and show that despite some distortion and narrowing of the pore inner diameter, the simulation structures remain in agreement with NMR-derived distance constraints. The evolution of the structure of the lipids around the peptide pore was also investigated by these authors [76], a point that is explored in greater detail with more recent molecular dynamics simulations in chapter 5. Finally, the simulations of Langham et al showed that ions and water can move readily through protegrin pores, which likely represents a key disruptive feature of protegrins. We therefore explore the ion conductance characteristics of protegrin pores in greater detail in this chapter, and the effects of such pores on an entire cell in the following chapter.

Prior to the NMR study of Mani et al [54] and the molecular dynamics simulations of Langham et al [76], early evidence of the ability of protegrin pores to

conduct ions was provided by the experiments of Mangoni et al [59] and Sokolov et al [58], in which the conductance characteristics of protegrin-treated membranes were measured. Such knowledge of the nonequlibrium ion flow through protegrin pores may be closely related to their mechanism of action, since the unrestricted flow of ions through the membrane could result in potentially lethal membrane depolarization. Mangoni and coworkers conducted voltage clamp experiments in *Xenopus laevis* oocyte membranes treated with protegrin-1 and several analogues [59]. They found that protegrins form weakly anion selective pores in the presence of several different salts, with KCl solutions exhibiting almost no selectivity. Furthermore, they found that the conductance of such pores does not exhibit any voltage dependence over a voltage range of -100 to +30 mV. Sokolov and coworkers [58] carried out conductance measurements across several different types of planar phospholipid bilayers treated with protegrin-1 as well as protegrin-3 (PG3, amino RGGGLCYCRRRFCVCVGR; note the only difference between PG-1 and PG-3 is the substitution of a glycine residue at the fourth position in PG-3 instead of the arginine residue in PG-1). These authors found that both protegrin analogues form weakly anion selective channels in mixed phospholipid bilayers, and moderately cation-selective channels in bilayers containing negatively charged bacterial lipopolysaccharide (LPS). They reported a voltage-dependent single-channel conductance in the range of 40-360 picoSiemens (pS), depending on the peptide used, the lipid bilayer composition, and the applied voltage.

We therefore set out to explain and quantify the conductance behaviour of protegrin pores in terms of structural information from the molecular dynamics simulations of Langham et al [76] as well as from NMR experiments [54]. In particular, we explore the connection between structural features such as the size of the pore opening and the magnitude of the conductance, the effects of different charged groups, as well as the surprising experimental finding of both Mangoni et al [59] and Sokolov et al [58] that the protegrin pore is only slightly anion selective, despite having a total charge of +56. Our investigation is based on the Poisson-Nernst-Planck (PNP) theory, a continuum method

of calculating non-equilibrium ion concentrations and fluxes around a fixed structure in the presence of an applied electrical voltage. In our model, we simulate a voltage across a protegrin pore embedded in a lipid bilayer patch, and measure the resulting current. Since the PNP model requires a rigid structure, we perform the calculations using several snapshots from the MD simulations of Langham and coworkers [76]. We did not repeat these calculations for structures form the MD simulations discussed in chapter 5, which were only very recently completed. However, considering the relatively small discrepancies among the structures tested herein, which also include toroidal pore structures from earlier simulations, we do not expect that the more recent simulations would yield significant differences with respect to the types of results that follow.

3.2 Methods

This section outlines the methods employed in the modeling of steady-state ion transport through protegrin pores. For methodological details related to the molecular dynamics simulations from which structures were extracted, the reader is referred to the work of Langham et al [76], as well as the Appendix at the end of this document.

3.2.1 Governing equations and boundary conditions

The underlying theory and governing equations are based on the Poisson-Nernst-Planck (PNP) set of equations, which describe non-equilibrium ion distributions in the presence of charged media with varying dielectric properties. In PNP theory, ion flux $\boldsymbol{J}_i$ is modeled according to the Nernst-Planck equation, which includes a Fickian diffusion term and a drift term to account for the effects of the electrostatic field:

$$\boldsymbol{J}_i = D_i \nabla c_i + \frac{D_i q_i e}{k_B T} c_i \nabla \varphi \qquad (3.1)$$

Here, the index *i* corresponds to each of N diffusing ionic species (e.g. Na^+, K^+, Cl^-), c_i is the concentration of species *i*, φ is the electrostatic potential, D_i is the (spatially-

dependent) diffusion coefficient of species *i*, q_i is the charge of species *i*, e is the elementary charge, 1.60217 X 10^{-19} C, k_B is Boltzmann's constant, 1.38065 X 10^{-23} J/K, and T is the temperature in Kelvins. At steady state conditions, mass continuity yields:

$$\nabla \cdot \boldsymbol{J}_i = 0 \tag{3.2}$$

Substituting equation (3.1) above for the species flux $\boldsymbol{J}_i$ yields:

$$\nabla \cdot \left(D_i \nabla c_i + \frac{D_i q_i e}{k_B T} c_i \nabla \varphi \right) = 0 \tag{3.3}$$

This is the steady-state Nernst-Planck equation, which determines ion concentrations given an electrostatic potential field. It is identical in structure to the steady-state Smoluchowski equation, where the concentration can be equivalently interpreted as a probability of finding a particle (ion) at a particular location. We are interested in the steady-state case because this is the physical situation we are trying to model; the implicit assumption in this type of modeling is that the dynamic fluctuations of the peptide structure do not have a significant effect on the ion current.

The electrostatic potential field, φ, depends on any fixed charges arising from macromolecular structures (e.g. peptides or lipids) as well as the mobile charge arising from the space-dependent ion concentrations through the Poisson equation of electrostatics:

$$\nabla \cdot (\varepsilon \nabla \varphi) = -\rho_f - \sum_{i=1}^{N} q_i c_i \tag{3.4}$$

Here, ρ_f represents the fixed charge density and ε is the (space-dependent) dielectric constant. The coupled equations (3.3) and (3.4) above represent the Poisson-Nernst-Planck system of equations that describe the steady-state (but not necessarily equilibrium) distribution of ions around a set of fixed charges.

The coupled PNP system (equations (3.3) and (3.4)) is solved on a three-dimensional domain defined by the protegrin pore structure. The fixed charge density ρ_f is determined from the positions and charges of the protein and lipid atoms in the pore structure. The boundary conditions in the plane of the bilayer are periodic for all ion concentrations as well as for the electrostatic potential. In the direction perpendicular to the bilayer, the electrostatic potential is set to zero far away from the pore on the side of the protegrin β-sheet turns, and to the applied voltage ($V_{applied}$) far away from the pore on the opposing side, nearest to the protegrin termini (refer to Figure 1-5). Similarly, ion concentrations are set to their bulk values far away from the pore in the direction perpendicular to the bilayer. Additionally, there is a no-flux boundary surrounding the peptide and lipid atoms that prevents ions from penetrating through the region occupied by the peptides and lipids. This can be physically interpreted as an approximate way to account for the short-range van der Waals repulsion between ions and peptide or lipid atoms.

Throughout the remainder of this chapter, the *z*-direction will refer to the direction along the axis of the channel, perpendicular to the plane of the bilayer. The applied voltage values correspond to the value of the potential on the side of the pore nearest to the termini of the peptides. Letting L_x, L_y and L_z represent the length of the computational domain, we can summarize the above boundary conditions as:

$$\begin{aligned}
&\varphi(x=0,y,z)=\varphi(x=L_x,y,z);\ \varphi(x,y=0,z)=\varphi\big(x,y=L_y,z\big)\\
&\varphi(x,y,z=0)=\ 0\ ;\ \ \varphi(x,y,z=L_z)=V_{applied}\\
&c_i(x=0,y,z)=c_i(x=L_x,y,z)\ ;\ \ c_i(x,y=0,z)=c_i\big(x,y=L_y,z\big)\\
&c_i(x,y,z=0)=c_{i,bulk,trans}\ ;\ \ c_i(x,y,z)=L_z=c_{i,bulk,cis}\\
&\boldsymbol{n}\cdot\boldsymbol{J}_i=0 \text{ along the surface surrounding peptide and lipid atoms}
\end{aligned} \tag{3.5}$$

After the PNP equations are solved, the electrical current across the pore can subsequently be calculated from the solution for the concentrations (c_i) and the electrostatic potential (φ_i) as:

$$I_z = \sum_{i=1}^{N} q_i \iint_{x,y} D_i \left(\frac{\partial c_i}{\partial z} + \frac{q_i e}{k_B T} c_i \frac{\partial \varphi}{dz} \right) dxdy \qquad (3.6)$$

Equation (3.6) above can be applied at any z-position along the pore axis, and shows only minor differences in the computed current values I_z due to numerical inaccuracies. In most cases presented here, these variations are on the order of ~2%.

3.2.2 Validity of PNP equations

Since the characteristic dimensions of a typical transmembrane protegrin pore are commensurate with the size of ions, it behooves us to discuss the validity of a continuum theory in studying such systems. The PNP theory has been shown to overestimate channel conductance in an OmpF porin model by approximately 50% relative to a more rigorous Brownian dynamics (BD) approach [77]. Considering that the narrowest constriction in an OmpF porin channel has an area of 15 $Å^2$, or an effective diameter of ~4.4 Å [78], this agreement is surprisingly good for a theory that treats ions in a purely continuum representation. In another study designed to test the validity of the PNP theory, Corry and coworkers found that conductance values computed by PNP theory and BD simulations become equivalent for cylindrical pores with a diameter of 32 Å, with PNP results improving in charged channels [79]. Additionally, several studies have used PNP theory to model ion conduction through the α-hemolysin channel, which, at its narrowest constriction point, has an open cross-section with a diameter of approximately 15Å [80]. Noskov and coworkers [80] as well as Dyrka and coworkers [81] found that the PNP theory overestimated experimental channel conductance values by approximately 30-60%. Furini and coworkers applied PNP modeling to a KcsA channel, which has a mean diameter of 2.8 Å in one region and 10-14 Å in another [82]. They obtained good agreement with experimental results, albeit by using an adjusted diffusion coefficient. The narrowest constriction point of the protegrin channel that we are interested in has a diameter ranging from 6 - 15 Å, depending on the conformation used.

Considering the range of applications described above that have employed PNP theory and met with reasonable success, we believe the protegrin pore is a good candidate for such a study. More rigorous approaches such as Brownian dynamics or molecular dynamics with applied voltages are far more computationally demanding (particularly the latter), and would not have allowed us to explore an extensive parameter space.

3.2.3 System setup

The charge density ρ_f was obtained by assigning to each atom location a point charge corresponding to the partial charge of that atom in the CHARMM27 force field [83]. Charges were distributed to the nearest grid points using a trilinear interpolation scheme. The ion accessibility region, and the surface for which $\boldsymbol{n} \cdot \boldsymbol{J}_i = 0$ (see equation (3.5)), was defined by excluding all of the space occupied by peptide or lipid atoms. This space consists of a sphere with a radius equal to the sum of the atom's van der Waals radius (as given in the CHARMM27 force field) and the probe radius. For all of the results presented in the current work, a value of 1.4 Å was used for the probe radius, roughly corresponding to the effective radius of a water molecule. Probe radii corresponding to the radii of the ions were also tested and found to have relatively minor effects on conductance characteristics. The dielectric constant profile was obtained by assigning an appropriate value to the space surrounding each atom. Values of ε=2 were assigned to lipid tail atoms, and values of ε=5 were assigned to peptide and lipid head group atoms. Regions not occupied by any of the channel's atoms were assumed to be occupied by water, and assigned a relative dielectric constant of ε=78. A few other schemes of assigning the dielectric constant were also tested, and found to have relatively minor effects on overall conductance characteristics. In all cases, the dielectric constant profile was numerically smoothed with a Gaussian convolution filter of width σ=1.5 Å. This improved numerical stability, and likely corresponds to a more realistic physical situation. Conductance characteristics were found to be insensitive to the choice of the smoothing parameter within a reasonable range.

Four different methods were used to assign the space-dependent diffusion coefficient. The resulting diffusion coefficient profiles were dubbed D1-D4, shown in Figure 3-1. Profile D1 consists simply of assigning the bulk literature values in all regions of the channel for both ions. These are 0.203 $Å^2$/ps for chloride and 0.196 $Å^2$/ps for potassium [84]. Since ion diffusion coefficients are likely to change in confined geometries [85], this profile likely represents a generous upper bound estimate. Profile D2 is based on measuring the diffusion coefficient of chloride and sodium ions from molecular dynamics simulations carried out by Langham et al [76]; however, in these simulations, there were only 2 events of sodium ions crossing the pore, and the sampling was simply not adequate to obtain reliable diffusion coefficient measurements inside the channel. Furthermore, the TIP3P water model [86] along with the CHARMM27 force field [83] used in the simulations is known to overestimate the diffusion coefficient of water by a factor of 2 [87]; as such, direct measurements of the diffusion coefficients of ions is not likely to yield accurate results. Instead, we looked at the approximate scaling of the diffusion coefficients in the MD simulations as compared to the measured bulk value, and applied the same scaling to the literature bulk value to create profile D2 for both ions. This amounts to a smooth decrease in the diffusion coefficient from the bulk value at the entrance of the channel to approximately 1/3 of the bulk value at the channel's centre.

Profile D3 is based on considering hydrodynamic effects of narrow constrictions on spherical particles. Based on the work of Paine and Scherr [88], Noskov et al [80] used the following correlation to calculate the diffusion coefficient of ions as a function of channel radius:

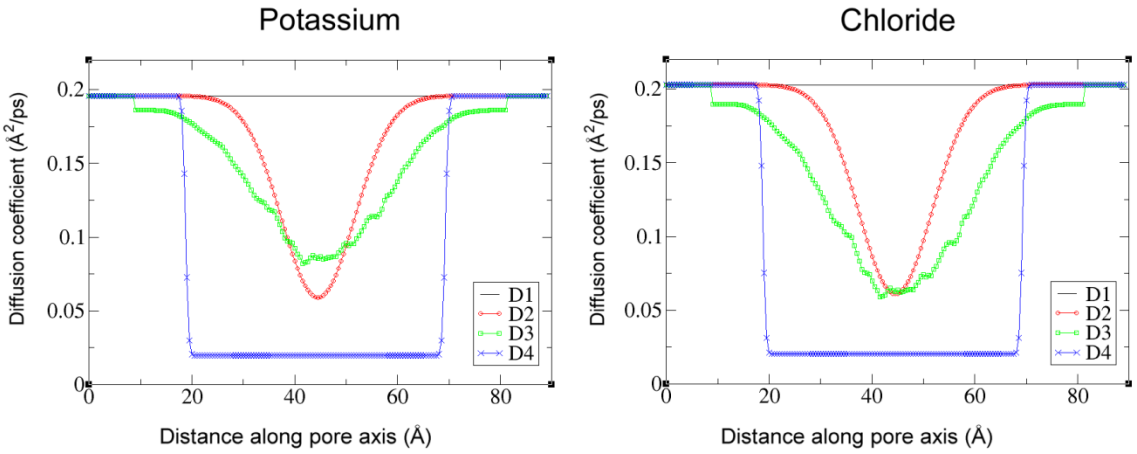
$$D(r) = \frac{D_{ionbulk}}{A + B\exp(\beta/C) + D\exp(\beta/E)} \tag{3.7}$$

Where $\beta = R_{ion}/R_{pore}$, and the empirical fitting parameters have the values given in Noskov et al (A=0.64309, B=0.00044, C=0.06894, D=0.35647, E=0.19409). The ion radii used in our model are the van der Waals radii specified in the CHARMM27 force field (R_{Cl-}

=2.27 Å, R_{K+}=1.764 Å), and the radius of the pore is calculated as an effective radius corresponding to the cross-sectional area found from a grid search. The resulting diffusion coefficient profiles for a snapshot at 93.5 ns are shown in Figure 3-1 (profile D3).

Profile D4 corresponds to setting the diffusion coefficients of both ions to their bulk literature values outside the channel, and to 10% of these values inside the channel. For chloride, this translates to a diffusion coefficient of 0.203 $Å^2$/ps in the bulk region, and 0.0203 $Å^2$/ps in the channel region. For potassium, the diffusion coefficient is 0.196 $Å^2$/ps in the bulk region, and 0.0196 $Å^2$/ps in the channel region. The channel region here is defined as the ion-accessible region bounded at 25 Å above and below the centre of the channel (i.e. $-25 < z < 25$, with z=0 corresponding to the channel centre). Similar scaling was used by Kurnikova et al [89] and Furini et al [82] in their PNP modeling of gramicidin S and KcsA, respectively, and was found to yield good agreement with experimental results. Although 10% of the bulk value likely underestimates the diffusion coefficient of ions inside the channel [90], given the success of this scaling, it may well be that this represents an effective correction of the tendency of PNP theory to overestimate ionic currents due to overestimation of the screening effects [79].

Figure 3-1: Ion diffusion coefficient profiles



3.2.4 Numerical algorithm

The numerical algorithm used in the present work closely follows the finite difference (FD) method developed by Im and Roux [77] based on the algorithm proposed by Kurnikova et al [89]. In fact, the code we employed for the PNP solver is built closely on the main solver module of the 3D-PNP code provided on the website of Dr. Benoit Roux: http://thallium.bsd.uchicago.edu/RouxLab/pbpnp.html

Briefly, the numerical scheme consists of first converting the Nernst-Planck equation (3.3) to a Laplace-like form with the substitution

$$\psi_i = c_i \exp(-q_i e\varphi / k_B T) \tag{3.8}$$

This is known as a Slotboom transformation [91], and converts the Nernst-Planck equations to the form:

$$\begin{aligned} &\nabla \cdot \left(\varepsilon_{eff,i} \nabla \psi_i\right) = 0 \\ &\text{where } \varepsilon_{eff,i} = D_i \exp\left(-q_i e\varphi / k_B T\right) \end{aligned} \tag{3.9}$$

The finite difference representation of equation (3.9) above leads to the following expression for a given point in the finite difference grid (corresponding to the index 0) in terms of its neighbors (corresponding to the index $j = 1,..,6$).

$$\psi_0 = \frac{\sum_{j=1}^{6} \psi_j \varepsilon_{eff,0-j}}{\sum_{j=1}^{6} \varepsilon_{eff,0-j}} \tag{3.10}$$

where ψ_j represents the effective concentration (vis. eq. (3.8)) at neighboring point *j*, and $\varepsilon_{eff,0-j} = (2\varepsilon_{eff,0}\varepsilon_{eff,j})/(\varepsilon_{eff,0}+\varepsilon_{eff,j})$.

The no-flux boundary condition is implemented by setting the value of ε_{eff} to zero for all points in the ion-inaccessible region. This has the same effect as setting the components of the flux perpendicular to the bounding region of the ion-inaccessible surface to zero.

Ion concentrations are also initially set to 0 inside the ion-inaccessible region, and do not change due to the implementation of the no-flux boundary condition.

Similarly, the Poisson equation (3.4) has the following finite-difference representation of the central point in terms of its neighbors:

$$\varphi_0 = \frac{\sum_{j=1}^{6} \varepsilon_{0-j}\varphi_j + \rho_{f,0} + \sum_{i=1}^{2} q_i c_{i,0}}{\sum_{j=1}^{6} \varepsilon_{0-j}} \tag{3.11}$$

Here, the summation over the index j represents summation over neighbouring grid points, while summation over i refers to summation over the ionic species, in this case only potassium and chloride; $\rho_{f,0}$ refers to the fixed charge density, and $\varepsilon_{0-j} = (2\varepsilon_0\varepsilon_j)/(\varepsilon_0+\varepsilon_j)$.

Equations (3.10) and (3.11) above form the basis of successive overrelaxation (SOR) schemes for solving the Nernst-Planck and Poisson equations, respectively. The Poisson equation is solved first, using either zero everywhere or the solution of the linearized Poisson-Boltzmann equation as an initial guess for the potential. The computed electrostatic potential is then used to construct the transformed concentration ψ_i and $\varepsilon_{eff,i}$ profiles, and this alternate form of the Nernst-Planck equation is solved using the FD scheme above (3.10). The ion concentrations are then passed back to the Poisson equation, and a new potential is computed. The process is repeated until convergence of both equations is obtained.

In order to achieve stability of the solver, it is necessary to mix the solution from any given iteration of the Poisson equation with the solution from the previous iteration:

$$\varphi_{new} = \lambda\varphi_{new} + (1-\lambda)\varphi_{old} \tag{3.12}$$

Here, φ_{new} represents the electrostatic potential computed at a given iteration of the Poisson equation, and φ_{old} represents the potential at the previous iteration. The mixing parameter λ is typically set to 0.03, and increased during the iterative process whenever the concentration profiles do not change appreciably from one iteration of the Nernst-Planck equation to the next. Additionally, due to the high charge of the protegrin pore, we occasionally encountered numerical instabilities early on in the solution process. We employed a continuation strategy based on solving the equations to a low tolerance with a scaled down charge density or a high dielectric constant, then gradually increasing the charge density or decreasing the dielectric constant to their real values after achieving convergence at each intermediate value.

Numerical tolerances were typically set to 10^{-14} Å^{-3}, or 1.7×10^{-11} M for the concentrations, and 10^{-8} e^-/Å, or 1.4×10^{-4} mV for the electrostatic potential. The grid resolution used for all runs was 0.5 Å/grid point. A difference of around 17% in the measured current was observed between a resolution of 1.0 Å/grid point and the 0.5 Å/grid point resolution. Although we did not explore finer grid resolutions, we do not expect any significant impact on the results to arise from higher resolution grids. The measured conductance characteristics were found to be insensitive to dimensions larger than 115 × 115 Å in the bilayer plane, and 80 Å in the direction perpendicular to the bilayer (corresponding to a buffer of ~15Å of water on each side of the bilayer patch).

3.3 Results and Discussion

In all of the results presented herein, we have modeled electrodiffusion of ions through a protegrin pore bathed in KCl solutions under a constant applied voltage. This corresponds to the experimental conditions in the work of Sokolov et al [58].

3.3.1 Assessing the validity of the PNP equations in the present case

The solution of the PNP equations yields the nonequilibrium concentration profiles of both ions and the electrostatic potential profile. Figure 3-2 shows a plot of the concentration profiles of both potassium and chloride ions as a function of location along

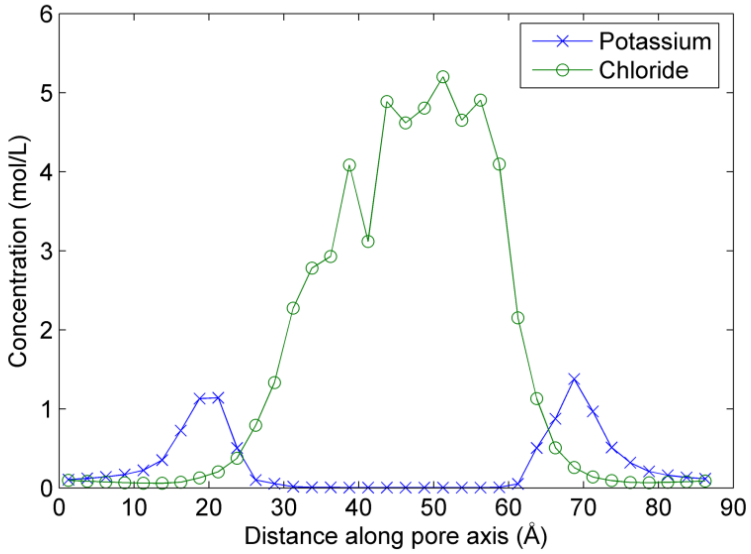
the pore axis for typical model parameters. The values shown represent average concentrations in small planar slabs along the pore axis. Due to the high concentration of positive charge in the pore structure, the concentration of potassium ions inside the pore is extremely low, while the concentration of chloride is extremely high in order to neutralize the arginine charges. Although concentration values as high as 5 mol/L exceed the saturation concentration of KCl, one must keep in mind that these are only local values in small regions, and cannot be strictly interpreted as concentrations in the context of solution thermodynamics. The Nernst-Planck equation can be alternatively cast as a Smoluchowski-type equation, in which case the concentrations represent relative probabilities of finding an ion at a certain location, which is perhaps a more relevant way to interpret these results. However, such high concentrations may also approach the limit of the applicability of PNP theory, in particular with regards to the lack of direct ion-ion correlations and ion-ion van der Waals interactions. We note that the high concentrations of 5 mol/L are still significantly lower than the maximum possible packing of chloride ions (~30 mol/L, based on the density of crystalline NaCl). One interesting area for future work is to expand the model to include activity coefficients based on a hard sphere repulsion model, similar to that discussed in chapter 2. However, this would make the numerical solution of (3.3) and (3.4) significantly more difficult.

The effects of such inaccuracies cannot be easily ascertained. However, we note that previous work [79] criticized the accuracy of PNP models in narrow channels due to the overestimation of screening effects that arises from the continuum approximation, in which an ion is locally screened by its counterion, regardless of geometric restrictions on the presence of both ions. In the case of a pore as highly charged as the protegrin pore, where the concentration of chloride inside the pore is orders of magnitude higher than potassium, such overestimations of screening effects would not arise. Corry and coworkers [79] indeed found that the accuracy of PNP models improved in moderately charged pores. In our system, the extent to which inaccuracies arising from the lack of ion-ion correlations in PNP models are offset by improvements in accuracy that arise

from the presence of only the chloride ion in the channel is not clear. This should be kept in mind as a possible caveat to the applicability of PNP theory to highly charged channels in the discussions that follow.

Figure 3-2: Concentration profiles of potassium and chloride

The values shown are averages over the x and y directions within the ion-accessible region. These profiles correspond to the PNP model applied to a snapshot at 93.5 ns of the NPT segment of the MD simulations of Langham et al, with an applied voltage of -20 mV, 0.1 M KCl bath concentrations, and diffusion coefficient profile D4.

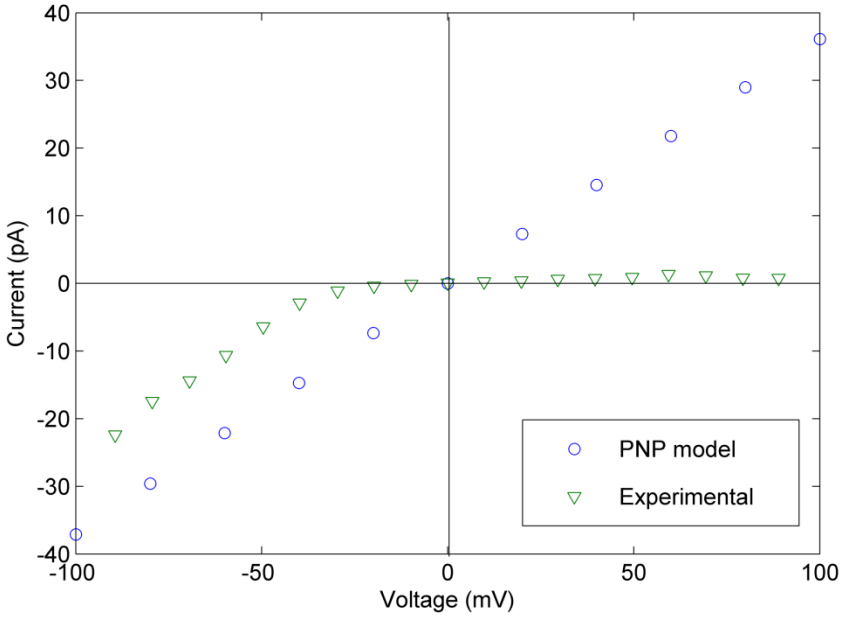


3.3.2 Current-voltage (I-V) relationships and conductance

From the concentration and electrostatic potential profiles, the net current can be obtained using equation (3.6) (see section 3.2.1). Repeating this for multiple voltage values yields a current-voltage (I-V) relationship that can be compared to experimental results. Figure 3-3 shows the I-V curve obtained for a snapshot at 93.5 ns of the NPT segment of the MD simulations of Langham and coworkers [76], along with experimental data from Sokolov and coworkers [58]. In both cases, these data correspond to 100 mM KCl solutions on both sides of the lipid bilayer.

Figure 3-3: Comparison of I-V curves from the PNP model and experiments [58]

All data correspond to a snapshot at 93.5 ns of the NPT segment of the simulations of Langham et al [76], using diffusion coefficient profile D4 (see section 3.2.3). KCl salt bath concentrations were set to 0.1 M on both sides of the membrane in order to match experimental conditions.



The most notable difference in this figure is the lack of voltage-dependent behaviour in our model, whereas the experimental data show a clear increase in conductance at negative voltages. The sign of the voltage in the experiments of Sokolov et al refers to the voltage value on the side opposite (*trans*) of the solution containing protegrin; in our model, the value of the voltage corresponds to the side with the protegrin termini, while virtual ground is defined on the side of the hairpin turns. The strong voltage dependence observed experimentally in Sokolov et al [58] is most likely attributable to major conformational changes of the channel caused by the applied voltage, rather than asymmetry in ion flux due to inherent pore structure asymmetry. At negative *trans* voltages, the channel is likely formed and open, while at low negative voltages and positive *trans* voltages, the channel does not form. This behavior cannot be captured in the PNP model, which uses the same rigid pore structure at all voltages (for a

given snapshot). Given the good agreement between the conductance at negative voltages obtained in the simulation and that in the experiments, we propose that the pore structure used here corresponds closely to the open pore structure formed at negative voltages in the experiments. As mentioned earlier, Mangoni and coworkers [59] found no voltage dependence of the conductance in their experiments with *Xenopus laevis* oocyte membranes treated with PG-1, which supports the hypothesis that the voltage-dependence is a function of membrane composition and the ability of protegrins to form conductive pores, rather than an inherent characteristic of the pore structure.

Sokolov and coworkers [58] report single-channel conductance values in the range of 50-100 pS for PG-1 in a black lipid bilayer, 40-100 pS for PG-3 in the same composition bilayer, and 80-360 pS for PG-3 in bilayers containing either lipopolysaccharide (LPS) or Lipid A, both of which are anionic lipids. The conductance presented above (the slope of the I-V curve in figure 3) is around 367 pS, close to the experimental values observed by Sokolov and coworkers in their anionic bilayers. Even considering the limitations and typical overestimations that sometimes arise from the use of PNP theory, the lower conductance values of 50-100 pS observed in neutral black lipid bilayers do not seem to be compatible with the structure extracted from the equilibrated MD simulations of Langham and coworkers which were conducted in a 3:1 POPE:POPG lipid bilayer [76] (and even less compatible with the structure directly proposed by the NMR experiments of Mani and coworkers [54], as discussed below).

From the experimental evidence discussed, it is clear that the exact membrane composition strongly affects the ability of protegrins to form pores, and likely the nature of the pores themselves. However, for anionic membrane compositions, our data along with the experimental observations discussed seem to suggest that pores are similar in structure across different types of anionic bilayers. This would explain why the pore structure from the equilibrium simulations in reference [76] in a 3:1 POPE:POPG bilayer yields a conductance comparable (within the limits of PNP theory) to what was measured by Sokolov and coworkers at a negative voltage in negatively charged lipid bilayers

composed of black lipids and LPS. Our data suggest that the low conductance values of 50-100 pS observed by Sokolov and coworkers in neutral black lipid membranes likely correspond to an altogether different pore structure, or a different conduction mechanism. The presence of anionic lipids could affect not only the ability of protegrins to form pores, but the types of pores formed, as well as having significant direct effects on ion permeation characteristics. We explore this last point in more detail below by repeating the calculations with the lipid charges turned off.

3.3.3 Ion selectivity

Perhaps the most significant and surprising difference between the experimental findings of Sokolov and coworkers [58] and the results of our model lies in the selectivity of the pore. The anionic current to cationic current ratios (I_{Cl^-}/I_{K^+}) in our simulations are typically on the order of 10^2-10^3, indicating a very strong anion selectivity of the protegrin pore. However, Sokolov and coworkers [58] found that protegrins generally form channels in black lipid membranes that are only 'weakly anion-selective'. The addition of LPS to membranes treated with PG-3 resulted in membranes that were moderately cation-selective. In black lipid membranes, Sokolov and coworkers have reported that the reversal potential in a 10:1 KCl gradient is on the order of 40 mV, which corresponds to a permeability ratio of P_{Cl^-}/P_{K^+} of around 10. In order to facilitate an easier comparison, we carried out simulations with a 10:1 (1.0 M to 0.1 M) KCl bath concentration ratio, and found the reversal potential to be between -49.8 and -61.4 mV, depending on the structure tested. From the Goldman-Hodgin-Katz equation [92], this corresponds to a permeability ratio of P_{Cl^-}/P_{K^+} in the range of 18 to 3200:

$$V_{rev} = \frac{k_B T}{e} \ln \frac{P_K [C]_0 + P_{Cl} [C]_i}{P_K [C]_i + P_{Cl} [C]_0} \tag{3.13}$$

Where $[C]_i$ and $[C]_0$ represent the concentrations of KCl in the baths surrounding the membrane, and P_K and P_{Cl} represent the permeabilities of K^+ and Cl^-, respectively. This

can be re-arranged to solve for the permeability ratio as a function of the salt concentration gradient and the reversal voltage:

$$\frac{P_K}{P_{Cl}} = \frac{\frac{[C]_i}{[C]_0} - \exp\left(\frac{eV_{rev}}{k_B T}\right)}{\frac{[C]_i}{[C]_0}\exp\left(\frac{eV_{rev}}{k_B T}\right) - 1} \tag{3.14}$$

The high anion selectivity that we obtain from our model reflects the intuitive notion that a highly cationic protegrin pore would almost completely reject cations (with a charge of +7 on each peptide, the total charge of the protegrin pore is +56). The high anion selectivity is also corroborated by the simulations of Langham and coworkers [76], where only two events of sodium ions crossing the pore are observed in 150 ns, as compared to 55 events of chloride ions crossing (similar data were obtained for a simulation with KCl; unpublished results). This discrepancy between the PNP model and the experimental data cannot be explained simply based on differences in membrane composition. The much higher anion selectivity in our model may be due to a shortcoming of PNP theory in highly charged channels. However, it might also be the case that the lipids, rather than being almost aligned with the pore axis as they are in our model, orient in a perpendicular fashion to line the pore and form what is known as a toroidal pore. Recent NMR experiments suggest that this might indeed be the case [55-57]. In a toroidal pore, there would be a much higher likelihood of arginine sidechains forming persistent guanidinium-phosphate bonds with the surrounding lipid phosphate groups, which would partially neutralize the charged arginine moieties. To test this scenario, we also solved the PNP system for two snapshots from an MD simulation of a toroidal pore (unpublished results; note that this is not the simulation discussed in chapter 5). As shown in, the high anion selectivity persists even in the toroidal pore models, ruling out this possible explanation for the discrepancy in selectivity.

3.3.4 Effects of variations in pore structure

The results of the PNP model solved for various structures are summarized in Table 3-1. All data except the reversal potential (see section 3.3.3 below) correspond to a symmetric 100 mM KCl solution with an applied voltage of -20 mV.

Table 3-1: Conductance characteristics measured for a variety of structures.

The nanosecond values in the structure description column indicate the simulation time from the start of the constant pressure and temperature (NPT) segment of the MD simulations of Langham and coworkers [76]

Structure description	Effective diameter of narrowest constriction* (Å)	Average effective diameter of channel (Å)	Total conductance (pS)	Current ratio (I_{Cl-}/I_{K+}) at -20mV	Reversal potential in a 10:1 KCl gradient
60 ns	11.3	28.9	427	374	-57.4
65 ns	7.51	26.9	361	2640	-60.7
70 ns	6.36	27.0	384	2660	-49.8
80 ns	5.97	23.7	286	6860	-61.4
81.5 ns	10.0	25.2	405	585	-60.2
85.5 ns	7.92	24.6	361	953	-60.1
87.5 ns	8.11	25.8	364	731	-59.0
91.5 ns	7.80	25.4	387	564	-58.7
93.5 ns	9.44	24.4	367	338	-58.7
NMR structure from [54]	14.5	31.5	679	389	-57.8
Toroidal pore after heating	12.8	33.2	664	330	-59.2
Toroidal pore after 50 ns	12.7	35.8	479	333	-56.6

*Effective diameters are based on a grid search of the ion-accessible area, as determined using a probe radius of 1.4Å (see Methods section for further details).

None of the structures described showed any significant voltage-dependence in an applied voltage range of ±100 mV. The conductance values obtained based on structures extracted from the molecular dynamics simulations are generally in good agreement with the experimentally observed values in anionic lipid bilayers (60-360 pS). The diameter of the smallest channel constriction in these structures ranges from 6-10 Å, which is approaching the limit of the applicability of PNP modeling (see discussion in section 3.2.2), but nonetheless sufficiently large to yield meaningful results. The variations that exist among different snapshots are the results of minor conformational changes caused by thermal fluctuations, in particular the motion of side chains into and out of the pore opening. Although these effects are noticeable, they are not drastic, and not as large as the differences between the MD structures from 60 – 93.5 ns and the other three structures. It is also worth noting that the diameter of the narrowest constriction, while it does somewhat correlate with conductance, is not the sole factor influencing it. For instance, the snapshot at 93.5 ns has a larger constriction diameter than either the 85.5 or 91.5 ns snapshots by almost 2 Å, and yet shows the same conductance as the 85.5 ns snapshot, and a lower conductance than the 91.5 ns snapshot. Thus, while steric effects clearly affect ion permeation, there are other conformational changes that have significant influence on pore conductance. These are likely electrostatic effects that arise from conformational changes of the charged side chains (e.g. arginine). These are explored in further detail below.

The results for the last three structures shown in Table 3-1 correspond most closely to the structure proposed by Mani and coworkers [54], which was used to start the MD simulations. This structure has a larger opening than the molecular dynamics simulations, and yields much higher conductance values. Even considering the inaccuracies in our model, these values are not in good agreement with experimentally measured values, which may suggest that the corresponding structures are not the dominant conformers. Since the NMR experiments do not directly resolve the positions of all side chains, we posit that the narrower opening observed in molecular dynamics

simulations, which is primarily due to the motion of side chains into the interior of the pore, may in fact represent a more realistic structure. However, as discussed below, there are several other considerations that can significantly affect the conductance values obtained from PNP modeling.

3.3.5 Influence of different charged groups

In order to quantify the relative importance of certain charged moieties within the structure, we analyzed the effects of removing the positive charges of the arginine residues of the peptides in one case, the negative charges in the POPG lipids in another case, and both in a third case, all the while keeping the rest of the system unchanged. The results are summarized in Table 3-2.

Table 3-2: Effects of various charged moieties in the system on conductance behaviour.

Model	Total conductance (pS)	Current ratio (I_{Cl^-}/I_{K^+}) at-20 mV
All charges present	286	6860
All positive arginine charges removed	34.5	0.030
Negative POPG charges removed	328	8480
Both positive arginine charges and negative POPG charges removed	16.5	0.17
Arg-1 positive charge removed	332	289
Arg-4 positive charge removed	204	94
Arg-9 positive charge removed	225	172
Arg-10 positive charge removed	260	244
Arg-11 positive charge removed	271	252
Arg-18 positive charge removed	231	135

As shown in Table 3-2, the positively charged arginine residues, which impart the pore a total charge of +56, are by far the dominant factor influencing conductance behaviour. Removal of all of the arginine charges decreases conductance by a full order of magnitude, and changes the selectivity of the pore from strongly anionic to cationic. If the POPG charges are also removed, the selectivity becomes only slightly cationic, and the total conductance is even lower. Turning off the charges in the POPG bilayer while retaining the arginine charges does not result in a significant difference in the total conductance or the current ratio. This suggests that in the presence of charged arginine residues, charges in the lipid bilayer do not have an appreciable direct effect on conductance characteristics in a formed pore, as the former clearly dominate all electrostatic behavior (however, the POPG charges clearly have a significant impact on the ability of the peptides to form a pore in the first place, as discussed in section 0). Although complete removal of the arginine charges would not happen in any realistic physical situation, the results in this section show what a drastic effect these charges have.

In order to investigate the effects of individual arginine residues, we solved the PNP PNP model for six additional situations, each corresponding to turning off the positive charges on the arginine residues for all peptides at a particular position. The calculations discussed were all performed on the structure corresponding to a snapshot at 93.5 ns of the NPT segment of the MD simulations of Langham and coworkers [76], using diffusion coefficient profile D4. As shown in Table 3-2

Table 3-2, the most significant effect is attained by turning off the charge at arginine position 4, which results in the largest drop in the conductance as well as the selectivity. This corroborates the hypothesis that the arginine residues at position 4 interact more strongly with the pore interior, likely due to their physical location within the pore. As such, while the other arginine residues likely play a larger role in peptide insertion, it would seem that the purpose of arginine-4 is primarily to effect ion passage through the pore by attracting chloride ions. It is also noteworthy that Arg-4 is the residue

that is absent in the experiments of Sokolov et al [58], who used PG-3 instead of PG-1, which may explain the differences in pore selectivity observed herein. However, it should be noted that the data in Table 3-2 represent only the conductance characteristics of one particular snapshot. Nonetheless, considering the consistency with the findings discussed in the next section, this appears to be a worthwhile hypothesis that warrants further future investigation.

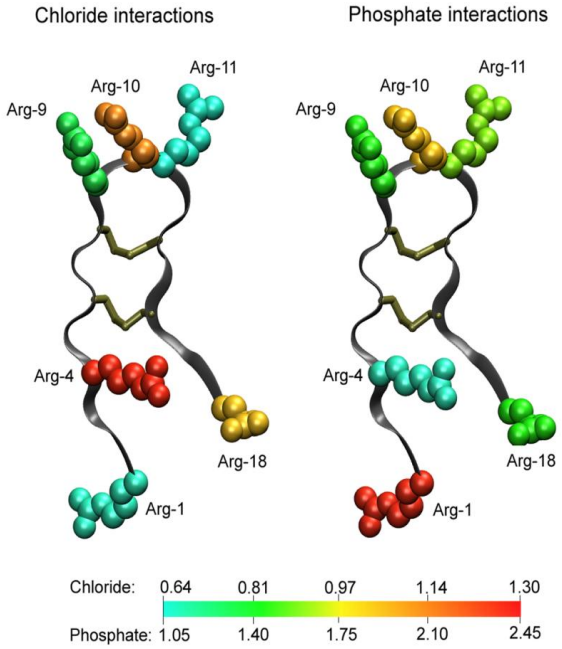
3.3.6 Specific interactions of arginine side chains

In order to explain the conductance behavior observed in our model in terms of more specific structural features, we have analyzed the molecular dynamics simulations of Langham and coworkers [76] in greater detail. The high selectivity of the pore for anions that we observe in PNP calculations is likely due to electrostatic attraction between the positively charged guanidinium groups in the arginine side chains and negatively charged chloride ions. However, competing with this interaction is the attraction of arginine side chains to negatively charged phosphate groups in the lipid head groups. In order to investigate the relative importance of these interactions, we have computed radial distribution functions for both chloride and phosphate groups from the ζ-carbon of all arginine side chains. For brevity, the resulting 48 plots are not included. Instead, we confine our discussion to averaged data for each arginine position, which is justified by the axial symmetry of the pore and the equivalence of all eight peptides (except perhaps with respect to dimer pairings).

Figure **3-4** shows the number of chloride and phosphate atoms within 7.5 Å of the arginine ζ-carbon at each position, averaged over time as well as over all eight peptides. The value of 7.5 Å was selected to include the first large peak in all of the radial distribution functions.

Figure 3-4: Interactions of arginine residues with phosphate groups and chloride ions

The numbers on the colour scales correspond to the number of phosphate groups/chloride ions within 7.5 Å of the ζ-carbon of the respective arginine residue. Data are averaged over the last 50 ns of the molecular dynamics simulation, as well as over the eight peptides.



As shown, arginine side chains at all positions interact fairly strongly with both phosphate groups and chloride ions. Arginine at position 1 exhibits a much stronger interaction with phosphate groups, indicated by an average value of 2.45 phosphate groups, as compared to 0.64 chloride ions. Also noteworthy is the difference at position 11, where phosphate interactions are again significantly stronger than chloride interactions.

Another notable feature in

Figure **3-4** is the stronger interaction of arginine at position 4 with chloride as compared to the other arginine residues. Although the difference is not drastic (1.31 chloride ions for position 4 compared to 0.64 for position 1, or 1.14 for position 10), this may represent a significant electrostatic feature of the protegrin pore, particularly considering that the data represent an average over all eight peptides. The arginine side chain at position 4 is in a favorable location to be interacting with the aqueous pore interior, where it can have the largest effects on conduction characteristics. Considering the location of the other arginine residues near the termini and the β-sheet turn, as well as the discussion of phosphate interactions in the preceding paragraph, one could hypothesize that the purpose of arginine at position 4 is primarily to attract anions through the pore. In contrast, the remaining arginine side chains interact more strongly with lipid head groups in order to facilitate peptide insertion and stabilize the pore structure. This is consistent with the fact that arginine residues at position 4 have the weakest interactions with phosphate groups.

3.3.7 Effects of the diffusion coefficient profiles

One of the most important parameters required for successful PNP modeling is the space-dependent diffusion coefficient of all diffusing ionic species. The diffusion coefficient is often empirically adjusted to account for the approximate nature of the PNP theory. The value of the diffusion coefficient is expected to be different from the bulk literature values in the constricted molecular geometry of the pore interior [85, 90]. Unfortunately, the diffusion coefficients inside the pore are also one of the most difficult parameters to assign. As far as we are aware, no experimental techniques are capable of measuring the diffusion coefficient of ions within a molecular pore to a high accuracy and with sufficiently high spatial resolution; certainly no such data are available for the protegrin pore. We have therefore used four different empirically-derived diffusion coefficient profiles, denoted as D1-D4, as discussed in section 3.2.3 and illustrated in Figure 3-1. The resulting conductance characteristics obtained using the pore structure at 93.5 ns are tabulated in Table 3-3. The rationale for each diffusion coefficient profile is discussed in greater detail in the methods section.

Table 3-3: Conductance characteristics as a function of diffusion coefficient profile.

Diffusion coefficient profile	Total conductance (pS)	Current ratio (I_{Cl-}/I_{K+}) at -20mV
D1 - bulk values everywhere	2310	220
D2 - scaled from bulk values as per MD simulation results	1550	314
D3 - hydrodynamic model based on channel radius profile (based on [80, 88])	1280	242
D4 - 10% of bulk values everywhere inside the channel	367	564

In all cases, the diffusion coefficient was assumed to be isotropic and invariant in the *x* and *y* directions (in the plane of the bilayer). The PNP system was solved for all four diffusion coefficient profiles (D1-D4) using the structure at 80 ns of the MD simulation. None of the cases tested showed any significant voltage-dependence for applied voltages in the range of +/- 100 mV. As already mentioned, the experimentally reported conductance is 50-100 pS in black lipid membranes and 60-360 pS in anionic membranes [58]. Clearly, diffusion coefficient profiles D1 and D2 result in unacceptably high conductance values, indicating that the diffusion coefficient inside the channel should be significantly lower. Using the hydrodynamic model of [88], which was also employed successfully in references [80-81] in PNP models of α-hemolysin channels, yields a slightly lower value, but still outside of the range of the experimentally determined conductance. The diffusion coefficient profile D4 yields the best agreement with experimental conductance data. Although setting the diffusion coefficient to 10% of the bulk value may seem a low estimate, similar scaling was used by Furini and coworkers [82] and Kurnikova and coworkers [89] in successfully modeling KcsA channels and Gramicidin A channels, respectively. Considering the differences between our model and the experimental systems, variations due to minor structure fluctuations, as

well as the approximate nature of our model, profile D4 appears to be a reasonable estimate for the diffusion coefficients, and can be treated as an effective fitting parameter.

3.4 Conclusions

We have applied the PNP theory to predict the conductance characteristics of several variations of a protegrin pore structure. We found that the best agreement with the experimentally measured net conductance was obtained using structures extracted from the latter part of molecular dynamics simulations of Langham and coworkers [76]. Even considering variations in structures due to thermal fluctuations, all of these structures yielded conductance values near the experimental upper bound of 360 pS. The structure of the pore from NMR experiments has a larger opening at its narrowest constriction point than the structures from MD simulations, primarily due to differences in the orientation of amino acid side chains in the pore interior. PNP models based on this NMR structure yielded conductance values that are significantly higher than the experimentally measured values, even considering the expected typical errors arising from PNP theory. This suggests that the narrower structures observed in molecular dynamics simulations may be more realistic. In all of the structures tested, we observed an extremely high anion selectivity, which is to be expected for such a cationic pore, but surprisingly disagrees with the experimental findings of Sokolov and coworkers [58]. No significant voltage dependence of the conductance was observed in any of our models, corroborating the hypothesis that the voltage-dependence observed by Sokolov and coworkers [58] is a result of voltage-dependent channel formation, rather than an inherent feature of the protegrin pore structure. Due to the sensitivity of PNP calculations to the diffusion coefficient profile, we have tested several alternatives, and found that a diffusion coefficient set to 10% of the bulk value inside the channel yielded the best agreement with experimental data. Even with these and other limitations of PNP theory in mind, we believe the general trends observed herein form a useful connection between the structural features and conductance characteristics of protegrin pores. This work has been published in a similar form in reference [93].

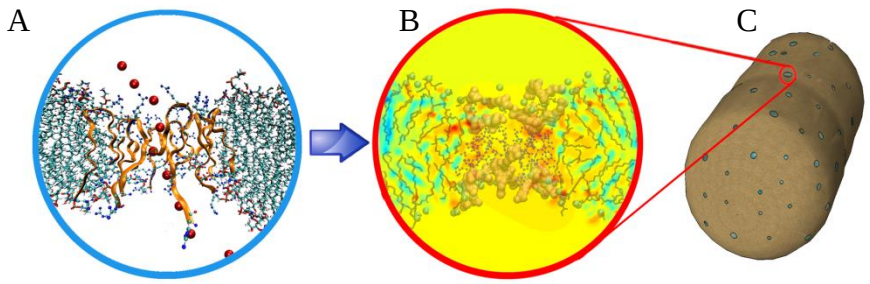
4 Modeling the time-dependent effects of protegrin pores on bacterial cells

4.1 Introduction

Encouraged by our success in modeling single-pore conductivity discussed in the previous chapter, we set out to determine the effects of multiple protegrin pores on an entire bacterial cell with regards to ion transport. The work described in this chapter was the result of a collaboration with the group of Dr. Robert Lehrer at UCLA, and has been published in a similar form in the journal Peptides [38]. Experimental work was carried out by Dr. Ehsan Hazrati, and is described herein in its entirety as it is essential to the present work. Through a combination of experimental and theoretical techniques, we investigate the effect of protegrin pore formation on *E. coli* cultures. We investigate how molecular level events result in antimicrobial function, focusing in particular on potassium efflux from a bacterial cell, transmembrane potential decay and cell volume changes. While many authors have suggested that membrane permeabilization by AMPs is an important part of the bacterial killing mechanism [9, 19, 25], the highly coupled effects of such permeabilization have not been quantified. Furthermore, no attempts have been made that we are aware of to establish a quantitative link between molecular knowledge of AMP structure and physiological effects measured on whole bacteria. We develop a model that uses single-pore permeabilities obtained from the PNP models discussed in chapter 3 as inputs to a larger scale, transient model of ion electrodiffusion from an entire cell. We also account for water transport and changing cell volume, which is a result of the osmotic pressure differences that arise as ions are exchanged between the cell interior and its environment. Our multiscale approach is illustrated graphically in Figure 4-1, and is described in greater detail in the section 4.2 below.

Figure 4-1: The multiscale modeling approach.

Atomistic molecular dynamics simulations (panel A) provide the high-resolution structures and charge distributions needed for continuum, nonequilibrium electrodiffusion modeling (see chapter 3) of the protegrin pore (panel B). Ion permeabilities are calculated based on these results, and transient ion transport is modeled for entire bacterial cells (panel C), allowing us to estimate the total number of pores



.

4.2 Methods

4.2.1 Experimental methods and materials

As already mentioned, the experimental work herein was performed primarily by Dr. Ehsan Hazrati in the laboratory of Dr. Robert Lehrer at UCLA. The details of the experimental work were provided to us by him, and are included here for completeness.

Synthetic PG-1 (RGGRLCYCRRRFCVCVGR-amide), ~96.5% pure, was purchased from the SynPep Corporation (Dublin, CA), diSC3-5 (3,5-dipropylthiacarbocyanine) was purchased from Molecular Probes (Eugene, OR) and diisodecyl phthalate and dibutyl phthalate were purchased from Pfaltz and Bauer (Waterbury, CT) and from Eastman Kodak (Rochester, NY), respectively. All other chemicals were purchased from Sigma. *E. coli* strain ML-35p [94-95] was used in all experiments. Light scattering assays were performed in 96 well, flat-bottom polystyrene tissue culture plates, with a final volume of 200 µL/well. In all cases, peptide or an equal

volume of buffer was added to washed *E. coli*, final concentration of 10^8 colony-forming units (CFU) /ml, in a solution that contained 100 mM NaCl and 10 mM sodium phosphate (pH 7.4). Optical density measurements at 350 nm were taken every 6 seconds with a SpectraMax 250 spectrophotometer (Molecular Devices, Sunnyvale, CA). Unless otherwise noted, all experiments were performed at 37°C.

In order to confirm the relationship between optical density changes and volume changes, we performed ^3H_2O scintillation counts to determine the bacterial water content following protegrin exposure. Stationary-phase *E. coli* ML-35p was washed three times with the usual buffer and resuspended in the same buffer containing ^3H_2O. Protegrin or an equal volume of its vehicle was added and the mixtures were incubated for 12 or 30 min at 37 ^oC. Then, the bacterial suspensions were layered over 0.3 ml of a cushion composed of a 5:3 (v/v) mixture of dibutyl phthalate:diisodecyl phthalate and centrifuged at 14,000 × g for 2 minutes. Because the phthalates are immiscible with water, this procedure rapidly separated the bacteria from the medium. After carefully removing the supernatant and phthalate mixture, the base of the tube was amputated with a razor blade and the stub, which contained the bacteria, was recovered for liquid scintillation counting.

Colony count assays were performed to determine the effects of protegrin on bacterial viability. *E. coli* ML-35p cells were washed with buffer and incubated at 37 °C for 10 minutes (exponential phase) or 60 minutes (stationary phase) with the desired concentration of protegrin or an equivalent volume of buffer . Aliquots were removed at intervals, and subjected to two sequential 100-fold serial dilutions with protegrin-free incubation medium. After further 1:50 dilution into 5 ml of a 43 ^oC solution of 1 % agarose in 100 mM NaCl, 10 mM sodium phosphate and 1% trypticase soy broth, the entire 5 ml sample was poured over a conventional trypticase soy agar underlay, and colonies in the overlay were counted after an overnight incubation.

A previously described procedure [96] was used to monitor potassium loss from *E. coli* ML 35p cells. In brief, a potassium-selective electrode was used to monitor the conductance of a solution containing 10^8 CFU/mL in the medium described above and suspended in a 250μL incubation tube. The K^+ concentration was recorded every second. The total bacterial K^+ content ($c_{K+,total}$) was measured after subjecting the bacteria to prolonged (3 min) sonication. All data are reported as the percent of the total K^+ released.

The relative transmembrane potential decay was measured using a general procedure recently described by Wu et al [32] using a self-quenching fluorescent dye, diSC3-5. Cells were incubated with 0.4 μM diSC3-5 for 30 minutes in the same buffer described above. After the fluorescence had stabilized at about 10% of its initial value, PG-1 (final concentration, 25 μg/ml) or its vehicle was added and fluorescence was monitored every 8 seconds using excitation and emission wavelengths of 622 and 670 nm, respectively.

Scanning electron microscopy (TEM) was carried out to inspect the effects of protegrin on *E. coli* ML-35p. Bacteria were washed twice in the usual buffer and then incubated at a concentration of 10^8 CFU/mL with a PG-1 concentration of 25 μg/mL. The organisms were fixed by adding fresh glutaraldehyde (2% final concentration) and allowed to adhere to mixed cellulose ester membrane filters (Millipore VCWP01300) for 20 minutes. The glutaraldehyde-fixed samples were then washed in phosphate buffered saline, postfixed in 1% osmium tetroxide for 45 minutes at room temperature, dehydrated through ethanol in propylene oxide, embedded in Epon 812 (Electron Microscopy Services, Hatfield, PA), thin-sectioned, and stained with uranyl acetate and lead citrate.

4.2.2 Modeling methods

The model system consists of a representative bacterial cell separated from the bathing solution by a thin membrane region. The volume of the entire system, bacterium and bath, is set to a constant value of 10^{-14} m^3, based on experimental conditions. Since

diffusion of ions is limited by the membrane, the two compartments can both be considered well-mixed volumes. Mass balances on the diffusing ionic species and a volume balance yield the following set of ordinary differential equations:

$$\frac{dn_i}{dt} = -j_i A \qquad i = Na^+, K^+, Cl^- \tag{4.1}$$

$$\frac{dV_{aq}}{dt} = -J_V A \tag{4.2}$$

where n_i is the number of moles of species i inside the cell, j_i is the flux of species i , J_V is the volume flux, A is the cell membrane area and V_{aq} is the cell aqueous volume.

The volume flux is driven by the pressure difference and the osmotic pressure difference across the membrane [97]:

$$J_V = -L_P\left(\Delta P - RT\sum_i \sigma_i\left(c_i^o - c_i^i\right) + RTc_f\right) \tag{4.3}$$

Here, L_P is the membrane hydraulic permeability, $\Delta P = P_{outside} - P_{inside}$ is the pressure difference across the membrane, R is the gas constant, T is the temperature, σ_i is the reflection coefficient (a measure of the membrane ideality; see [97-98]), c_i^o and c_i^i represent the concentration of ionic species i outside and inside of the cell, respectively, and c_f represents the concentration of osmotically active but impermeable solutes present in the cell interior. The summation is over the three ionic species considered, namely K^+, Na^+ and Cl^-. The model assumes that these are the only relevant osmolytes that cross the cell membrane as a result of PG-1 permeabilization. Intracellular osmolytes such as nucleic acids and proteins that do not readily cross the membrane are accounted for through c_f. Assuming that 30% of the total bacterial volume consists of solids, the total cell volume is found according to:

$$V_c = V_{aq} + V_S = V_{aq} + \frac{0.3V_{aq,t=0}}{0.7} \tag{4.4}$$

where V_{aq} is the aqueous cell volume.

To calculate the ion fluxes, we have used the integrated form of the one-dimensional Nernst-Planck equation under the assumption of a spatially constant field, which has been shown to yield good results even in cases where this is not necessarily the case [99]:

$$j_i = P_i q_i u \frac{\left(c_i^o - c_i^i e^{q_i u}\right)}{1 - e^{q_i u}} \qquad i = Na^+, Cl^- \tag{4.5}$$

where j_i is the ion flux in mol/m^2/s, q_i is the ion charge, P_i is the (effective) ion permeability, and $u = FV/RT$, where V is the value of the electrostatic potential in the cell interior and F is Faraday's constant. We have also added an active potassium uptake term based on available kinetic data for the four main K^+ uptake systems in *E. coli* (Kdp, TrkA, TrkD and TrkF) [100]:

$$\begin{aligned} j_{K^+} = P_{K^+} u \frac{c_{K^+}^o - c_{K^+}^i e^u}{1 - e^u} - \Big(&\frac{V_{max,Kdp} c_{K^+}^o}{K_{m,Kdp} + c_{K^+}^o} + \frac{V_{max,TrkA} c_{K^+}^o}{K_{m,TrkA} + c_{K^+}^o} \\ &+ \frac{V_{max,TrkD} c_{K^+}^o}{K_{m,TrkD} + c_{K^+}^o} + K_{m,TrkF} \end{aligned} \tag{4.6}$$

The kinetic parameters V_{max} and K_m for all three uptake systems were obtained from Rhoads et al [100].

The transmembrane potential is found by imposing a zero-current restriction (required to maintain bath electroneutrality in both the cell interior and the surrounding bath):

$$\sum_i q_i j_i = 0 \tag{4.7}$$

Due to the active K^+ uptake terms, equation (4.7) is solved numerically at each time step to obtain *V*.

The pressure difference across the membrane, ΔP, is calculated as a simple linear relation with the volume:

$$\Delta P = \Delta P_0 + \frac{m(V_c - V_{c,0})}{V_{c,0}} \tag{4.8}$$

where ΔP_0 is the initial pressure difference across the membrane, m is an effective volume modulus, V_c is the cell volume, and $V_{c,0}$ is the initial cell volume. Although the mechanics of the peptidoglycan sacculus are known to be much more complex [101-103], this particular approximation is sufficient for the present model, and has been employed by other authors in similar modeling [104].

4.2.2.1 Ion permeabilities, P_i.

In order to estimate the permeabilities of ions through the bacterial membrane region (P_i in equations (4.5) and (4.6)), it is assumed that the most significant resistance to ion diffusion arises from the inner cytoplasmic membrane (a safe assumption given the inherently high permeabilities of the periplasmic space and outer bacterial membrane). Permeabilities are then calculated assuming parallel passive diffusion through pores and the bilayer region:

$$P_i = (1 - \chi)P_{i,bilayer} + \chi P_{i,pore} \tag{4.9}$$

Here, χ is the fraction of the membrane area covered by protegrin pores. It is related to the number of pores, N_{pores}, according to: $\chi = N_{pores} \times A_{pore}/A_{cell}$, where A_{pore} is

the total area of the membrane that can be designated as being occupied by the pore. This was estimated to be 1120 $Å^2$ according to structures extracted from the molecular dynamics simulations of Langham et al [76].

The effective permeabilities of ions through the pore, $P_{i,pore}$, are estimated based on the numerical solution of the 3D Poisson-Nernst-Planck (PNP) equations for the structure of a single protegrin pore embedded in an anionic lipid bilayer, discussed in chapter 3. The total flow rate of each ion through the pore, F_i, was calculated from the 3D PNP solution and substituted in the integrated 1-D Nernst-Planck equation to obtain the effective ionic permeabilities:

$$P_{i,pore} = \frac{\left(\frac{n_i}{A_{pore}}\right)}{\frac{q_i FV}{RT}\frac{c_i^o - c_i^i e^{\frac{q_i FV}{RT}}}{1 - e^{\frac{q_i FV}{RT}}}} \tag{4.10}$$

The boundary concentrations of a particular species, c_i^o and c_i^i, are the boundary conditions used at the top and bottom of the pore in the 3D PNP model, while the transmembrane potential V corresponds to the boundary value of the electrostatic potential. The 3D-PNP model was originally solved only for potassium and chloride ions. Since the only distinction in the 3D-PNP model between potassium and sodium is the diffusion coefficient, the calculations were not repeated for sodium. Instead, the permeability of sodium ($P_{Na+,pore}$) was calculated by scaling the permeability of the potassium ion with the ratio of the diffusion coefficients of Na^+ and K^+.

The permeabilities of ions through the bilayer region, $P_{i,bilayer}$, were estimated based on literature values [105-106]. For K^+, this corresponds to approximately 3.0×10^{-14} m/s, for Na^+ 1.0×10^{-13} m/s, and for Cl^- 2.3×10^{-11} m/s.

4.2.2.2 Hydraulic permeability, L_P.

As in the case of the ion permeabilities, the hydraulic permeability is calculated on the assumption of parallel flow through the pore and bilayer regions of the membrane.

$$L_P = (1 - \chi)L_{P,bilayer} + \chi L_{P,pore} \tag{4.11}$$

The hydraulic permeability through the pore, $L_{P,\ pore}$, was calculated based on three-dimensional pore structures obtained from the molecular dynamics simulations of Langham and coworkers [76], assuming Poisseuille flow in series along the pore axis [107-108]. Values of $L_{P,pore}$ obtained this way do not vary significantly with the pore structure used; a typical value, which was used for all subsequent calculations, was 4.1×10^{-9} m^3/N/s. Literature values for the value of $L_{P,\ bilayer}$ range significantly depending on the type of bilayer used in the measurements [105, 108]. We have used a value of L_P=7.0×10^{-14} m^3/N/s, well within the range of the relevant literature values.

4.2.2.3 Initial conditions and other parameters.

The initial aqueous cell volume, $V_{aq,t=0}$, was set to a value of 4.18×10^{-18} m^3, based on the mean volume of cellular water measured in exponential-phase *E. coli* using ^3H_2O measurements. Similarly, the initial K^+ intracellular concentration was set to 137.9 mM, based on the total concentration of K^+ measured after sonication of a 10^8 CFU/mL bacterial suspension. The initial concentrations of Na^+ and Cl^- in the cell interior were set to 6mM and 5mM respectively, based on values obtained from the CyberCell Database [109]. This implicitly assumes the existence of charged structures (proteins, nucleic acids, etc.) in the cell interior to achieve electroneutrality; see also discussion in section 4.2.2. The initial pressure difference across the bacterial cell wall was taken to be 4 bar [109-110]. The effective volume modulus of the bacterial envelope, *m* in equation (4.8), was taken to be 9 bars. [110]. The parameter c_f was set so that the initial conditions yield an osmotic equilibrium. The values of all reflection coefficients σ_i were set to unity,

justified because the vast majority of the membrane (the lipid bilayer region) is practically impermeable to ions in comparison to the permeability of the pores.

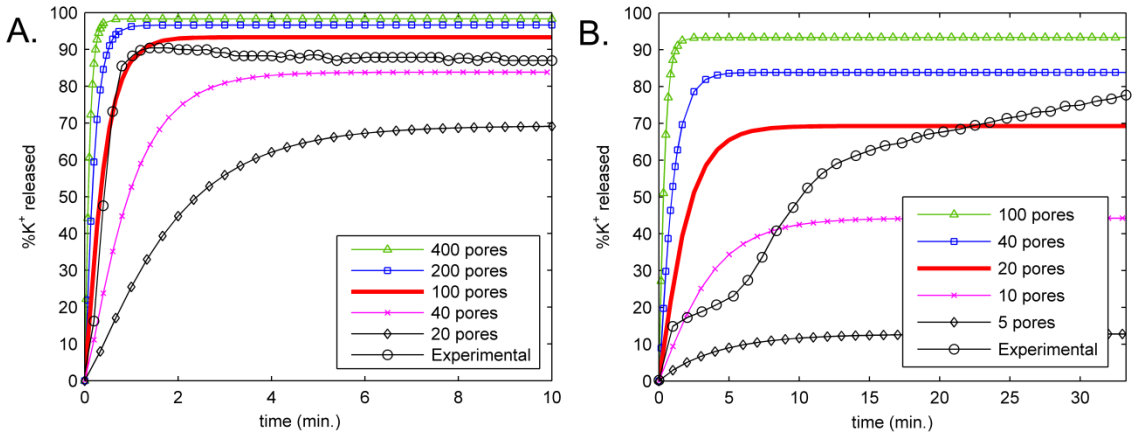
4.3 Results and Discussion

4.3.1 Ion diffusion and potassium release.

Exposure of bacteria to 25 μg/ml of PG-1 caused a rapid release of potassium from both exponential phase (Figure 4-2 A) and stationary phase *E. coli* (Figure 4-2 B). In both cases, potassium release began within a few seconds and continued until about 90% of the total bacterial K^+ was lost. The K^+ release curve for exponential phase bacteria can be closely matched by the model by setting the number of protegrin pores to approximately 100; the slower release observed in stationary phase organisms is more closely matched with 10-20 pores. It should be noted that these and all other values of the number of pores mentioned always represent an effective, average number of open pores. In reality, the pores are transient structures that likely form and collapse frequently on the timescale of the experimental observations. The data in correspond to a typical choice of ion permeabilities through pores obtained based on the 3D-PNP calculations ($P_{K+,pore}$=3.06×10^{-4}, $P_{Na+,pore}$=2.34×10^{-4}, $P_{Cl-,pore}$=0.0906 m/s), as described in the Methods section. Although some variation exists in these parameters depending on the bath boundary concentrations and the particular snap shot from the molecular dynamics simulations that is used to obtain the detailed pore structure, the results are not substantially affected. The variation in the number of pores is between 40 and 200 in the case of exponential-phase organisms for all of the 3D-PNP conditions and structures tested.

Figure 4-2: Potassium release curves

Exponential phase (A) and stationary phase (B) *E. coli*. The figure shows release curves obtained from the model for several values of the number of pores, along with the experimental data for both growth stages.



Both the experiment with exponential phase bacteria and the model show a plateau in the K^+ release curve after approximately 90% of the bacterial K^+ has been released (Figure 4-2 A), less than two minutes after protegrin treatment. In the model, this plateau occurs due to the existence of the active uptake mechanisms (see equation (4.6)). As K^+ is released, the driving force for flux out of the cell diminishes, and eventually the K^+ uptake mechanisms are able to match the outward electrodiffusive flux of K^+, corresponding to the observed steady state concentration values. Since the electrodiffusive flux takes place primarily through the pores, the height of the plateau decreases with a decreasing number of pores.

The potassium release curve for stationary phase bacteria cannot be matched exactly by the model. The experimental curve shows a two-stage release process (see

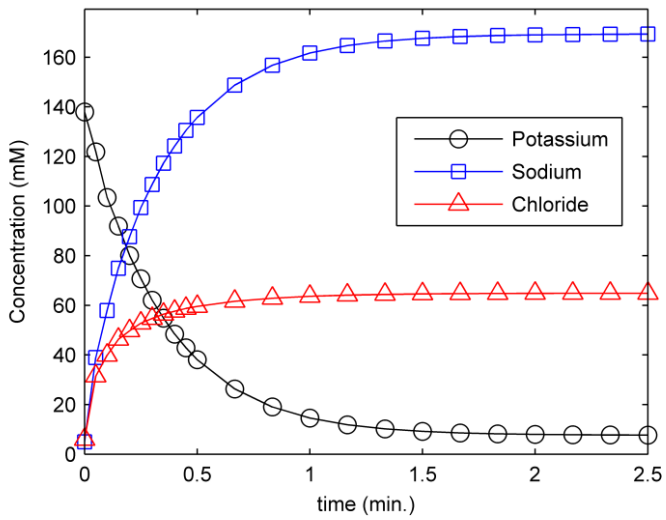
Figure 4-2 B). In the first stage, approximately 25% of the total bacterial K^+ content was quickly released (within 2 minutes of exposure to PG-1), after which the release rate slowed considerably. At 7 minutes, the release rate increased until 15 minutes, at which point the release rate slowed down once again. Along with the experimental data, Figure 4-2 B shows several K^+ release curves computed for different values of the number of pores. A value between 10 and 20 pores yields the closest match to the experimental data: the rate of release shown by the model (the slope of the curve) is similar to the release rate in the two fast stages shown in the experiment, and a plateau is observed between 50%-70% K^+ release, somewhat lower than the experimental result. The slowing of the release rate starting at 2 minutes and ending at 7 minutes observed experimentally is not captured by the model. One possible explanation for the two-stage behaviour may be that the first stage corresponds to the leakage of K^+ from the periplasm following outer membrane permeabilization by protegrin, while the second stage corresponds to leakage from the cytoplasm following inner membrane permeabilization. However, the outer membrane is known to be inherently highly permeable to small solutes such as ions, so it is unlikely that a significant concentration gradient exists across this membrane to drive a high rate of potassium leakage. Additionally, our model does not account for the kinetics of protegrin diffusion to the inner bacterial membrane, or the kinetics of pore formation, which may also explain the S-shaped curve in Figure 4-2 B; however, this does not explain why the difference arises between exponential and stationary phase bacteria, as there is no reason to believe that protegrin diffusion and pore formation would be affected so drastically by the bacterial growth stage. A more probable explanation is that the intermediate slowdown observed in the case of stationary phase bacteria represents the activation of a bacterial response mechanism, either a temporary increase in the potassium uptake rate, or a temporary closure of potassium channels.

In addition to the percentage of potassium released, the model offers the possibility of monitoring the interior concentration of the other diffusing ions, namely sodium and chloride. Figure 4-3 shows the cytoplasmic concentration of all three ionic

species as a function of time. The data shown correspond to the typical ion permeabilities already mentioned and 100 pores (N_{pores} = 100). As potassium leaks from the cytoplasmic interior, sodium and chloride enter the cell from the higher concentration exterior environment. Interestingly, sodium accumulates inside the cell to a higher concentration than the 110 mM exterior sodium concentration. This is because the zero-current condition (equation (4.7)) requires the baths to remain at a constant total charge. Since the baths must be neutral, this further implies that a non-permeating, negatively charged species exists in the cytoplasm to achieve electroneutrality of the cell interior for the given initial conditions. This anionic species corresponds to the cumulative negative charges of proteins, nucleic acids and small organics. As sodium replaces the potassium that is released, it also enters alongside the chloride ions that are driven into the cell by concentration gradients. The implications of a high concentration of sodium in the cell interior could be significant, as sodium is known to compete with potassium for many enzyme binding sites involved in vital metabolic processes [111-112].

Figure 4-3: Intracellular ion concentrations

A value of 100 pores was used for these data, corresponding to K^+ release kinetics observed in exponential phase bacteria.

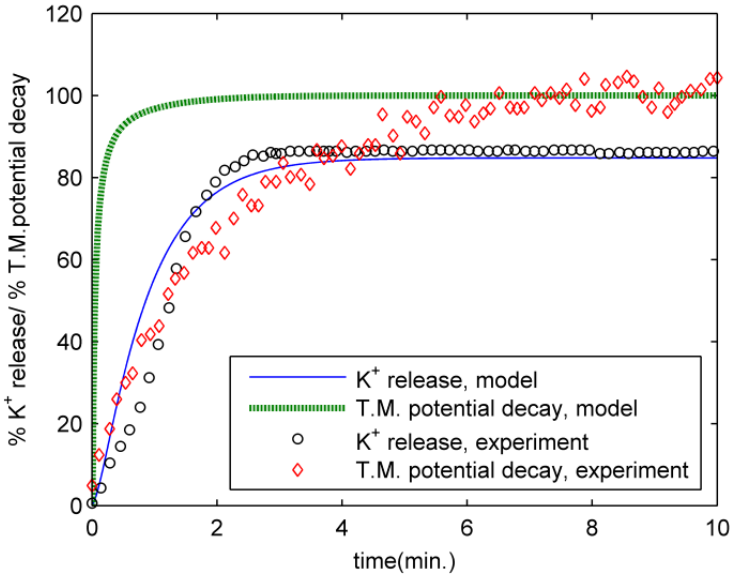


4.3.2 Transmembrane potential

The relative fluorescence of diSC3-5 measured after treatment of exponential phase *E. coli* with 25 μg/mL of PG-1 is shown in Figure 4-4, along with the corresponding K^+ release curve. The equivalent results obtained from the model are shown alongside the experimental data; the number of pores was adjusted to 43 in order to reproduce the slower K^+ release observed in this particular experiment. The transmembrane potential change is scaled as a percentage of complete depolarization, based on the initial value and the steady state value obtained at long times. As expected, the permeabilization of the membrane by PG-1 allows ions to exchange freely, thereby upsetting the concentration gradients across the bacterial membrane that contribute to the transmembrane potential. The higher rate of transmembrane potential decay observed in the model as compared to the experiments could be partially explained by the delay associated with the diffusion of diSC3-5 in response to changes in the transmembrane potential, whereas the model provides direct and instantaneous transmembrane potential data. In the model, the transmembrane potential initially has a value of approximately -75 mV (cell interior negative), and decays to a value of around -12 mV at steady state. These values are well within the range of physiologically relevant potentials measured in bacteria (commonly between -130 and -150 mV) [17]. The nonzero value obtained for the steady-state transmembrane voltage results from the Donnan equilibrium associated with the presence of a nonpermeating, anionic species in the cell interior (see discussion near the end of section 4.3.1), which arises as a result of the electroneutrality condition.

Figure 4-4: Comparison of potassium release and transmembrane potential decay between model and experiment.

Experimental measurements were carried out on exponential phase *E. coli*. The experimental transmembrane potential decay data corresponds to the relative fluorescence of diSC3-5 observed. The model data corresponds to setting the number of pores to 43 in order to match the K^+ release data for this particular experiment.



4.3.3 Light scattering and volume changes.

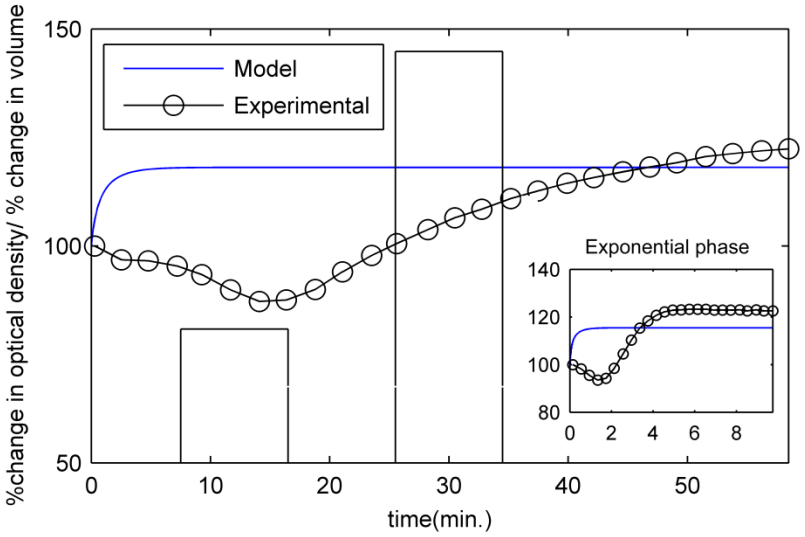
The time-dependent cell volume calculated from the model is shown in

Figure 4-5 along with the experimental light scattering data and ^3H_2O measurements for exponential and stationary phase organisms, respectively. All values are scaled as a percent change from the initial values, which correspond to untreated bacteria. In both exponential (fig. 5 inset) and stationary-phase bacteria (Figure 4-5, main graph), the number of pores was set to achieve the best agreement with the experimental K^+ release (100 pores for exponential-phase bacteria and 20 pores for stationary-phase bacteria). With the parameter choices described in the *Methods and Materials* section, the trend of an initial decrease followed by a subsequent increase in

volume is not even qualitatively reproduced for either bacterial growth stage. In both cases, the model shows a rapid increase in cell volume that occurs much sooner than the measured increases in optical density.

Figure 4-5: Optical density measurements and volume changes for stationary phase *E. coli* (N_{pores}=20)

The bars indicate data obtained from ^3H_2O scintillation counts. The inset shows optical density changes and computed volume changes (N_{pores}=100) for exponential phase *E. coli*. All data are shown as a percent change from the initial value.



The volume behaviour in the model is caused by an osmotic imbalance that results from ion transport. Although potassium leaks out of the cell, sodium and chloride enter it in larger quantities, increasing the osmolarity of the cell interior, thereby providing a strong driving force for water influx. Due to the relatively high permeability of the cell membrane, water enters the cell very quickly in response to this osmolarity change, resulting in a sharp increase in the cell volume. As water flows into the cell, it dilutes its contents and increases the hydrostatic pressure, thereby reducing the driving force for osmotic water influx. An osmotic equilibrium and an equilibrated cell volume are reached shortly after ion diffusion ceases. Since the response to osmolarity changes

is very fast, the slow volume changes observed cannot be a delayed response to osmolarity changes caused by ions, but must be the result of the transport of other osmolytes. Experiments aimed at measuring the kinetics of volume changes in *E. coli* [113-114] and in *Saccharomyces cerevisiae* [115] in response to changes in external medium osmolarity suggest a time scale for volume changes on the order of milliseconds to seconds. As such, the time scale of light scattering changes (and presumably volume changes) observed in the present experiments cannot be a delayed response to changes in ion concentrations, since ion equilibrium is attained well before the volume has stopped changing.

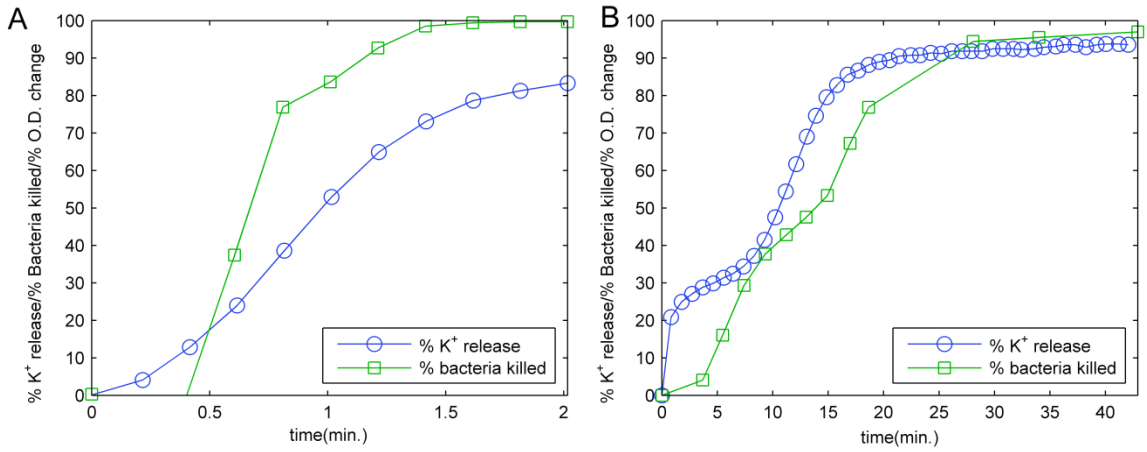
Considering the ambiguities involved in relating light scattering data to volume changes, as well as the simplifying assumption made in the model that K^+, Na^+ and Cl^- are the only relevant permeating osmolytes, the inconsistencies between the experiment and the model in this area are not surprising. Very likely, bacteria are able to adopt several additional strategies to regulate their volume, such as the release of other small molecules (e.g. glycine betaine, glutamate, trehalose, and proline), synthesis of additional osmolytes [116], and active transport of ions [117]. Despite these shortcomings in the model, one notable result that emerges from both the experiments and the model is that exposure to protegrin ultimately causes a significant increase in bacterial volume, which could very well compromise the structural integrity of the sacculus and lead to osmotic lysing of the bacteria. This in turn could be an integral part of the mechanism of action of protegrin-1, as well as other pore-forming antimicrobial peptides.

4.3.4 Bacterial viability loss.

The viability of bacteria was determined by extracting aliquots from each experiment at different times and performing colony counts after overnight incubation. The percentage of bacteria killed after different exposure times and the K^+ release curve are shown for exponential phase bacteria treated with 25 μg/mL of PG-1 at 23°C in

Figure 4-6 A, and for stationary phase bacteria treated in the same conditions in Figure 4-6 B.

Figure 4-6: Concurrent measurements of potassium release and bacterial viability for (A) exponential phase and (B) stationary phase *E. coli*.



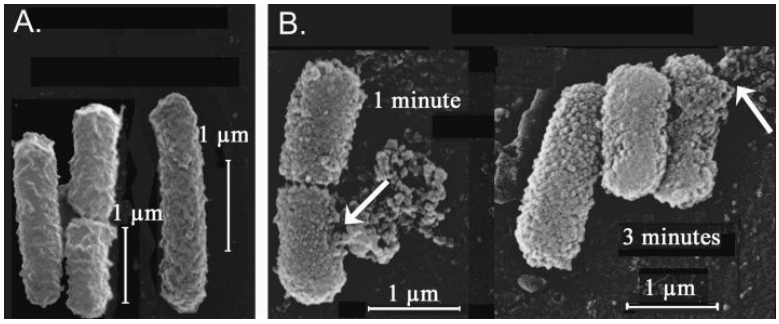
The figures show the bacterial killing kinetics to be similar to the potassium release kinetics, which have been shown above to be concurrent with transmembrane potential decay. This does not necessarily imply that potassium release or membrane depolarization are sufficient to cause cell death; ionophores such as valinomycin can cause complete membrane depolarization that effectively 'stuns' bacteria but does not kill them - once the valinomycin is removed, the transmembrane potential is reestablished, and bacteria remain viable [32]. Nonetheless, the rapid loss of potassium accompanied by the increase in intracellular sodium concentrations could have a strong contribution to cell death, as the exchange of potassium for sodium is likely to inhibit numerous metabolic functions. At the same time, the loss of the transmembrane potential will severely compromise normal membrane function. Although the light scattering data were collected in a different experiment (Fig. 5), they do not appear to have a clear relationship to the bacterial killing kinetics. Since most of the bacteria are rendered nonviable well before the experimentally observed rise in light scattering, any osmotic lysing that occurs

due to cell volume expansion represents an overkill mechanism. Given the regenerative abilities of many bacterial species, this aspect of the killing mechanism of PG-1 may prove useful to the host.

The effects of a large volume increase on bacterial cells are seen most dramatically in microscopy images of PG-1 treated bacteria, once again provided by Dr. Robert Lehrer as part of this collaborative effort. Figure 4-7 shows scanning electron microscopy images of *E. coli* strain ML-35p after several minutes of exposure to PG-1 at a concentration of 25 μg/mL PG-1. A rupture in the outer envelope of the organism in figure 8 leads to extrusion of the cellular membrane through the cell wall, followed by massive leakage of cell contents.

Figure 4-7: Scanning electron microscopy images of *E. coli* treated with 25 μg/mL of PG-1

The arrows in panels A and B point to rupture sites in the bacterial membrane, where the cellular membrane along with intracellular contents are being extruded through the peptidoglycan layer.



4.4 Conclusions

Both the experiments and models suggest that even a relatively small degree of bacterial membrane permeabilization by protegrin can cause rapid potassium leakage and decay of the transmembrane potential. The model presented herein uses as input the ion

permeabilities from the numerical study of three-dimensional electrodiffusion through the octameric pore structure (chapter 3), which in turn is based on a molecular dynamics study of the protegrin pore [76]. According to the present results, we estimate that the number of protegrin pores required to reproduce the experimentally measured K^+ release curve is around 100 for exponential-phase bacteria, and between 10 and 20 for stationary-phase bacteria. It should be noted once again that these numbers represent the average, apparent number of fully formed, open pores, rather than the number of pores or membrane-bound protegrin molecules at any particular instance. The differences between the K^+ release behaviour of organisms in different growth stages are not readily explained by our model; most likely, active transport and release mechanisms that are not accounted for in the model play a significant role.

Considering the rapid potassium release observed for exponential-phase bacteria and its agreement with the present model, we posit that the exponential case more accurately represents a situation in which passive diffusion is the dominant mechanism of potassium release; it appears the bacterial response in the exponential growth stage is too slow to compensate for the passive ion exchange process. In both exponential and stationary phase organisms, the model clearly shows a significant increase in the sodium intracellular concentration. Not surprisingly, the decay of the transmembrane potential occurs in parallel with the potassium release and sodium influx, a trend that is observed both experimentally and in the model.

Although it is difficult to ascertain a clear relationship between light scattering and bacterial volume changes, the light scattering data combined with the ^3H_2O measurements presented herein suggest that in many cases, the volume decreases initially, and increases following the plateau in the K^+ release. This behaviour is not well captured by the model, suggesting that osmolytes other than the three ions considered here are involved. Nevertheless, both the model and the experiments suggest that sufficiently high concentrations of PG-1 cause a significant volume increase, which leads to catastrophic membrane rupture, as observed by electron microscopy. Comparing the

light scattering data to the kinetics of bacterial killing, it appears that the optical density rise is not directly linked to any causes of death, but rather represents an effect of bacterial killing by protegrin. The cause of death, although not conclusive from the present study, coincides with, and may well be a direct result of concurrent potassium release, sodium influx and transmembrane potential decay. As the result of this damage unfolds, the bacterial volume expands, leading to membrane rupture at sufficiently high PG-1 concentrations.

Since different concentrations of protegrin result in different magnitudes of volume increase (Figure 4-5 A), it appears that osmotic lysing by protegrin is not merely a consequence of bacterial death, but rather a characteristic behaviour that occurs only with high concentrations of PG-1. This therefore represents an additional overkill mechanism that protegrins are able to deploy against pathogen targets to render them irreparable. Given the robustness and regenerative abilities of most bacteria, this mechanism would prove useful to the protegrin host.

5 Molecular dynamics simulations of toroidal protegrin pores

5.1 Introduction and Background

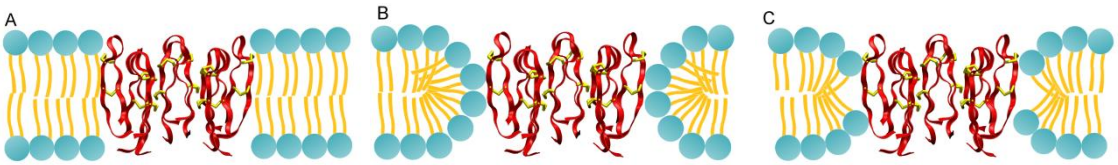
As discussed in the previous two chapters, the key structural unit for the activity of protegrins is a transmembrane pore structure. In section 1.5.2 and in the introduction to chapter 3 (section 3.1), we described some of the previous work pertaining to the investigation of the atomistic features of these structures. In particular, NMR experiments first confirmed the presence of octameric pores in 3:1 POPE:POPG lipid bilayers[54], and subsequent molecular dynamics simulations performed by Langham and coworkers [76] provided a dynamic picture of these structures with unprecedented detail. As part of their work, these latter authors reported on the evolution of the lipid structure surrounding the peptide pore, which constitutes the focus of this chapter.

The simulations of Langham et al were started from a configuration referred to as a barrel-stave pore, where the lipids are aligned parallel to the membrane normal, and the exterior surfaces of the peptide pore are in contact with the hydrophobic tails of the lipids (see Figure 5-1 A). However, the balance of more recent evidence [55-56] suggests that the pores formed by protegrins adopt a toroidal conformation, where some lipids tilt fully from the lipid membrane normal, so that lipid head groups line the exterior of the peptide pores (see Figure 5-1 B). In the simulations of Langham et al, the lipids were observed to adopt a highly tilted conformation after a total of 150 nanoseconds, resulting in local thinning of the bilayer near the pore. The resulting structure was dubbed a semi-toroidal pore (see Figure 5-1 C), as it resembles the toroidal pore with respect to the tilting of lipid molecules, but the extent of this tilting is not sufficient to result in the exterior of the pore being lined entirely by lipid head groups. However, a comparison of average distances separating ^{31}P atoms in the phosphate groups of lipids and ^{13}C atoms in various peptide residues leads to some disagreement between these simulation results and the NMR experiments of Tang and coworkers [56], as discussed in section 0 below. We therefore set out to repeat such simulations starting from a toroidal configuration (Figure 5-1 B) to

see how this impacts the results of the simulations, with a particular focus on the lipid structure and lipid dynamics. The work presented in this chapter is very recent, and data herein are preliminary as of the time of the writing of this document.

Figure 5-1: Different pore topologies suggested for protegrins

A. Barrel-stave pore, in which lipids are aligned parallel to the lipid bilayer axis. **B.** Toroidal pore, in which lipids near the center are perpendicular to the lipid bilayer normal, resulting in a continuous leaflet and the peptide exterior being completely lined by lipid head groups. **C.** The semi-toroidal pore, a topology in which the lipid bilayer thins around the region of the peptide pore, but two distinct leaflets remain.



The construction of the initial configuration for a lipid bilayer is not trivial, especially when peptides are inserted in the lipid bilayer. In particular, the toroidal lipid conformation presents some unique challenges due to the complex arrangement of the lipid molecules surrounding the protegrin pore. A major goal of this work was to develop a rigorous, robust protocol for the construction of toroidal pore structures in MD simulations. This method is not only useful in the present context, but can be applied to simulations of other protein-membrane systems where the toroidal topology is believed to be relevant. Furthermore, the simulation of protein-membrane systems in general is problematic due to controversy regarding the choice of ensemble [118-120]. Most lipid bilayer simulations with the CHARMM27 force field are typically performed in the NP_zAT ensemble, where the area of the simulation box remains constant; otherwise, this force field has a tendency to equilibrate to artificially low values of the area per lipid. In the present case, where the pore is laterally flexible, compression of the lipids under area constraints can result in undesirable and artifactual disruption of the pore. This led Langham et al to carry out their simulations in the NPT ensemble, which allows the lateral area of the simulation box to fluctuate, even though this is not generally

recommended for lipid bilayer simulations employing the CHARMM27 force field [118-120]. Shortly after we began this work, an update was released for the CHARMM force field specifically to address this issue, dubbed CHARMM36 [121]. We therefore decided to carry out simulations using both of these force fields, and explore the differences between the resulting behaviors in the context of toroidal pores.

5.2 Methods

In order to achieve a physically realistic initial configuration, the lipid bilayer construction method developed by Roux and coworkers [122] was adopted, with appropriate modifications made for the toroidal geometry. Briefly, the method involves first constructing a bilayer composed of spherical particles of dimensions similar to the lipid head groups around a pre-equilibrated peptide pore structure. These surrogate lipid particles are constrained to surfaces corresponding to the locations of lipid bilayer head groups, but are otherwise free to move, and short dynamics simulations are performed with the peptides fixed in order to achieve reasonable packing of the spherical particles around the pore and in the simulation box. The spherical particles are then replaced by lipid molecules taken from a library of conformers of the lipids of interest. Such libraries are available through the CHARMM-GUI web-based series of programs (URL: www.charmm-gui.org) [123]. Water and ions are then added to the top and bottom of the bilayer, and a series of minimizations and short dynamics runs are performed to equilibrate the system.

5.2.1 Defining the toroidal surface

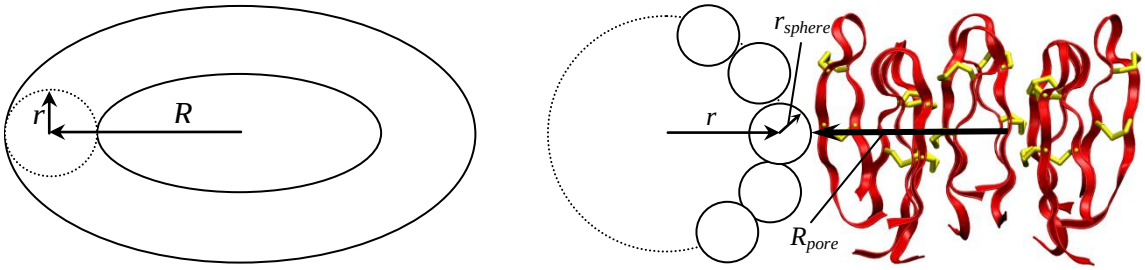
Clearly, the case of the toroidal pore system presents unique challenges due to the complex arrangement of lipids around the peptides. In order to apply the same methodology discussed above using the spherical surrogate particles, we developed and implemented a restraining potential that allows particles to move on a toroidal surface

built around the peptides. This restraining potential U_{res} is based on the equation for a torus, and has the following mathematical form:

$$U_{res} = k\left[\left(R(\theta) - \sqrt{x^2 + y^2}\right)^2 + z^2 - r^2\right]^2 \quad (5.1)$$

Here, k is a force constant that determines the strength of the harmonic potential restraining each particle to the surface, x, y, z are the co-ordinates of the spherical particle being restrained, and r and R are the minor and major radii of the torus respectively, depicted in Figure 5-2 A below. As discussed below, R may have a functional dependence on θ, the angular coordinate describing the perimeter of the peptide pore.

Figure 5-2: Defining the toroidal surface



Since we are interested only in the inner surface of the torus (compare to Figure 5-1 B), this restraint is only active for $x^2 + y^2 \leq R^2$. The appropriate values of r and R are determined based on the pore geometry (see Figure 5-2 B). Clearly, r is easily obtained based on the length of the lipids, in this case the tail-to-phosphate distance, which is 19 Å for POPE/POPG lipids:

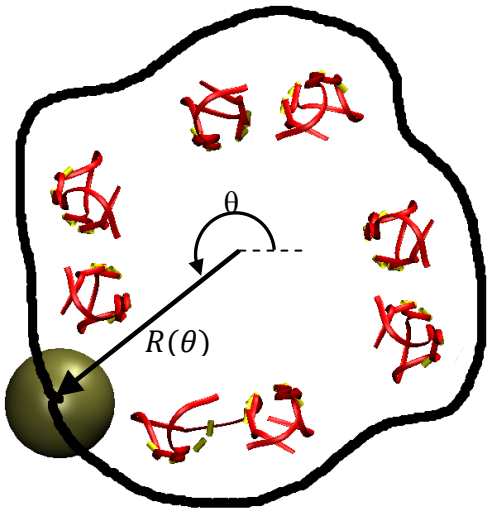
$$r = L_{lipid} = 19\text{Å} \quad (5.2)$$

The value of R, the major radius of the torus, will be the sum of the radius of the protegrin pore (R_{pore}), the spherical particles used as lipid head group surrogates, and the lipid tails (again, refer to Figure 5-2 B):

$$R = L_{lipid} + r_{sphere} + R_{pore} \quad (5.3)$$

The precise method in which the outside radius of the peptide pore (R_{pore}) is determined is important, as it dictates the packing of lipids around the protegrin pore. In our approach, we allow R to have an angular dependence, with different values of R at different locations around the peptide pore circumference (see Figure 5-3). This functional dependence is obtained by running a probe particle around the outside of the peptide pore, and noting the optimal R value at each θ interval. This optimal value is determined by the minimum van der Waals interaction energy between the peptides and the probe particle.

Figure 5-3: The angular dependence of the toroidal major radius, R



The functional dependence of R on θ is then fit using a polynomial, and the restraint described by equation (5.1) is implemented through the user energy module of CHARMM. In addition to the toroidal surface restraints, the particles are also restrained at surfaces $z = \pm 19$Å for cases where $x^2 + y^2 > R^2$. This corresponds to the usual upper and lower leaflet restraints outside of the toroidal surface region.

5.2.2 Determining sphere sizes and numbers of lipids

The sphere sizes are set based on the mean area per lipid of a 3:1 POPE:POPG lipid bilayer, which is assumed to be a weighted average of the area per lipid (A_{lipid}) of the two pure lipids, namely 63 Å^2 per lipid for POPE and 62 Å^2 per lipid for POPG [123]. Given the similar values of the area per lipid for these two lipids, we expect the approximation of the weighted average of 62.75 Å^2/lipid to be adequate. The sizes of the spheres are then set based on equating the projected area of the sphere to this value:

$$r_{sphere} = \sqrt{\frac{A_{lipid}}{\pi}} = \sqrt{\frac{62.75\ \text{Å}^2}{\pi}} \approx 4.47\ \text{Å} \tag{5.4}$$

In order to determine the number of lipids in the toroidal region, we find the maximum number of lipids that could be packed into the corresponding volume. The volume of a lipid in this case corresponds to the packing volume below the head group phosphate location (z<19), which is simply the area per lipid multiplied by this height. The volume of the half-torus region is found using a standard geometric relation, the derivation of which is omitted:

$$V_{toroid} = \frac{\pi r^2}{2} \int_0^{2\pi} R(\theta) d\theta \tag{5.5}$$

Finally, the number of lipids in the toroid region is found according to a simple packing argument:

$$N_{lip} = \frac{V_{toroid}}{V_{lip}} = \frac{V_{toroid}}{A_{lipid} r} = \frac{\pi r}{2A_{lipid}} \int_0^{2\pi} R(\theta) d\theta \tag{5.6}$$

5.2.3 Simulation setup

The lipid surrogate particles were placed randomly so that N_{lip}=135 particles were in the toroidal region, while another 254 particles were placed outside of this region, in a hexagonal simulation box of dimensions 132.4 Å × 132.4 Å × 72 Å. This number was selected to create approximately 3 layers of lipid molecules between the edge of the toroidal pore region and the edge of the simulation box. The particles were allowed to equilibrate under the constraints described in sections 5.2.1 and 5.2.2 as well as periodic boundary conditions in a hexagonal unit cell during a short Langevin dynamics (LD) simulation in CHARMM [124]. The peptides were then introduced, and another short LD simulation was carried out with the peptides fixed to equilibrate the lipid particles around the exterior of the pore. The lipid particles were then replaced with POPE and POPG lipids drawn randomly from libraries of single POPE/POPG lipid structures [123], and aligned so that the phosphorus atoms of the lipids coincided with the centers of the spheres, and the long axes of the lipid molecules were perpendicular to the toroidal surface, with the tails pointing to the bilayer interior. Two starting structures were created, dubbed V1 and V2, corresponding to slightly different distributions of the POPE and POPG lipids. In structure V1, a 3:1 ratio was maintained over all the lipids with no constraints on the compositions of the toroidal and non-toroidal regions, while in structure V2, the 3:1 ratio was enforced separately for the toroidal and non-toroidal regions. The effect of these differences is to create two different initial POPE/POPG lipid distributions; since diffusion and mixing of lipid molecules is typically much slower than simulation time scales, we are interested in seeing how differences in lipid distributions affect simulation results.

Water molecules that were included in the lipid structures and had any overlaps were removed, as were any water molecules between $-12 < z < 12$ (where z is the coordinate along the membrane normal). The peptide pore was previously simulated in an aqueous environment, wherein the peptides were kept fixed, and water and ions were allowed to equilibrate around the structure. From these simulations, the equilibrated

solvent (water and ions) that remained inside the pore were copied to the lipid bilayer simulation in order to fill the pore with solvent. Additional water layers 15 Å in thickness were placed on top and below the lipid bilayer. Sixty-one K^+ counterions were added to balance the net negative charge of the lipids, peptides and chloride ions associated with the peptides, as well as an additional 40 K^+ and 40 Cl^- ions to simulate a solution with an ionic strength of 0.1 M. This corresponds to the ionic strength in several relevant ion conductance experiments [58], as well as a typical concentration of K^+ in bacterial cells. The ions were placed according to a simple Metropolis Monte Carlo scheme, where the electrostatic component of the interaction between the ion being moved and the remainder of the system (excluding water) was used as the energy criterion.

The solvated system was then equilibrated using the NAMD software package [125] in four stages: in the first stage, all atoms except the solvent were fixed, and the system was minimized for 7500 steps, then heated gradually in 5,000 step/10 K increments to 310.15 K (37°C) and equilibrated for 100 ps. In the second stage, the solvent atoms were fixed to their locations at the end of the first stage, and the procedure was repeated for the lipid atoms (due to the fixed solvent, this mostly affects the lipid tails). In the third stage, only the peptides and atoms within 12 Å of the peptides were restrained, while the remainder of the system was heated and equilibrated under the same conditions. In the fourth stage, only lipid phosphorus atoms and peptide α-carbons were restrained, and the system was minimized, heated and equilibrated again. The peptide backbone atoms were additionally weakly restrained for the first 15-20 ns of production runs to their initial locations.

5.2.4 Production runs

All production runs were carried out with the NAMD software package version 2.7b2 [125]. We employed the CHARMM 27 force field [83] with CMAP corrections [126] for proteins, and the recently released CHARMM 36 parameter set [121] for lipids as well as the original CHARMM27 lipid parameter set. The CHARMM36 recent release improves the simulation of lipid bilayers in the constant temperature and pressure

(NPT) ensemble, which is implemented in NAMD using the Nose-Hoover-Langevin [127-128] piston method. The simulations based on the CHARMM27 force field will be referred to as V1-27 and V2-27, while those carried out using the CHARMM36 force field are denoted V1-36 and V2-36. In all cases, the pressure was set to 1 atmosphere, the pressure piston period was set to 200 fs and the piston decay was set to 100 fs. Electrostatic interactions were modeled using the particle mesh Ewald summation technique [129-130], with a fast Fourier transform grid with a spacing of approximately 1 grid point per Å, while van der Waals interactions were smoothly switched off between 8 and 12 Å. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm [131], which allowed for an integration time step of 2 fs. A more general background to molecular dynamics simulations that explains some of these methods in more detail is provided in the Appendix at the end of this document.

5.3 Results and Discussion

The results obtained for four simulations using the toroidal pore geometry (V1-36, V2-36, V1-27, V2-27) and the simulations of Langham et al [76] which were started from a barrel-stave geometry are all compared in the context of lipid bilayer structure, peptide pore structure, and ion transport through the pore. The total simulation times are listed in Table 5-1 below; note that the simulations of Langham et al were additionally equilibrated for 50 ns in the NP_zAT ensemble prior to the production runs described in the analysis below. For more details about this simulation, the interested reader is referred to [76].

Table 5-1: Simulation lengths for all cases discussed

Simulation:	V1-36	V2-36	V1-27	V2-27	Langham et al
Total time (ns):	98.5	101	62.4	66.1	103*

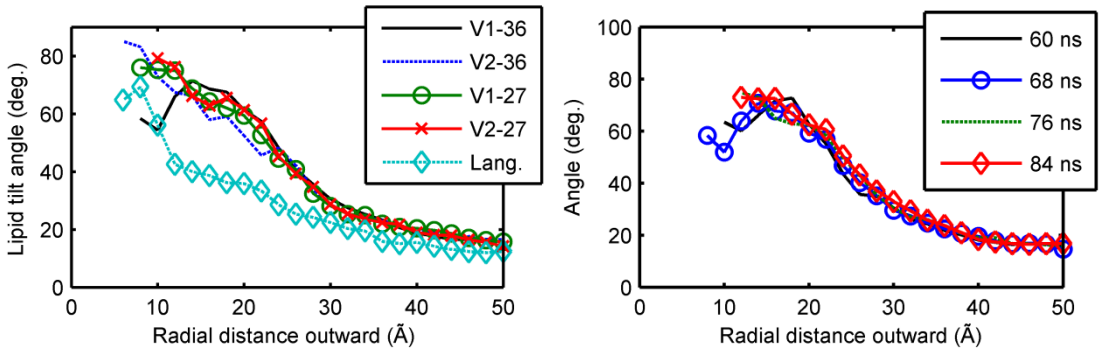
* The simulations of Langham et al were preceded by 50 ns of equilibration in the NP_zAT ensemble for a total of 153 ns. Only the last 103 ns were used for the present analysis.

5.3.1 Lipid bilayer structure

In order to analyze the structure of the lipid bilayer, we have measured the tilting of lipids from the membrane normal as a function of radial distance from the peptide pore center. The resulting data are plotted in Figure 5-4 below for all cases, along with time-dependent profiles for simulation V1-36. The lipid tilt angle is defined as the angle between the vector connecting the lipid phosphorus atom to the center of mass of the C-16 and C-18 atoms in the terminal methyl groups of the lipid tails. Note that due to the double bond in the oleoyl 18-carbon chain of both POPE and POPG (refer to Figure 1-3 and Figure 1-4), a kink exists in this chain, so that this vector is not fully parallel to the membrane normal even in a flat lipid bilayer. In all cases, r=0 corresponds to the center of the peptides; the lack of data for $r < \sim 8\text{Å}$ is a result of the exclusion of lipids from the peptide pore, and the noise in the range $8\text{Å} < r < 12\text{Å}$ is a result of inherently poor sampling, as very few lipid molecules spend a significant fraction of the simulation time in this region.

Figure 5-4: Lipid tilt angle as a function of radial distance outward from the pore center

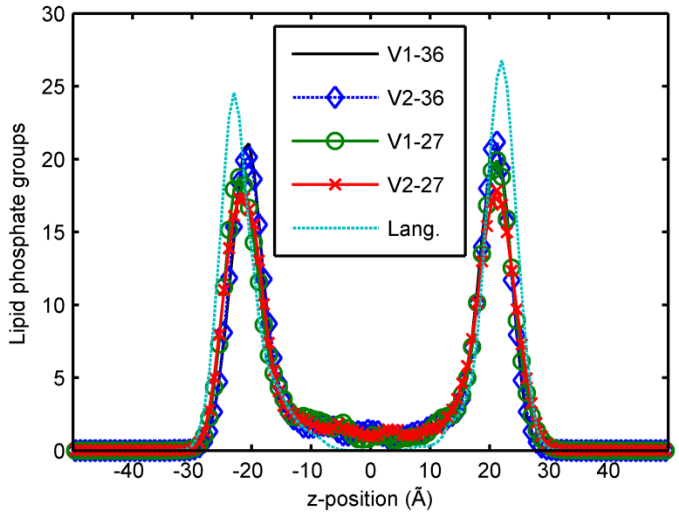
Left: data for all simulations, averaged over the last 20 ns in each case. **Right:** Several profiles averaged over consecutive 8 ns intervals near the end of simulation V1-36. Each line corresponds to a different 8-ns time interval.



As expected, notable differences exist between the toroidal pore simulations and the simulations of Langham et al. As already noted by these authors, the lipids in their simulation tilt somewhat towards the peptides, forming what they dubbed a semi-toroidal pore [76]. However, as shown in Figure 5-4, this tilting is still not nearly as large as in fully toroidal simulations, where lipids near the peptide exterior ($r \approx 12$Å) are nearly perpendicular to the membrane normal. As already mentioned, the high values observed for all simulations for $8\text{Å} < r < 12\text{Å}$ are not particularly meaningful, as they represent a poorly sampled region. Furthermore, the toroidal lipids appear to be stable, as indicated in the right panel of Figure 5-4, which shows that the same plot for four consecutive 8-ns intervals near the end of simulation V1-36 does not change appreciably as a function of time. Overall, these data provide a quantitative measure of the toroidal nature of the pore, and confirm that a) it is distinctly different in all four toroidal pore simulations as compared to the barrel-stave simulations of Langham et al, and b) the toroidal lipids appear to be stable on the time scales of these simulations, and do not show any strong tendencies to reorient towards a barrel-stave configuration.

A similar picture is provided by Figure 5-5, which shows the distributions of lipid phosphate groups as a function of distance along the membrane normal (z) in all simulations.

Figure 5-5: Distribution of phosphate groups in all simulations

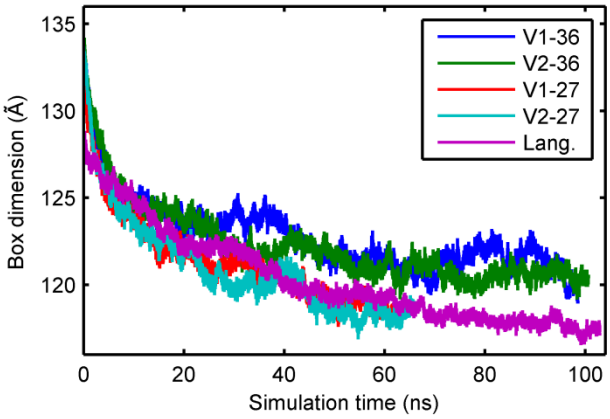


Lipid phosphate groups are present near z=0 in all of the toroidal pore simulations, but are absent for $-5 < z < 5$ in the simulations of Langham et al (the respective line is partially obscured by the abscissa). In all cases, the large peaks at $z \approx \pm 20$Å correspond to the upper and lower leaflets of the flat portions of the bilayer, which have the highest density of lipid phosphate groups. In the simulations of Langham et al, these peaks are larger due to the larger fraction of lipids in the flat portion of the bilayer (as opposed to the toroidal portion). The peaks are also slightly farther apart than in the other simulations, likely due to the slightly higher degree of lateral compression of the bilayer observed in this simulation, which results in an increased thickness, reflected by the spreading of the phosphate group z-distribution. The same effect is noted in simulations V1-27 and V2-27, which are based on the same force field; the effect is less notable in these cases, likely as a result of the shorter simulation times.

We explore this point further by plotting the variations in the simulation cell x-dimension, which in the case of a hexagonal simulation unit cell corresponds to the distance between the centers of periodic images in the x-y plane. As shown in Figure 5-6, the simulations of Langham et al, as well as V1-27 and V2-27, all experience a slightly

higher degree of compression than the simulations employing the updated CHARMM36 lipid force field.

Figure 5-6: Simulation box dimensions for all cases as a function of simulation time



Note that the simulations of Langham et al were previously equilibrated for 50 ns with the area constrained to a fixed value, whereas the data for the toroidal simulations represent the cell dimension immediately following the initial minimization and heating cycles. The initial compression seen in all of the toroidal pore simulations is likely to a large extent a result of the compression of the toroidal pore section of the lipid bilayer as well as the peptide pore. The differences in the simulation box compression between the simulations employing the CHARMM27 lipid force field (Lang., V1-27 and V2-27) and those based on the CHARMM36 force field (V1-36 and V2-36) is on the order of only ~3Å, which translates to ~2.5% in the present case. We therefore posit that while the updated CHARMM36 force field yields slightly higher areas per lipid, this effect is relatively small for simulations with large numbers of lipids (in the present case, ~200 lipids/leaflet, or lateral dimensions of ~120Å).

5.3.2 Comparison to experimental data

The experimental investigations of Tang et al [56] provide a good point of comparison for our simulations. These authors have investigated the interaction of protegrin with 3:1 POPE:POPG lipid bilayers using solid-state magic angle spinning NMR, and reported distances between various peptide moieties and lipid phosphorus atoms. We measured the same distances in our simulations, and have compiled the data along with the experimental data in Table 5-2 below. The simulation data in all cases correspond to the distance between a particular peptide atom and the nearest lipid phosphorus atoms, averaged over all eight peptides and over the last 10 ns of the simulation. In all cases, the distances are between the listed atomic site on the peptides and the nearest phosphorus atom.

Table 5-2: Comparison of inter-atomic distances between simulations and experiments[55].

The values shown correspond to distances between respective peptide sites and the nearest lipid phosphorus atom. All data are in Å.

Site	V1-36	V2-36	V1-cr	V2-cr	Langham et al.	Experimental
Arg_4-Cζ	7.9	7.2	6.1	7.9	6.8	5.7
Arg_4-Cα	**9.0**	**7.8**	**7.7**	**7.5**	**8.2**	**6.5**
Arg_4-CO	8.9	7.9	8.0	7.6	8.8	6.2
Arg_{11}-Cζ	8.5	6.3	6.5	8.2	5.4	5.1
Arg_{11}-Cα	8.1	7.4	7.6	8.0	6.5	5.5
Arg_{11}-CO	8.6	7.8	7.9	8.6	7.1	5.6
Leu_5-Cα	**8.3**	**8.1**	**8.5**	**7.5**	**10.0**	**6.5**
Val_{16}-CO	8.9	7.7	7.1	7.5	9.5	6.2

We note that in all cases, the simulations overestimate the distances measured experimentally. We conjecture that this systematic deviation arises primarily because the experimental data correspond to solid-state experiments conducted at 253 K, whereas the simulation data reflects the solution-state structure at 310 K. As a result, the flexibility and fluctuations of the lipids and peptides can be expected to be much larger in simulations, resulting in higher overall densities, and weaker specific interactions of

peptide residues with phosphate groups. This is particularly true for the Cζ atoms of Arg_4 and Arg_{11}, which are located at the end of long, flexible arginine side chains (refer to Figure 1-5 (a); Cζ atoms are located near the end of the blue arginine segments), and are known to form relatively strong non-covalent bonds with phosphate groups through guanidinium-phosphate interactions. These interactions become less favorable at higher temperatures due to the associated entropic cost, which may partially explain why the simulation distances are consistently higher than the experimental data. As for differences among simulations, it is difficult to ascertain a clear general pattern. The simulations of Langham et al. tend to be closer to experimental values for the Arg_{11} atoms, but are decidedly farther from experimental data for Leu_5 and Val_{16} than the toroidal simulations. Since Arg_{11} is located in the β-hairpin turn region, it is not surprising that a barrel-stave geometry places this residue closer to the lipid head groups (refer to Figure 5-1). While the experimental constraints would still allow Arg_{11} to interact closely with lipid phosphate groups in a toroidal geometry, this close binding is rendered even more unfavorable than in a barrel-stave geometry at higher temperatures.

Another important difference in the toroidal simulations as compared to those of Langham et al concerns the distance between Leu_5-Cα and the nearest phosphate groups. Tang et al, the authors of the NMR experiments cited in Table 5-2 [56] point out that a barrel-stave pore geometry and a thinned, semi-toroidal geometry can both be ruled out by the fact that Arg_4-Cα and Leu_5-Cα have identical distances to phosphate groups (6.5Å in their experiments). Since these atoms are constrained to a minimum distance of ~3.6 Å from one another by covalent bonds, they must be interacting with different phosphate groups, or else with a phosphate group that resides above the location of Arg_4. Since neither case is compatible with a barrel-stave or thinned/semi-toroidal geometry, the authors conclude that the fully toroidal geometry is the most probable one [56]. We note that in all of the toroidal pore simulations, the Arg_4-Cα and Leu_5-Cα to phosphate distances are extremely similar to one another, consistent with experimental observations. However, this is not the case in the simulations of Langham et al, where a difference of

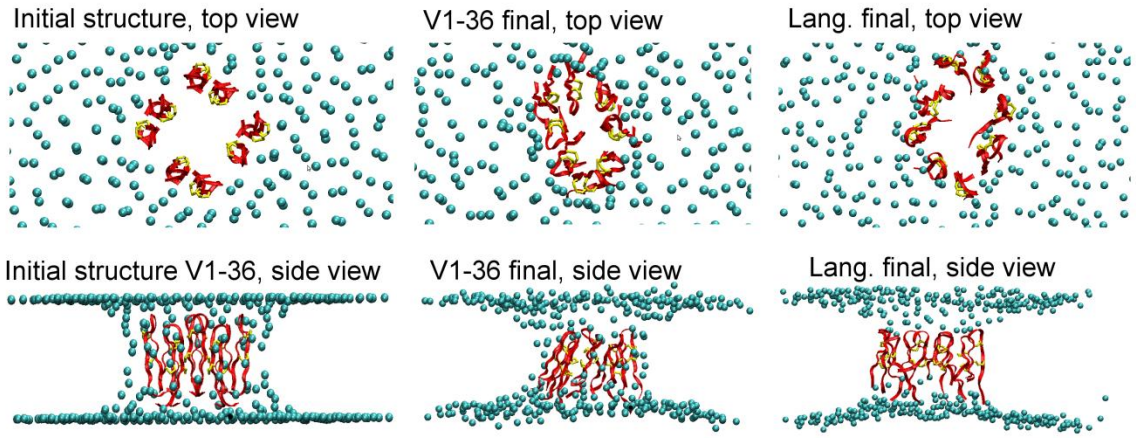
1.8Å is observed, and an unusually large discrepancy of 3.5Å exists between the experimental and simulation Leu_5-Cα to phosphorus distances. In light of all these considerations, we posit that the toroidal pore simulations provide a superior match to experimental data as compared to the barrel-stave pore simulations of Langham et al.

5.3.3 Peptide structure

Figure 5-7 shows snapshots from the end of simulation V1-36 and the simulation of Langham et al, as well as the starting peptide pore structure of V1-36. The initial structure corresponds to the NMR structure suggested by Mani et al [54]. The lipid atoms are perfectly aligned initially because no heating or equilibration has been carried out yet.

Figure 5-7: Snapshots of the initial and final structures of the pore from simulations

The backbones of protegrin peptides are shown as red ribbons, with yellow bonds representing the disulfide bonds. Phosphorus atoms are shown as cyan spheres.

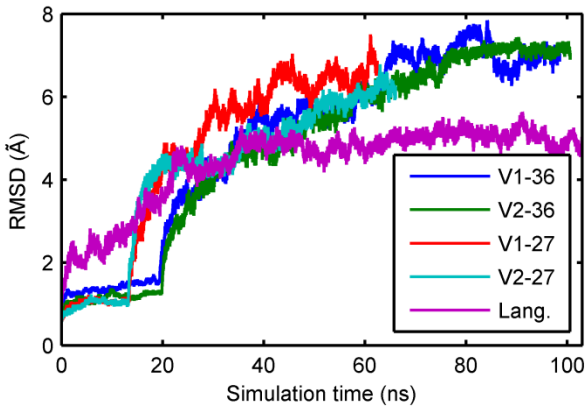


Both simulation snapshots show a significant distortion of the peptides from the starting structure, with the effect more pronounced in the toroidal pore simulation (V1-36). In this case, peptides translate slightly along the curved toroidal pore towards the β-hairpin tail region (bottom of the figure), causing them to tilt inward and narrow the pore opening as

compared to the simulations of Langham et al. The same slight translation of the peptides towards the β-hairpin tail region is observed in both simulations, but no tilting occurs in the simulations of Langham et al due to the lack of curvature in a barrel-stave pore. Also of note is the persistence of the toroidal pore structure in simulation V1-36 (i.e. the presence of lipid head groups covering the exterior of the peptide structure), and the tendency in the simulations of Langham et al for lipids in the initially barrel-stave configuration (not shown from simulations; refer to Figure 5-1 A) to tilt towards the bilayer center, increasingly approaching a toroidal configuration.

We can further quantify the distortion of the peptide structure by the root mean square deviation (RMSD) of all backbone atoms from the initial structure as a function of time, shown in Figure 5-8 for all simulations.

Figure 5-8: RMSD plots for all simulations



The RMSD data show that the distortion from the starting structure is larger for all of the toroidal simulations. The flat portion in the first ~20 ns of these simulations corresponds to the equilibration stage in which the peptide backbone atoms were weakly restrained to their starting structures. Also recall that the simulations of Langham and coworkers were equilibrated for an additional 50 ns in the NP_zAT ensemble, and the RMSD shown in Figure 5-8 is based on the initial structure in the 103 ns NPT segment of the simulation, not the starting structure of the entire simulation. However, the barrel-stave simulations

of Langham and coworkers appear to have stabilized with regards to peptide structural changes, whereas this is not completely the case for any of the toroidal simulations. It may therefore be beneficial to run all of the toroidal simulations even longer in order to ensure that a stable, equilibrated peptide structure is indeed achieved.

Following Langham et al, we also compare two important inter-peptide distances to the NMR experiments of Mani et al, who first proposed the octameric protegrin pore structure [54]. In their experimental work, Phe_{12} and Tyr_7 were replaced with 4-^{19}F-Phenylglycine (F-Phg), an artificial amino acid that includes a fluorine atom attached to the fourth carbon of the phenyl ring, required for ^{19}F CODEX NMR experiments. The equivalent distances for our simulations are the distances between the H_ζ atoms of Phe_{12} in intra-dimer pairs (the atom that is replaced with ^{19}F in F-Phg), and the oxygen atom in the hydroxyl group of Tyr_7 for inter-dimer pairs. The average distances are shown in Table 5-3 below for all simulations, as well as the experimental data.

Table 5-3: Comparison of peptide-peptide distances across all simulations and experimental data.

All entries are in Å. Sampling was carried out for the last 10 ns of all simulations, every 2 ps. See text for further explanations.

Distance	V1-36	V2-36	V1-27	V2-27	Langham et al	Experimental
Phe_{12}-Phe_{12} (intra-dimer)	10.8	6.8	8.1	9.8	8.5	9.0
Tyr_7-Tyr_7 (inter-dimer)	10.4	12.4	12.7	8.8	10.8	6.5

Despite the distortion of the peptides from the initial structure shown by the RMSD plots as well as simulation images, these experimental distances are satisfied reasonably well in most cases, particularly considering the chemical differences between the peptides in the experiment and in the simulations, as well as the low temperatures under which experiments were carried out (in this case, 225-233 K, as compared to 310K for simulations; see earlier discussion in section 0). We thus conclude that in all cases,

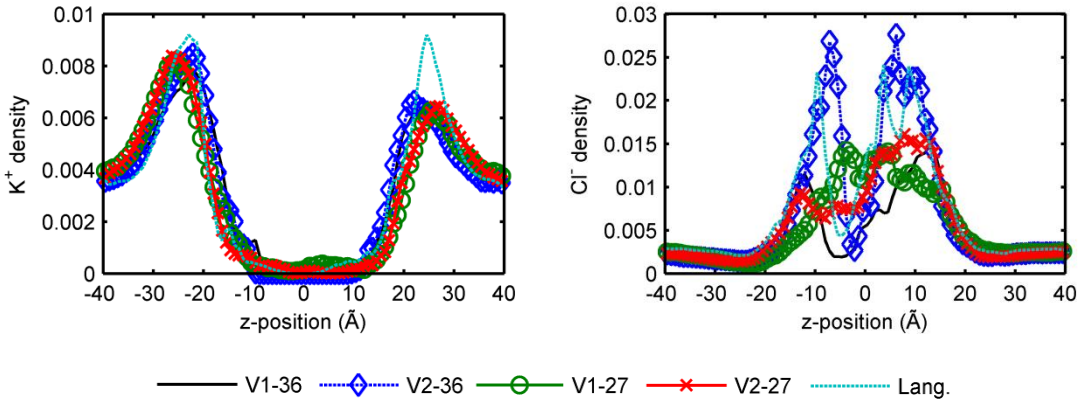
the peptide pore structure remains compatible with experimental NMR constraints despite the apparent distortions from the starting structure. Considering the variation in these numbers, we cannot ascertain that any particular membrane topology, composition or force field is notably more successful than the alternatives as far as matching the experimental data.

5.3.4 Ion transport through the peptide pore

An important property of protegrin pores is their ability to facilitate unrestricted ion transport across cell membranes, which may play an integral role in bactericidal activity. We therefore briefly report some of the simple ion transport characteristics of these structures that emerge from the present simulations. Figure 5-9 shows the distributions of K+ (Na+ in the case of the simulation of Langham et al) and Cl^- ions as a function of distance along the membrane normal (z).

Figure 5-9: Ion density profiles across the membrane

Density values are normalized for water density, leading to units of ions/water molecule. In the simulations of Langham et al, Na+ ions were used instead of K+ ions.



All distributions were normalized based on the distribution of water as well as differences in the total salt concentration between the toroidal and barrel-stave simulations, so that the ordinate axis represents an effective concentration, in units of

ions per water molecule. Not surprisingly, cations are not present near the membrane center region (near z=0) nearly as much as Cl^- ions, since the latter are much more likely to reside inside the highly cationic peptide pore environment. A significant surplus of cations also exists immediately outside of each membrane interface. This is a result of the exclusion of K^+ from the pore and, perhaps more importantly, the presence of negatively charged lipid head groups (POPG) at the interface. The noise in the chloride density profiles inside the pore ($-15 < z < 15$) is likely a result of poor sampling inherent due to the narrow geometry of the pore, and does not necessarily represent meaningful trends. However, it appears that the openings of the pore contain a higher concentration of chloride ions, likely due to the fact that the cationic arginine residues are primarily located at the hairpin turn regions and the tail region (i.e. the openings of the pore).

Following Langham et al, we also track the number of times that Cl^- and K^+ ions cross the pore. A pore crossing event is considered when an ion that originates from $z > 0$ reaches a location corresponding to $z < -15$, or an ion originating from $z < 0$ reaches a location corresponding to $z > 15$. The results for all simulations are tabulated below, along with an approximate rate of ion crossing events determined based on the length of each simulation.

Table 5-4: Summary of ion crossing events in all simulations

Simulation	K^+ crossing events	Cl^- crossing events	K^+ crossing frequency (events/ns)	Cl^- crossing frequency (events/ns)
V1-36	0	30	0	0.30
V2-36	0	35	0	0.35
V1-27	2	50	0.03	0.80
V2-27	0	44	0	0.67
Langham et al*	2	55	0.02	0.53

*In the simulations of Langham et al, Na^+ ions were used instead of K^+.

Given the number of samples, there is little that we can infer about ion transport properties other than the fact that anions cross readily, while cations do not. As expected,

no significant patterns in these ion transport data are observed with regards to membrane topology, choice of force field, or lipid distribution, as ion transport is most strongly affected by the structure and dynamics of the peptides rather than the lipids.

5.4 Conclusions

We have presented a rigorous methodology for the construction of lipid bilayers containing toroidal pores, and have applied the method to investigate the protegrin pore structure. We find that the lipid toroidal structure persists for the duration of the present simulations (~100 ns), and does not show any significant tendency to rearrange into a fully lamellar, barrel-stave topology. We observe a systematic tendency for simulations to overestimate peptide-phosphate distances, which we attribute to the significantly higher temperatures at which simulations were carried out. Nonetheless, the experimental distances appear to be more consistent with simulation results obtained in the toroidal pore simulations than those of Langham et al [76] which were started from a barrel-stave conformation. In particular, the distances from Leu_5-C_α and Arg_4-C_α to phosphate groups were found to be identical in experiments, which was cited by Tang and coworkers as primary evidence for the formation of toroidal pores. These distances are extremely close to one another in all of the toroidal simulations, but differ significantly in the barrel-stave simulations of Langham et al [76]. As such, we conclude that the present toroidal pore simulations provide a superior match to experimental data.

With regards to peptide pore stability, we note significant distortions from the initial structures in all simulations. Based on RMSD plots, these distortions appear to be less pronounced in the barrel-stave simulations of Langham et al [76] than in any of the toroidal pore simulations. We attribute this additional distortion to the faster dynamics of the peptides due to higher solvent exposure, as well as increased lateral pressure from the toroidal pore structure. Nonetheless, two important distances between different peptide sites reported from NMR experiments [48, 54] are matched reasonably by all of the simulations, particularly considering the chemical differences and temperature

differences between experimental and simulation systems. The RMSD plots also suggest that the peptide structures in the toroidal pore simulations have not fully stabilized, which suggests that even longer simulation times may be in order.

Confirming the findings of Langham et al, we note a high tendency for ions and water to cross through the pore. In particular, chloride ions cross the pore readily, while very few potassium ion crossing events are observed. This suggests a high anion selectivity for the pore, which is not surprising given the highly cationic nature of the protegrin octameric pore (eight peptides carry a total charge of +56). The transport of ions through the pore may be an essential feature in the bactericidal mechanism of protegrin, as it would cause the collapse of the transmembrane potential in bacteria, leading to the loss of vital membrane functions and eventually cell death.

Two additional goals of our simulations were to isolate the effects of the new CHARMM force field release for lipids, dubbed CHARMM36[121] as well as the effects of minor differences in the distribution of the two types of lipids (POPE and POPG). With regards to the effect of the updated force field, we notice a less pronounced tendency for the lateral area of the lipid bilayer to shrink, which also compresses the lipid bilayer and causes its thickness to increase slightly. However, the differences in this effect are relatively minor, and no other systematic deviations between the two force fields (i.e. between simulations V1-27 and V2-27 on one hand and V1-36 and V2-36 on the other) were noted for any of the properties described above. This is good news for the lipid bilayer simulation community, as it suggests that while the updated force field is preferable, it does not render previous simulation efforts using the NPT ensemble and the CHARMM27 force field invalid, and likely only has a minor effect on the properties ultimately extracted from simulations. This may not be the case if the focus of the simulations is to examine the lipid properties in detail (e.g. tail order parameters, electron density profiles, etc.), but this is not the case here.

Likewise, no systematic differences were noted in the properties analyzed due to minor differences in the distributions of POPE/POPG (simulations V1-27 and V1-36 on the one hand, and V2-27 and V2-36 on the other). Once again, this is good news, as it shows that many simulation results are in fact not sensitive to the random and arbitrary distribution of lipids that is typically employed in simulations of heterogeneous lipid bilayer systems. However, even simulations extending to hundreds of nanoseconds are insufficient to capture the lateral redistribution of lipids, so we cannot discount the possibility that POPE or POPG-rich domains do in fact form and affect the behavior of the pore; in particular, it is entirely possible that regions near the cationic peptide pore would be rich in anionic POPG lipids. Given that this has not been confirmed experimentally, and we have no clear basis for assigning a particular lipid distribution, we elected to follow convention and start all of our bilayer simulations with a random distribution of lipids.

Overall, the work presented in this chapter provides a useful framework for the study of protegrin pores, the structural units believed to be most important for bactericidal activity. Recent experimental evidence suggests that the toroidal lipid arrangement is the dominant topology in the case of protegrin pores. As such, this work represents the most current and accurate atomistic simulations of protegrin pores reported to date, while our approach to the construction of toroidal pores represents the most sophisticated and rigorous method for the assembly of such pores that we are aware of. We hope that these methods can find application to simulations of toroidal pores in general, whether in the case of other AMPs or other protein-membrane systems where such topologies are relevant.

6 Conclusions

As discussed in the first chapter, antimicrobial peptides represent an extremely promising class of compounds for the development of lasting, effective therapeutic agents in the fight against antibiotic resistant bacteria. Among the well-studied types of AMPs, protegrin has emerged as an important model peptide, particularly as a representative of β-sheet structured and pore-forming AMPs. The current body of knowledge of the mechanism of action of protegrin, while more expansive than for almost any other AMP, still lacks a quantitative, physics-based perspective. In particular, while biophysical studies have isolated and elucidated several important phenomena (e.g. dimer formation, membrane association, pore formation, ion conduction), few attempts have been made to provide a unified analysis that connects these different important aspects of the protegrin mechanism of action. This work addresses this through a series of biophysical modeling strategies that attempt to bridge the time and length scales that separate these phenomena.

We first develop and implement a modeling framework to analyze the key states of protegrin that precede the formation of pores, particularly monomers and dimers in the bulk phase, bound to the membrane surface, and inserted in the membrane hydrophobic core. We present a model that provides a unified description of the coupled equilibria of these processes using as input the potentials of mean force calculated based on all-atom molecular dynamics simulations in earlier work in our group [67-69]. Our model also approximately accounts for non-ideal crowding effects in the membrane due to the presence of high surface densities of peptides. We present isotherms for all membrane-associated species using the insertion free energy as a variable parameter, and find that inserted species become dominant for free energies more negative than around -5 kcal/mol. Using the free energy of insertion computed from MD simulations results in levels of peptide insertion that most likely imply membrane disintegration, or else membrane expansion and rupture in a biological context. This is interesting in light of the experimental evidence discussed subsequently in chapter 4, which suggests that at

sufficiently high protegrin concentrations, an overkill mechanism is activated that causes complete structural failure of the bacterial cell membrane.

Based on earlier work in our group that investigated the structure of protegrin pores [76], we set out to quantify the ion conductance of such pore structures in greater detail, as this represents one of the key features responsible for bactericidal activity in protegrin and other pore-forming peptides. Unfortunately, all-atom molecular dynamics simulations cannot currently access the time scales required to obtain sufficient sampling of ion crossing events; furthermore, simulations with applied voltages are problematic, and extensive sampling would need to be carried out at each applied voltage. We therefore turned to an approximate, continuum model of ion transport, based on the Poisson-Nernst-Planck (PNP) equations. The method is described in detail in chapter 3, and we show that despite the inherent approximations in this approach, we are able to match experimentally determined single-pore conductance values of ~300 pS. We find that in the context of our model, both the high conductivity and selectivity of protegrin pores towards anions are primarily a result of the cationic arginine side chains. Our model also suggests that arginine at position 4 is particularly important for ion transport, as it resides close to the pore interior, where it interacts more strongly with ions than any other peptide residues.

Our success in modeling the single-channel conductance of protegrin pores led us to ask what the effects of such pores would be on an entire cell, and whether we could use potassium leakage data from experiments previously carried out by our collaborators as a means to connect single-pore properties to measurable macroscopic data. In chapter 4, we present a modeling framework that attempts precisely this. By adjusting only the number of pores in our model, we were able to match experimentally measured potassium leakage data from live *E. coli* cells, and thus provide the first estimate of the number of pores required to kill a bacterium – approximately one hundred, provided they are all stable and open for several minutes. More importantly, this work provides a quantitative connection between molecular level information in the form

of pore structures, single-pore conductance characteristics and macroscopic, biologically relevant data. Some of the other significant results from this work show that exponential-phase and stationary-phase bacteria exhibit fundamentally different behavior with respect to protegrin-induced potassium release, both quantitatively and qualitatively; we hypothesize this to be a result of either differences in cell membrane structure that result in much slower diffusion of peptides to the inner membrane, or the presence of much stronger active potassium transport mechanisms in stationary-phase bacteria. We also find that our model fails to capture the changes in bacterial cell volume indicated by optical density measurements, which suggests that osmolytes other than sodium, potassium and chloride are important in the final stages of bacterial killing. However, since bacterial killing kinetics appear to correlate closely with potassium leakage rates and cell volume changes take place after bacterial death, we posit that in capturing potassium release kinetics, our work succeeds at modeling the process most relevant to bacterial killing, and connecting it to molecular information in the form of pore structures.

In chapter 5, we present recent efforts to refine and improve upon the molecular dynamics simulations of protegrin pores carried out by Langham et al [76], with a particular focus on incorporating an updated force field [121] and studying an alternate lipid topology. Recent experiments suggest that the lipid bilayer forms a toroidal structure around the protegrin pore, wherein lipids tilt so that their long axes are perpendicular to the bilayer normal, and their lipid heads completely line the pore exterior (refer to Figure 5-1). This lipid arrangement could have a significant effect on the structural and dynamic behavior of the peptides. As such, we have developed a sophisticated method for the construction of toroidal lipid structures, described in section 5.2. We report the results of molecular dynamics simulations based on this model construction method, and find that the resulting structure yields overall better agreement with experimental NMR data [56] than previous simulation work conducted in our group, which was started from a barrel-stave configuration. We find that the toroidal lipid

structure is stable, at least on simulation time scales of ~100 ns. However, we do observe significant distortion of the peptide pore throughout the simulation, likely a result of faster dynamics and increased lateral pressure due to the toroidal lipid structure. In concurrence with earlier simulations performed in our group, we observe a significant tendency for ions to pass through the protegrin pore, with a strong preference for anions.

The work we discussed herein has all been carried out in the context of the antibacterial mechanism of action of protegrins. Of equal interest is the investigation of the toxicity mechanism of protegrins, for which experimental investigations suggest a fundamentally different biophysical mechanism that does not involve the formation of transmembrane pores [54]. This is a fortuitous finding for drug design efforts, as it potentially enables peptide modifications to be made that diminish toxicity without affecting antibacterial efficacy. From the perspective of simulations, it also presents an excellent opportunity for future work. The models presented in chapter 2 can be particularly useful in this context, as they allow a direct comparison of the membrane-associated states in different lipid bilayers. Additional molecular dynamics simulations of protegrin peptides in the presence of zwitterionic, mammalian-mimicking bilayers can provide important insights into how protegrins interact with and disrupt such membranes, while potential of mean force calculations can be used in a similar manner to quantify these interactions and extract molecular-level thermodynamic quantities. Other areas of interest for future work include the augmentation of the equilibrium models of chapter 2 to account for non-idealities caused by long-range electrostatic interactions, as well as the development of kinetic models that can model the time-dependent behavior of the different processes discussed, and eventually connect these to the dynamic ion transport models discussed in chapters 3 and 4.

The multi-scale analysis presented herein makes a number of significant contributions to the current understanding of the mechanism of action of protegrins, an important class of antimicrobial peptides. We have also introduced a number of novel modeling methods that can easily be applied to the study of other antimicrobial peptides,

as well as protein-membrane systems in general. We hope that this work stands as proof of the usefulness of quantitative, physics-based approaches in the investigation of complex biological systems, particularly their ability to connect distinct but coupled phenomena across disparate time and length scales.

7 Bibliography

1. Spellberg, B., et al., *Trends in antimicrobial drug development: implications for the future.* Clin Infect Dis, 2004. **38**(9): p. 1279-86.
2. Pinner, R.W., et al., *Trends in infectious diseases mortality in the United States.* JAMA, 1996. **275**(3): p. 189-93.
3. Boucher, H.W. and G.R. Corey, *Epidemiology of methicillin-resistant Staphylococcus aureus.* Clin Infect Dis, 2008. **46 Suppl 5**: p. S344-9.
4. Boucher, H.W., et al., *Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America.* Clin Infect Dis, 2009. **48**: p. 1-12.
5. Hogberg, L.D., A. Heddini, and O. Cars, *The global need for effective antibiotics: challenges and recent advances.* Trends Pharmacol Sci, 2010.
6. *Bad bugs, no drugs: as antibiotic discovery stagnates, a public heatlh crisis brews.* Infectious Diseases Society of America, 2004.
7. Morell, V., *Antibiotic resistance: road of no return.* Science, 1997. **278**(5338): p. 575-6.
8. Jenssen, H., P. Hamill, and R.E. Hancock, *Peptide antimicrobial agents.* Clin Microbiol Rev, 2006. **19**(3): p. 491-511.
9. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**(6870): p. 389-95.
10. Ge, Y., et al., *In vitro antibacterial properties of pexiganan, an analog of magainin.* Antimicrob Agents Chemother, 1999. **43**(4): p. 782-8.
11. Wimley, W.C., *Describing the mechanism of antimicrobial Peptide action with the interfacial activity model.* ACS Chem Biol, 2010. **5**(10): p. 905-17.
12. Hancock, R.E. and H.G. Sahl, *Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies.* Nat Biotechnol, 2006. **24**(12): p. 1551-7.
13. Trotti, A., et al., *A multinational, randomized phase III trial of iseganan HCl oral solution for reducing the severity of oral mucositis in patients receiving radiotherapy for head-and-neck malignancy.* Int J Radiat Oncol Biol Phys, 2004. **58**(3): p. 674-81.
14. Giles, F.J., et al., *A phase III, randomized, double-blind, placebo-controlled, study of iseganan for the reduction of stomatitis in patients receiving stomatotoxic chemotherapy.* Leuk Res, 2004. **28**(6): p. 559-65.
15. Jenssen, H., P. Hamill, and R.E.W. Hancock, *Peptide antimicrobial agents.* Clinical Microbiology Reviews, 2006. **19**(3): p. 491-511.
16. Zhang, L. and T.J. Falla, *Antimicrobial peptides: therapeutic potential.* Expert Opin Pharmacother, 2006. **7**(6): p. 653-63.
17. Yeaman, M.R. and N.Y. Yount, *Mechanisms of antimicrobial peptide action and resistance.* Pharmacol Rev, 2003. **55**(1): p. 27-55.

18. Matsuzaki, K., et al., *Interactions of an antimicrobial peptide, magainin 2, with outer and inner membranes of Gram-negative bacteria.* Biochim Biophys Acta, 1997. **1327**(1): p. 119-30.
19. Brogden, K.A., *Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?* Nat Rev Microbiol, 2005. **3**(3): p. 238-50.
20. Boman, H.G., J. Marsh, and J.A. Goode, *Antimicrobial peptides, Ciba Foundation Symposium, vol. 186*. 1994, Chichester: Wiley and Sons.
21. Bechinger, B., M. Zasloff, and S.J. Opella, *Structure and orientation of the antibiotic peptide magainin in membranes by solid-state nuclear magnetic resonance spectroscopy.* Protein Sci, 1993. **2**(12): p. 2077-84.
22. Steiner, H., D. Andreu, and R.B. Merrifield, *Binding and action of cecropin and cecropin analogues: antibacterial peptides from insects.* Biochim Biophys Acta, 1988. **939**(2): p. 260-6.
23. Pouny, Y., et al., *Interaction of antimicrobial dermaseptin and its fluorescently labeled analogues with phospholipid membranes.* Biochemistry, 1992. **31**(49): p. 12416-23.
24. Huang, H.W., *Molecular mechanism of antimicrobial peptides: the origin of cooperativity.* Biochim Biophys Acta, 2006. **1758**(9): p. 1292-302.
25. Hancock, R.E., *Peptide antibiotics.* Lancet, 1997. **349**(9049): p. 418-22.
26. Matsuzaki, K., et al., *Molecular basis for membrane selectivity of an antimicrobial peptide, magainin 2.* Biochemistry, 1995. **34**(10): p. 3423-9.
27. Epand, R.M. and H.J. Vogel, *Diversity of antimicrobial peptides and their mechanisms of action.* Biochim Biophys Acta, 1999. **1462**(1-2): p. 11-28.
28. Hancock, R.E. and R. Lehrer, *Cationic peptides: a new source of antibiotics.* Trends Biotechnol, 1998. **16**(2): p. 82-8.
29. Kokryakov, V.N., et al., *Protegrins: leukocyte antimicrobial peptides that combine features of corticostatic defensins and tachyplesins.* FEBS Lett, 1993. **327**(2): p. 231-6.
30. Hale, J.D. and R.E. Hancock, *Alternative mechanisms of action of cationic antimicrobial peptides on bacteria.* Expert Rev Anti Infect Ther, 2007. **5**(6): p. 951-9.
31. Hancock, R.E. and A. Rozek, *Role of membranes in the activities of antimicrobial cationic peptides.* FEMS Microbiol Lett, 2002. **206**(2): p. 143-9.
32. Wu, M., et al., *Mechanism of interaction of different classes of cationic antimicrobial peptides with planar bilayers and with the cytoplasmic membrane of Escherichia coli.* Biochemistry, 1999. **38**(22): p. 7235-42.
33. Wimley, W.C., M.E. Selsted, and S.H. White, *Interactions between human defensins and lipid bilayers: evidence for formation of multimeric pores.* Protein Sci, 1994. **3**(9): p. 1362-73.
34. Lehrer, R.I., A.K. Lichtenstein, and T. Ganz, *Defensins: antimicrobial and cytotoxic peptides of mammalian cells.* Annu Rev Immunol, 1993. **11**: p. 105-28.

35. Mani, R., et al., *Membrane-disruptive abilities of beta-hairpin antimicrobial peptides correlate with conformation and activity: a 31P and 1H NMR study.* Biochim Biophys Acta, 2005. **1716**(1): p. 11-8.
36. Lai, J.R., et al., *Roles of salt and conformation in the biological and physicochemical behavior of protegrin-1 and designed analogues: correlation of antimicrobial, hemolytic, and lipid bilayer-perturbing activities.* Biochemistry, 2006. **45**(51): p. 15718-30.
37. Chen, J., et al., *Development of protegrins for the treatment and prevention of oral mucositis: structure-activity relationships of synthetic protegrin analogues.* Biopolymers, 2000. **55**(1): p. 88-98.
38. Bolintineanu, D., et al., *Antimicrobial mechanism of pore-forming protegrin peptides: 100 pores to kill E. coli.* Peptides, 2010. **31**(1): p. 1-8.
39. Fahrner, R.L., et al., *Solution structure of protegrin-1, a broad-spectrum antimicrobial peptide from porcine leukocytes.* Chem Biol, 1996. **3**(7): p. 543-50.
40. Qu, X.D., et al., *Protegrin structure and activity against Neisseria gonorrhoeae.* Infect Immun, 1997. **65**(2): p. 636-9.
41. Ostberg, N. and Y. Kaznessis, *Protegrin structure-activity relationships: using homology models of synthetic sequences to determine structural characteristics important for activity.* Peptides, 2005. **26**(2): p. 297-306.
42. Hancock, R.E., *Cationic peptides: effectors in innate immunity and novel antimicrobials.* Lancet Infect Dis, 2001. **1**(3): p. 156-64.
43. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.
44. Gidalevitz, D., et al., *Interaction of antimicrobial peptide protegrin with biomembranes.* Proc Natl Acad Sci U S A, 2003. **100**(11): p. 6302-7.
45. Mani, R., et al., *Solid-state NMR investigation of the selective disruption of lipid membranes by protegrin-1*. Biochemistry, 2004. **43**(43): p. 13839-48.
46. Gottler, L.M., et al., *Using fluorous amino acids to probe the effects of changing hydrophobicity on the physical and biological properties of the beta-hairpin antimicrobial peptide protegrin-1*. Biochemistry, 2008. **47**(35): p. 9243-9250.
47. Ishitsuka, Y., et al., *Insertion selectivity of antimicrobial peptide protegrin-1 into lipid monolayers: effect of head group electrostatics and tail group packing.* Biochim Biophys Acta, 2006. **1758**(9): p. 1450-60.
48. Mani, R., et al., *Membrane-bound dimer structure of a beta-hairpin antimicrobial peptide from rotational-echo double-resonance solid-state NMR.* Biochemistry, 2006. **45**(27): p. 8341-9.
49. Buffy, J.J., A.J. Waring, and M. Hong, *Determination of peptide oligomerization in lipid bilayers using 19F spin diffusion NMR.* J Am Chem Soc, 2005. **127**(12): p. 4477-83.
50. Buffy, J.J., et al., *Immobilization and aggregation of the antimicrobial peptide protegrin-1 in lipid bilayers investigated by solid-state NMR.* Biochemistry, 2003. **42**(46): p. 13725-34.

51. Yamaguchi, S., et al., *Solid-state NMR investigations of peptide-lipid interaction and orientation of a beta-sheet antimicrobial peptide, protegrin.* Biochemistry, 2002. **41**(31): p. 9852-62.
52. Heller, W.T., et al., *Membrane thinning effect of the beta-sheet antimicrobial protegrin.* Biochemistry, 2000. **39**(1): p. 139-45.
53. Heller, W.T., et al., *Multiple states of beta-sheet peptide protegrin in lipid bilayers.* Biochemistry, 1998. **37**(49): p. 17331-8.
54. Mani, R., et al., *Membrane-dependent oligomeric structure and pore formation of a beta-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR.* Proc Natl Acad Sci U S A, 2006. **103**(44): p. 16242-7.
55. Tang, M., A.J. Waring, and M. Hong, *Arginine dynamics in a membrane-bound cationic beta-hairpin peptide from solid-state NMR.* Chembiochem, 2008. **9**(9): p. 1487-92.
56. Tang, M., A.J. Waring, and M. Hong, *Phosphate-mediated arginine insertion into lipid membranes and pore formation by a cationic membrane peptide from solid-state NMR.* J Am Chem Soc, 2007. **129**(37): p. 11438-46.
57. Tang, M., et al., *Effects of guanidinium-phosphate hydrogen bonding on the membrane-bound structure and activity of an arginine-rich membrane peptide from solid-state NMR spectroscopy.* Angew Chem Int Ed Engl, 2008. **47**(17): p. 3202-5.
58. Sokolov, Y., et al., *Membrane channel formation by antimicrobial protegrins.* Biochim Biophys Acta, 1999. **1420**(1-2): p. 23-9.
59. Mangoni, M.E., et al., *Change in membrane permeability induced by protegrin 1: implication of disulphide bridges for pore formation.* FEBS Lett, 1996. **383**(1-2): p. 93-8.
60. Yang, L., et al., *Crystallization of antimicrobial pores in membranes: magainin and protegrin.* Biophys J, 2000. **79**(4): p. 2002-9.
61. Lam, K.L., et al., *Mechanism of supported membrane disruption by antimicrobial peptide protegrin-1*. J Phys Chem B Condens Matter Mater Surf Interfaces Biophys, 2006. **110**(42): p. 21282-6.
62. Khandelia, H. and Y. Kaznessis, *Structure of the antimicrobial beta-hairpin peptide protegrin-1 in a DLPC lipid bilayer investigated by molecular dynamics simulations.* Biochem. Biophys. Acta Biomembranes, 2007. **1768**: p. 509-520.
63. Roumestand, C., et al., *Oligomerization of protegrin-1 in the presence of DPC micelles. A proton high-resolution NMR study.* FEBS Lett, 1998. **421**(3): p. 263-7.
64. Tang, M., A.J. Waring, and M. Hong, *Intermolecular packing and alignment in an ordered beta-hairpin antimicrobial peptide aggregate from 2D solid-state NMR.* J Am Chem Soc, 2005. **127**(40): p. 13919-27.
65. Li, S., et al., *Water-protein interactions of an arginine-rich membrane peptide in lipid bilayers investigated by solid-state nuclear magnetic resonance spectroscopy*. J Phys Chem B, 2010. **114**(11): p. 4063-9.

66. Wi, S. and C. Kim, *Pore structure, thinning effect, and lateral diffusive dynamics of oriented lipid membranes interacting with antimicrobial peptide protegrin-1: 31P and 2H solid-state NMR study.* J Phys Chem B, 2008. **112**(36): p. 11402-14.
67. Vivcharuk, V. and Y. Kaznessis, *Free energy profile of the interaction between a monomer or a dimer of protegrin-1 in a specific binding orientation and a model lipid bilayer.* J Phys Chem B, 2010. **114**(8): p. 2790-7.
68. Vivcharuk, V. and Y. Kaznessis, *Thermodynamic Analysis of Protegrin-1 Insertion and Permeation through a Lipid Bilayer.* In preparation, 2010.
69. Vivcharuk, V. and Y.N. Kaznessis, *Dimerization of protegrin-1 in different environments.* Int J Mol Sci, 2010. **11**(9): p. 3177-94.
70. Darve, E., D. Rodriguez-Gomez, and A. Pohorille, *Adaptive biasing force method for scalar and vector free energy calculations.* J Chem Phys, 2008. **128**(14): p. 144120.
71. McQuarrie, D., *Statistical Mechanics*. 2000, Sausalito, California: University Science Books.
72. Reiss, H., H.L. Frisch, and J.L. Lebowitz, *Statistical mechanics of rigid spheres.* J Chem Phys, 1959. **31**(2): p. 369-380.
73. Hill, T.L., *Statistical Mechanics*. 1956, New York: McGraw-Hill Book Company, Inc.
74. Boublik, T., *Two-dimensional convex particle liquid.* Molecular Physics, 1975. **29**(2): p. 421-428.
75. Talbot, J., *Molecular thermodynamics of binary mixture adsorption: A scaled particle theory approach.* journal of chemical physics, 1997. **106**(11): p. 4696-4706.
76. Langham, A.A., A.S. Ahmad, and Y.N. Kaznessis, *On the nature of antimicrobial activity: a model for protegrin-1 pores.* J Am Chem Soc, 2008. **130**(13): p. 4338-46.
77. Im, W. and B. Roux, *Ion permeation and selectivity of OmpF porin: a theoretical study based on molecular dynamics, Brownian dynamics, and continuum electrodiffusion theory.* J Mol Biol, 2002. **322**(4): p. 851-69.
78. Im, W. and B. Roux, *Ions and counterions in a biological channel: a molecular dynamics simulation of OmpF porin from Escherichia coli in an explicit membrane with 1 M KCl aqueous salt solution.* J Mol Biol, 2002. **319**(5): p. 1177-97.
79. Corry, B., S. Kuyucak, and S.H. Chung, *Tests of continuum theories as models of ion channels. II. Poisson-Nernst-Planck theory versus brownian dynamics.* Biophys J, 2000. **78**(5): p. 2364-81.
80. Noskov, S.Y., W. Im, and B. Roux, *Ion permeation through the alpha-hemolysin channel: theoretical studies based on Brownian dynamics and Poisson-Nernst-Plank electrodiffusion theory.* Biophys J, 2004. **87**(4): p. 2299-309.
81. Dyrka, W., A.T. Augousti, and M. Kotulska, *Ion flux through membrane channels-An enhanced algorithm for the Poisson-Nernst-Planck model.* J Comput Chem, 2008. **29**(12): p. 1876-1888.

82. Furini, S., F. Zerbetto, and S. Cavalcanti, *Application of the Poisson-Nernst-Planck theory with space-dependent diffusion coefficients to KcsA.* Biophys J, 2006. **91**(9): p. 3162-9.
83. MacKerell, J., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R.L.;Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D.T.; Prodhom, B.; Reiher, III, W.E.; Roux, B.; Schlenkrich, M.; Smith, J.C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., *All-atom empirical potential for molecular modeling and dynamics Studies of proteins.* Journal of Physical Chemistry B, 1998. **102**: p. 3586-3616.
84. Hille, B., *Ionic Channels of Excitable Membranes*. 1984: Sinauer Associates.
85. Allen, T.W., Kuyucak, S., Chung, S.H., *The effect of hydrophobic and hydrophilic channel walls on the structure and diffusion of water and ions.* J. Chem. Phys., 1999. **111**: p. 7985-7999.
86. Jorgensen, W.L., et al., *Comparison of simple potential function for simulating liquid water.* J Chem Phys, 1983. **79**: p. 926-935.
87. Levitt, M.H., M.; Sharon, R; Laidig, K. E.; Daggett, V., *Calibration and Testing of a Water Model for Simulation of the Molecular Dynamics of Proteins and Nucleic Acids in Solution.* J. Phys. Chem. B, 1997. **101**: p. 5051-5061.
88. Paine, P.L. and P. Scherr, *Drag coefficients for the movement of rigid spheres through liquid-filled cylindrical pores.* Biophys J, 1975. **15**(10): p. 1087-91.
89. Kurnikova, M.G., et al., *A lattice relaxation algorithm for three-dimensional Poisson-Nernst-Planck theory with application to ion transport through the gramicidin A channel.* Biophys J, 1999. **76**(2): p. 642-56.
90. Allen, T.W., S. Kuyucak, and S.H. Chung, *Molecular dynamics estimates of ion diffusion in model hydrophobic and KcsA potassium channels.* Biophys Chem, 2000. **86**(1): p. 1-14.
91. Slotboom, J.W., *Iterative scheme for 1- and 2- dimensional d.c.-transistor simulation.* Electron. Lett., 1969. **5**: p. 677-678.
92. Hodgkin, A.L. and B. Katz, *The effect of sodium ions on the electrical activity of the giant axon of the squid.* J Physiol, 1949. **108**(1): p. 37-77.
93. Bolintineanu, D.S., et al., *Poisson-Nernst-Planck Models of Nonequilibrium Ion Electrodiffusion through a Protegrin Transmembrane Pore.* PLoS Comput Biol, 2009. **5**(1): p. e1000277.
94. Lehrer, R.I., et al., *Interaction of human defensins with Escherichia coli. Mechanism of bactericidal activity.* J Clin Invest, 1989. **84**(2): p. 553-61.
95. Lehrer, R.I., A. Barton, and T. Ganz, *Concurrent assessment of inner and outer membrane permeabilization and bacteriolysis in E. coli by multiple-wavelength spectrophotometry.* J Immunol Methods, 1988. **108**(1-2): p. 153-8.
96. Orlov, D.S., T. Nguyen, and R.I. Lehrer, *Potassium release, a useful tool for studying antimicrobial peptides.* J Microbiol Methods, 2002. **49**(3): p. 325-8.

97. Kedem, O. and A. Katchalsky, *Thermodynamic analysis of the permeability of biological membranes to non-electrolytes.* Biochim Biophys Acta, 1958. **27**(2): p. 229-46.
98. Katchalsky, A. and P.F. Curran, *Nonequilibrium Thermodynamics in Biophysics*. Harvard Books in Biophysics. 1965, Cambridge, Massachusets: Harvard University Press.
99. Goldman, D.E., *Potential, Impedance and Rectification in Membranes.* Journal of General Physiology, 1943. **27**: p. 37-60.
100. Rhoads, D.B., F.B. Waters, and W. Epstein, *Cation transport in Escherichia coli. VIII. Potassium transport mutants.* J Gen Physiol, 1976. **67**(3): p. 325-41.
101. Koch, A.L. and S. Woeste, *Elasticity of the sacculus of Escherichia coli.* J Bacteriol, 1992. **174**(14): p. 4811-9.
102. Yao, X., et al., *Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy.* J Bacteriol, 1999. **181**(22): p. 6865-75.
103. Vollmer, W., D. Blanot, and M.A. de Pedro, *Peptidoglycan structure and architecture.* FEMS Microbiol Rev, 2008. **32**(2): p. 149-67.
104. Klipp, E., et al., *Integrative model of the response of yeast to osmotic shock.* Nat Biotechnol, 2005. **23**(8): p. 975-82.
105. Paula, S., et al., *Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness.* Biophys J, 1996. **70**(1): p. 339-48.
106. Pagano, R. and T.E. Thompson, *Spherical lipid bilayer membranes: electrical and isotopic studies of ion permeability.* J Mol Biol, 1968. **38**(1): p. 41-57.
107. Levitt, D.G., *Kinetics of diffusion and convection in 3.2-A pores. Exact solution by computer simulation.* Biopolymers, 1973. **13**(2): p. 186-206.
108. Finkelstein, A., *Water Movement Through Lipid Bilayers, Pores and Plasma Membranes. Theory and Reality*. 1987, New York: John Wiley & Sons.
109. Sundararaj, S., et al., *The CyberCell Database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate in silico modeling of Escherichia coli.* Nucleic Acids Res, 2004. **32**(Database issue): p. D293-5.
110. Koch, A.L., *Shrinkage of growing Escherichia coli cells by osmotic challenge.* J Bacteriol, 1984. **159**(3): p. 919-24.
111. Shabala, L., et al., *Ion transport and osmotic adjustment in Escherichia coli in response to ionic and non-ionic osmotica.* Environ Microbiol, 2009. **11**(1): p. 137-48.
112. Chebotareva, N.A., B.I. Kurganov, and N.B. Livanova, *Biochemical effects of molecular crowding.* Biochemistry (Mosc), 2004. **69**(11): p. 1239-51.
113. Matts, T.C. and C.J. Knowles, *Stopped-flow studies of salt-induced turbidity changes of Escherichia coli.* Biochim Biophys Acta, 1971. **249**(2): p. 583-7.
114. Alemohammad, M.M. and C.J. Knowles, *Osmotically induced volume and turbidity changes of Escherichia coli due to salts, sucrose and glycerol, with particular reference to the rapid permeation of glycerol into the cell.* J Gen Microbiol, 1974. **82**(1): p. 125-42.

115. de Maranon, I.M., P. Gervais, and P. Molin, *Determination of cells' water membrane permeability: Unexpected high osmotic permeability of Saccharomyces cerevisiae.* Biotechnol Bioeng, 1997. **56**(1): p. 62-70.
116. Record, M.T., Jr., et al., *Responses of E. coli to osmotic stress: large changes in amounts of cytoplasmic solutes and water.* Trends Biochem Sci, 1998. **23**(4): p. 143-8.
117. Record, M.T., Jr., et al., *Biophysical compensation mechanisms buffering E. coli protein-nucleic acid interactions against changing environments.* Trends Biochem Sci, 1998. **23**(5): p. 190-4.
118. Feller, S.E. and R.W. Pastor, *On simulating lipid bilayers with an applied surface tension: periodic boundary conditions and undulations.* Biophys J, 1996. **71**(3): p. 1350-5.
119. Jensen, M.O., O.G. Mouritsen, and G.H. Peters, *Simulations of a membrane-anchored peptide: structure, dynamics, and influence on bilayer properties.* Biophys J, 2004. **86**(6): p. 3556-75.
120. Chiu, S.W., et al., *Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane.* Biophys J, 1995. **69**(4): p. 1230-45.
121. Klauda, J.B., et al., *Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types.* J Phys Chem B, 2010. **114**(23): p. 7830-7843.
122. Woolf, T.B. and B. Roux, *Structure, energetics and dynamics of lipid-protein interactions: a molecular dynamics study of the gramicidin A channel in a DMPC bilayer.* Proteins: Struct. Funct. Genet., 1996. **24**: p. 92-114.
123. Jo, S., et al., *CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes.* Biophys J, 2009. **97**(1): p. 50-8.
124. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program.* J Comput Chem, 2009. **30**(10): p. 1545-614.
125. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD.* J Comput Chem, 2005. **26**(16): p. 1781-802.
126. MacKerell, A.D., Jr., M. Feig, and C.L. Brooks, 3rd, *Improved treatment of the protein backbone in empirical force fields.* J Am Chem Soc, 2004. **126**(3): p. 698-9.
127. Feller, S.E., et al., *Constant pressure molecular dynamics simulation: The Langevin piston method.* J. Chem. Phys., 1995. **103**(11): p. 4613-4621.
128. Hoover, W.H., *Canonical dynamics: Equilibirum phase-space distributions.* Phys Rev A, 1985. **31**: p. 1695-1697.
129. Darden, T., D. York, and L.G. Pedersen, *Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems.* J Chem Phys, 1993. **98**(12): p. 10089-92.
130. Darden, T., et al., *Particle-Mesh Based Methods for Fast Ewald Summation in Molecular Dynamics Simulations.* 1998: p. 1-7.

131. Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendesen, *Numerical integration of the cartesian equations of motion for a system with constraints: molecule dynamics of n-alkanes.* J Comp Phys, 1977. **23**: p. 327-341.
132. Andersen, H.C., *Molecular dynamics simulations at constant pressure and/or temperature.* J Chem Phys, 1980. **72**: p. 2384-2393.
133. Nose, S., *Constant-Temperature Molecular-Dynamics.* Journal of Physics-Condensed Matter, 1990. **2**: p. Sa115-Sa119.
134. McQuarrie, D.A., *Statistical Mechanics*. 1976, New York: Harper Collins.
135. Hill, T.L., *An Introduction to Statistical Thermodynamics*. 1986, New York: Dover Publications Inc.
136. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy, minimization, and dynamics simulations.* J Comp Chem, 1983. **4**: p. 187-217.
137. MacKerell, J., A. D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* Journal of Physical Chemistry B, 1998. **102**: p. 3586-3616.
138. Neria, E., S. Fischer, and M. Karplus, *Simulation of activation free energies in molecular systems.* J. Chem. Phys., 1996. **105**: p. 1902.
139. Allen, M.P. and D.J. Tildesley, *Computer simulation of liquids*. 1987, Oxford: Oxford University Press.
140. Steinbach, P.J. and B.R. Brooks, *New spherical-cutoff methods for long-range forces in macromolecular simulation.* J Comput Chem, 1993. **15**(7): p. 667-683.
141. Ewald, P., *Die Berechnung optischer und elektrostatischer Gitterpotentiale.* Ann. Phys., 1921. **64**: p. 253-287.
142. deLeeuw, S.W., J.W. Perram, and E.R. Smith, *Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants.* Proceedings of the Royal Society of London. A. Math. and Phys. Sci., 1980. **373**(1752): p. 27-56.
143. Eastwood, J.W., R.W. Hockney, and D. Lawrence, *P3M3DP - the three-dimensional periodic particle-particle/particle-mesh program.* Comput. Phys. Commun., 1980. **19**: p. 323-365.
144. Feller, S.E., et al., *The Effect of Force Truncation on MD Simulations of Bulk Water Transport Properties and Interfacial Water Structural Properties.* J. Phys. Chem., 1996. **100**: p. 17011.

A Appendix: molecular dynamics simulation methods

This appendix presents several fundamental aspects pertaining to the method of molecular dynamics simulations implemented in chapter 5 of the present work. Other details specific to those simulations are given in section 5.2.

All computer simulations and theoretical developments discussed in this appendix are exclusively classical mechanical in nature; no quantum mechanical phenomena are explicitly considered. Although this is not fully justified on purely theoretical grounds (i.e. there are bond vibrations in the systems of interest with characteristic energies *hv* higher than mean molecular thermal energy, k_BT), over three decades of research and applications of classical mechanical molecular dynamics with empirical potential functions have shown this approximation to be valid for a wide range of biomolecular systems. It is worth mentioning that the empirical potential function used in classical MD simulations is strongly based on quantum mechanical calculations, so that many such effects are implicitly accounted for.

A.1 Overview

The method of molecular dynamics involves simultaneously solving Newton's equations of motion for all particles in a system, thus propagating the system forward in time. The forces acting on the different particles arise from interactions quantified by an *empirical* potential function. In fully atomistic simulations, each particle is represented as a point mass corresponding to an atom in the system of interest. Any given atom can undergo a variety of interactions, arising from covalent bonds, electrostatics, or van der Waals dispersion forces. These are all encompassed in the empirical potential function, which will be discussed in section A.4 below. The method of molecular dynamics can be summarized mathematically as follows:

Newton's equations of motions for all *N* atoms in the system are:

$$\boldsymbol{F}_i = m_i \ddot{\boldsymbol{r}}_i \quad i = 1, \ldots, N \tag{A.1}$$

where i is an index running over all N atoms, $\boldsymbol{F}_i$ is the force acting on atom i, m_i is the mass of atom i, $\ddot{\boldsymbol{r}}_i$ is the acceleration of atom i, $\ddot{\boldsymbol{r}}_i = d\boldsymbol{r}_i/dt$.

In a conservative system, we have:

$$\boldsymbol{F}_i = -\frac{\partial U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)}{\partial \boldsymbol{r}_i} \qquad i = 1, \ldots, N \tag{A.2}$$

Here, U is the total system potential energy, a function of the positions of all the atoms. The empirical form of U used in molecular dynamics simulations in the present work is discussed in section A.4. For numerical solution convenience, this second order system is reduced to a doubly sized system of first order equations:

$$\begin{aligned} -\nabla U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) &= m_i \dot{\boldsymbol{v}}_i \\ \boldsymbol{v}_i &= \dot{\boldsymbol{r}}_i \end{aligned} \tag{A.3}$$

Here, $\mathbf{v}_i$ is the velocity of atom i, and the single dot indicates first derivative with respect to time. This is an initial value problem consisting of system of $2N$ coupled ODEs (with the coupling arising from the dependence of U on the atomic positions, $\mathbf{r}_i$). The systems of interest in the present work typically consist of 80,000 atoms or more (N>80,000); additionally, any realistic potential function U entails significant algebraic complexity (see section A.4). Thus, the integration of the above system of equations is no simple task, and requires significant algorithmic developments as well as extensive computational resources. Several numerical schemes have been employed for the solution of . For the present work, we have employed an explicit first order scheme known as the leapfrog integration algorithm. Given positions at time t and velocities at time $t+\Delta t/2$:

$$\boldsymbol{r}_i(t + \Delta t) = \boldsymbol{r}_i(t) + \boldsymbol{v}_i\left(t + \frac{\Delta t}{2}\right)\Delta \mathrm{t} \tag{A.4}$$

$$\boldsymbol{v}_i\left(t+\frac{3\Delta t}{2}\right)=\boldsymbol{v}_i\left(t+\frac{\Delta t}{2}\right)+\frac{-\nabla U\left(\boldsymbol{r}_1(t+\Delta t),\ldots,\boldsymbol{r}_N(t+\Delta t)\right)}{m_i}\Delta t$$

Provided the time step Δt is sufficiently small, this method constitutes an integration scheme that is both fast and accurate. As with all explicit integration methods, numerical stability cannot be guaranteed. In order to circumvent this, the time step is chosen to be significantly smaller than the characteristic time of the fastest motion expected in such systems. Typically, these are the oscillations of any covalent bonds involving hydrogen, with frequencies on the order of 10^{14} Hz, which dictates a time step no greater than 1 femtosecond (1 fs = 10^{-15}s). However, it is common practice to apply constraints to all bonds involving hydrogen, which allows for a safe integration time step of 2 fs. This value is somewhat flexible, but has been selected in accordance with the vast majority of studies employing the same empirical energy function for similar systems. Numerical instability can still occur if the system is initialized in a physically unrealistic state, and so great care must be taken in the initial set up of the system. It should also be noted that while the equations of motion for constant temperature and pressure (see section A.2) differ significantly from equations (A.3), the preceding discussion of integration routines still applies.

A.2 Simulations in other ensembles

Clearly, a system of several thousand atoms as discussed above is not at the thermodynamic limit. Statistical mechanics offers a connection between such microscopic systems and the observable properties of the macroscopic systems they represent. Solution of the above equations for a system at a constant volume results in sampling of the phase space trajectory of a closed, isolated system (commonly referred to as NVE, where N stands for constant number of atoms, V for constant volume, E for constant energy)[71]. In order to sample a system that more closely resembles

experimental and physiological conditions, additional controls must be implemented. The most appropriate statistical mechanical ensemble for the present work is the isothermal-isobaric ensemble, denoted as NPT (N for constant number of atoms, P for constant pressure, T for constant temperature). It should be noted that in the limit of large N (the thermodynamic limit), all ensembles are equivalent; however, the NPT ensemble has the advantage that one knows a priori the appropriate pressure and temperature values to use (i.e. ambient, physiological conditions) in order to observe the system behaviour of interest. The equations of motion must be modified to simulate an NPT system. The basic idea is to introduce additional degrees of freedom in the system, representing the kinetic and potential energy of a virtual piston, and couple these to the particle positions and velocity in a way that allows the volume to change in response to pressure fluctuations. The quantity analogous to the position of the piston is taken to be the system volume V (i.e. the volume of the central simulation box), which makes the rate of volume expansion (dV/dt) analogous to the "velocity" of the piston. The natural choice for the piston potential energy is the product of the volume with its conjugate variable, the pressure:

$$U_{piston} = P_{ext}V \tag{A.5}$$

Here, P_{ext} is the pressure set point for the controller. The particle positions and velocities are then scaled according to $V^{1/3}$, which represents the coupling between particles and the additional degrees of freedom of the piston. This development is largely due to Andersen[132].

Similarly, temperature control is achieved by the incorporation of an additional set of degrees of freedom into the equations of motion. A thermal bath is taken to represent an additional element in the system, and the coupling to particles appears only through the scaling of velocity. The thermal bath degrees of freedom are less physically intuitive quantities: W represents the thermal inertia that is analogous to the mass of the thermal bath, while s is a time-scaling parameter, which effectively represents the "positional" degree of freedom of the thermal bath and couples this to the particle

velocities. The potential energy term associated with the thermal bath was shown by Nosé [133] to be of the form:

$$U_{thermalbath} = (f+1)k_B T_{eq} \ln s \tag{A.6}$$

Here, f is the total number of degrees of freedom in the system, k_B is Boltzmann's constant, and T_{eq} is the controller's temperature set point.

To incorporate both temperature and pressure controls and simulate the NPT ensemble, the equations of motion are derived based on the Lagrangian equations of motion. The derivation is omitted here for brevity, but the final expressions are as follows:

$$\begin{aligned}
\dot{\boldsymbol{r}}_i &= \frac{\boldsymbol{p}_i}{m_i} + \frac{\dot{V}}{3V} r_i \\
\dot{\boldsymbol{p}}_i &= \frac{\partial U(\boldsymbol{R})}{\partial \boldsymbol{r}_i} - \left(\frac{\dot{V}}{3V} - \zeta\right)\dot{\boldsymbol{p}}_i \\
Q\ddot{V} &= 9V^2(P - P_{ext}) - \dot{\zeta}\dot{V}Q + \frac{Q\dot{V}^2}{V} \\
W\ddot{\zeta} &= \sum_{i=1}^{N} \frac{\|\boldsymbol{p}_i\|^2}{m_i} + \left(\frac{\dot{V}}{3V}\right)^2 Q - (f+1)k_B T_{eq}
\end{aligned} \tag{A.7}$$

Here, V is the system volume, $\boldsymbol{r}_i$, m_i and $\boldsymbol{p}_i$ represent the position, mass and momentum of particle *i*, *Q* is the effective pressure piston mass, *P* is the instantaneous pressure, and ζ is the thermal bath friction coefficient, defined as $\zeta = s/\dot{s}$. In all cases, the subscript $\dot{}$ denotes the first derivative with respect to time, while $\ddot{}$ denotes the second derivative with respect to time.

A.3 Connection to thermodynamics

Statistical mechanics tells us that the macroscopic, observable value of a thermodynamic property of interest, *M*, can be calculated from a weighted average of all

possible microscopic states of a system – known as an ensemble average. For the isothermal-isobaric ensemble, in the classical approximation, this can be written as [134-135]:

$$\langle M \rangle_{NPT} = \frac{1}{\Delta_{NPT}} \int_0^\infty \int_V \dots \int_{-\infty}^\infty M(\boldsymbol{r}^N, \boldsymbol{p}^N) \exp\left(-\frac{H(\boldsymbol{r}^N, \boldsymbol{p}^N) + PV}{k_B T}\right) d\boldsymbol{p}^N d\boldsymbol{r}^N dV \quad \text{(A.8)}$$

where Δ_{NPT} is a quantity known as the isothermal-isobaric partition function, given by:

$$\Delta_{NPT} = \frac{1}{h^{3N} N! V_0} \int_0^\infty \int_V \dots \int_{-\infty}^\infty \exp\left(-\frac{H(\boldsymbol{r}^N, \boldsymbol{p}^N) + PV}{k_B T}\right) d\boldsymbol{p}^N d\boldsymbol{r}^N dV \quad \text{(A.9)}$$

In equation (A.8), M is any experimentally observable property of the system, and $\boldsymbol{r}^N$ and $\boldsymbol{p}^N$ represent coordinate and momentum vectors of all atoms in the system, respectively ($\boldsymbol{r}_1, \dots, \boldsymbol{r}_N$ and $\boldsymbol{p}_1, \dots, \boldsymbol{p}_N$). The integrals over momenta run from -∞ to ∞, and the integrals over the positions are carried out over the bounds of the system. N is the number of particles in the system, h is Planck's constant, V_0 is an arbitrary unit of volume introduced for dimensional consistency, P and T are the fixed system pressure and temperature. H is the (classical) system Hamiltonian, defined as the sum of the kinetic and potential energies, and is a function of the momenta and coordinates of all atoms. For all but the simplest, idealized systems, the potential energy (and thus the Hamiltonian) depends on atom positions in a complex manner. This dependence renders equation (A.8) intractable to analytical evaluation, and is the reason for much of current research in equilibrium statistical mechanics [135].

The type of average that can be obtained from a molecular dynamics simulation is a time average. The ergodic hypothesis, a fundamental postulate of statistical mechanics, equates the ensemble average to a time average in the limit of very long time:

$$\langle M \rangle_{ensemble} = \lim_{t \to \infty} \frac{1}{t - t_0} \int_{t_0}^{t} M(t) dt = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} M_i \qquad \text{(A.10)}$$

The integral representation corresponds to continuous time sampling, whereas the summation represents discrete sampling, as is the case in computer simulations (here, *N* represents the number of discrete samples). Clearly, a computer simulation cannot sample an infinitely long period of time. As such, any estimates of experimentally observed properties obtained from molecular dynamics simulations are liable to suffer from sampling and convergence issues. Depending on the property of interest, convergence can be obvious or extremely difficult to determine.

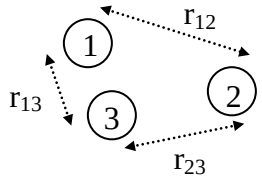
A.4 The CHARMM potential function

The approximations arising from classical treatment and finite time simulation of atomic systems are often less limiting than those arising from the use of an empirical potential function. In the present work, we have employed the Chemistry at Harvard Molecular Mechanics (CHARMM) software suite [136] and potential energy function [137], which is described next in some detail. This potential function is the result of over two decades of research by numerous investigators, combining experimental data, first principles quantum mechanical computations, and extensive testing. It is one of the most widespread potential functions used for biomolecular systems, and has been successfully applied to systems containing proteins, lipids, nucleic acids, small molecule drugs, various solvents and electrolytes. For an overview of the main strategies used in obtaining the extensive parameter set associated with this potential function, the reader is referred to [137].

In the version of the CHARMM potential function we employ, all atoms are treated explicitly. That is, every atom of every molecule in the system is represented by a single point mass with appropriate parameters, which govern its interactions with all other atoms. A vital approximation common to most potential functions used in biomolecular studies is the assumption of pairwise additivity. This treats the *direct*

interactions between any given pair of atoms as independent of the presence of any other atoms. Thus, for a simple system containing three atoms, the total potential energy is the sum of three pairwise interactions:

Figure A-1: **Illustration of the concept of pairwise interactions**



$$U_{total} = U_{12}(r_{12}) + U_{13}(r_{13}) + U_{23}(r_{23})$$

More generally, for N atoms, the total potential energy is given by:

$$U_{total} = \sum_{i=1}^{N} \sum_{j>i}^{N} U_{ij}(r_{ij}) \tag{A.11}$$

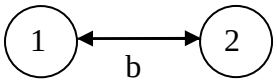
where r_{ij} is the separation distance between atoms *i* and *j*.

Pairwise interactions arise either from covalent bonding interactions, or non-bonded, long-range interactions. The typical version of the CHARMM potential [137] has the following functional form; the individual terms are explained below.

$$\begin{aligned} U(\boldsymbol{r}) = & \sum_{dihedrals} K_\chi(1 + \cos(n\chi - \delta) + \\ & + \sum_{\substack{Urey-Bradley \\ pairs}} K_{UB}(S - S_0) + \sum_{\substack{improper \\ dihedrals}} K_\varphi(\varphi - \varphi_0) + \\ & + \sum_{\substack{nonbonded \\ pairs}} \frac{q_i q_j}{\varepsilon_1 r_{ij}} + \sum_{\substack{nonbonded \\ pairs}} \varepsilon \left[\left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{12} + \left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{6} \right]^{12} \end{aligned} \tag{A.12}$$

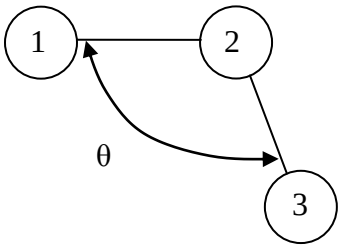
A.4.1 Bonded terms

Covalent bonds (commonly referred to as 1,2 interactions) are modeled as simple Hookean springs, with a force constant K_b and a minimum energy bond length b_0, both of which depend on the atom types involved in the bond (for example, the C-H bond from the α-carbon of an amino acid will have a different set of K_b and b_0 parameters than the C-H bond in an aliphatic lipid tail).



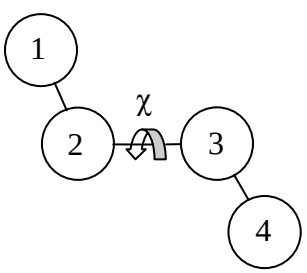
$$U = K_B(b - b_0)^2$$

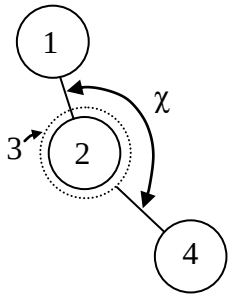
Angles between atoms separated by two covalent bonds (1,3 interactions) are modeled as simple angular harmonic oscillators, with a force constant K_θ and an optimal angle θ_0, which again depend on the specific atoms involved at positions 1, 2 and 3.



$$U_\theta = K_\theta(\theta - \theta_0)^2$$

Dihedral angles represent torsions about bonds, and are modeled by a periodic potential. K_χ is the dihedral angle force constant, which effectively controls the stiffness of rotation. χ is the variable dihedral angle (see figure and attached explanation), n is the periodicity number of the torsion, and delta is the phase shift corresponding to the equilibrium value.

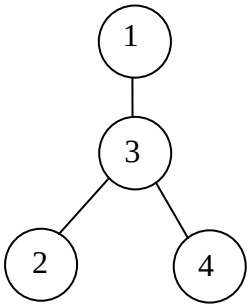




$$U_\chi = K_\chi(1 + \cos(n\chi - \delta))$$

Side view View along 2-3 bond

Other bonded potential terms include 1,3 Urey-Bradley interactions, modeled as harmonic oscillators. This potential term is only included as a minor correction for a limited set of atom pairs to optimize fits to vibrational spectra [137]. CHARMM additionally employs an improper dihedral angle term for certain atom bonding configurations, which is modeled as a harmonic oscillator. The improper dihedral angle term is used to maintain planarity about sp^2 hybridized carbons, and to prevent chirality switches about chiral centres. As with regular dihedral angles, four atoms are used to define the angle – the geometric angle is then the angle between the normal of the plane containing atoms 124 and that containing atoms 234 (see figure below). More recent enhancements to the CHARMM force field [126] also include the addition of empirically-derived cross-terms for protein backbone dihedral angles to improve the match to experimentally determined secondary structural motifs.



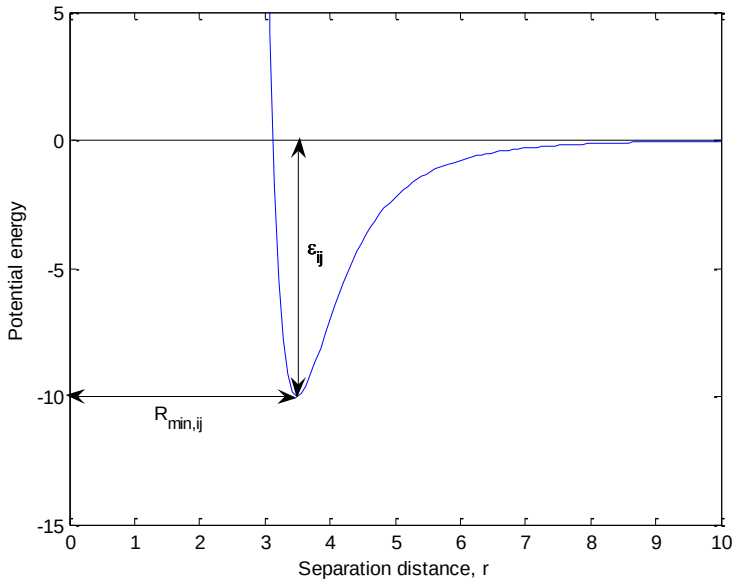
$$U_\varphi = K_\varphi(\varphi - \varphi_0)$$

The angle is defined as the angle between the normal of the plane containing atoms 1,2,4 and that containing atoms 2,3,4. In this case, planarity can be maintained around atom 3 by having a φ_0 value of 0 (i.e. atoms 1, 2, 4 and 2, 3, 4 are coplanar at the improper dihedral angle energy minimum).

A.4.2 Non-bonded terms

By far the most computationally demanding aspect of a classical molecular dynamics simulation is the evaluation of non-bonded energy terms. This arises from the much higher number of pairwise nonbonded interactions that any given atom can participate in as compared to its bonded interactions. The two phenomena responsible for long-range interatomic forces are coulombic attractions and van der Waals dispersion forces. The former is simply Coulomb's law for two point charges, and the latter is modeled via a 6-12 Lennard-Jones potential, depicted in Figure A-2 below (see last two terms of equation (A.12) above). Point charges are parametrized individually for each atom in a given molecule, and are usually based on quantum mechanical electronic structure calculations. The dielectric constant used for explicit solvent simulations is unity, since the simulations represent a fully atomistic system with vacuum between atoms.

Figure A-2: The 6-12 Lennard-Jones potential used to model van der Waals dispersion forces.



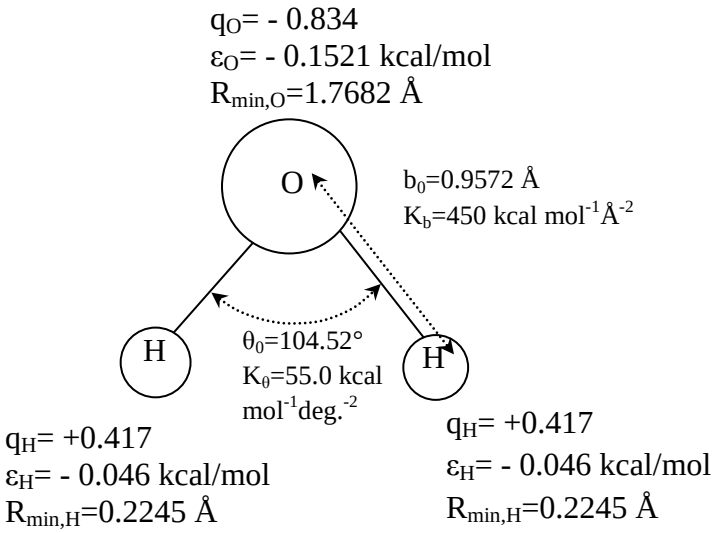
A given atom type *i* will have a unique pair of ε_i, $R_{min,i}$ values. The Lennard-Jones constants for the interaction of a pair of unlike atoms (say, *i* and *j*) are obtained from the Lorentz-Berthelodt combination rules, wherein $\varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2}$, $R_{min,ij} = (R_{min,i} + R_{min,j}) / 2$.

Non-bonded interactions are computed for all atoms separated by three or more covalent bonds (1,4 interactions or higher, including dihedral angles). Atoms directly bonded (1,2) or separated by two covalent bonds (1,3) interact only through their bonded terms. Hydrogen bonding is not explicitly included in the latest versions of the CHARMM potential function (as well as most modern biomolecular force fields), since appropriately parametrized Coulomb and van der Waals energy terms for the relevant acceptor and donor atoms capture these effects sufficiently well [137-138].

A.4.3 TIP3P – the explicit atom water model used in CHARMM

Numerous water models have been used in explicit simulations. The model recommended for the CHARMM force field is a three-site model due to Jorgensen [86], with each site representing a single atom. The O-H bond lengths are constrained using the SHAKE algorithm [131], as is the H-O-H angle, so that rigid body rotation and translation are the only degrees of freedom available. Although this model is not ideal for predicting the dynamic properties of pure water (for instance, it overestimates the water self-diffusion coefficient by a factor of 2), it has been found to be satisfactory in molecular simulations of biomolecular systems. Its widespread use can be largely attributed to its simplicity and the resulting computational savings. This model for the simplest molecule in an explicit solvent simulation is shown below to illustrate some of the parameters discussed. With the use of SHAKE constraints [131], a fictitious H-H bond is introduced to constrain the H-O-H angle to the equilibrium value dictated by the geometry shown.

Figure A-3: TIP3P, the three-site water model used with the CHARMM potential function



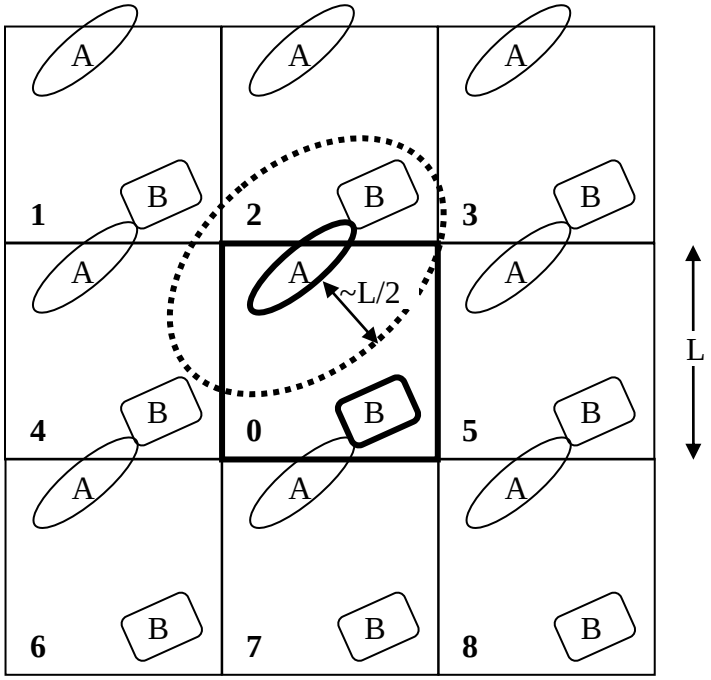
A.5 Periodic boundary conditions

In any real biomolecular situation, the molecule(s) of interest are surrounded by an effectively infinite amount of solvent, so that the boundaries of the container (be it living tissue or a beaker) do not directly influence molecular behaviour. In fully atomistic molecular simulations, it is generally not feasible to incorporate sufficient explicit solvent molecules to reproduce this effect, and some boundary must be introduced relatively close to the molecule(s) of interest. If this is a wall (rigid or soft), its presence will have an undesirable influence on molecular behaviour; if no boundary is used (vacuum), solvent molecules near the edge may effectively evaporate, and ambient pressure cannot be maintained. One of the most widely adopted solutions to this problem is the use of periodic boundary conditions (PBC). In the PBC approach, the simulation system consists of the biomolecule(s) of interest with sufficient solvent placed in a central box (for simplicity, a rectangular box is used here; other shapes are possible and

indeed advantageous in many cases). The central box is then surrounded with identical images of itself, as shown in Figure A-4 below.

Figure A-4: Periodic boundary conditions.

The central box is replicated and translated in all directions to completely fill the space around it. For simplicity, a 2D representation with 8 image boxes is shown here (a three-dimensional orthorhombic PBC setup requires 26 image boxes).

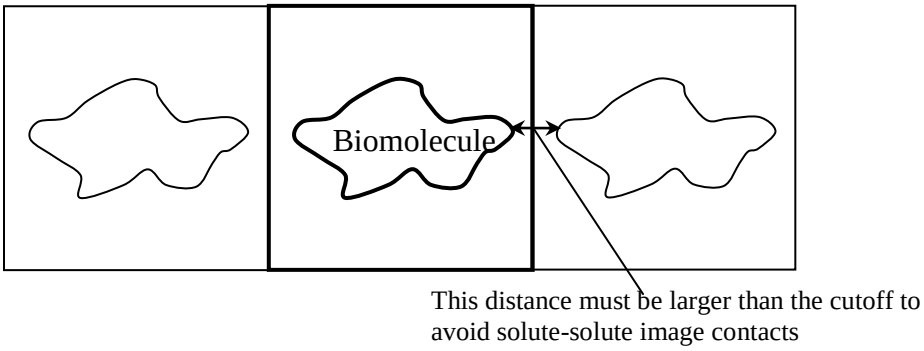


A molecule at the edge of the primary box thus interacts with molecules in the adjacent image box, which are identical to molecules at the opposite side of the primary box (for instance, molecule A interacts with B in image box 2, which is the image replica of B in the primary box). In order to avoid aphysical behaviour due to the enforced system periodicity, it is generally required that no molecule interact with image copies of itself, or with multiple image copies of another molecule (molecule A interacts with B in image box 2, but not B in the primary box or any other images of B; also, it does not interact

with any of its own images). For point charges, this can be achieved by truncation of the long-range forces (van der Waals and Coulombic) at a distance no greater than half the box length (L/2). In the illustration above, this cutoff range is shown for molecule A as a dotted perimeter. This is commonly referred to as the minimum image convention [139].

In the case of simulations of large biomolecules consisting of many atoms, it is generally desirable to avoid having the central biomolecule(s) (henceforth referred to as the solute) come into contact with any parts of its own images. This corresponds to a system at infinite dilution, in that the behaviour of the solute molecule being studied is unaffected by the presence of other solute molecules. In practice, the truncation distances of long-range forces are dictated by physical arguments, and specific values are recommended for the empirical force field. Typical biomolecules have dimensions exceeding the cutoff values, so that the requirement for avoiding solute-solute image contacts is the presence of a solvent layer between the primary solute and its nearest image of thickness greater than the cutoff distance. This is illustrated in the figure below.

Figure A-5: Determining box size for a given solute



The size of the solute molecule and the cutoff values thus determine the amount of solvent required and the size of the primary box. Note that because a solute molecule can

and will rotate during a simulation, it is its largest dimension that determines all three dimensions of the primary box:

Box dimension = largest dimension of solute molecule + cutoff distance + ~3Å buffer

An additional "safety buffer" of solvent is typically added to accommodate conformational changes of the solute and fluctuations in box size that occur during constant pressure simulations. The value of this buffer depends on the extent of elongation expected for the solute, as well as box size fluctuations in constant pressure simulations.

In our simulations of anisotropic systems (e.g. peptide-bilayer systems), we have used a hexagonal prism as a central box shape. This offers a significant advantage in that less solvent can be used to provide the same solvent "buffer" between the primary solute and its images, allowing for a smaller system and increased computational efficiency. By the same argument, a rhombic dodecahedral box shape is useful in simulations of isotropic systems.

A.6 Treatment of long-range forces

The choice of cutoff distances entails a crucial compromise between accuracy and computational speed, usually made based on force field recommendations and experience. Simple truncation schemes involve setting both the electrostatic and the van der Waals forces to zero at a specified distance. However, this introduces a slight discontinuity at the cutoff distance, which results in an artificial heating effect as atoms move back and forth across cutoff distances. For a review of the numerous options available in the treatment of long-range forces, the reader is referred to [140]. We will now only describe the methods employed in the current work.

A.6.1 Van der Waals forces

For the treatment of van der Waals forces, we have used a potential switching function, which smoothly switches the potential to zero over a specified distance near the cutoff. This approach circumvents the artificial heating effect described above, while maintaining an accurate potential. Since the van der Waals potential term decays as r^{-6}, the truncation method for this term is much less critical. Typical values are 8 Å for the beginning of the switching domain (r_{on}=8) and 11 Å for the end (r_{off} = 11). Mathematically, the switch-truncated van der Waals potential term can be expressed as follows:

$$
\begin{aligned}
U_{vdW,i-j} &= \varepsilon\left[\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{6}\right] && when\ r_{ij} \le r_{on} \\
&= \frac{\varepsilon\left[\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{6}\right]\left(r_{off} - r\right)^{2}\left(r_{off} + 2r - 3r_{on}\right)}{\left(r_{off} - r_{on}\right)^{3}} && when\ r_{on} < r \le r_{off} \\
&= 0 && when\ r > r_{off}
\end{aligned}
\tag{A.13}
$$

While similar schemes are still being used in electrostatic potential truncation, it has been found that the Ewald summation algorithm provides a superior envelope of accuracy and computational efficiency. The relative complexity of this algorithm is justified, since the electrostatic potential decays as r^{-1}, much more slowly than the van der Waals interactions. Furthermore, the importance of electrostatic effects in biomolecular interactions cannot be overstated (particularly for AMP systems); thus, an algorithm for the accurate and efficient computation of electrostatic forces is indispensable.

A.6.2 Ewald summation

The Ewald summation was originally developed by Peter Ewald in 1921 [141] to calculate the electrostatic energy of the unit cell of an ionic crystal. Due to the use of

periodic boundary conditions, as described in section A.5 above, the systems in the present work can be treated similarly, with the central box equivalent to the crystal unit cell. The basic problem is to formulate an expression for the total potential energy of the central unit cell due to a *fixed* periodic charge distribution. The derivation starts with a general expression, which can in principle be evaluated using Coulomb's law for all pairs over the entire lattice to determine the electrostatic potential:

$$U(\boldsymbol{r}) = \sum_{i=1}^{N} q_i \varphi(r_i) = \frac{1}{2}{\sum_{\boldsymbol{n}}}' \sum_{i=1}^{N} \sum_{j=1}^{N} q_i q_j / (4\pi\varepsilon_0 \left| \boldsymbol{r}_i - (\boldsymbol{r}_j + \boldsymbol{n}) \right| \qquad \text{(A.14)}$$

The indices *i* and *j* represent the particles in the unit cell, φ is the electrostatic potential, ε_0 is the permittivity of free space, and $\boldsymbol{r}$ is the position vector. The sum over $\boldsymbol{n}$ indicates a sum over all lattice basis vector combinations (for a cubic system of side length L, the vectors are (L,0,0), (0,L,0),(0,0,L)). The factor of 1/2 is used to compensate for double-counted pairs (*i-j, j-i*), and the ` symbol indicates that the terms with **n**=0 and i=j are neglected. As it stands, this sum is conditionally and very slowly convergent. The basic strategy of the Ewald summation is to convert this slowly convergent sum into two sums, one in real space and the other in Fourier space; with appropriate parameter choices, both sums can be made to converge rapidly, resulting in an algorithm that is both fast and accurate. This procedure can be carried out in a purely mathematical fashion, as has shown by deLeeuw et al [142]. We only briefly outline an alternative approach, based on more physical arguments.

The basic idea behind the Ewald summation technique is to add two charge densities (i.e. spatially-distributed electrostatic charges) to the original charge distribution of the system, denoted ρ_1. The first of these, denoted ρ_2, consists of Gaussian functions with variance κ^2 equal in magnitude and opposite in charge compared to the original charge distribution, while the second (ρ_2) consists of Gaussian charges that are equal in magnitude to the original. The net effect of these added distributions cancels by the

superposition principle of electrostatics, so that the augmented system is equivalent to the original one. However, the ingenuity in the Ewald summation lies in noting that the electrostatic potential due to ρ_1 and ρ_2 is short-ranged due to mutual screening effects, which allows for a relatively simple solution based on the Poisson equation. Meanwhile, a solution in Fourier space can be obtained for the electrostatic potential due to ρ_3. The final expression for the electrostatic potential energy of the unit cell is:

$$U(\boldsymbol{r}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sum_{\boldsymbol{n}}{}' \frac{q_i q_j \,\mathrm{erfc}(\kappa|r_i - (r_j + n)|)}{4\pi\varepsilon_0|r_i - (r_j + n)|} + \sum_{\boldsymbol{k}\neq 0}\sum_{j=1}^{N}\frac{4\pi q_i q_j}{\varepsilon_0 V_0 k^2}\exp\left(-\frac{k^2}{4\kappa^2}\right)\exp\left(i\boldsymbol{k}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j)\right)\right) - \frac{2\pi}{3V}\left|\sum_{i=1}^{N} q_i \boldsymbol{r}_i\right|^2 - \frac{\kappa}{2\pi^{3/2}\varepsilon_0}\sum_{i=1}^{N} q_i^2 \qquad \text{(A.15)}$$

As usual, forces can be easily obtained by differentiation of the above with respect to particle position, $\mathbf{r}_i$. Note that the parameter κ governs the relative convergence of the real-space and Fourier space summations [139]. A large value of κ, corresponding to sharp Gaussians, results in a quickly converging real space sum (since erfc(x) → 0 quickly as x becomes large). However, this requires more terms in the Fourier sum (since $\exp(-k^2/4\kappa^2)$ requires a higher k value to vanish if κ is large). In practice, the value of κ is chosen so that the real space summation is sufficiently accurate when truncated at the same distance as the van der Waals terms (~11Å), and the Fourier summation limit is adjusted accordingly. This optimization is relatively straightforward, and can be carried out to adequate precision with a simple line search. The NAMD software package [125] which employs the CHARMM force field includes an automatic Ewald summation optimization algorithm.

A further improvement on the Ewald summation algorithm came with the introduction of the paricle-mesh Ewald summation (PME) technique [143-144]. This technique treats the real-space part of the Ewald sum in the same manner, but employs several modifications for the Fourier-space summation. First, the canceling charge

distribution (ρ_3) is interpolated onto a regularly-spaced grid in the simulation box. This allows the use of the fast Fourier transform (the discrete version of the Fourier transform) to solve Poisson's equation for the canceling distribution, and obtain the required electrostatic potential. Forces are then computed using discrete derivatives. This method thus eliminates the need for the $O(N^2)$ summation of the long-range potential in equation A.14 above. Due to advances in computational algorithms for grid interpolation and fast Fourier transforms, the PME technique has become the method of choice for electrostatic calculations in molecular dynamics simulations.