

ELUCIDATING SELF-ASSEMBLY
AND ANTIMICROBIAL STRATEGIES
OF SYNTHETIC PEPTIDES:
AN IN SILICO INVESTIGATION

Irene Marzuoli

RANDALL CENTRE FOR CELL AND MOLECULAR BIOPHYSICS
KING'S COLLEGE LONDON



THIS DISSERTATION IS SUBMITTED FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

SEPTEMBER 2019

Declaration

This dissertation describes work I have carried out between October 2016 and September 2019 at the Randall Centre for King's College London, under the supervision of Professor Franca Fraternali (first supervisor) and Dr. Chris D. Lorenz (second supervisor).

This dissertation contains material appearing in the following articles:

- ...

In addition to the above, I have contributed to the following publications during the course of my PhD:

- ...

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. It has not been submitted in whole or in part for any degree or diploma at this or any other university.

Irene Marzuoli
September 2019

Acknowledgements

...

Summary

Elucidating self-assembly and antimicrobial strategies of
synthetic peptides an in silico investigation

Irene Marzuoli
King's College London

... ...

Contents

1	Introduction	11
1.1	Antimicrobial resistance	14
1.2	Alternative antibiotic strategies: antimicrobial peptides	19
1.3	Gene therapy	33
1.4	Delivery of therapeutic material	35
1.5	Closing the circle: an antimicrobial drug delivery vehicle	39
1.6	A computational approach to understand capzip	46
2	Methods	48
2.1	Algorithms for Molecular Dynamics	49
2.2	The force field problem	52
2.3	Beyond the classical framework	61
2.4	MD and experiments: ensembles versus averages	63
2.5	MD simulations: successes	65
3	Capzip simulations	77
3.1	Modelling the assembly	77
3.2	Simulations details	81
3.3	Analysis	88
3.4	Results: capsule in solution	91
3.5	Results: peptide-membrane interactions	100
3.6	Rest	104
3.7	Additional material	115
4	Lipid parametrisation	117
4.1	Additional material: effects of membrane undulation	117
	Appendices	170

A.1	On the derivation of the GP predictive distribution	170
-----	---	-----

List of Figures

1.1	Graphical abstract of introduction	13
1.2	Mechanisms of antimicrobial resistance to small drugs	16
1.3	Modes of action of antimicrobial peptides	20
1.4	Structure of Gram-positive and -negative cell envelope.	21
1.5	Structure of some known AMPs.	24
1.6	Principles of gene therapy	34
1.7	Cazip molecule	40
1.8	Experimental results on capzip	43
2.1	MD force field classifications	55
2.2	SIRAH force field amino acid and DNA description	58
2.3	MARTINI force field mapping for amino acids, lipids and water .	60
2.4	Disordered pore model	69
2.5	Polyarginine translocation	71
2.6	Peptide amphiphiles assembly through MARTINI simulations .	75
3.1	Hydrogen bonds between RRWTWE β -sheet	78
3.2	Building blocks of capzip assembly	79
3.3	Cohesion measures on the pentagonal subunit	80
3.4	Snapshot of simulations of capzip on model membranes	82
3.5	Atomistic run of buckyball in solution: final configuration	92
3.6	Structural measures on buckyball in solution	93
3.7	Correlation of motion between molecules of the buckyball	105
3.8	Branch pairing during simulations of the buckyball	105
3.9	Contacts between molecules during simulations of the buckyball .	106
3.10	Hydrogen bonds in the buckyball molecule	107
3.11	SASA per residue of a buckyball in solution	107
3.12	Non-bonded protein energy contribution to capsule structures .	108

3.13	Comparison of monolayer and bilayer structural properties	109
3.14	Snapshot of SIRAH monolayer simulation	110
3.15	SASA per residue of monolayer and bilater	110
3.16	ApL, D and S_{CD} for membrane simulations	111
3.17	Hexagonal order parameter of lipids in protein-lipid simulation . .	112
3.18	Snapshot from relevant membrane-peptide simulations	114
4.1	Comparison between methods to compute ApL	168
4.2	Control computation on bilayer electron density calculation	169

Chapter 1

Introduction

SCIENCE is a theoretical framework that explains the reality of nature. The scientist's vocation is to navigate this huge gap, in an effort to bring theory and reality closer, by allowing the understanding of the world where we live. Out of the many science fields at service of human wealth, the most challenging and still far away from being mastered is the comprehension and manipulation of the human body and mind. It is striking how we are finally scratching the understanding of these two entities, mind and body, in the same historical moment, at a point where computers start imitating human reasoning [Editor BBC, 2016; Google, 2016; Alom et al., 2019] and biological materials are turned into semi-functional organs [Rossi et al., 2018].

The process of understanding is intermingled with the process of creation, as scientific knowledge funded the field of technology. In the past two centuries we witnessed a technology evolution that condensed the collective expertise we have accumulated so far into objects belonging, by now, to our daily life. Thus, while we were deconvolving the mysteries of nature, human inventions have become increasingly convoluted.

This boost of technology allowed for the development of novel instruments, and in turn for the accumulation of a broader scientific knowledge. Consequently, the times of all-round scientists left the way to an era in which specialisation is necessary to master and understand a subject, in the hope and trust that piecewise knowledge builds an organic corpus once the effort of many scientists is joined together.

The increasing understanding of how we work - or fail to do so sometimes - together with the discovery of bacteria and the symbiotic or disruptive rela-

tionship we have with them, poses novel questions to science: why the body, which is carefully design to function effectively, fails sometimes? And why in these cases it defends itself also against beneficial drugs? But mostly, what can we do to repair such faults?

Our biological and emotional push to change the course of nature to improve our lives lead to the development of “medical technologies” such as drugs and treatments to aid the body in fighting the agents (pathogens) which attack it. Historically, this proceeded from a trial and errors process of natural substances, to an informed synthesis of artificial drugs to repair the body [Wishart et al., 2018], and disinfectants [City of Vienna Climate Protection Programme, 2012] and antibiotics [The Johns Hopkins University, 2017] to fight pathogens: all together, we are closer and closer to the *magic bullet* envisioned a century ago by Nobel Prize Paul Ehrlich, who dreamed of a “personalised and tailored drug” able to target specific molecular defects while being beneficial to other cells [Strebhardt and Ullrich, 2008]. Such success will be a life saving technology, and condense in a tiny space centuries of efforts to understand nature.

Given the vast amount of knowledge on the topic, every research task focusses on single, simplified questions to complete the jigsaw. While new problems arise and are answered, new experimental techniques are developed to investigate them. However, the difficulties of studying micrometric systems as cells and bacteria are manifold, as we often don not possess instruments to look at them with the desired level of detail, or without perturbing their natural conditions. Therefore, in the last decades a new investigation approach emerged, proposing to model the systems of interest from the theoretical knowledge gathered so far, in terms of their structure, behaviour and properties. Computational biology is the field that proposes to do that, implementing those models first and then querying them for properties which are still unknown. At times, this proves to be the only method possible to answer the question posed [Lee et al., 2009; Dror et al., 2012; Chou, 2015; Leipzig, 2016].

This thesis aims at employing techniques belonging to Computational Biology to answer the specific question of how an artificial molecule behaves within selected biological environments. Thus, while studying a structural biology problem, it places itself at the boundary with the fields of Medicine

and Bioengineering: the following introduction is meant to give an overview of the many different challenges these fields have faced in recent years and the solutions found to those challenges, motivating the interest in developing novel molecules as the one described in this work. In particular, two problems are described: the insurgence of drug resistance, and the problem of drug delivery, as they are both addressed by the system in exam. Figure 1.1 provides a work flow of this introductory chapter, to help the reader in identifying the sections of interest.

1.1 Antimicrobial resistance

For most of the last century, the development of new drugs rotated around the paradigm that a drug is a small inorganic compound (of mass up to 900 Da), which intervenes on a specific target of a mammal or bacterial cell. Very often the targets of interest are (intracellular) proteins: out of the 695 small drugs approved by FDA (the American Food and Drug Administration agency) to target human molecules, 667 acts on proteins. Similarly, 189 of the 198 approved to treat pathogens have a protein as their target (with all the caveats coming from the challenges of identifying an unambiguous target, especially when the drug binds to a protein complex or to a number of closely related gene products [Santos et al., 2017]).

In presenting the aforementioned figures, the data were naturally split among the drugs which act on human molecules, “repairing” some faulty process in the human body, or the ones active against bacteria (commonly named as antibiotics) which “disrupt” the bacterium life cycle in order to kill or prevent the reproduction of the pathogen. It appears evident that the pool of drugs available to the second purpose are in substantially lower number than the ones addressing human molecules. This comes from the nature of the action they perform: molecules targeting human proteins need to be highly specific to avoid interference with other proteins or with healthy cells, and in a sufficient number to address the variety of diseases affecting the human body. Antibiotic must be non-toxic for human cells as well, i.e. their target must not be shared between mammal and bacterial cells, but there is a less stringent requirement on their selectivity against different bacterial species. On the contrary, it is often useful to have a broad-spectrum compound. This cross-species efficacy

Motivations of the work: a graphical abstract

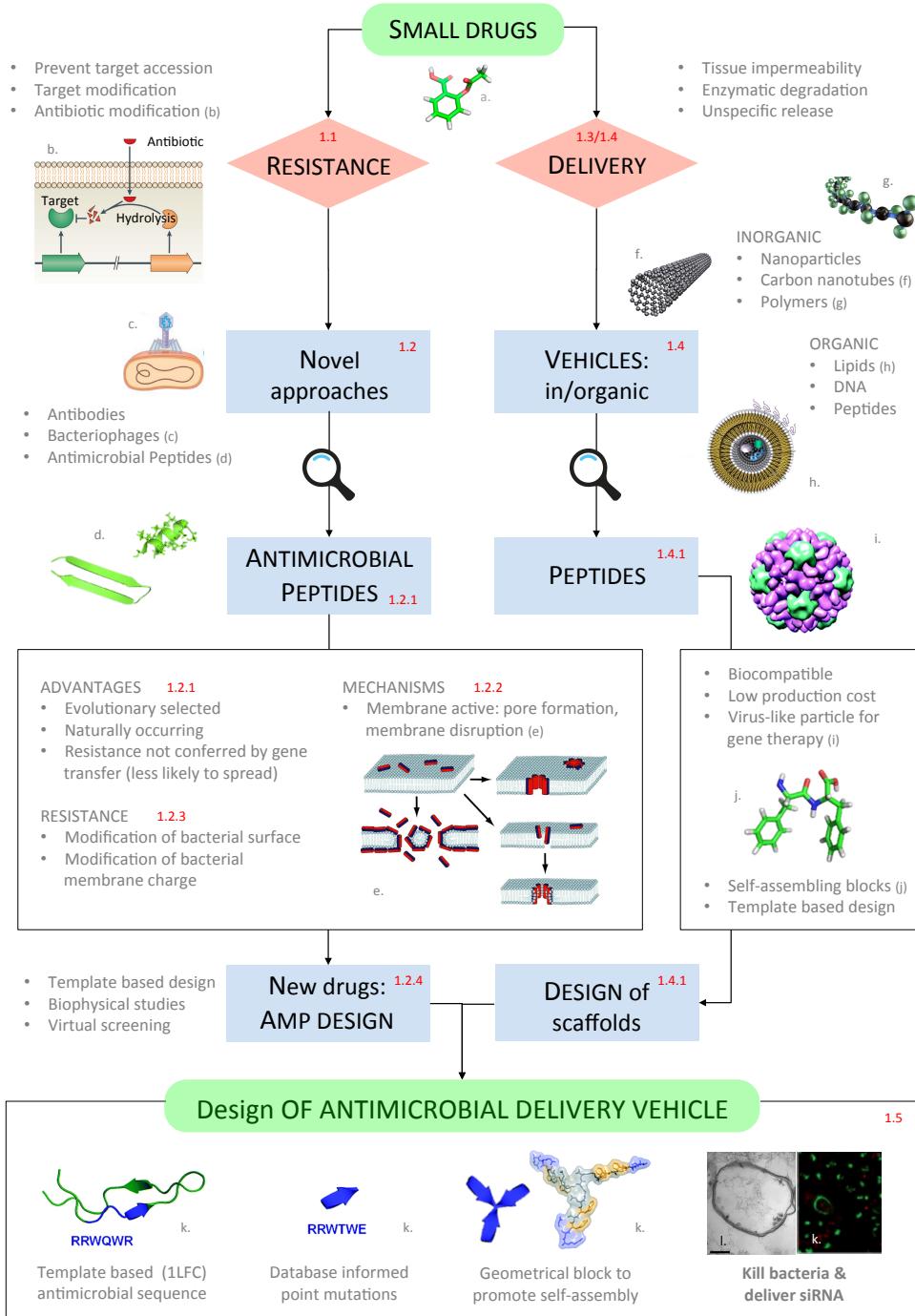


Figure 1.1: Figures a. Acetylsalicylic acid and j. Diphenyl-alanine in bond representation. Remaining figures adapted from: b. [Blair et al., 2014]; c. [Kim, 2017]; d. [Torres et al., 2019]; e.[Nguyen et al., 2011]; f. [nan, 2019]; g. [Mettler Toledo, 2018]; h. [Wikipedia, 2015]; i. [Schoonen and van Hest, 2014]; k. [Castelletto et al., 2016]; l. in house data.

and at the same time non-toxic property is obtained thanks to the evolutionary relationship among bacterial species, and between bacteria and humans: while the first are closely related, and therefore share homologous proteins with very similar structures, humans have less folds in common with them, allowing for a resilience against bacteria-targeting drugs. To be precise, the set of bacterial species is very diverse and the cross-species effectiveness of some drugs does not extend to the whole bacterial population. This is actually demonstrated to be a positive feature, given the large amount of beneficial bacteria that live in symbiosis with the human body (especially in the gut) and that must be preserved for an optimal wellness.

In the framework described above, it is understandable that first-time research on antibiotics was satisfied with the development of a handful of potent, broad-spectrum compounds. Penicillin, the first of them to be synthetically produced, was isolated from a mould in 1928 by Alexander Fleming. It acts inhibiting the formation of a cross-links between particular molecules (peptidoglycan) in the bacterial cell wall, binding to the enzyme responsible for its catalysis, and thus preventing the wall complete formation [Gordon et al., 2000] (for further details on bacterial cell membrane structure see Section 1.2.1). As foreseen from Fleming himself in his Nobel Prize acceptance speech, some species of bacteria quickly became immune to penicillin, and this was achieved in many ways: either by production of penicillase, an enzyme that degrades penicillin, or by subtle changes in the structure of the penicillin-binding proteins to prevent penicillin binding, or again by removal of the drug outside of the cell through specially re-purposed efflux pumps [Lobanovska and Pilla, 2017].

The mechanisms just outlined are not an exceptional characteristic of penicillin, and many drugs lost their effectiveness against some bacteria since their discovery till nowadays, urging the research of new ones on a constant basis. By now, a broad knowledge has been gathered on how bacteria escape the action of a drug: this understanding helps interpreting the pitfalls of existing drugs and identifying the characteristics sought in the development of new compounds.

1.1.1 Mechanisms of antimicrobial resistance to small drugs

Antimicrobial resistance can manifest through many different mechanisms, which can be grouped in three main classes, in line with the three processes mentioned in the example of the penicillin resistant bacteria.

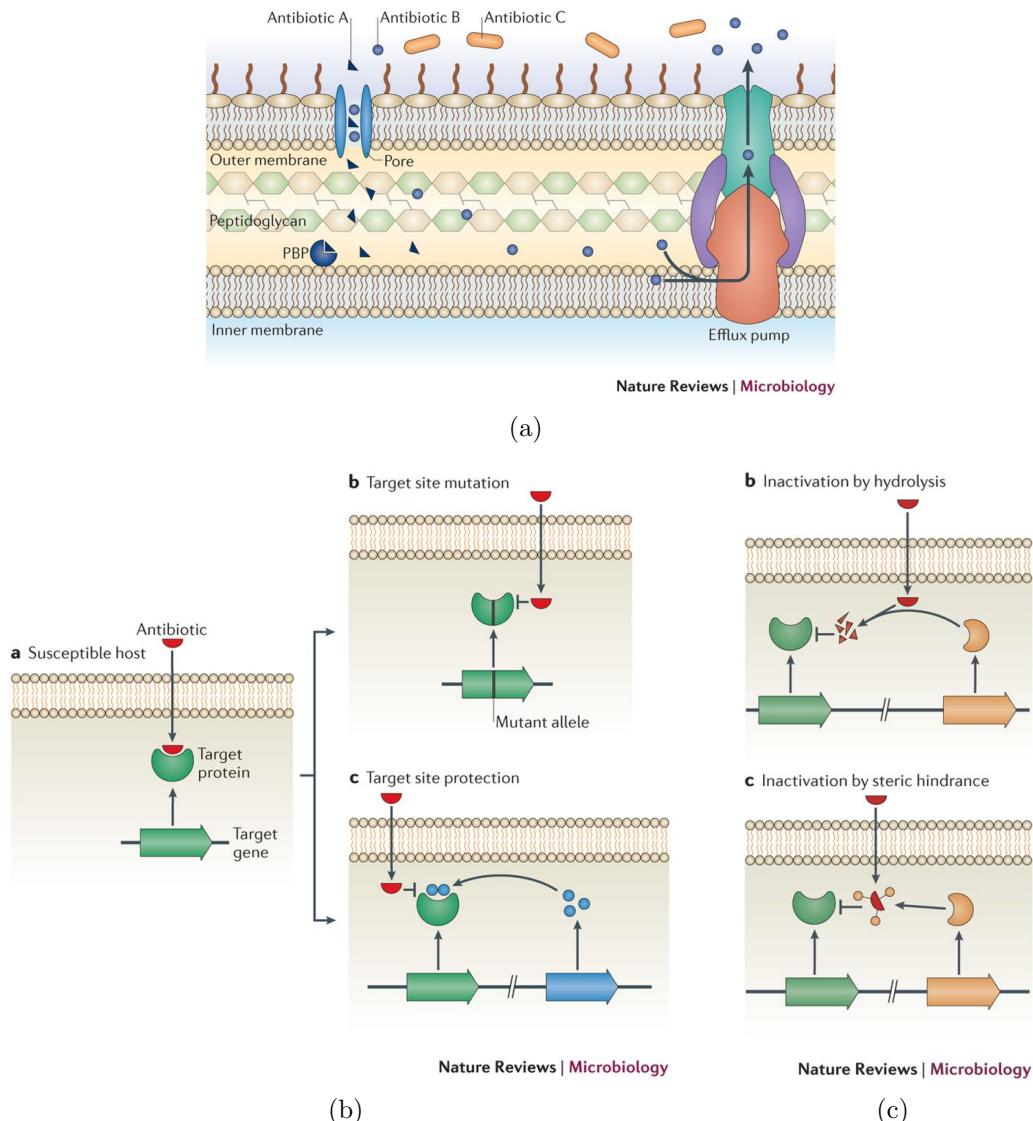


Figure 1.2: Mechanisms of antimicrobial resistance to small drugs. (a) Intrinsic mechanisms of resistance (removal of antibiotic B by efflux pump and inaccessibility of antibiotic C to the Penicillin Binding Protein target because of membrane impermeability). (b) Target site change via mutation or protection. (c) Direct interactions with antibiotics causing its disruption or structural modification. Reproduced from Blair et al. [2014].

Prevention of access to target A first class of resistance mechanisms aims at minimising the intracellular concentration of the antibiotic preventing its penetration or maximising its efflux in the eventuality it has entered the cell (Figure 1.2(a)). Not all the molecules can enter the cell permeating the membrane, and this holds particularly for hydrophilic antibiotics tackling Gram-negative bacteria which are intrinsically less permeable than Gram-positive ones because of the additional presence of the outer membrane [Delcour, 2009] (see Section 1.2.1). These molecules must then be imported into the cell through outer-membrane porin proteins [Vargiu and Nikaido, 2012; Kojima and Nikaido, 2013]. Resistance arise when porins are either replaced with more selective channels, which prevent the antibiotic penetration, or down regulated so that the internal concentration of the drug does not reach a critical concentration [Lavigne et al., 2013]. Porin-coding genes can also accumulate multiple mutations, to acquire the selectivity they lack in their wild type [Poulou et al., 2013].

A complementary strategy to prevent drug influx is to employ bacterial efflux pumps. Some of them are denominated multidrug resistance (MDR) efflux pumps for their effectiveness in the task and are produced by many bacteria [Floyd et al., 2010; Ogawa et al., 2012], but, additionally, can be transferred via plasmids to other bacterial species [Dolejska et al., 2013]. Indeed, bacteria are able to exchange genetic material with other individuals via small rings of DNA in a process called conjugation [Sørensen et al., 2005], so that the advantageous resistant genotypes can spread quickly across species. Over-expression of such efflux pump is observed in multidrug-resistant bacteria, triggered by exposition to the drug, and proceeding via mutation in the relative regulatory network, [Abouzeed et al., 2008], or simply as a response to environmental signals [Nikaido et al., 2011].

Change or modification of the antibiotic target The second class of resistance mechanisms works modifying the antibiotic target: most antibiotics bind to their substrate with high affinity and specificity, thus small modifications in the target structure can disrupt an efficient binding, still allowing the target to maintain its normal function (Figure 1.2(b)).

Mutations of some residues in the binding pocket (upon mutation in the gene coding for it) or its post-translational protection via addition of chemical

groups are equally wide spread mechanisms. Notable examples of the first include the development of methicillin (an antibiotic in the penicillin class) resistant strains of *S. aureus* [Shore et al., 2011; Billal et al., 2011]. Again, it is interesting to notice that several of these mutations are acquired by horizontal gene transfer from other bacterial species, so that resistance development in one specie promotes quickly its insurgence in other ones. For the second process, the most relevant mechanism of chemical group addition is methylation, which, for example, is very common when the drug target are rRNA subunits [Long et al., 2006]).

Direct modification of antibiotics Finally, bacteria can destroy drugs, usually by hydrolysis, or modify them by transfer of a chemical group (Figure 1.2(c)). The first historical discovery of drug-degrading enzyme is penicillinase (a β -lactamase) which destroys penicillin [Abraham and Chain, 1988; Lobanovska and Pilla, 2017]. Since then, thousands of similar enzymes have been identified that can modify antibiotics of different classes [Livermore, 2008; Nordmann et al., 2011]: these enzymes co-evolves with newly developed drugs, to include in their spectrum of action new compounds of composition similar to the ones they were originally effective on [Woodford and Johnson, 2013].

Antibiotics constituted by large molecules with many exposed hydroxyl and amide groups are instead particularly susceptible to addition of chemical groups. Many enzymes are responsible for this, and according to the chemical moiety added they are grouped in acetyltransferases, phosphotransferases and nucleotidyltransferases [Wright, 2005].

All together, the recent progress in understanding the mechanisms of antimicrobial resistance has helped in directing the development of new drugs, in particular the modification and the improvement of existing compounds to escape the resistance developed by bacteria. To ultimately employ the existing drugs at best, it is important to understand also the dynamics of AMR, beyond its molecular mechanisms, to device the most effective strategies of drug administration.

1.1.2 Course of antimicrobial resistance

In the first stages of the insurgence of antimicrobial resistance (AMR) against a given drug, given strains of bacteria are not damaged by the standard doses of the drug as they came to possess some natural occurring mutations in their genome which promote an escape mechanism invalidating the drug effectiveness [Kapoor et al., 2017; Blair et al., 2014]. Usually, only a small population of bacteria is resistant in the first instance, however the resistant population will replicate faster than the peers of the same species because it is more fit in an environment challenged by the presence of the drug. It is noteworthy that this fitness might not be optimal in a natural drug-free environment, and indeed the wild population has not been selected for that genotype. To prevent this insurgence, high doses are often used, to kill the bacteria in a short time and decreases the chances of resistance development [Baker et al., 2018]. However, high drug doses are not always applicable due to the severe side effects they are connected to, creating favourable conditions for resistant strains.

As already mentioned, the spread of resistance between bacterial cells and even between species is very effective as bacteria exchange genetic material through conjugation. Therefore, despite AMR is an evolutionary mechanism, the fast pace at which bacteria replicates, their enormous population (in terms of individuals), and the relative easy horizontal gene transfer, place the insurgence of resistance well within the human lifespan time scale [O’neill, 2016].

It is then clear that this complex problem depends on many variables: the casual appearance of resistant individuals, the transfer of information between them, the relative gain in fitness that the mutation implies, but also the dosage and time line of the drug administration. Many mathematical models have been implemented to understand the issue [Birkegård et al., 2018; Niewiadomska et al., 2019], and it is known that some strategies of drug administration are actually promoting the proliferation of so called “super” bugs. One example is the underdosage of antibiotics which is likely to harm but not to kill pathogens, rather promoting the fitness of resistant ones, or their abuse, as it puts an high pressure on the pathogenic populations, which is desirable but can sometimes induce a faster emergence of escape mechanisms [O’neill, 2016].

In this context it must be noticed that many drugs are bacteriostatic agents as opposed to bactericidal: i.e. they prevent the bacterium growth rather than

kill it, as they are meant to slow down the damage while host defence mechanisms eradicate them. Thus if a high dosage of a bactericidal agent may extinguish the bacterial population and eradicate the disease, bacteriostatic drugs allow bacteria to start again the reproduction cycle once removed (if the host defence could not properly work).

It is then clear that the antibiotic landscape is dynamically regulated: newly discovered drugs enter, while others exit after having been exploited for years, therefore we need constant input from research to find new treatments to old diseases. The severity of the AMR threat is such that it has been raised to the status of national emergency in several countries, including UK. Indeed, strict regulations on the health, agricultural and food industry sector must be taken to prevent the misuse of antibiotics, as we are leaving the century in which antibiotics were discovered, to enter a phase in which we count the number of the ones losing efficacy [O’Neill, 2016].

1.2 Alternative antibiotic strategies: antimicrobial peptides

In the landscape sketched above, it is evident that the development of novel drugs is of crucial importance. Even more beneficial would be to have at disposal a new paradigm for their design, in order to attack pathogens in a completely novel way, avoiding to target pathways which are known to lead easily to the development of antimicrobial resistance. Several novel materials have been developed for the task, not to rely on small molecules and exploit different mechanisms of action, for example antibodies, bacteriophages or antimicrobial peptides [Mantravadi et al., 2019].

While the use of pathogen-specific antibodies relies on mechanisms of the host immune system, bacteriophages therapy employs viruses which infect bacteria and archaea rather than eukarya. But antimicrobial peptides are the main focus of this thesis: indeed some peptides can have an active role against bacteria, if their sequence possesses specific characteristics, and are thus referred to as antimicrobial peptides. The following subsections will explore their characteristics, modes of action and the response of bacteria against them. It is indeed crucial to understand the knowledge available versus the questions that are still open in order to direct the efforts of future research. This holds in par-

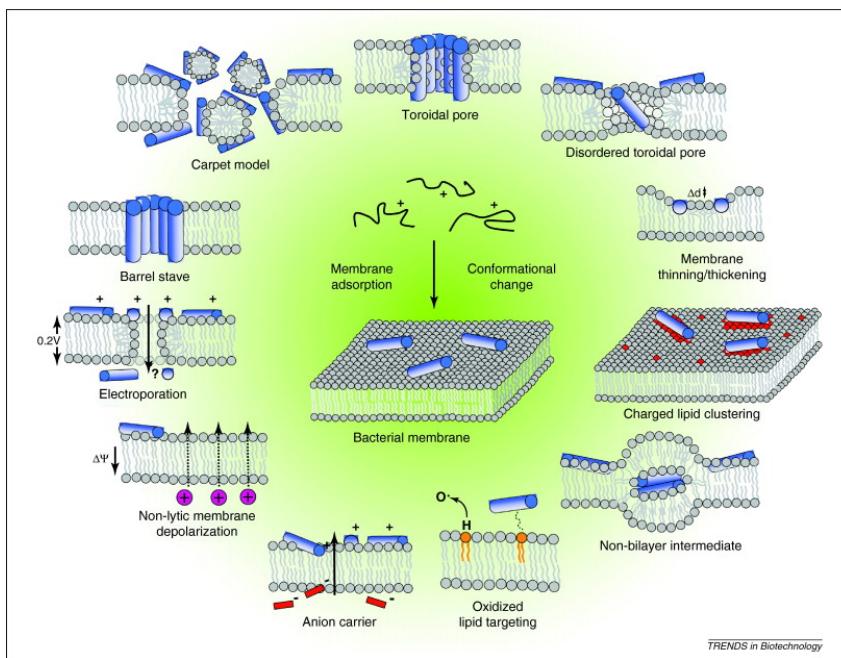


Figure 1.3: Events occurring at the bacterial cytoplasmic membrane following initial antimicrobial peptide (AMP) adsorption. Reproduced from Nguyen et al. [2011].

ticular when the investigation proceeds by the use of simplified models, as the ones employed in Computational Biology, as meaningful results can proceed only if such modelling is performed in a sensible and informed fashion.

1.2.1 Membrane active peptides

Antimicrobial peptides (AMPs) are naturally produced by eukarya, either as stand-alone sequences or embedded in larger proteins, as a first, weak, and broad-spectrum defence against bacteria [Nguyen et al., 2011]. This pool of molecules has been selected through evolution to be active against pathogens, suggesting that they will most likely not cause resistance in the near future.

To exploit their potential and engineer AMP-like molecules, a careful characterisation and classification of such peptides must be done. This task has been carried on throughout the past decades but because of its complexity at present there are many peptides with ascertained antimicrobial activity for which the mode of action is still not fully understood [Ebbengaard et al., 2015]. However, some general characteristics of these sequences and some of

the mechanisms they employ have emerged. Unsurprisingly, AMPs are heterogeneous in sequence, structure, targets and modes of action, to tackle the different challenges bacteria pose. Their size can vary between 6 and 59 amino acids [Brogden, 2005]: despite being small with respect to the average size of a protein in the human body, these macromolecules are hundreds of times larger than small molecule drugs and as such they penetrate and act on bacteria differently with respect to small compounds.

The most common target of AMPs is the bacterial membrane. Many of them cause disruption of the physical integrity of the microbial membrane while others translocate into the cytoplasm to act on intracellular targets, and the combination of the two is not uncommon either [Hancock and Sahl, 2006] (Figure 1.3). In general, it is widely accepted that membrane interaction and its relative consequences are a key factor for the antimicrobial activity of AMPs [Nguyen et al., 2011].

As such, we propose a brief overview of the structure of the bacterial membrane [Silhavy et al., 2010], and of its differences with the one of mammalian cells, to better understand how AMPs can be effective and selective against bacteria at once.

Structure of bacterial membrane The determinant driving the interaction between AMPs and bacterial membranes is the positive charge that many AMPs present, opposed to the negative charge of the latter [Nguyen et al., 2011; Mahlapuu et al., 2016]. It is striking that such simple mechanism, based on the presence of a certain number of negatively charged lipids, holds across many bacterial species despite the great variability found in their membrane composition. Indeed, based on the differences in their cell envelope structure, bacteria are classified into two macro families, Gram-positive and Gram-negative. In Gram-positive bacteria, the cytoplasmic membrane is surrounded by a thick peptidoglycan layer, while for Gram-negative bacteria this membrane (which assumes the name of internal one) is surrounded by a thin peptidoglycan layer as well as an outer membrane [Silhavy et al., 2010; Lin and Weibel, 2016] (Figure 1.4).

Starting from the inside and proceeding outwards (from bottom to top of Figure 1.4), the cytoplasmic membranes of both Gram-positive and Gram-negative bacteria are rich in phospholipids like phosphatidylethanolamine,

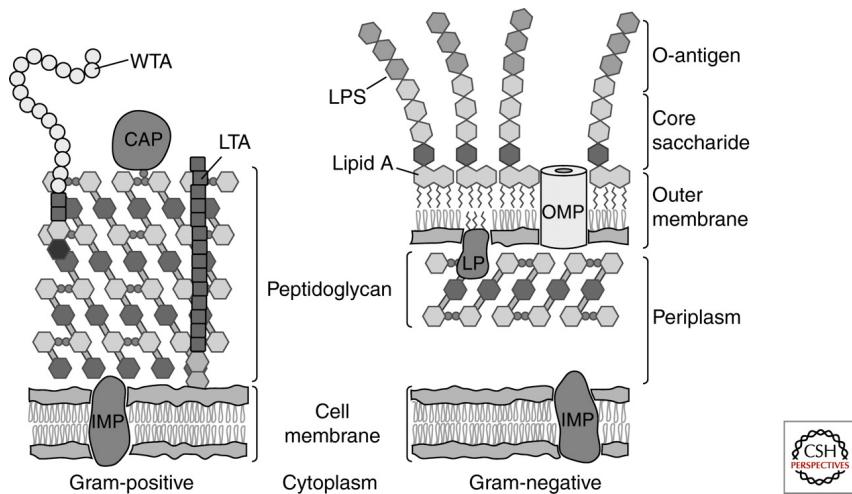


Figure 1.4: Structure of Gram-positive and -negative cell envelope. IMP: integral membrane protein; CAP: covalently attached protein; LTA: lipoteichoic acid; WTA: wall teichoic acid; LP: lipoprotein; OMP: outer membrane protein; LPS: lipopolysaccharide. Reproduced from Silhavy et al. [2010].

which is neutral, and phosphatidylglycerol, cardiolipin, and phosphatidylserine, which have negatively charged headgroups, highly attractive for positively charged AMPs [Silhavy et al., 2010; Lin and Weibel, 2016]. This is often sufficient to promote the preferential interaction between this membrane and the peptides - provided they get into its proximity. Perturbation of this membrane is highly disruptive as many functions are associated to it: as bacteria do not possess organelles, all the membrane-related proteins reside and perform their function within the inner membrane.

In the case of Gram-negative bacteria (Figure 1.4, right), the inner membrane, together with the outer one, delimits the periplasm space, an aqueous cellular compartment, which allows the sequestration of harmful substances and the transport of nutrients. Proceeding outward, inside the periplasm is situated the peptidoglycan cell wall. This substance is made of a disaccharide cross-linked by penta-peptide side chains (from which the name), and these repeated units constitute the rigid skeleton of Gram-negative bacterial cells [Gan et al., 2008]. It is fundamental for cell life as damage in the peptidoglycan layer results usually in living but not viable cells [Joseleau-Petit et al., 2007].

Grafted to the cell wall through Brauns lipoproteins (or LPP) [Asmar and Collet, 2018] is the outer membrane. This membrane presents an asymmetric

structure: phospholipids are present in the inner leaflet, while the outer one is composed of glycolipids, mainly lipopolysaccharides (LPS) [Silhavy et al., 2010]. These complex molecules consist of lipid A, which presents multiple fatty acids, and a polysaccharide [Raetz and Whitfield, 2002]. The polysaccharide is made of an inner core, covalently bond to the lipid, an outer core, and finally a repetitive glycan polymer (O-antigen). The O-antigen (top right in Figure 1.4) is the molecule exposed by Gram-negative bacteria to the external environment and thus is the target of antibodies recognition.

Given the complexity of the Gram-negative cell envelope, and especially the presence of the LPS layer, these bacteria are particularly impermeable to hydrophilic molecules, which are usually imported within the cell through porins and similar transmembrane proteins.

For Gram-positive bacteria (Figure 1.4, left) the inner membrane is enveloped in a thick peptidoglycan layer. If its thickness in Gram-negative bacteria reaches few nanometers, in Gram-positive ones it spans from 30 to 100 nm. Its composition is similar to the one described above, with some variations present among different bacteria, for example in the nature of the peptidic linker or in its precise position [Vollmer et al., 2008]. This thick layer is threaded by long anionic polymers (the teichoic acids), mainly composed by glycerol phosphate, glucosyl phosphate, or ribitol phosphate repeats [Swoboda et al., 2009]. Disseminated in this layer there are several surface proteins with various functions, among which adhesins, which attach to components of the host extracellular matrix.

Gram-positive membranes are generally more permeable because they do not possess a double-membrane structure, nevertheless the peptidoglycan later they are coated with constitute a challenge for drug delivery.

Comparison with mammalian membrane The fact that AMPs tackle negatively charged membranes is crucial for their selectivity, i.e. that they are harmless for the mammalian cells they are produced from [Glukhov et al., 2005]. This is possible because mammalian cells have a different membrane composition. They present a single membrane, which is rich in proteins (up to 50% of its volume) and in lipids, with a small percentage of carbohydrates, mainly embedded in glycoproteins, which promote cell-cell recognition.

The lipidic component is abundant in zwitterionic phospholipids such as

phosphatidylethanolamine, phosphatidylcholine, and sphingomyelin, providing a neutral net charge. [Spector and Yorek, 1985; van Meer et al., 2008]. Furthermore, the mammalian cell membrane has a high content of cholesterol [Yeaman and Yount, 2003; Lai and Gallo, 2009], a sterol fat, which is proposed to stabilise the membrane regulating its fluidity across the differences of physiological temperatures, and it is also thought to favour a better accommodation of the perturbations caused by AMPs [Zasloff, 2002]. Strictly speaking, some negatively charged lipids are present in a few mammal cell types, however they are located in the inner leaflet, while the zwitterionic phospholipids are more abundant in the outer leaflet, in an asymmetric composition [van Meer et al., 2008; Matsuzaki, 2009]. This structure promotes weaker interactions between AMPs and mammalian cells with respect to bacterial ones, as the former is driven mainly by hydrophobic interactions, while the latter by electrostatic ones.

Another relevant difference between bacterial and mammalian cells is that the first ones have typically a higher transmembrane potential - the difference of electrostatic potential between the inside and the outside environment: for bacteria it falls between -130 and -150 mV, while for mammalian cells between -90 and -110 mV [Yeaman and Yount, 2003; Matsuzaki, 2009; Ebenhan et al., 2014]. Given that a potential generates an electric field across the membrane, the higher this is, the higher the resulting electric field pointing from outside to inside the cell will be. A field in such direction pushes cationic compounds on the outside of the membrane toward the membrane itself. Therefore the stronger bacterial transmembrane potential may promote an enhanced - and thus disruptive - interaction of AMPs with the cell, contributing to the AMPs selectivity between bacteria versus mammals [Yeaman and Yount, 2003].

1.2.2 Common mechanisms of action of AMPs

Investigating the perturbation and disruption of a bacterial membrane by antimicrobial peptides is a key point of this work, therefore it is important to highlight the mechanisms known so far through which AMPs reach this outcome. As already mentioned, many AMPs have a positive charge which facilitates the binding to the membrane via charge-charge recognition; accordingly, Arginine and Lysine residues are usually abundant in AMPs sequences. How-

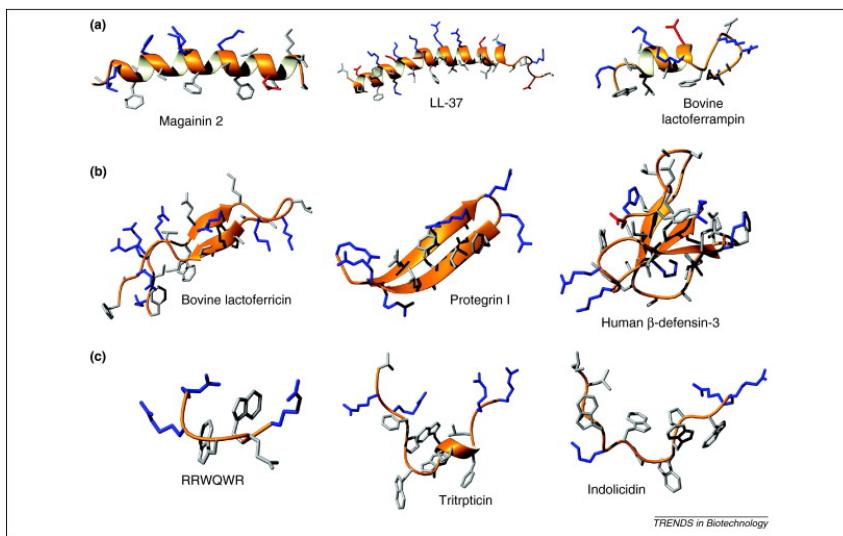


Figure 1.5: Structure of some known AMPs. Reproduced from [Nguyen et al., 2011].

ever, the disruptive action takes place through the interaction of the AMP with the hydrophobic core of the membrane, therefore their sequences contain also hydrophobic aromatic residues, especially Tryptophan, which favours the anchoring to the lipid core [Chan et al., 2006]. Overall, AMPs resort often to adopt an amphiphatic structure to segregate the hydrophilic from the hydrophobic amino acids and thus act at the interface between membrane and solution. It is interesting to notice that some of them fold into the active structure only in the nearby of the membrane, as they can expose their hydrophobic components to face its core, while in solution these ones are preferentially buried inside to be screened from the solvent [Nguyen et al., 2011]. Common folds adopted by AMPs are both α -helix or β -sheet rich structures. Amphiphatic α -helices present a charged side which is tailored to face towards the phospholipid head groups and an hydrophobic ones which is favourably buried into the acyl chains core, and a similar arrangement is found for structures rich in β -sheets include β -hairpins (Figure 1.5).

Membrane disruption Several models have been proposed to describe the exact mechanisms of AMPs penetration after they bind to the cytoplasmatic membrane, and how their combined action leads to membrane permeabilization (Figure 1.3) [Brogden, 2005; Nguyen et al., 2011].

For example, for a single copy of a amphiphatic helical AMP, the proposed mechanism of action suggests that initially the peptide is attracted with its charged side to the membrane and lies parallel to its plane, with the hydrophobic side unfavourably exposed in solution. Then the helix rearranges to have the two faces in the respective favourable regions. Subsequently, the helix axis starts to form an angle with the membrane plane, and finally inserts deeper into the lipid core, often spanning the full membrane thickness [Ebenhan et al., 2014]. Similarly, for β -sheet rich structures, it is suggested that they insert within the membrane after a first flat approach. The final insertion arrangement depends on the peptide characteristics and length, the presence of kinks in its structure (in case of helices), and the interactions with other copies of the peptide.

The picture becomes more complex for oligomer-mediated insertion, i.e. when the action is triggered by the combined action of many copies of the peptide. At low peptide to lipid ratio, the favourable configuration is represented by peptides lying parallel to the membrane plane as described previously [Yang et al., 2001], but an increase in peptide concentration triggers the transition to an inserted state where the main axis of the AMP is perpendicular to the membrane. The organisation of AMPs inside the membrane core can assume different configurations, as described below [Brogden, 2005; Nguyen et al., 2011; Ebenhan et al., 2014; Mahlapuu et al., 2016] (see Figure 1.3).

The “barrel-stave” model proposes that AMPs insert perpendicularly into the bilayer. Recruitment of peptides in the same area results in the formation of a transmembrane pore with a central lumen. The walls of the pore are constituted by the hydrophilic face of the peptides, while their hydrophobic side is interacting with the lipid tails around the pore. This model is adopted for example by the α -helical AMP alamethicin, which forms voltage-dependent ion channels by aggregation of four to six molecules [Spaar et al., 2004].

In the “toroidal” pore model instead, the insertion of peptides forces the phospholipid to bend continuously from one leaflet to the other. The toroidal model differs from the barrel-stave model as the peptides are always associated with the lipid head groups even when they are perpendicularly inserted in the lipid bilayer. Toroidal pores are induced by α -helical magainins, protegrins and melittin [Yang et al., 2001; Matsuzaki et al., 1996; Hallock et al., 2003], and lead to membrane perturbation which extends further away from the pore than in

the barrel-stave case, as lipids must rearrange around them: as a comparison, alamethicin induced barrel-stave pores have an inner and outer diameters of 1.8 nm and 4.0 nm respectively [Spaar et al., 2004], while magainin-induced toroidal pores have variable sizes, with an inner diameter of 3.0-5.0 nm and an outer diameter of 7.0-8.4 nm [Matsuzaki et al., 1997].

Finally, in the “carpet” model, the accumulation of AMPs on the surface of the membrane, laying parallel to it, causes tension in the bilayer and the membrane is then disrupted by peptides in a detergent-like manner, leading to the formation of micelles. The critical threshold concentration triggers a cascade effect, in which formation of the first disruption allows the penetration of AMPs in the inner side of the bilayer. The cooperation between peptides on both sides of the lipid membrane enhances the AMP-induced curvature causing accelerated disruption. The “carpet” model mechanism is observed for peptides presenting an α -helical structure (like melittin [Ladokhin and White, 2001]) or several helices connected by short loops (ovispirin [Yamaguchi et al., 2001]).

The prevalence of examples with an helical structure for the above models derives from the fact that the understanding of how helical AMPs function is often easier than the one of β -sheet rich structures. Indeed, helices have a well defined fold (at least nearby the membrane environment), a compact structure, and often a clear segregation of complementary patches that can attract other copies of the peptide and thus promote the self-assembly process necessary for the pore formation.

On the contrary, many β -sheet AMPs have a more flexible structure, and more diversified mechanisms of action [Nguyen et al., 2011; Mahlapuu et al., 2016]. AMPs rich in β -sheets can be divided into β -hairpins and peptides from the defensin family [Nguyen et al., 2011]. Many representative of the former class disrupt bacterial membranes via formation of toroidal pores: as an example, porcine peptide protegrin I triggers the toroidal pore formation assembling into a β -barrel structure when in contact with anionic membranes (while it folds into β -sheet aggregates on the surface of cholesterol containing membranes, thus acting selectivity on bacterial membranes only [Tang and Hong, 2009]). In the case of defensins, many mechanisms are known according to the specific member of the family [Lehrer, 2004; Kagan et al., 1990; Takeuchi et al., 2004]. Although various descriptions of membrane damage have been

reported, and include ion channels, transmembrane pores and extended rupture of the membrane, they are likely related, being a modulation of a similar acting principle.

Alternative mechanisms of action Finally, many non-lytic mechanisms are suggested for AMPs, especially for β -sheet structures: defensin A from *P. terramoviae* reduces the cytoplasmic potassium concentration [Brogden, 2005], partially depolarising the inner membrane; tachyplesin from horseshoe crabs is able to bind to the minor groove of DNA, interfering DNAprotein interactions [Yonezawa et al., 1992], and bovine lactoferricin can act synergistically with other antimicrobial agents by affecting the transmembrane potential and proton-motive force, resulting in inhibition of ATP-dependent multi-drug efflux pumps [Gifford et al., 2005]. Moreover, after translocation within the cell, bovine lactoferricin can also inhibit DNA, RNA and protein synthesis.

Section 1.5.1 will treat in detail the functioning of this AMP, distinguishing its role as membrane active peptide as opposed to intra-cellular targeting compound: indeed, many works have focussed on lactoferricin antimicrobial processes versus locating the section of the sequence performing the membrane disruptive activity [Tomita et al., 1994; Schibli et al., 1999], to understand whether it retains the efficacy regardless of the fold. These type of investigations performed on many AMP sequences provided the discovery of first minimal functioning antimicrobial blocks, which promoted the understanding of how AMPs work in general, and boosted the design of synthetic AMPs tailored for specific functions.

1.2.3 Mechanisms of resistance to AMPs

Antimicrobial peptides were introduced here as a class of new drugs and a possible solution to the crisis of antimicrobial resistance. Any new drug entering the pool of the clinically approved compounds is (at least temporary) a solution to the problem of resistance to known antibiotics; but it must be clarified that bacteria can develop resistance to AMPs too. Nevertheless, the resistance to their action is generally not based on dedicated genes that are conferred by horizontal gene transfer, as in the case of many antibiotics resistance mechanisms [Peschel and Sahl, 2006]. Because of that, a certain increase

of resistance after exposure to the drug is to be expected, but it is less likely to spread quickly to other species.

Some of the mechanisms of AMPs resistance are similar to the ones employed by bacteria to counteract small molecule drugs, for example over-expression of efflux pumps to dispose of AMPs, degradation of the peptide by extracellular enzymes and sequestration by the bacterial or biofilm matrix to prevent accession to the target [Peschel and Sahl, 2006].

Differently with respect to antibiotics hydrolysis, AMPs proteolitic degradation is operated by proteases, secreted on the extracellular side of the membrane specifically to destroy other proteins. Linear AMP are more prone to this type of degradation [Sieprawska-Lupa et al., 2004], as opposed to the ones presenting disulfide bonds [Peschel and Sahl, 2006], such as defensins, which nevertheless can be hydrolysed by more specific proteases enzymes [Nelson et al., 2011].

Bacteria can also enhance their resistance to AMPs organising into specialised structures known as biofilms [Costerton et al., 1999]. Biofilm bacteria adhere to a surface and secrete an extracellular matrix with adhesion and protection functions [Jolivet-Gougeon and Bonnaure-Mallet, 2014], which effectively repels and/or captures AMPs. For example the polysaccharide intercellular adhesin (PIA) present in the extracellular matrix of *S. aureus* and other bacteria can be deacetylated to increases its positive net charge, thus repelling more efficiently cationic AMPs (like HBD-3, LL-37), and increasing sequestration of the anionic ones (like dermcidin) at the same time [Vuong et al., 2004a,b]. AMPs are nevertheless a promising alternatives in the treatment of biofilm-associated infections. In this type of infections bacteria are growing slowly, so it is advantageous to have bactericidal agents such as AMPs, as opposed to bacteriostatic ones such as traditional antibiotics which target fast-growing bacteria only [Strempele et al., 2014].

But the most specific mechanism of AMPs resistance concerns modifications of the bacterial cell envelope: bacteria modify the characteristics of their surface to prevent the efficient binding of an AMP, even in the eventuality that the peptide reaches the bacterial envelope intact. The target of such modifications are different for Gram-positive and Gram-negative bacteria, according to their distinct cell envelopes. Gram-positive bacteria change the structure of their teichoic acids (TA): for example, D-Alanylation of TA observed in *S. Au-*

reus adds a positive charge to it, reducing the attraction of cationic AMPs and in turn increasing the cell wall density, so reducing the surface permeability [Saar-Dover et al., 2012]. Alternatively the bacterial peptidoglycan precursor, lipid II, can be modified: it has a key role in the formation of the cell wall ad is often targeted by AMPs. For example, the replacement of its terminal D-alanine with D-lactate or D-serine [Bugg et al., 1991] avoids the action of the glycopeptide vancomycin, which binds to the D-Ala-D-Ala terminal motif of the precursor.

In Gram-negative bacteria a positive charge can be added to lipopolysaccharides (LPS) by addition of amine-containing molecules [Moskowitz et al., 2004] or by removing phosphate lipids (which have a negative charge) from lipid A [Wang et al., 2006b], to obtain the same effects as for the increased charge in TA. Importantly, the resistant Gram-negative bacteria can also enhance the rigidity of the outer membrane to reduce permeability to AMPs via addition of extra acyl chains into lipid A [Guo et al., 1998]. Moreover, they can act at the cytoplasmic membrane level, as this is the final target of many antimicrobial peptides: in the eventuality that AMPs successfully reach this membrane, they are attracted to its surface by the negative charge of the lipids composing it, in particular phosphatidyl-glycerol (PG) and diphosphatidylglycerol (DPG, also called cardiolipin). Their negative charge can be masked by amino-acylation of the PG head group, so that the final compound repels AMPs through electrostatic interaction [Peschel et al., 2001]; alternatively the overall rigidity of the cytoplasmic membrane can be enhanced, by an increase in saturated acyl chains which has been proven to confer resistance [Kumariya et al., 2015]. To be noticed that resistant bacteria often employ many of the aforementioned strategies at the same time [Band and Weiss, 2014].

1.2.4 Principles of AMP design

The study and classification of AMPs provide knowledge on the characteristics a sequence must have to perform an antimicrobial function. As discussed in Section 1.2.2, there are some features which, comprehensively, help in discriminating AMPs against non antimicrobial peptides. The constantly increasing amount of data available is gathered in several curated databases [Wang et al., 2016a; Pirtskhalava et al., 2016; Jhong et al., 2019; Lata et al., 2010; Bhadra et al., 2018], which catalogue AMPs (or subclasses of them, like membrane ac-

tive, biofilm active or haemolytic peptides) based on such features to promote future data-driven prediction of the antimicrobial character of a sequence.

We recapitulate below a few key characteristics which are peculiar of AMPs. To be noticed that while some are easily retrievable from the sequence of the peptide, others imply direct experimental measures:

- **Structure:** both α -helical and β -sheet rich AMPs exist, as well as mixed structures. Short helix (~ 22 amino acids) and short β -sheet (~ 10 amino acids) are particularly common. When screening a potential AMPs, it must be considered that the fold adopted by the peptide close to the membrane environment might be different with respect to the one in solution.
- **Charge:** AMPs are charged moieties, usually cationic (up to $\sim +10\ e$), with fewer anionic examples (like dermcidin). Among cationic ones, not all the positive amino acids have equal role, for example Arginines are more effective than Lysins [Chan et al., 2006]. The potency, but also their haemolytic activity, are often directly related to the amount of charge [Jiang et al., 2011].
- **Hydrophobicity:** AMPs contain also hydrophobic residues, usually with abundance of aromatic chains and specifically Tryptophan, as they must insert and anchor into the lipid core of membranes, which is an hydrophobic environment.
- **Amphipathicity:** to host both the charged and hydrophobic residues, most AMPs organise themselves in an amphiphatic structure, i.e. the two types of amino acids side chains are located on the opposite sides of the peptide.
- **Solubility** AMPs need good solubility to prevent aggregation in the aqueous environment they float in before arriving to the membrane. Aggregation would most likely impede their optimal interaction with the membrane.

A part from predictions regarding existing sequences, the knowledge of AMPs sequence-activity and structure-activity relationships is beneficial to find new, better performing ones. The design of new AMP sequences aims at producing peptides with improved characteristics:

- **specificity** against particular bacterial species;
- **stability** against the action of proteases, thus allowing a longer residence time in the body;
- **low cytotoxicity** at the therapeutic dose required (so an high therapeutic index).

The need for such improved peptides lies in the fact that their natural form constitutes a first broad spectrum defence our body employs against infectious bacteria and thus AMPs are often of mild potency. However, foreseeing their application as future drugs, it is desirable to tailor them to fulfil different criteria according to the infection to treat. At the present state of the art, a golden rule for the design of such sequences is still missing, however several methodological approaches to AMP design have been explored, and they can be grouped in three main lines: template based studies, biophysical studies and virtual screenings [Fjell et al., 2011].

Template based studies The main idea behind template based methods consists in modifying existing antimicrobial sequences in the direction of the desired characteristics. The most widely explored templates are helical peptides, for their short sequences and because several of them (cecropin, magainin and protegrin) have been well characterised [Wang et al., 2015].

Ideally, an amino acid scanning of all the residues in an AMP provides information on the role of each of them, prompting at the most suitable mutations. High-throughput methods for the synthesis and characterisation allow nowadays for such thorough investigation in the case of short AMPs [Hilpert et al., 2005], but similar, less resource consuming Alanine scannings can still point at the most important residues for the antimicrobial activity [Migoń et al., 2018]. Alternatively, simpler approaches aim at designing peptides with enhanced charge and amphiphilicity, as these characteristics are deemed crucial for their effectiveness (see the paragraph above) [Wang et al., 2015].

However, all the above methods focus on single amino acids and can not take into account the interplay between residues, nor the three dimensional structure of the peptide. Without such information, it is difficult to extract general rules on why some mutations work better, and often the results of these studies give indeed enhanced AMPs, but cannot be generalised to other

sequences. Therefore, recent works sought to integrate structural information on template based models, successfully designing peptides active against many bacterial lines at the same time [Liu et al., 2018] or enhancing the selectivity of some sequences for bacterial membranes [Jiang et al., 2011].

A complementary approach to single point mutations on known peptides consists in designing minimal antimicrobial blocks: several investigations proved the importance of single residues and their intercalated pattern in natural and designed AMPs. For example, natural AMPs are rich in Tryptophan and Arginine residues [Chan et al., 2006], and synthetic ones have been produced with only Lysines-Leucine, or Arginine-Valine combinations to produce amphipathic helices [Deslouches et al., 2005]. An effort to extract principles from these examples is represented by text based models where amino acids constitute the letters and patterns occurring in natural AMPs are the grammar rules [Loose et al., 2006; Cipcigan et al., 2018; Spänig and Heider, 2019].

In general, the advantage of template based methods is in the reduced number of sequences to test, with decreased cost, as only a subspace of them is explored, namely the ones close to the original template.

Biophysical studies Biophysical studies aim at understanding the functioning of AMPs investigating their structure. Free energy perturbation, Molecular Dynamics (MD) simulations and thermodynamics calculations can all provide knowledge on how the three dimensional arrangement of residues is important to allow their functional role. Contrary to sequence based methods, these techniques give an insight into the mechanism of action of an AMP but their drawback lays in the high computational cost. For example simulations can approach systems with a limited size, and only on short (microseconds) time scales, preventing the reproduction of phenomena of the order of millisecond (a detailed overview of the state of the art, advantages and drawback of MD simulations will be given in Chapter 2). For these reasons, such techniques have been applied to fewer systems in comparison to sequence based screenings, and only few mutations have been tested and compared *in silico*.

The strength of biophysical studies lay instead in the fact that they exploit the whole information available on a system (sequence, structure, chemistry), so they can single out the interactions that are crucial for a mechanism, clarifying whether they can be transferred to a different environment, or again

they can discriminate cases in which similar sequences behave differently due to the environment around them. In this respect, they provide a generalisable knowledge applicable to different systems and thus to the design of novel AMPs at the atomistic level. Extensive examples of such workflow are given in Section 2.5.

Virtual screenings Contrary to biophysical assays, virtual screening methods are employed to analyse a large number of sequences, when an experimental or computational test of all of them would be prohibitive. The concept of these methods consists in identifying descriptors which allow to predict the potency of the sequence: from the analysis of a database of AMP with known activity, a model is created and used to score novel synthetic sequences [Fjell et al., 2011; Kleandrova et al., 2016].

The recent evolution of machine learning (ML) techniques, artificial neural network in particular, gave a great impulse to virtual screening of AMPs (for a historically informed review see Table 1 in Ref. [Fjell et al., 2011; Veltri et al., 2018]). Machine learning appears particularly suitable to the task as the potency of AMPs is determined by the combination of many factors, the relative weight of which can be difficult to identify. Moreover, these approaches can help in the identification of relevant features traditionally overlooked.

Practically, ML algorithms are trained on a set of AMPs labelled by their potency and characterised by different properties (features) along the line of the ones mentioned at the beginning of the section: sequence, partial charge, hydrophobicity, etc; but also experimental measures (pK measures, nuclear magnetic resonance data, octanol-water partition fraction and so on). The more the input properties to consider, the more expensive is to train the model, but the higher the accuracy that can be meet. At the same time, the output descriptor is increasingly complex (i.e. a combinations of many features) and thus of difficult interpretation. In principle, this is not a problem (provided there is no overfitting) as, rather than guide the design of AMPs from first principles, descriptors can be used to score combinatorial sequences of the desired length to identify the best ones. In practice however this is often difficult because of their exponentially growing number.

The second obstacle to ML procedures is given by the fact that the more features one wants to consider, the more sequences need to be given as input

to the algorithm, i.e. need to be experimentally tested. However, modern high-throughput synthesis methods, together with surrogate measures of bacterial killing are allowing it nowadays, as shown by Cherkasov et al. [2009], who assessed the antimicrobial properties of thousands of 9 residues sequences and trained a neural network on the outcome, to then score novel sequences with good accuracy. This is a first step toward an automated and general procedure for AMP design.

1.2.5 Clinical applications

Antimicrobial peptides have been studied for many years, however the push to capitalise them to get compounds viable for the clinical stage has been delayed by many factors, including production costs, and lack of interest in the face of more potent small molecules which were deemed more economically advantageous by pharmaceutical companies. The constant increase of AM resistance has resulted in more effort focusing on AMPs, mainly from small biopharmaceutical companies, and at present several of these preparations are in clinical trials, either in phase 1 or 2 [Naafs, 2018].

The two major problems encountered so far for AMPs sequences in trial are the liability to proteolytic degradation, and the unknown toxicology profile when administered systemically [Hancock and Sahl, 2006]. For the last reason in particular, many of them in trial for topical use against skin infections only, while they are deemed unsuitable for internal administration. Design of novel AMPs can be tailored to improve the liability to degradation, for example introducing D-amino acids, non natural amino acid analogues of opposite chirality, which, with appropriate formulations, are mimetic to the immune system [Wipf et al., 2009]. Moreover, machine Learning protocols can help in pre-screening their toxicity through virtual screening methods.

Overall, antimicrobial peptides remain a promising tool to counteract infections and, as their design is still - comparatively - in its infancy, there is room to explore novel applications and synthesise improved sequences apt to get to the clinical stage.

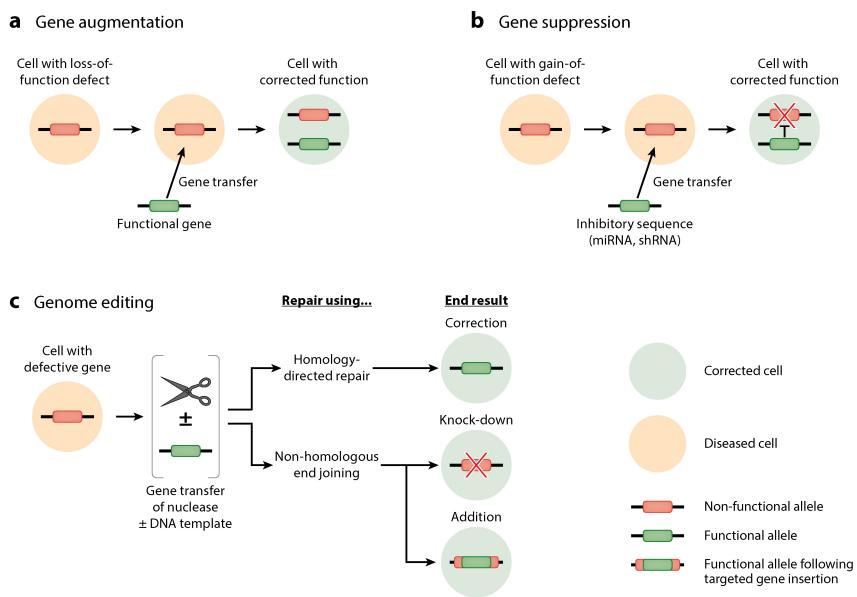


Figure 1.6: Principles of gene therapy. Reproduced from Anguela and High [2019].

1.3 Gene therapy

Alongside the new compounds used to counteract bacterial infections, we want to bring the reader's attention to another class of therapies which is relevant for the work of this thesis and has been developed in the last decades for the treatment of non infectious diseases: gene therapy. In recent years it has greatly evolved and gained attention for the treatment of tumours, genetic diseases and complex acquired disorder [Anguela and High, 2019]. The key concept is the delivery of genetic material to cells of diseased state which possess a faulty copy of a gene, to influence its expression. Such fault can result in lack of synthesis of the protein of interest or in its misfold and/or malfunction. The correction can be performed in three different ways (Figure 1.6).

Augmentation gene therapy introduces an healthy gene copy to restore the normal functionalities of the protein of interest and thus of the cell: it usually consists in the delivery of a DNA strand, which in turn can be internalised in the genome and thus spread when the cell replicates, or not internalised and thus can influence the functionalities of the particular transfected cells only

[Anguela and High, 2019]. Suppression gene therapy suppresses a detrimental gene, and this is particularly useful in the case of cancer, to impede cancer cells replication. Usually this strategy employs RNA interference, delivering miRNA (microRNA) or siRNA (small interfering RNA) strands which repress the transcription of the problematic RNA sequence. Finally, gene-editing, the most recent advance in the field, overlooks the possibility of correcting base pairs mutations to restore the original healthy sequence, and is often done through the functionalisation of the CRISPR-Cas9 technology, a mechanism found in prokaryotic organism as defence against viruses [Barrangou, 2015]: CRISPR (clustered regularly interspaced short palindromic repeats) is a library of DNA fragments from viruses that have previously infected the prokaryote, and the Cas9 enzyme (CRISPR-associated protein 9) uses these sequences to recognize and cleave strands of DNA complementary to the CRISPR sequences to block subsequent infection preventing the viral replication. Research has been able to engineer the CRISPR-Cas9 technology to edit (rather than simply cleave) genes within eukaryotic organisms [Zhang et al., 2014a], thus performing a therapeutic role.

Despite the challenges posed by the development of genome editing tools, and the risk associated to them (for example the possibility of deleterious insertional mutagenesis or deleterious immune responses), at present six gene therapies have received approval in the Western world [Anguela and High, 2019], with many more undergoing regulatory review.

One of the main problems in the development of such therapies lies in the identification of a suitable vector: delivery of free genome in solution results in poor internalisation and low therapeutic effect. Thus nowadays the outlook of gene therapy research lies not only in improving specific cargos, but also in the research of appropriate vectors with low toxicity, low induced immune response and high delivery efficiency. Viruses can be used, modifying their genome to include the necessary sequence and remove the ones promoting viral replication [Naldini, 2011; Mingozi and High, 2011], but synthetic vectors are now investigated for a virus-free delivery strategy. The system studied in this thesis proposes, among its other functions, to deliver genetic material into human cells.

1.4 Delivery of therapeutic material

The problem of gene delivery sets a parallel with the small drug one, introduced at the beginning of the chapter. Indeed small molecules need delivering agents to be efficiently internalised in the cells, in the same manner that genes do.

To reach the aimed organ, therapeutic molecules must be compatible with the different cellular environments they cross, but be preferentially retained, and act only on the ones they are designed for. This implies a subtle balance between a invasive activity on one side, and mimesis on the other, to minimise the possibility that the compound is recognised as dangerous and disposed by the immune and reticuloendothelial systems.

As an example, one can think of the trip of a standard, orally administered drug, passing through the digestive system, with its challenging acidic environment and limited permeation across the intestinal epithelium, and from there to the blood stream [Masaoka et al., 2006; Mitragotri et al., 2014]. At this point, the drug is generally coated by a protein corona based on its shape and charge [Krol, 2012]. The nature of this coating is difficult to predict and can disrupt or decrease significantly the efficacy of a compound as it modifies the way it is recognised and absorbed. From the blood the drug diffuses in the tissues flanking the blood vessels naturally depleting its concentration downstream [Krol, 2012], so that regions further away from the injection region have less chances of getting a sizeable dose. Moreover some specific tissues are highly impermeable: the blood brain barrier (BBB) for example allows the passage of small molecules with high lipid solubility only [Pattni and Torchilin, 2015; Krol, 2012]; while tumoural tissues are instead poorly vasculated, reducing the chances of delivery at their interior [Pattni and Torchilin, 2015].

For all the above reasons, research has focused on developing systems to assist the delivery of drugs. A mimetic carrier can not only improve delivery, but also be designed to selectively bind to particular tissues, or to trigger the drug release after a delay, or only upon changes in environmental variables (for example pH) to reduce drug concentration in non targeted regions. A stand alone field of research has then focused on the development of delivery vehicles irrespective from the quest for new drugs. The optimised products of the two separate efforts can then be paired according to the condition to address, to give a successful therapy.

At present, many molecules have been successfully employed to build drug vehicles, both organic and inorganic, to offer a range of different physico-chemical characteristics useful to target different cells [Hughes, 2005] (Figure ??). We briefly list them to point out the variety and exotoxicity of structures which are useful, sometimes unexpectedly, to the medical world, and we then focus on peptidic delivery vehicles: once more this class of molecules can offer a solution to a therapy-related task.

Inorganic materials for small drugs delivery Many inorganic compounds have been used to transport drugs or to enhance their biocompatibility. A first class is constituted by metal nanoparticles, with golden nanoparticles covering a major role. Thanks to their metallic nature, these materials can be customised in shape and size (down to a 10 nm radius), and possess optical properties that allow to track them inside the body or to thermally stimulate them to trigger drug release [Boisselier and Astruc, 2009]. They can also be coated with biologically active moieties to enhance their mimesis [Singh et al., 2018]. For these reasons, they are used to selectively treat tumoural cells, but up to now only a few golden nanoparticle based compounds have made to the clinical stage, as there are mixed evidence about their toxicity [Boisselier and Astruc, 2009].

Similarly, materials made from carbon, especially carbon nanotubes, can be used for biomedical applications as they have a high loading efficiency thanks to the high surface area and easy interaction with biomolecules through van der Waals forces, π - π stacking or hydrophobic effect [Erol et al., 2017]. They can be conjugated to extra organic groups to increase their biocompatibility and have potential for targeted drug release upon change in environmental pH [Depan et al., 2011].

Finally polymers are a class of inorganic molecules already well validated as drug excipients. The most notable example is Polyethylene glycol (PEG): thanks to its high hydrophilicity, it is widely used to coat structures (e.g. inorganic nanoparticles, peptides) which in turn carry a drug [Lammers et al., 2009]; or as a stand alone carrier system with high drug payload [Liechty et al., 2010]. The great strength of polymers is their flexibility: as each of their constituent monomers can be either hydrophilic or hydrophobic, they can be engineered to assemble in many different structures, to swell slowly in water

triggering a sustained drug release [Nicolas et al., 2013], or to undergo sol-gel phase transition upon specific changes in the environment [Liechty et al., 2010].

Organic materials for delivery: lipids and DNA A somehow opposite approach for designing drug vehicles consists in using molecules similar to the ones present in the body, such as lipids, DNA and peptides, in an effort to exploit already available biocompatible materials and reduce toxicity [Yoo et al., 2011].

A great variety of lipids is present in the cell membrane, and the chemical space is further enlarged by the production of synthetic ones. The components selected for drug delivery are usually taken from the biological lipidome, modifying the composition and possibly including synthetic molecules to tune the release and robustness to degradation [Yingchoncharoen et al., 2016]. Being amphiphatic, lipids can encapsulate efficiently both hydrophobic or hydrophilic drugs, arranging respectively in micelles (monolayer spheres with the hydrophobic tails facing the interior) or in liposomes (bilayer spheres with a water filled core) [Bunker et al., 2016], and by now, many lipids are approved as delivery agents for cancer and infection drugs [Pattini et al., 2015; Jain and Pillai, 2017].

DNA scaffolds are instead a more novel tool: DNA origami is nowadays an established technique to build three dimensional customised solids [Linko et al., 2015], and the nanometric knowledge about their constituents allows to fine tune them for a triggered release of the content [Douglas et al., 2012]. First studies proved them successful in delivering anticancer agents [Zhang et al., 2014b; Jiang et al., 2012], however they are very sensitive to different cellular environments which challenge their stability. This, united with high production costs and the relative young developments in their manipulation, prevented them to constitute a viable class of carriers so far, but at the same time holds promise for future improvements and applications.

1.4.1 Peptidic scaffolds

Another widely used and trustworthy mimetic vehicle comes, quite surprisingly, from the world of pathogens: viruses have co-evolved with humans, to be able to penetrate into cells where they complete their reproductive cycle

[Lobo et al., 2009]. Therefore their capsid, the peptidic shell encapsulating the viral genome, is highly suitable for cell penetration. The first application sought historically was to employ genome free viruses to stimulate and train the natural immune response against the respective genome-loaded ones, creating viral vaccines - in a concept similar to the inoculation of dead bacteria to counteract the infections caused from them [Lauer et al., 2017]. Later in the history, their potential as cargo carrier was pursued by first modifying the original genetic material to include sequences beneficial for the host cell, and inactivate the ones promoting the infectious duplication at the same time [Daya and Berns, 2008]. To fully exploit the potential of a peptidic carrier many efforts have focused on synthesising *in vitro* gene-free capsids, either as they appear in nature [Wu et al., 2009] or designing artificial building blocks, which assemble in so called Virus-Like particles (VLPs) which trigger lower immune response. Similarly to other delivery vehicles, the surface of VLPs can be functionalised with additional molecules to improve the target selectivity and increase biocompatibility, while the peptidic scaffold grants robustness to the structure. Therefore, VLPs loaded with drugs can be tuned for an efficient intra cellular release [Ma et al., 2012].

A step further in engineering peptidic structures is represented by the design of self-assembling functional blocks from first principles. Indeed, self-assembling peptides can form nanostructures ranging from nanoparticles to nanotubes, nanofibers, nanorods and hydrogels [Fan et al., 2017]. Among their advantages, peptides present biocompatibility and a tunable bioactivity thanks to their chemical diversity, which helps in tailor the assembly toward the target of interest [Fan et al., 2017]. Moreover, the variety of amino acid available makes possible to load peptidic structures with both hydrophilic and hydrophobic drugs, according to their amino acid composition [Ma et al., 2012]. The peptidic self-assembly is modulated by the peptide length and its hydrophobic or hydrophilic character, given by its amino acid composition: simple phenylalanine dipeptides were designed with inspiration from a pathogenic pathway of molecular self-assembly and were shown to self-assemble in a multi-scale process producing nanotubes able to load drug molecules [Silva et al., 2013]. The relatively small diphenylalanine building block is non the less complex as it bears two charged termini (as the process is observed at neutral pH), and two aromatic hydrophobic rings, so that the dipeptide is driven towards

assembly by the hydrophobic forces acting on the phenylalanine side chains and the complementary charges of the termini.

In a different approach, longer sequences can be employed to guide the formation of the local structure, as they organise spatially in well studied secondary structures with known interactions among themselves. This knowledge is possible as proteins are a fundamental component of the human body and as such an updated database of their structure is available (the Protein Data Bank [Berman et al., 2000]) and can be queried to understand how small peptides hierarchically assemble into larger units. Moreover, the vast literature on their interactions with membranes, cell receptors and in general biological components, can inspire the design of building blocks sensible to particular triggers within the body. From this background, the outlook of protein design often goes in the direction of surpassing natural limitations, synthesising exotic, non natural, geometries [Yeates, 2019; Malay et al., 2019] for multi-functional materials.

1.5 Closing the circle: an antimicrobial drug delivery vehicle

Twice in this introduction peptide design has been brought to the reader's attention. First, it can produce antimicrobial peptides with improved potency or selectivity, or reduced toxicity. Second, design can engineer self-assembling building blocks for the formation of delivery scaffolds. As design is not bound to natural rules, it can foresee and imagine multifunctional materials which are not observed in nature. In particular, the introduction above poses the question of whether it is possible to engineer peptides able to perform both an antimicrobial and a delivery function at once (either of drugs or genetic material).

Such self-assembling, antimicrobial compounds would have a twofold interest for medical applications. First of all, self-assembly is functional to the antimicrobial activity: many AMP sequences have a weak potency, and only a high (critical) concentration can trigger the bactericidal mechanism. This is intuitive in the case of the carpet model strategy (see Section 1.2.2), where AMPs lay homogeneously on the surface of the bacterial membrane and breaks it upon sufficient coverage of its area. Also the barrel-stave and toroidal pore

models rely on the mutual interactions between peptides to maintain the pore edges. Generally, as AMPs are positively charged, the localised presence of many copies of a sequence enhances the local electric field and charge imbalance, which are critical to the membrane stability. Second, in order for the assembly to be able to perform the additional delivery function, it must be able to either organise in a tailored structure (for example a capsule able to host a drug), rather than an amorphous aggregate, or to co-assemble with the cargo of interest.

Out of all the possible applications, the most promising is perhaps the use of such vehicles to deliver drugs to treat metabolic or genetic diseases: while the cargo tackles a defect of the host system, the vehicle can counteract the proliferation of bacteria. This is particularly important in situation where the host immune response is weakened and thus infections normally harmless can spread and cause damage. However, it must be noticed that the cargo is not bound to be a small molecule, as long as it can effectively co-assemble with the peptidic carrier. As mentioned in the previous section, gene therapy is also an actively expanding field which looks with interested at the development of vehicles for genetic material. Given that viruses have been the first choice for DNA/RNA delivery so far, peptidic carriers seem a natural evolution of them.

Given the above premises, it is evident the importance of pursuing the research on novel multifunctional peptidic materials. As mentioned when discussing AMPs design, to understand such systems, each of them must be characterised by itself, as a generalised knowledge is still lacking. With this aim, this thesis proposes to elucidate the behaviour of a specific synthetic self-assembling peptide, suitable for antimicrobial activity and gene delivery strategies. Its full characterisation will complete the knowledge on its mechanisms of action and complement the broader information already known on the class of such functional building blocks. This will be crucial to engineer new synthetic blocks with improved characteristics, either regard their antimicrobial activity, assembly performances, or tailored cargo delivery.

1.5.1 Capzip

The molecule capzip has been designed to perform the functions mentioned above at once. To recapitulate, the properties it possesses are:



Figure 1.7: Capzip molecule scheme and bond representation. [TO BE IMPROVED] Adapted from [Castelletto et al., 2016].

1. assembly into nanoscale virus-like capsules with and without nucleic acids. This ensures that the vector can autonomously form and thus there is flexibility in the choice of the cargo;
2. antimicrobial activity of the molecule itself and of the capsule on a time scale useful for therapeutic applications;
3. promotion of gene transfer into mammalian cells when the peptide is co-assembled with the RNA strands, without causing cytotoxic and haemolytic effects.

Furthermore, the design effort aimed at building a template structure of minimal complexity, in order to reduce the synthesis effort to a short sequence. Arguably, short sequences are also more flexible in their assembly: it is thus important to explore them and prove whether even small blocks can form ordered structures.

Based on the above requirements, two design principles emerged: first the employment of a non-linear structure. There is indeed some evidence suggesting that non linear peptides are more prone to assemble in three dimensional structures, opposed to planar ones [?], and this holds in particular for short sequences which do not fold into a defined secondary structure. The second principle consists in the use of a template antimicrobial sequence which is short and has proved potency. Given that AMPs are usually anionic, the co-assembly with anionic RNA sequences is arguably inherited by consequence.

To satisfy the above guidelines, a short peptidic scaffold constituted by a β -Alanine and two Lysins has been engineered. Three identical copies of the antimicrobial sequence of choice are covalently bonded to the N-terminus of the scaffold sequence and to the nitrogen atom of the Lysin residues side

chains (Figure 1.7). Regarding the antimicrobial sequence selected, it has been derived from the antimicrobial peptide bovine lactoferricin, which is in turn a portion of the Lactoferrin protein.

Lactoferrin Lactoferrin is an iron binding protein present in milk (in which it is most abundant, hence its name), saliva and other secretions, as well as in polymorphonuclear leukocytes. It works as an iron binder and provides a natural defence against bacteria and fungi [Sánchez et al., 1992; Arnold et al., 1977, 1980; Kirkpatrick et al., 1971; Jahani et al., 2015], constituting a first defence for infants.

Lactoferrin contributes to bacterial suppression in several ways. At present, its known modes of action fall in three categories: first, thanks to its iron sequestering capabilities, it removes essential substrate required for bacterial growth [Farnaud and Evans, 2003]; second, it interacts with bacterial membranes and in particular binds to the lipopolysaccharides of bacterial walls, oxidising them and affecting the membrane permeability with consequent cell lysis [Farnaud and Evans, 2003]; finally it is implicated in the stimulation of different immunological cells (killer cells [Shau et al., 1992], polymorphonuclear leukocytes, and macrophages [Gahr et al., 1991]). The peptide fragment responsible for binding of lactoferrin to the bacterial membrane, named lactoferricin (Lfcin), has been identified near its N-terminus and found to have a more potent bactericidal effect than intact lactoferrin on a wide range of bacteria [Gifford et al., 2005; Bellamy et al., 1992; Tomita et al., 1994; Wakabayashi et al., 1996]. Similarly, a synthetic short peptide derived from a subsequence of human lactoferricin has been proven effective against bacteria as it depolarises the cytoplasmic membrane decreasing the pH gradient [Aguilera et al., 1999].

The bovine homolog of lactoferricin (LfcinB) has a higher bactericidal potency than human lactoferricin on several bacteria [Cochran et al., 2001] and therefore has been more extensively studied. Its active core LFC is a 25-amino acid sequence which adopts a helical conformation in the full structure but, once isolated, crystallises in a β -hairpin with a disulfite bridge nearby the terminals which stabilises the fold, but was shown to be not essential for bactericidal activity [Cochran et al., 2001]. Further experiments on LfcinB subsequences identified an even shorter antimicrobial core, constituted by the six amino acids RRWQWR [Schibli et al., 1999]. This core presents a charac-

teristic Tryptophan zipper motif WTW, which appears very often in nature in β -turn and β -sheet conformations, paired to another copy of the same motif, so that Tryptophan rings from facing strands are packed tightly against each other in an alternated way [Cochran et al., 2001] (Figure –). In general, the six amino acid sequence contains both charged and hydrophobic residues, in line with the usual composition of antimicrobial peptides. Accordingly, its antimicrobial action is likely derived from the interaction with biological membranes through charge recognition first and aromatic rings insertion in a second moment.

To further elucidate this mechanism, several experimental investigations have been carried both on LfcinB and its subsequences. First, the structure of LfcinB in solution has been investigated by NMR (Nuclear Magnetic Resonance), resulting very flexible [Hwang et al., 1998]. Then, the binding of its antimicrobial core to sodium dodecyl-sulfate micelles was studied [Schibli et al., 1999], suggesting a favourable interaction of aromatic residues with the micelles surface. Similar experiments were performed on large unilamellar vesicles, constituted by lipids modelling biological membranes: ePE:ePC was chosen as a model of a mammal membrane, and ePE:ePG or ePC:ePG for a bacterial one. The experiments showed preferential binding to the latter ones, based on Tryptophan fluorescence [Nguyen et al., 2005], suggesting a selective antimicrobial action on anionic membranes. Additional experiments have been performed on the full sequence or mutated subsequences [Tsutsumi et al., 2012; Arseneault et al., 2010] to investigate the binding to other different model membranes but, as the systems investigated are slightly different, as well as the experimental conditions, therefore it is difficult to relate them and give a unified interpretation of the modes of action of lactoferricin derived peptides.

Finally, an alanine scanning has attempted to clarify the role of each amino acid in the antimicrobial activity of the LFC peptide [Strøm et al., 2002]. The results suggested a binding function for the Tryptophan residues, in line with one of the roles Tryptophan can assume in antimicrobial peptides [Chan et al., 2006]. Other possible roles involve its propensity to form hydrogen bonds, in which case the residue would position itself at the interface between solution and membrane, rather than inside the latter (which happens instead when Tryptophan residues have a binding role).

AFM/TEM	cryo-em
fluo hollow capsule	fluo RNA uptake

Figure 1.8: ... Reproduced from [Castelletto et al., 2016]

The designed block From the active core of LFC (of sequence RRWQWR), a mutated sequence was obtained to comply the design criteria of a self-assembling building block. Two mutations were introduced to favour the assembly of arms belonging to different molecules in an antiparallel fashion. Specifically, given that the original RRWQWR sequence is found in a β -sheet (at least in the crystal lattice), the mutations aim at promoting a similar structure. Therefore, the Glutamine residue and the C-terminal Arginine of the lactoferrin motif were replaced with Threonine and Glutamic acid residues to have a self-complementary sequence RRWTWE: the pairing is promoted by the attraction of opposite charges at the ends of the sequence. Three copies of this sequence were thus covalently bonded to the scaffold described previously and shown in Figure 1.7, to obtain a self-assembling molecule in a three dimensional shape, hosting multiple copies of an actively antibacterial sequence.

1.5.2 A viable systems: experimental background and question

Many experiments have been performed to verify that capzip had the characteristics it was designed for. The set of experimental results obtained on the molecule has been published in Reference [Castelletto et al., 2016], while more recent results extend and consolidate the previous findings.

Experimental results First, the assembly ability have been tested: the peptide does not show assembly in pure water (as verified by Dynamic Light Scattering), while in biological buffer (MOPS, 150 mM) at physiological pH of 7.0 it forms capsules with dominating size range of 20-200 nm. This is confirmed by images of the capsules obtained with multiple techniques, namely transmission electron microscopy (TEM) (Figure 1.8), atomic force microscopy (AFM), and cryo-scanning electron microscopy (SEM). The fine structure of these assemblies appears irregular to the resolution power of such techniques. Some insight into the details of the assembly is given by Circular Dichroism

(CD) spectra, which show a profile characteristic of β -turns and contain elements of a β -sheet structure and of indole rings, with minima at $\lambda \sim 200$ nm and 214 nm. Complementary evidence about the overall shape of the assembly was provided by the cross-sectional analysis of the assembled capsules by fluorescence microscopy using fluorescein to label capzip. The signal comes from the wall of the capsule only, showing an inner cavity (Figure 1.8). Finally, small angle X-ray scattering (SAXS) measurements were consistent with compact capsules interfacing with solvent.

The assembly process is also tested and monitored in combination with small interfering RNAs (siRNA): as mentioned in section 1.3, these sequences are a promising tools for RNA interference techniques which aim at inhibiting the expression of specific genes, however, they are easily degradable and thus difficult to deliver to the target cell without an appropriate vehicle. The co-assembly of a 21 base pairs duplex with the peptide shows the formation of structures similar to the ones formed by the stand alone peptide only: CD spectra highlight the helical signal from RNA together with the features proper of the peptide.

These co-assembled structures were tested for siRNA delivery in HeLa cells, showing that the presence of the peptide favours the internalisation of the genetic material with respect to the transfection results of a pure siRNA control. The delivery of fluorescent siRNA (Figure 1.8) showed that the internalisation occurred within the first hours from the transfection in localised regions of the cytoplasm, suggesting an endocytic uptake. This distribution was stable over the first five hours of incubation after which the fluorescence signal decayed. Flow cytometry assays quantified the increase in siRNA uptake levels due to the presence of capzip, confirming that this molecule is competitive with other commercial transfection reagents (like Lipofectamine[®], unpublished results).

To further quantify the level of RNA internalisation, a mRNA knockdown experiment was performed on a HeLa cell line with two housekeeping genes, ACTB (β -actin, targeted) and GAPDH (reference) [Crombez et al., 2009]. The silencing of β -actin mRNA was detected 22 ± 2 hours after transfection; and its knockdown “fitness” was expressed relative to cells treated with siRNA alone (background) and normalised against viable cell counts (Figure 1.8). Capzip fitness was lower than Lipofectamine[®] one, however cells treated with capzip remained viable after 24 or 48 hour, resulting in higher cell counts than

the samples treated with the commercial reagent, suggesting that capzip has little cytotoxicity. The experiment above was performed at neutral to positive charge ratio close to one (where each siRNA molecule has a -42e charge and capzip a $+6\text{e}$ charge), as test experiments performed at higher peptide-to-siRNA ratio showed no improved uptake.

Finally, the peptide does exert an antimicrobial function: the non-assembled peptide has shown to be effective against both Gram positive and negative bacteria (*E. coli*, *P. aeruginosa* and *S. aureus*), with no haemolytic effects and minimum inhibitory concentrations typical of other antimicrobial agents. On Supported Lipid Bilayer with negative total charge (mixed DLPC and DLPG, 3:1 ratio), the capsules create localized pores within minutes, as proven by AFM experiments repeated in time. The pore depth ranges between 1.4 and 2.2 nm, which is smaller than the radii of the capsules, however it is sufficient to disrupt the structure of the membrane. Finally, to prove the viability of capzip as antimicrobial agent *in vivo*, it was used to counteract methicillin-resistant *S. aureus* (MRSA) infections in *G. mellonella* larvae. The particular bacterial strain used was susceptible to vancomycin, which could be used as control: the larvae treated with capzip showed survival rates significantly higher than the untreated control, and comparable to those treated with high doses of vancomycin (unpublished results).

Open questions Despite the success of the experiments mentioned above, there is much information still to be uncovered on the precise mechanism of action of such peptide.

Specifically, both the assembly process and the antimicrobial mechanisms contain some unknown: regarding the former, it is important to understand which amino acids or sub-structures allow the pairing of molecules, whether such pairing is specific or not, how reversible it is, and how rigid the final structure is. Regarding the latter, it must be highlighted what molecules in the membrane the peptide binds to, and how this binding affects the full membrane structure. Finally, as there is evidence that the assembled molecule is a more powerful antimicrobial compound than the single molecule, it is interesting to understand whether any cooperative action is taking place or the enhanced antimicrobial power of the assembly is due only to the localised higher concentration.

Even if further experiments or future improvements in the techniques already employed might tackle some of the aspects above in a near future, arguably no experimental outcome can provide an atom-by-atom knowledge of the processes of interest in any time soon. Ideally though, one would like to track each of them, i.e. the processes happening in any the environments capzip has been exposed to (physiological solution, supported lipid bilayers, bacterial extracellular matrix, mammal cell membrane and cytoplasm) both in space and time with the finest level of details, and the impossibility of that leaves large gaps in the understanding of the system.

1.6 A computational approach to understand capzip

The gaps mentioned in the characterisation of the systems prompts for new investigations in order to complement the knowledge already provided.

Beside the quest to enrich the fundamental knowledge on self-assembling peptides and antimicrobial ones, the understanding of this very system is crucial for its further development. We outlined already in Section 1.2.4 how antimicrobial peptide design can proceed from already viable templates and empirical principles, when first principles are not available. Similar rules hold for designing self-assembling peptidic materials, to obtained tailored delivery vehicles (see Section 1.4). Therefore, a full knowledge of the interactions between peptides and between their assembled structures and the membrane, i.e. of the mechanism of its functions, will drive the engineering of new likewise peptides. A knowledge-driven design would hopefully provide new blocks suitable for a double action as the one capzip performs, and this in a shorter amount of time than a research based on less information or on a trial-and-error procedure of mutations in the chemical composition or in the architecture of the molecule. A few examples of possible knowledge-related improvements include the following:

- the knowledge of capzip binding mode to the bacterial membrane might suggest its suitability as a broad range spectrum compound or the possibility of tuning its action against specific pathogens;
- understanding the molecule-molecule interactions classifies the robustness of the assembled structure and the possibility of designing block

which disrupts under particular chemical conditions only;

- querying the electrostatic profile of the assembled structure suggests which type of molecules, other than siRNA, could be efficiently co-assembled and thus delivered.

In recent years, computational techniques are becoming increasingly accessible and sophisticated and can usefully complement the incomplete experimental knowledge to give more insights into how biological systems work. For this reason, it seems natural to query such techniques to study the capzip system as well. Zooming into the details of the interactions can be performed via a theoretical modelling of the system in time, and thus through the simulation of its evolution, starting from few basic principles and the knowledge of the chemical composition of its parts. The technique this work focuses on is Molecular Dynamics simulations, which aims at reproducing the behaviour of a system of atoms in a semi-classical description using basic physical laws, as it will be described in details in the next chapter.

Thus it is the aim of this thesis to prove that Molecular Dynamics simulations can clarify further details on the assembly mechanisms of capzip and on its interactions with biological membranes, in order to gather more information on the system and contribute in the future to the designed of new molecules with enhanced functional capabilities.

Chapter 2

Methods

MOLECULAR DYNAMICS simulations have been rightly defined as the ‘Computational Microscope’ [Lee et al., 2009; Dror et al., 2012] as they offer otherwise inaccessible insights into the molecular details underlying conformational changes of proteins and nucleic acids. Computational methods and tools based on MD are routinely applied in structural biology to quantitatively characterise the dynamics and thermodynamics of proteins and their complexes. The increasing computational power available, and the flexibility of the algorithms which can be implemented on different platforms, have made possible to access molecular dynamical properties inaccessible to experiments. These techniques and the associated force fields are also commonly used in the process of structure determination from NMR data and theoretical structure prediction [Vogel and Huster, 2017; Heo and Feig, 2018].

Modelling and simulating a biological system consists in describing its components and their mutual interactions, and implementing the laws of physics in the attempt to reproduce its natural evolution. A quantum mechanic description would be the most accurate but expensive to achieve for large systems. To facilitate the task, several simplified models have been devised, each most suitable to investigate particular cases. In particular, a large class of models focuses on a classical mechanics description of the dynamics: by increasing the size of the systems and with longer time spans described, the classical approximations will become more accurate, and objectively the only possible representation computationally affordable.

We briefly present here the core theory and implementation of classical MD simulations, together with a discussion of their strengths, limitations and

comparison with experiments. Understanding their methodology provides the interpretative key with which simulations must be designed, run and interpreted in each specific case [van Gunsteren et al., 2006]. A review of relevant successes of MD simulations will complete the chapter.

2.1 Algorithms for Molecular Dynamics

In this section, the core algorithm behind MD simulations is presented, together with the corrections and refinements implemented on it, as it sets the ground for the approximations to follow: in classical MD, the system will be processed by an engine based on classical mechanics rules and finite steps approximations, and the forces acting on it can thus be modelled with a similar degree of approximation.

2.1.1 The Newton's law

In the aforementioned framework, Newton's second law of motion rules the dynamics, stating that the acceleration \mathbf{a} that a particle is subject to at time t , depends on the total force \mathbf{F} acting on the particle itself and on its mass m (bold denotes vectorial quantities):

$$\mathbf{F}(t) = m \cdot \mathbf{a}(t). \quad (2.1)$$

As the acceleration $\mathbf{a}(t)$ is the second derivative of the position $\mathbf{r}(t)$ with respect to time, given the initial position and velocity of the particle $(\mathbf{r}(t_0), \mathbf{v}(t_0))$, their temporal evolution can be computed integrating $\mathbf{a}(t) = \mathbf{F}(t)/m$ as follow:

$$\mathbf{v}(t) = \mathbf{v}(t_0) + \int_{t_0}^t \frac{\mathbf{F}(t')}{m} dt'; \quad (2.2)$$

$$\mathbf{r}(t) = \mathbf{r}(t_0) + \int_{t_0}^t \mathbf{v}(t') dt' + \int_{t_0}^t \int_{t_0}^{t'} \frac{\mathbf{F}(t')}{m} dt'' dt'. \quad (2.3)$$

In the case of complex biomolecular systems with many particles and multiple interactions acting between them, it is impossible to integrate analytically Equations 2.2-2.3, while a different and feasible approach consists in discretising them. The idea is to consider very short time steps of length Δt , so that in such intervals the forces can be considered constant, and the integration of

Equations 2.2-2.3 becomes trivial. The simplest choice consists in considering for the integration both the acceleration and the velocity at time t_0 , producing the Euler algorithm:

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \frac{\mathbf{F}(t)}{m} \Delta t; \quad (2.4)$$

$$\mathbf{r}(t_0 + \Delta t) = \mathbf{r}(t_0) + \mathbf{v}(t_0) \Delta t + \frac{\mathbf{F}(t)}{m} \Delta t^2. \quad (2.5)$$

This procedure contains approximation of the order of $(\Delta t)^2$ that will accumulate step after step, biasing greatly the outcome. However, a careful choice of the values to integrate allows to reduce considerably this approximations. For example, choosing the acceleration value at time t_0 but the velocity at time $t_0 + \Delta t/2$ decreases the error down to orders of $(\Delta t)^4$. This framework is at the basis of the so-called leap-frog algorithm, which is used in the vast majority of MD engines:

$$\mathbf{v}\left(t_0 + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\mathbf{F}(t)}{m} \Delta t; \quad (2.6)$$

$$\mathbf{r}(t_0 + \Delta t) = \mathbf{r}(t_0) + \mathbf{v}\left(t_0 + \frac{\Delta t}{2}\right) \Delta t. \quad (2.7)$$

This algorithm can thus “solve” every possible Newton equation, at the expenses of some precision.

Considering that MD deals with bonded atoms belonging to multi-atoms molecules, the length of these bonds must be kept within a physically meaningful value. The approximate procedure mentioned above may however give raise to unphysical configurations, even if the forces acting between single atoms tend to bring them in the correct geometry. On the other hand, if the initial configuration was far from equilibrium, very abrupt changes in the bond length can arise in the initial steps, which are unphysical as well. To correct this, constraint algorithms are applied after the update of atoms positions, to limit the change in bond length. The ones used in this work are LINCS (Linear Constraint Solver) [Hess et al., 1997] and SETTLE [Miyamoto and Kollman, 1992]: the first finds a solution iteratively within an approximation tolerance, as the problem of constraints is hardly solvable analytically; the second is an exact implementation of the solution for rigid bodies of three elements only, and as such is useful for the treatment of water molecules (which have three

atoms in their atomistic description).

2.1.2 Thermostats and barostats

As an addition to the geometrical constraints algorithms mentioned above, other specific algorithms have been developed to more realistically reproduce the simulation's conditions of choice, for example Temperature and Pressure.

Most experiments are conducted under constant temperature, so that it is desirable to reproduce this condition in simulations. To set up a temperature, at the beginning of the simulations particles can be given random initial velocities sampled from the Maxwell-Boltzmann distribution, which describe velocities of atoms of a noble gas at temperature T . The velocity of each of them will be influenced by the specific interactions occurring in the system but, in a constant temperature environment, the total average kinetic energy (proportional to T) must remain constant. Even in absence of any dissipative term in the dynamics, the approximations performed in the MD algorithm make this quantity to drift away from its initial value, therefore to ensure temperature is maintained throughout the simulation, thermostat algorithms have been devised.

The principle behind a thermostat consists in rescaling the velocity of one or few selected particles at fixed interval of times, to restore the correct average kinetic energy. This fixed interval of time must not correspond to the timestep itself, as the goal of a thermostat is to maintain the average temperature, and not its value at all times, as fluctuations are expected in natural systems. Moreover, it is strongly recommended to couple solute and solvent to separate baths, to ensure that both maintain the correct temperature. Indeed, it is possible that the energy exchanged between solute and water (or other components) is not perfect, due to different conditions adopted for their simulations, like, for example, cut-offs or restraints. The most used thermostat algorithms are the Berendsen [Berendsen et al., 1984], Nosé-Hoover [Nosé and Klein, 1983; Hoover, 1985], Andersen [Andersen, 1980] and velocity rescale [Bussi et al., 2007], which differ in the quantities selected for the velocity rescaling.

Another macroscopic condition one wishes to maintain is either the volume or the pressure of the system. While maintaining the volume constant is easy (and, combined with constant temperature, gives the NVT ensemble), pressure

regulation (i.e. maintaining a NPT ensemble) requires a barostat. Pressure is directly proportional to the average quantity of motion exchanged between the particles and the walls of the box they are confined to, which depends on the frequency of collision and thus on the extent of the box. Barostats change the box size to regulate the pressure. It has to be noticed that most MD simulations are run under periodic boundary conditions, i.e. a particle which exits from the simulation box during a move is brought back on the opposite side of the box, leaving the box density constant. This mimics the presence of an infinite number of equivalent boxes one next to the other, and alleviates the finite size effects that arise when simulating small systems. In this scenario, particles are not bouncing on the box walls, rather a virtual pressure is computed from the velocities of the ones trespassing the box boundaries during a move. Similarly as for thermostats, the coupling frequency of barostats must be larger than the time step. Usually all the box dimensions are rescaled by the same amount. However, in the case of anisotropic systems like lipids, to maintain the correct pressure, the directions parallel to the membrane plane can be rescaled separately with respect to the one perpendicular to it. Also for pressure coupling several algorithms can be used: the Berendsen [Berendsen et al., 1984], Parrinello-Rahman [Parrinello and Rahman, 1981], or Martyna-Tuckerman-Tobias-Klein (MTTK) [Martyna et al., 1996] (to be used in conjunction with the Nosé-Hoover thermostat), are all barostats which differ in the way they approach the desired pressure (e.g. exponentially, in an oscillatory way, etc.).

2.2 The force field problem

Once the equations of motion and the control algorithms are set up, the next challenge is represented by modelling the forces and thus the potential energy function. Force fields for classical MD simulations usually rely on the breakdown of interactions into several independent, additive and derivable terms, identified on an empirical physical basis. We report here the functional form adopted for the GROMOS force field [Oostenbrink et al., 2004; Schmid et al., 2011] as implemented in the GROMACS MD engine [Berendsen et al., 1995; Abraham et al., 2015, 2018], as an example of the structure of a classical force field. Other force fields can have slightly different implementations, however the general classification of interactions and the type of functional forms used

are similar.

Covalent (bonded) interactions Covalent interactions are modelled with potential energy terms representing bond stretching, angle bending, improper and proper dihedral angles torsion. The functional form of the potential-energy for bonded interactions aims at a simplified, semi-classical description of the sub atomic motion of molecules, assuming harmonic-like vibrations around the equilibrium position of the bond, angle or dihedral in exam.

Specifically, in the GROMOS force field, a bond between atoms i and j is described by a fourth power potential (similar to a harmonic form, but computationally more efficient):

$$V_b(\mathbf{r}_{ij}) = \frac{1}{4} k_{ij}^b (|\mathbf{r}_{ij}|^2 - b_{ij}^2)^2 \quad (2.8)$$

where the force constant k_{ij}^b is given in kJ/mol/m² and b_{ij} is the equilibrium position of the bond between atom i and j .

The deviation of the angle between three atoms i , j and k from the preferred value θ_{ijk}^0 is implemented through a cosine based angle potential:

$$V_a(\theta_{ijk}) = \frac{1}{2} k_{ijk}^\theta (\cos(\theta_{ijk}) - \cos(\theta_{ijk}^0))^2 \quad (2.9)$$

$$\text{with: } \cos(\theta_{ijk}) = \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{r_{ij} r_{kj}} \quad (2.10)$$

with k_{ijk}^θ in kJ/mol.

Improper dihedrals are used to ensure ring planarity and control the chirality of some tetrahedral centres. They are described through a harmonic potential:

$$V_{id}(\xi_{ijkl}) = \frac{1}{2} k_{ijkl}^\xi (\xi_{ijkl} - \xi_{ijkl}^0)^2 \quad (2.11)$$

where the ξ values are given in degrees and the force constant in kJ/mol/rad². By convention, the improper dihedral for a set of four atoms i , j , k and l , is taken as the angle between the plane defined by atoms (i, j, k) and the one defined by atoms (j, k, l) .

Finally, the last bonded interaction is represented by proper dihedrals,

described through a periodic potential:

$$V_d(\phi_{ijkl}) = k_{ijkl}^\phi \left(1 + \cos(n\phi_{ijkl} - \phi_{ijkl}^0) \right) \quad (2.12)$$

following the convention that ϕ_{ijkl} is the angle between the (i, j, k) and (j, k, l) planes, with i, j, k , and l four subsequent atoms (for example along a protein backbone). A value of zero for a proper dihedral corresponds to a *cis* configuration; n denotes the number of equally spaced minima available for the dihedral in a 360° turn. The constant k_{ijkl}^ϕ is expressed in kJ/mol.

It must be noticed that these potentials can not model the rupture of a bond: for this, more sophisticated descriptions are needed.

Non bonded interactions Non bonded interactions include the short range Pauli repulsion, the “mid”-range van der Waals attraction, and the long range electrostatic term.

The first two can be modelled together by a Lennard-Jones potential. Its functional form, describing the interaction between two neutral atoms at distance r , models the long range dispersion with a r^6 behaviour typical of the dipole-dipole interactions found in noble gases (London dispersion forces), while the Pauli term is represented by a r^{12} behaviour to ease the computation in relation with the previous one:

$$V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]. \quad (2.13)$$

Two parameters, ϵ and σ , tune the interaction strength and the equilibrium distance. They are fitted against experimental data and are specific of each pair of atoms species.

The Coulomb energy between two charges q_1 and q_2 at distance r is represented by the Coulomb law:

$$V_C(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{\epsilon_r r_{ij}} \quad (2.14)$$

with ϵ_0 the dielectric constant of vacuum and ϵ_r the relative dielectric constant, introduced to properly take into account the screening provided by the material surrounding the object, as polarisability is not included in this description.

The treatment of non-bonded interactions requires particular care because

of their long range nature: in every point of the simulation box many forces from distant atoms are acting at the same time, making the prediction of the outcome difficult, moreover small shift in the parameters choice can give very different “macroscopic” results. The van der Waals forces decay fast, therefore the tail of their functional can be cut after a threshold distance with little impact on the outcome; while Coulomb interactions, with their slower decay, must be taken into account throughout the whole simulation box. Many algorithms have been devised to efficiently compute them, like the Particle Mesh Ewald [Essmann et al., 1995] or the Reaction Field [Tironi et al., 1995] approaches.

Finally, all biomolecular force fields, and in particular their van der Waals interactions, are parametrised to describe systems at room temperature, therefore simulations performed at substantially different temperatures must be interpreted carefully.

2.2.1 Force fields: classifications

Many force fields for classical MD simulations adopt a functional form equal or similar to the one described above. Their difference lies in the number of degrees of freedom modelled, in a hierarchy of descriptions proceeding from detailed to coarse (Figure 2.1). Three possible classes of descriptions are:

- all-atoms force fields, where all the atoms are present in the description, represented as spheres of variable size according to their van der Waals radius (e.g. proportional to σ in a Lennard-Jones model). Examples of all-atoms force fields are AMBER [Maier et al., 2015; Dickson et al., 2014; Wang et al., 2004], CHARMM [MacKerell et al., 1998; Klauda et al., 2010; Huang and MacKerell, 2013] and OPLS all-atom [Jorgensen and Tirado-Rives, 1988].
- united atoms force fields, similar to the previous ones but where non-polar hydrogens are incorporated in the heavy atom they are bonded to. The “united atom” is given a new σ parameter and increased mass according to how many hydrogen it includes. The GROMOS force field [Oostenbrink et al., 2004; Schmid et al., 2011] follows this philosophy, and the OPLS force field has also a united atom version [Jorgensen et al., 1996].

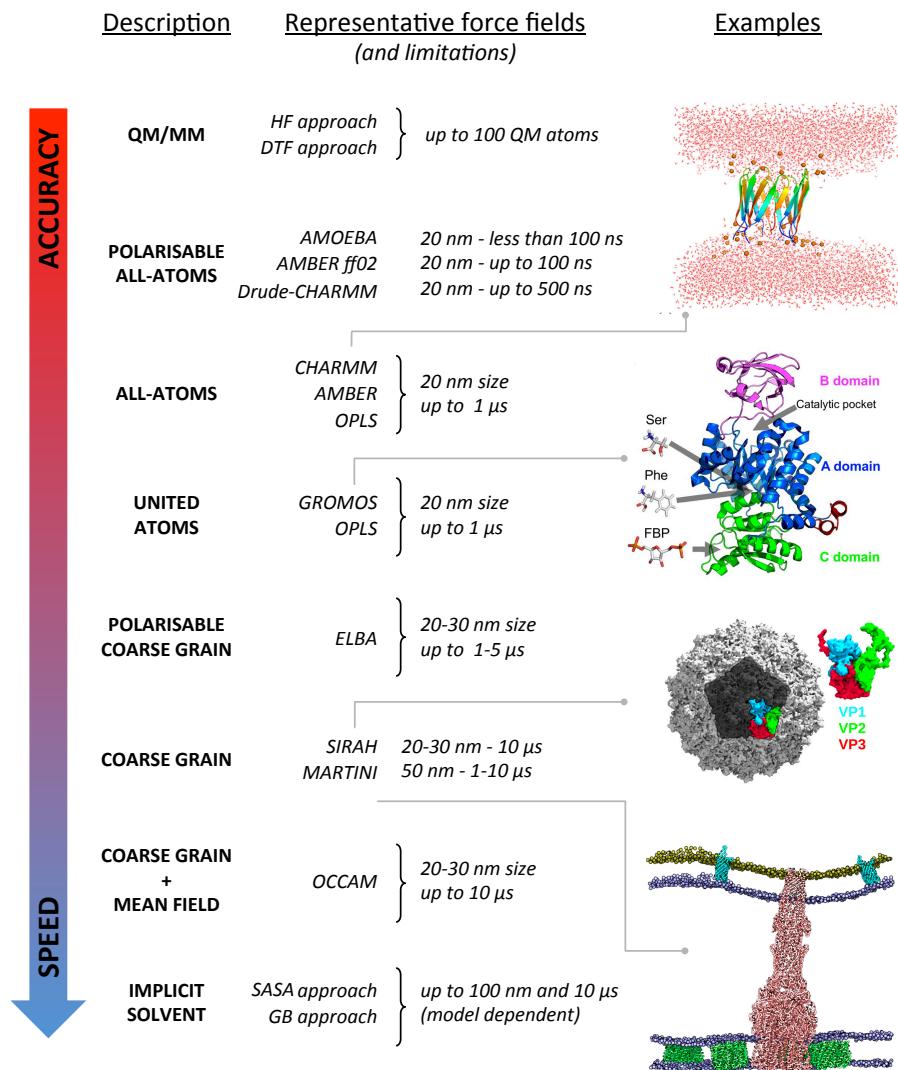


Figure 2.1: List of most popular simulation force field for biomolecules, ordered from detailed to coarse (references in Section 2.2.1). On the left, snapshot of notable systems simulated with the force fields CHARMM (adapted from Lipkin et al. [2017]); GROMOS (adapted from Macpherson et al. [2019]); SIRAH (adapted from Machado et al. [2017]) and MARTINI (adapted from Samsudin et al. [2017]).

- coarse-grain force fields, which group together in one unique bead few atoms, to reduce the number of variables to compute. The clustered atoms are such that their mutual distances are expected to vary little with respect to the ample movements of components of the system far away from each other (which will be grouped in different beads). The MARTINI [Marrink et al., 2007; Monticelli et al., 2008; de Jong et al., 2013] and SIRAH [Machado et al., 2018; Barrera et al., 2019] force fields belong to this category.

We now give a more detailed insight in the characteristics and parametrisation strategies of the atomistic and coarse-grain force fields which will be employed in the work of this thesis.

2.2.2 The GROMOS force field

All-atoms and united atoms force fields are parametrised against first-principle or experimental values. While for the all-atom force fields AMBER and CHARMM the parametrisation is based on ab initio quantum mechanics calculations refined against experimental data [Maier et al., 2015; Dickson et al., 2014; Wang et al., 2004; MacKerell et al., 1998; Klauda et al., 2010], the united atom GROMOS force field relies on the reproduction of heat of vaporization of small molecules and free enthalpies of solvation of small compounds in different solvents, at physiological temperatures and pressures [Oostenbrink et al., 2005; Schmid et al., 2011; Reif et al., 2013]. This procedure sets not only the constants of the bonded interactions, but also the partial charges of the atoms inside a molecule: as no electrons are included for the sake of efficiency, their redistribution across atoms which are bonded is modelled through fractional charges assigned to each atom (while the total charge of a molecule must sum to an integer). Moreover, it is assumed that the parametrisation performed for small moieties can be transferred to a larger compound including these moieties. This limits the number of chemical groups to be described in order to simulate biomolecules.

In every MD simulation, the description of water is crucial. Out of the many water model proposed, the GROMOS parametrisation has been performed with a flexible simple point charge (SPC [Berendsen et al., 1981]) model. This description represents water as a three atoms molecule, with a negative charge

on the oxygen and a positive complementary one on the two hydrogen atoms, and allows flexible hydrogen-oxygen bonds. This model reproduces correctly the density and dielectric permittivity of water. To be noticed that, computationally wise, water molecules are the vast majority of the particles involved in a simulation and thus a significant fraction of the computer time is spent in updating their positions and calculating solvent-solvent interactions.

The improvement of computational techniques and reparametrisation strategies prompts the periodical release of newer versions of force fields. In the present work, we employed version 53a6 of the GROMOS force field [Oostenbrink et al., 2004] for the set of simulations involving peptidic assembly in solution, while we switched to 54a7 [Schmid et al., 2011] and 54a8 [Reif et al., 2013] for the simulations involving biological membranes. While it is advisable to always have a coherent set of parameters across simulations, to compare their outcome in a consistent manner, when extending the system simulated to include membranes, we deemed the newer parameter sets more suitable because of the improvements introduced in the phosphocholine head parametrisation (see Chapter 4 for a complete discussion on lipid parametrisation in GROMOS).

2.2.3 The SIRAH force field

The first coarse grain force field we introduce groups multiple atoms in one bead but aims at maintaining chemical and structural details of the biomolecules described. As such it sets itself between the atomistic GROMOS description, and the coarse grain MARTINI force field [Marrink et al., 2007; Monticelli et al., 2008; de Jong et al., 2013] which will be introduced in the following.

Two approaches are possible to develop a coarse-grain description: parameters can be fit directly to global quantities derived from experiments, in a top-down approach as performed in the atomistic GROMOS parametrisation; or coarse-grain simulations results can be fit to outcomes from atomistic ones, in a bottom-up approach.

SIRAH [Machado et al., 2018; Barrera et al., 2019] is a top-down force field derived to fit structural properties of proteins. It aims at reducing the complexity of an atomistic description while still being able to reproduce the correct secondary structure of proteins across a wide variety of folds contained in the PDB, together with a correct representation of their dynamics.

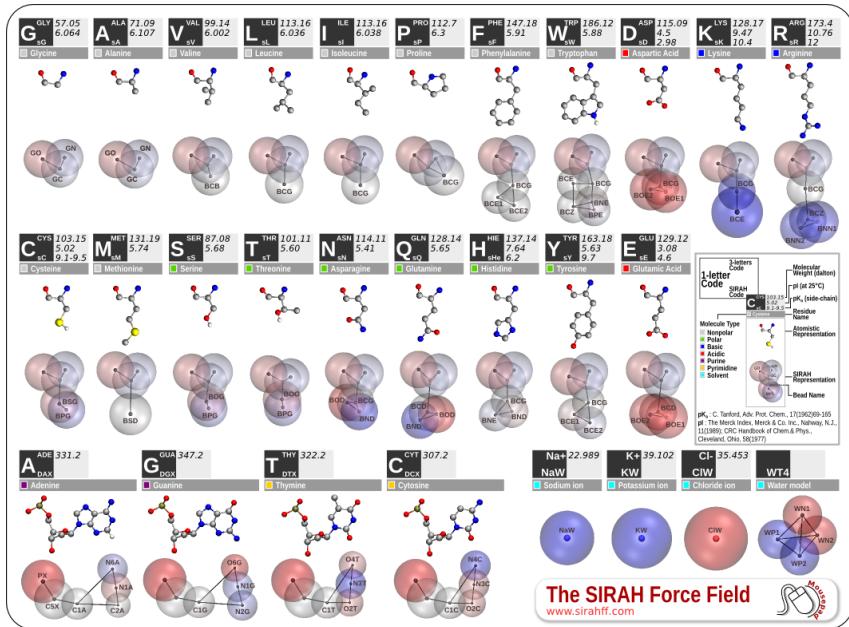


Figure 2.2: Description of amino acids and nucleic acids in the SIRAH force field. Reproduced from Group of BioMolecular Simulations [2019].

To obtain this, it opts for a non-uniform granularity, i.e. according to the region of interest a different number of heavy atoms is grouped in a bead, from a minimum of two up to four. In the case of proteins, it maintains the backbone flexibility by grouping NH, $C_\alpha H$ and CO in three different beads, while side chains are represented with less details, generally grouping three atoms together. A schematic of the mapping for each amino acid is shown in Figure 2.2. Contrary to force fields where the amino acid backbone is mapped to one bead only, the SIRAH description allows to reproduce secondary structures without recurring to additional constraints. The dual granularity approach is based on physico-chemical intuition, and is more difficult to generalise than a uniform one. Nevertheless, the force field has been recently extended to lipids [Barrera et al., 2019], while it comprised a parametrisation for DNA molecules since its infancy.

The modelling of water in a coarse-grain force field is also critical: usually, a few water molecules are grouped together in one bead. This has two implications: water particles are large and thus cannot solvate very narrow pockets; moreover, collapsing the molecules in one single point in space removes the separation of charges and the characteristic dipole every water molecule should have is lost. The dipole of water is responsible for hydrogen bonds forma-

tion and for the electrostatic screening observed in an aqueous solution. Such screening can be roughly modelled tuning the relative dielectric constant, but as this is a mean field approach, it cannot account for local effects. To partially obviate to that, SIRAH force field maps four waters to a tetrahedral molecule, with one bead on each vertex: all the bonds are rigid, and the structure serves the purpose of having a repartition of plus and minus charges, by assigning a positive charge to two vertices and the opposite charge to the other two, giving a polarisable structure. The geometrical arrangement reproduces the tetrahedral network of water molecules observed in its liquid state, which is characteristic of this fluid and tunes its remarkable properties.

Based on the above premises, SIRAH force field simulations of different peptides and proteins in solution proved to match the relative NMR results, showing a good reproduction of secondary structures; simulations of lipids randomly oriented in water showed the formation of an organised bilayer, and the expected behaviour of a few transmembrane proteins in model membranes was correctly reproduced [Machado et al., 2018; Barrera et al., 2019].

2.2.4 The MARTINI force field

The MARTINI force field is another popular coarse-grain description of biological molecules [Marrink et al., 2007; Monticelli et al., 2008; de Jong et al., 2013]: developed originally with a focus on lipids, much earlier than SIRAH, it has been then extended to include proteins, small ligands and DNA/RNA molecules.

MARTINI opts for a four-to-one approach, i.e. four heavy atoms are grouped in one bead, resulting in a uniform graining and a coarser description than the SIRAH one. The only exception to this scheme is represented by rings molecules, where a two-to-one approach is needed to maintain the circular topology (consequently, these beads have a reduced mass with respect to the others, all described with the same mass value).

The number of bead types has been kept to the minimum necessary to represent biological molecules. They are organised systematically in polar, non-polar, apolar, or charged, and each type has a number of subtypes with increasing polarity to differentiate the chemical nature of the underlying atomistic structures. This systematic parametrisation can be easily transferred to

new compounds, without the need of introducing new bead types, analogously to what GROMOS does when parametrising small moieties.

Similarly to GROMOS, the MARTINI force field chooses a top-down approach to parametrise non-bonded interactions, tuning them against experimental partitioning free energies between polar and apolar phases, while bonded interactions are derived from reference all-atom information, in a bottom-up approach. Specifically, they are designed so that the results from coarse-grain simulations match the structural data of the underlying atomistic geometry, derived either from available structures or atomistic simulations. For example the distribution of the length of a bond in a molecule would be mapped to its coarse-grain version (i.e. computing the corresponding bead-to-bead distance) and compared with the respective one obtained from coarse-grain simulations.

The four-to-one mapping implies that the amino acid backbone is represented by one bead only (Figure 2.3(a)), preventing the description of directional bonds which are key to reproduce the secondary structure. The bonded parameters partially account for this, favouring for each residue type the backbone conformation in which it is most likely found (as computed from the Protein Data Bank - PDB [Berman et al., 2000]). When this is not sufficient, the protein can be constrained around a given structure through an elastic network model approach (ElNeDyn [Periole et al., 2009]). However, both the backbone parametrisation and the use of ElNeDyn imply that conformational changes in the structure are penalised and therefore not well sampled in MARTINI simulations.

The MARTINI force field provides two water models. The standard one groups four water molecules in one bead only (Figure 2.3(c), left), losing the polarisability typical of water molecules, the effect of which is partially restored with the use of a high dielectric constant. The polarisable water model [Yesylevskyy et al., 2010] maps instead four water molecule to a three-beads molecule (Figure 2.3(c), right) with a positive and negative charge on two of them, which can account for the water dipole. This model allows to revert the dielectric constant back to a value closer to 1.

Overall, the MARTINI force field pushes the limits of simplification to enhance the simulations speed-up, with considerable gain in efficiency with respect to atomistic or even SIRAH simulations. Despite it can not capture some

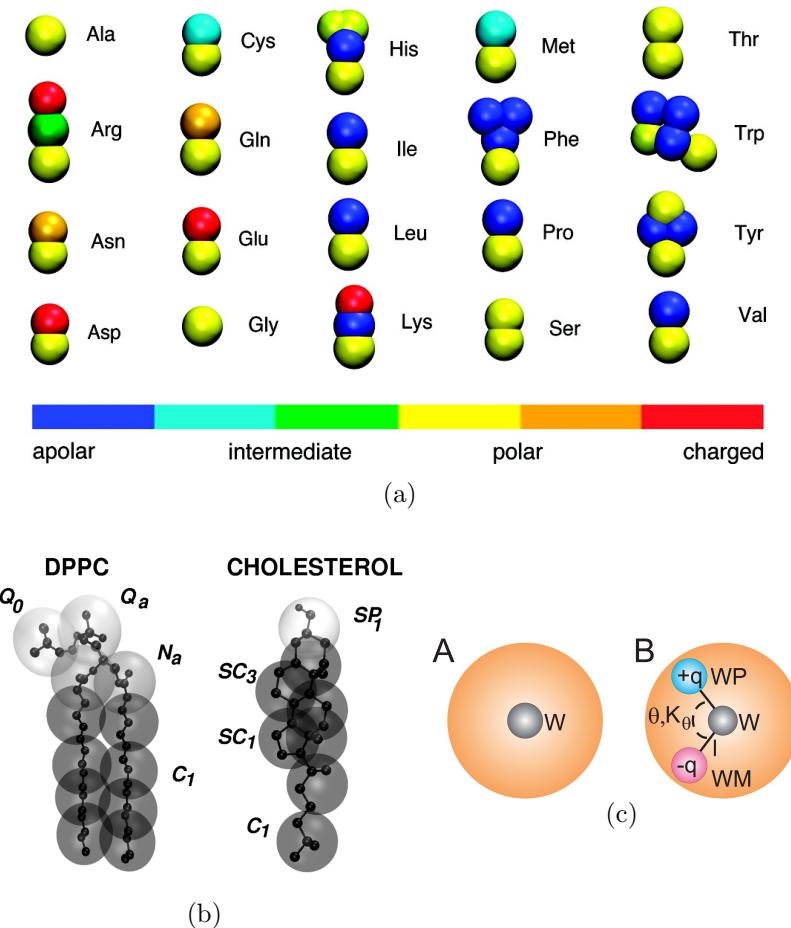


Figure 2.3: Description of amino acids (a), example lipids (b) and water models (c) in the MARTINI force field. In (c) the orange sphere represents the van der Waals radius of the central atom. Reproduced from Monticelli et al. [2008]; Biocomputing Group [2016]; Yesylevskyy et al. [2010]

fine details of the systems studied, it has been successfully applied to describe the behaviour of many biological membranes [Khalid et al., 2019; Samsudin et al., 2017], lipid self-assembly [Marrink et al., 2007] and peptide-membrane binding [Song et al., 2019]. The (re)introduction of a more detailed water model allowed the description of electroporation processes and the translocation of ions through bilayers [Yesylevskyy et al., 2010].

2.2.5 Backmapping techniques

Coarse-grain descriptions are very effective in reproducing long time scales; however, to retrieve finer details after such extensive exploration, backmap-

ping techniques have been designed to obtain atomistic configurations from the coarse-grain ones [Wassenaar et al., 2015]. These backmapped structures can in turn be simulated at the atomistic level to explore the short time scale movements around such interesting conformation. The easy conversion between the two resolution, gave rise to many multiscale studies applied to biomolecular systems [Lee et al., 2012].

2.3 Beyond the classical framework

Without entering into the details, it is important to mention that the classical MD framework can benefit of additional terms aimed at improving its accuracy, and/or of specific techniques aimed at improving the sampling within accessible computer time.

Polarisability and quantum effects The first refinement possible is the introduction of polarisability, i.e. the displacement of electrons with respect to the nucleus, as a consequence of the surrounding electrostatic environment. None of the force fields mentioned above accounts for that, because electrons and nucleus of an atom are modelled as a unique object. Specific force fields have been modelled to include this effect, either on top of atomistic descriptions, as in AMOEBA [Ren and Ponder, 2003; Ponder et al., 2010], Drude polarisable CHARMM [Anisimov et al., 2005] or AMBERff02 [Cieplak et al., 2001], or in combinations with a coarse-grain descriptions, as in the ELBA force field [Orsi and Essex, 2011]. Polarisability does improve the accuracy of simulations (see Chapter 4, where it is discussed in the context of lipid tails parametrisation), but it can significantly slow down simulations. Moreover, for biological processes governed by quantum mechanics - such as photosynthesis, DNA mutation processes or some enzymatic activities - many semi-classical hybrid techniques have been developed [Ahmadi et al., 2018]. They combine computational quantum mechanical modelling methods, such as Density Functional Theory (DFT) or Hartree-Fock computations (HF) [Shao et al., 2015], with classical Molecular Dynamics to gain the accuracy of a quantum description in the region of interest and the speed up of a classical one in the surrounding areas.

Reduction of the number of degrees of freedom At the other end of the spectrum, tackling instead efficiency issues, many models have been implemented to reduce the number of degrees of freedom to deal with, among which the coarse-grain force fields already mentioned. Another approach is constituted by the use of implicit solvent, where water is represented as a continuous medium, as opposed to explicit models which include all its particles [Kleinjung and Fraternali, 2014]. Models of implicit solvent can be based on different assumptions: for example the solute-solvent interactions can be taken as proportional to the solvent accessible surface area (SASA) of every particle of solute [Fraternali and van Gunsteren, 1996; Kleinjung et al., 2003, 2012; Fornili et al., 2012], or instead can be derived from a solution of the Poisson-Boltzmann equation governing the charge density in a material, for example in the form of the Generalised Born equation [Zhu et al., 2005] which is valid under particularly simple conditions. Another technique to alleviate the computational burden is constituted by hybrid particle-field algorithms. The idea is to treat non-bonded interactions through a mean field approach, where atoms/beads move in the field generated by the others. The field does not need to be updated at every time step, as it is a collective and thus slowly evolving variable; moreover, for each particle only the interaction with the field, and not with all the neighbouring particles needs to be computed, reducing the computations effort further. This approach has been employed with a coarse-grain description of polymers and biological molecules in the OCCAM software [Milano and Kawakatsu, 2009].

Enhanced MD techniques Finally, computational strategies have been devised to bias the natural evolution of the system in order to enhance the sampling of many configurations, i.e. to speed up its pace. This is particularly necessary when studying large-scale conformational transitions. This scenario corresponds to transitions between states separated by an energy barrier, but also matches cases in which we want to describe more accurately the phase space of a rugged landscape. Indeed very often, in the impossibility of understanding from first principles which states of a system are the most energetically favoured, and can therefore influence its macroscopic behaviour, MD simulations are run to explore the broadest possible set of them. The interpretation of such sampled states may become difficult and requires sophisticated

analysis techniques.

To be noticed that, as biomolecular systems have a complex landscape dense of energy barriers, a realistic, limited time simulation usually samples regions around the initial configuration only. In the simulations of proteins, this would often correspond to a structure derived from X-ray crystallography, which might not represent the native state of the protein in solution nor the functional form of interest. An alternative could be the use of different pre-modelled initial structures so to reduce the sampling of states far from the equilibrium.

The challenges mentioned above have promoted the development of enhanced MD techniques. As a non comprehensive list, we mention replica-exchange algorithms [Okamoto, 2004] which combine together multiple simulations held at different conditions, local potential-energy elevation (or metadynamics) [Huber et al., 1994; Laio and Parrinello, 2002] which avoids the resampling of already visited conformations adding an energy penalty to them, umbrella sampling [Torrie and Valleau, 1977] which reconstructs free-energy barriers from simulations held at specific values of the coordinate along which the barrier exists, or finally simply the use of higher temperature to overcome energy barriers [Kirkpatrick et al., 1983]. To be noticed that a coarse-grain description of a system, by reducing the number of degrees of freedom, discards the high-frequency or less interesting ones, giving a smoother energy surface, so that the search is speed up both by the reduced computational load and by the fact that the system is not trapped into local minima due to the landscape roughness.

2.4 MD and experiments: ensembles versus averages

The validation of MD simulations is performed by comparison with experiments: the properties obtained experimentally are computed from the MD trajectory as well, and the latter compared with the former. If these are correctly reproduced, it is usually assumed that the simulation is sampling the correct ensemble of states. This holds if the properties of the simulation are not drifting away, namely the system has reached equilibration and it is thus in a stationary state. Once the simulation has been validated, one can identify, from the conformations in the trajectory, the details of the processes

responsible for the experimental outcome of interest, as such information is not accessible by the experiment itself.

The comparison however is not always easy, because of a fundamentally different focus that experiments and simulations have. The former measures often an average quantity in time and/or space, while simulations access the (hopefully complete) *ensemble* of states that the system can visits and their occupancy, which in turn determines the macroscopic behaviour measured in the experiments.

In the case of a protein in solution, the *ensemble* corresponds to the variety of shapes it adopts, which can differ significantly from the crystal-structures available. This flexibility is confirmed by small angle X-ray scattering (SAXS) and nuclear magnetic resonance (NMR) experiments, the results of which cannot be explained by a single-conformation scenario [Bonomi et al., 2017; Kikhney and Svergun, 2015; Kleckner and Foster, 2011]. In this context, only simulations can help in deconvoluting the results to map them back to the conformations and their relative contribution responsible for the outcome, uncovering their relative importance.

However, it is still challenging to compare experimental and simulations derived ensembles. Following from the example above, many different combinations of structures can match a SAXS profile, so in each specific case it is important to understand which are the relevant properties playing a role in the measured ensemble before attempting comparisons.

In general, in the validation of MD outcomes, it is necessary to have a critical attitude both in the case of agreement and disagreement with the experiment, and to interpret the result within the validity of the approximations performed by both of them [van Gunsteren et al., 2008]. Indeed, agreement may arise from either a simulation that reflects correctly the experimental system; but also when the property examined is insensitive to the details of the simulated trajectory, or it can be a result of compensation of errors, which is more likely to occur for systems with a high number of degrees of freedom (as biomolecular ones). Similarly, disagreement may hint at an error in the simulation (either in the model, the implementation, or simply the estimated simulation's convergence) or an error in the interpretation and/or conditions of experimental set-up (either in the result itself or its interpretation), so that both must be carefully checked to improve a convergence in the agreement. In

this, some apparently negative results may suggest or stimulate new experimental settings to validate the hypotheses one was set to test [Gonçalves et al., 2013; Meißner et al., 2014].

Moreover, simulations still suffer from the limited computational time accessible to effectively simulate the system in study: most of the times, the experimental system is simply too large to be reproduced and the time scale of the process too long to be accurately sampled. Simulations are thus confined to explore a restricted space, implying that the initial conditions must be chosen carefully to optimise the search and avoid any bias which might persist for the whole length of the simulation. The use of enhanced MD techniques does increase the chances of sampling relevant states, however it introduces a bias which must be removed or properly accounted for in the interpretation of the results [Bernardi et al., 2015; Best et al., 2005; Barducci et al., 2010, 2011; Mills and Andricioaei, 2008]. Finally, one should keep in mind that the force fields used are far from optimal, partly because they rely on approximate functional forms, and partly because it is difficult to find experimental observables measured with the desired resolution able to discriminate between sets of parameters.

Despite the challenges outlined, Molecular Dynamics simulations have played a crucial role in disclosing important details behind biophysical processes and in unravelling molecular details not accessible to experiments. The following section will highlight a few of the many successes of MD simulations.

2.5 MD simulations: successes

Consistently with the focus of this thesis, we will privilege examples of simulations investigating antimicrobial peptides as well as self-assembling ones, showing how computational techniques can help the design of novel molecules with improved specific characteristics.

2.5.1 Simulations of antimicrobial peptides

MD simulations of antimicrobial peptides are quite well documented since the first developments of the technique. Such peptides are a suitable system for a computational investigation as, in most of the cases, their mechanisms

of action are not completely understood from the experimental information available (see Section 1.2.2). As experiments prove that even the mutation of one single residue in short AMPs can change remarkably the antimicrobial activity of the sequence (see Section 1.2.4), it is then clear that their action is governed by subtle atomic interactions, so that MD simulations, with their atomistic resolution, can help in understanding this aspect.

Systems As mentioned in the previous chapter, it has been proposed that most AMPs act through a process of attraction to the bacterial membrane, possible aggregation with other copies of the same sequence, insertion, and membrane lysis. The time scales of the overall process are accessible if using coarse-grain techniques, but not - or rarely - atomistic ones. For this level of description instead, the different steps are usually investigated separately, based on prior hypotheses: for example, the peptide can be positioned close to the membrane surface with an orientation known to promote binding (from experiments or based on energetic assumptions) [Wang et al., 2012b], or again can be placed directly within the membrane core with different insertion depths, tilt angles and oligomerisation states to verify which configurations are the most disruptive ones [Lipkin et al., 2017]. In this case, the full insertion process can only be reconstructed from a “stepwise” knowledge combining the different states sampled and further exploring the intermediate regions if necessary. For these reasons, the choice of the conformation to simulate, i.e. the initial conditions in terms of the mutual position of peptide and membrane, is crucial, as it likely biases the simulation towards the sampling of a particular subset of configurations, and this must be considered in the interpretation of the results. Recent advances are making possible the simulations of the full process even at the atomistic resolution for simple enough systems [Ulmschneider, 2017; Sun et al., 2015], as it will be shown in the following, nevertheless the “stepwise” approach is still common and the preferred one in case of complex AMP systems.

Model membranes The second important choice in the setup of a simulations of antimicrobial peptides concerns the model of the membrane to simulate. In an effort to keep complexity low, bacterial and mammal membranes can be modelled with a minimal number of lipids. Very often, models of bac-

terial membrane retain as only key characteristic an overall negative charge, with about 25% of the lipids being anionic ($-1e$ charge) and the rest being zwitterionic, i.e. neutral but with positively and negatively charged regions separate in space (see Chapter 4) [Lipkin et al., 2017; Wang et al., 2012b; Zhao et al., 2018; Chen et al., 2019]. For a model mammal membrane instead, only zwitterionic lipids are employed, with the occasional inclusion of cholesterol, as it is deemed important in describing more realistically their behaviour. [Lipkin et al., 2017; Wang et al., 2012b; Zhao et al., 2018; Chen et al., 2019; Risselada and Marrink, 2008]. Because of their simplicity, very similar or identical systems are used also in experiments [Castelletto et al., 2016; Tang and Hong, 2009; Glukhov et al., 2005], making possible a direct comparison with simulations. Therefore, even if these simple membranes don't model accurately the structure of the cellular envelope, simulations and experiments of these systems can provide a first explanation of some steps of the antimicrobial activity, with the two techniques complementing and validating each other.

For example, an *in silico* experiment simulated a dermicidin channel inserted into patches of phospholipids membranes with variable cholesterol content [Song et al., 2019]. Nine membrane compositions were tested overall resulting in different membrane thickness, thus in a different orientation of the dermicidin channels inserted into it, with consequent variance in the conductance of the channel itself. This structure-function relationship shows the importance of an accurate membrane model to fully capture all the aspects of transmembrane protein activity as every change influences it. Notably, the simulations were performed with the coarse-grain model MARTINI, showing that a supra atomistic view retains enough details to investigate such systems.

Nevertheless, attempts to model more accurately cell membranes have been pursued. This can be performed at the atomistic level [Piggot et al., 2011] but the task is especially suited for a coarse-grain description, as the inclusion of all the elements of the cell membranes results in quite large systems for which atomistic computations started only in recent years to be affordable. Accordingly, coarse-grain (MARTINI) simulations have been incorporating more and more components into model membranes, describing the bacterial inner membrane, the bacterial wall, and finally the combination of the two [Khalid et al., 2019] (see Figure 2.1, bottom). These large scale, coarse-grain simulations provide information on the mechanic characteristics of the system: for exam-

ple, simulation of the outer membranes of Gram negative bacteria combined with the peptidoglycan layer (which, in bacteria, is positioned between the two membranes) elucidated how the distance between the two is variable, thanks to the presence of Braun's lipoproteins [Asmar and Collet, 2018] which act as a bridge between them, and can bring them closer by bending and tilting. On the other hand, the permeability of membranes to ions and small compounds needs to be assessed at the atomistic level, and to access informative simulation time scales, smaller and simpler systems must be chosen for the task (e.g. the inner membrane only), often together with enhanced MD techniques such as metadynamics, umbrella sampling, and replica exchange umbrella sampling. [Sun et al., 2016; Piggot et al., 2011; Carpenter et al., 2016; Pokhrel and Maibaum, 2018].

Force field comparison Finally, simulations of the peptide interaction with a model membrane are clearly determined by the parametrisation of the force field employed for protein and lipids (and by their mutual consistency). There are multiple evidence suggesting that different force fields produce very different outcomes when simulating the same system, under the same conditions. This is also valid for simulations of pure lipid patches (see Chapter 4), resulting in incompatible values of area per lipid, organisation of the tails and energetic profiles across the membrane, and thus has an impact in the simulations of AMPs interacting with a membrane.

For example, Wang et al. [2014] run simulations of the antimicrobial peptide melittin with different force fields, namely CHARMM27 and 36 (for protein and lipids respectively) [MacKerell et al., 1998; Klauda et al., 2010], OPLS all atoms (for protein) and united atoms (for lipids) [Jorgensen et al., 1996] and GROMOS 53a6 [Oostenbrink et al., 2004] (and the TIP3P water model for all of them [Jorgensen et al., 1983]). Despite these parametrisation have similar values of partial charges on the different atoms, and similar bonded interactions at the protein level, the unfolding of melittin in the membrane was significantly different among them, with the CHARMM force fields suggesting an almost completely folded state for melittin bound to a membrane, in line with the NMR available results, while the other force fields promoted partial unfolding. Most likely this can be attributed to the fact that lipid and protein parametrisations are obtained separately and might present some

inconsistencies. The most evident example is the OPLS case, for which an all atom description of lipids is not available and a mixed description has been adopted. In the case of GROMOS, both components are parametrised at the united atom level, but their mutual consistency might be questioned as well, as fully explained in Chapter 4. In general, a united atom description is clearly less accurate than an all atom one, and in the case of lipids it has been postulated that it is not able to represent faithfully the dynamical processes happening in the hydrophobic tail region [Chowdhary et al., 2013].

A similar investigation has been proposed by Bennett et al. [2016], proving that the propensity of the synthetic AMP CM15 to form pores strongly depend on the force field used but also on some extent - at least at the time of the work - on the MD engine used (GROMACS compared to NAMD [Phillips et al., 2005]). This should not come as a surprise because the membrane characteristics emerge from the collective behaviour of lipids, so that a small difference in the way their interactions are treated might be amplified resulting in different macroscopic outcome for the simulations [Reißer et al., 2017].

A more systematic study on the topic has been performed by [Sandoval-Perez et al., 2017], focussing on the reproduction of membrane-protein interactions in different force fields (GROMOS 54a7 [Schmid et al., 2011], CHARMM36, Amber14SB/Slipids [Jämbeck and Lyubartsev, 2012] and Amber14SB/Lipid14 [Dickson et al., 2014]). All of them were able to reproduce the overall positioning of transmembrane proteins in the case studies tested, together with their α -helical and β -sheet content, while they showed discrepancies in the insertion angle for a short helical peptide (sybII) spanning the membrane. The amino acid side chains insertion depth was also tested: the CHARMM force field suggested a deeper insertion for hydrophobic amino acids but the other parametrisations gave not so clear a distinction. In general, with the GROMOS force field, a higher energy is required to insert amino acids to the bilayer centre. Interestingly, all parametrisations gave a very broad minimum for the insertion of Tryptophan, comprising the phosphate region of the phospholipid membrane tested, but also part of the tail and head regions. This is in line with the different interpretations given on the role of this amino acid for the action of antimicrobial peptides: either as an anchoring point positioned deeper in the hydrophobic region, or as a partner for hydrogen bonding with the hydrophilic heads. The subtle differences between parametrisations lead

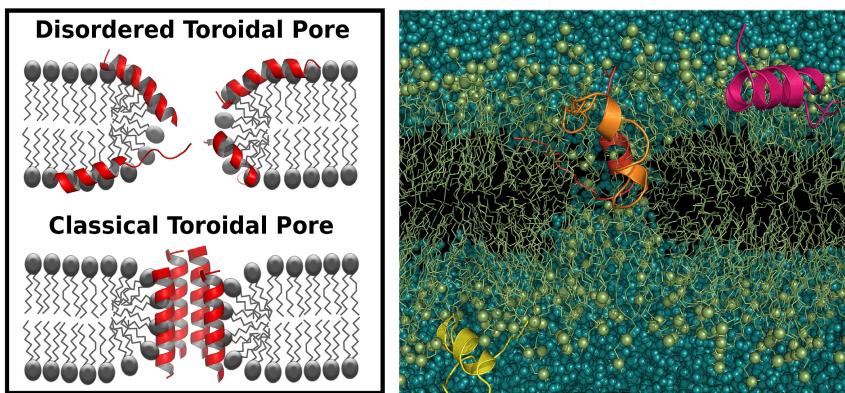


Figure 2.4: (Left) A cartoon image comparing the disordered toroidal pore state (lack of a well defined peptide orientation) to the traditional view. (Right) A snapshot of the disordered toroidal pore from simulations of melittin in DPPC. Peptides in cartoon, lipids in lines, water and ions in bead representation. Reproduced from Sengupta et al. [2008].

to the conclusion that for every particular system tested, the comparison with at least one experimentally measured quantity would be the only way to assess the simulation performance accurately.

Simulations of membrane-peptide interaction: examples Even in the context of a simplified model scenario and with the caveats coming from the chosen parametrisation, simulations of antibacterial peptides on a membrane have been successful in elucidating some of their mechanisms. The first important contribution consists in the introduction of the disordered toroidal pore concept: as explained in the previous chapter (Section 1.2.2), the models of membrane poration due to AMPs consist often in ordered structures (see Figure 1.3) where many peptides gather together to contour a pore, and they are either in contact with the hydrophobic tails of the lipids (barrel-stave model), or with their head, as lipid molecules bend around the pore to keep their tails screened from the outside environment (toroidal pore model). However, simulations of the short helical peptide magainin MG-H2 [Leontiadou et al., 2006a], among others, showed that a single copy of the helix, inserted at an angle with the normal to the membrane plane, was sufficient to displace the lipids around in a non organised manner and form a water-filled pore (Figure 2.4).

Regarding possible rearrangements of the antimicrobial peptide structure when interacting with a membrane, simulations of cathelicidin LL-37 on pure POPG (anionic) and POPC (zwitterionic) lipid patches showed that LL-37 has

a propensity to bind to the former, as expected due to the opposite charge that membrane has with respect to the cationic peptide [Zhao et al., 2018]. However the simulations highlighted also that, in contact with POPC, the helical secondary structure was lost, while the interaction with POPG preserved it, suggesting that the spatial arrangement of the residues, and not only the overall chemical character, is important for their action. Such type of information is hardly available to experiments or through a theoretical reasoning.

Further insights into the role of the secondary structure were obtained simulating the helical antimicrobial peptide CM15 nearby a POPC membrane, starting from a fully structured helix or from a coil configuration: Wang et al. [2012b] proved that the interaction with the lipids is stronger when the peptide approaches the membrane in its disordered form rather than in a fully formed helix. This happens because of the larger flexibility of the coil arrangement which allows for more residues to come in contact with the membrane at once. Notable, the α -helical fold binds as well to the membrane, but only after a time span larger than 100 ns, which would have been inaccessible up to a few years ago, potentially leading to wrong conclusions.

The improvements in computational resources is slowly removing some of these obstacles, pushing the extent of simulation time to the microsecond timescale. In a recent example, the translocation of the helical PGLa peptide through the membrane has been observed as a rare event, dependent on the concentration of the peptide, on the multi microsecond timescale without the formation of an organised pore [Ulmschneider, 2017]. The *in silico* experiment still benefited of an enhanced sampling in the form of a higher temperature used for the simulations, but no pre-insertion of the peptide was performed. This study shed light on a possible mechanism of permeabilisation which is usually overlooked in favour of processes involving organised channels and pores. The fact that no organised neither disorganised pore is observed matches the experimental results which can not identify such structures for the peptide considered.

Similarly, simulations were able to shed light on the mechanism of translocation of Arginine-rich peptides, proposing a mechanism of action on an otherwise puzzling problem [Sun et al., 2015]. These sequences have high positive charge, but despite this, possess a high propensity to penetrate membranes, overcoming the hydrophobic region represented by the lipid tails. Very similar

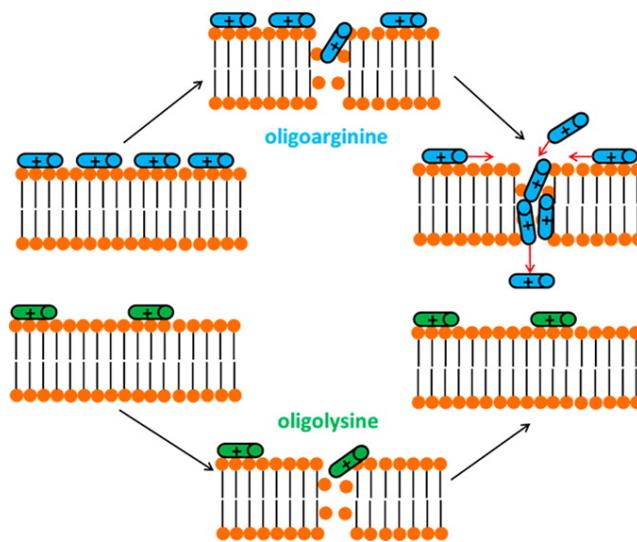


Figure 2.5: Cartoon representation of the different assembly mechanism that allow polyarginine but not polylysine to translocate through spontaneously formed pores. Reproduced from Sun et al. [2015].

peptides where the Arginines were swapped with Lysins showed no significant penetration. A commonly used explanation considers polyarginine translocation a quasi-equilibrium process, but this does not explain the selectivity against Lysins rich peptides. After extensive simulations of the two systems (multiple, hundreds of nanosecond long runs, with two different force fields), the proposed mechanism involves the spontaneous formation of thermal pores: in some of these rare events, the transient pore would be occupied by a peptide (a precursor), which slows down its dynamics and thus closure. In such situation, the translocation of other copies of the peptide is highly favoured if their concentration is sufficiently high. Indeed, other copies of the peptide are driven to aggregate with the precursor inside the membrane and are then pushed toward the opposite side as there is a lower charge density in that region. Differently from polyarginines, polylysins have a much lower aggregation propensity, so that the presence of a precursor peptide inside the membrane does not induce an enhanced insertion of further peptides.

The last two examples mentioned bring the attention on whether and for which peptides oligomerisation is necessary for an efficient antimicrobial activity. MD simulations can offer insights on this aspect as well. Contrary to oligomerisation in solution, which can happen on shorter time scales, the spontaneous aggregation of peptides on a membrane surface requires a long

time, as the structures must diffuse on the membrane to meet each other, and many other competing processes (such as insertion) are happening at the same time.

A recent example of how MD elucidated oligomerisation mechanisms comes from simulations of maculatin (an helical AMP), which showed that the pores it forms can include a variable number of helices and thus assume many different conformations [Wang et al., 2016b]. The suggested process of pore formation proceeds via insertion of a single residue, closely followed by other ones which are able to penetrate the membrane thanks to the lipid defects already created by the first peptide.

Similar investigations can be carried on also for other cell penetrating peptides, which are not antimicrobial: as such, some of them aim at inserting within cells without necessarily causing poration. One example is constituted by the influenza fusion peptides, which have been extensively studied with a simulation set up similar to the one mentioned for AMPs: a few copies of the peptide were positioned on a model membrane and their oligomerisation and insertion processes were followed in time, showing the formation of aggregates of different sizes which perturbed locally the membrane [Haria et al., 2014; Collu et al., 2015].

To be noticed that, when investigating oligomerisation, the size of the system must necessarily be increased to include all the copies necessary to form the aggregates observed experimentally. As such, with the present accessible computer time, not all the systems can be investigated from unbiased initial conditions. In the case of protegrin, a β -hairpin antimicrobial peptide which has been long thought to act through the formation of transmembrane β -barrels, many variables can influence the outcome of the unknown final structure. Even with the computer power available now, it is unlikely to sample all the possible conformations resulting in stable or transient β -barrel pores in simulations starting from a few peptides scattered on the surface, and this hinders the understanding of their relative importance. To overcome such problems, a semi-systematic investigation has been carried on by Lipkin et al. [2017], simulating different assembly (see Figure 2.1, top). Microsecond long simulations discriminated which ones of these initial configuration formed stable pores for the whole length of the simulations, and the ones which were disrupted. As in the previous example, several different possibilities were found stable in solu-

tion, suggesting that single AMPs might have multiple mechanisms of insertion into membranes.

Most of the examples above employ atomistic descriptions of the system. Similar investigations have been carried on also using the MARTINI force field, indeed the coarse-grain description does capture the pore-forming behaviour of some AMPs. As an example, simulations of maculatin and aurein on POPC membranes showed different propensities for pore formation versus aggregation, showing that the model retains enough details and chemical information to reproduce different membrane perturbing behaviours [Balatti et al., 2017]. Nevertheless, the developers of the MARTINI model themselves pointed out how some aspects of pore formation might not be captured in a satisfactory way [Marrink and Tieleman, 2013], for example the penetration of water can be misrepresented as can be intuitively expected from a model which clusters four water molecules together (indeed some pore conformations allow for the passage of fewer water molecule if not one at the time).

In general, the outlook of simulations of antimicrobial peptides interacting with membranes goes in the direction of reproducing longer time scales thanks to the enhanced computational power available, trying to match the experimental findings showing that many antimicrobial related processes happen at the microsecond scale or beyond. This enhanced power would also reduce the need to use biased initial conditions or higher temperatures to speed up the simulations. Moreover, gathering the contribution of the whole community, simulations will likely go in the direction of modelling more accurately the bacterial membrane, and while this is already at an advanced stage for coarse-grain simulations, it is still an ongoing process for atomistic ones. Finally, the force field issue must be solved in collaboration with experimentalists, finding new tests and experimental quantities to compare the computational outcome with and make the different parameters sets converge toward a similar description of the phenomena observed, which is consistent with the experimental results.

Simulation-aided AMPs design The role of simulations in aiding AMPs design has been briefly sketched in Section 1.2.4. As pointed out, MD simulations are hardly a tool to analyse large dataset, therefore a systematic analysis can be performed for very small systems only, or, alternatively, the investigation can focus on a few selected sequences.

As already mentioned, when classifying AMPs, simulations can be helpful in integrating structural information which is otherwise lacking, when no crystal structure of the peptide is available. Such approach was followed by Liu et al. [2018] to complement the chemical information available on a dataset of short AMPs, and the overall information was used to feed a predictor of AM activity of novel sequences. Preliminary results showed that such structural information of minimal AM sequences improved significantly the ability of the predictor to discriminate whether a new sequence was suitable for antimicrobial activity or not.

Another commonly followed approach consists in using simulations to elucidate the reasons why a particular mutation is important and effective in terms of increased activity or decreased toxicity. Indeed, for short sequences, such mutation screenings can be afforded experimentally, thus there is little need to predict whether they would be beneficial. Rather, once assessed they are, it is interesting to understand why: for example, simulations of ovispirin and a mutant peptide with reduced toxicity showed that the bend in the helix in the latter was responsible for mitigating the interaction with mammal membranes and thus reducing haemolysis [Khandelia and Kaznessis, 2005]. Again, the mechanism of lipid disordering and insertion by indolicidin was assessed through MD, and the amino acids responsible for each of them separately were identified, so that mutants could be designed with either reduced toxicity or enhanced potency [Tsai et al., 2009]. Finally, temporin and a derived sequence were investigate to discover that the mutant improved activity derived from a reduced aggregation propensity of the peptide in water, so that more copies were ready to bind to the membrane and thus disrupt it [Farrotti et al., 2017].

Many more examples can be listed, each with a slight different focus: the protocol of integrating simulations and design is usually customised according to the system in exam, as the field has not reached yet a systematic organisation. However, it is clear that simulations used in conjunction with experimental testing can be used to optimise already available AMP sequences, and thus contribute to device design rules for the creation of synthetic sequences with tailored properties.

2.5.2 Simulations of self-assembling peptides

Self-assembling peptides are another fascinating and challenging topic that MD simulations can help investigating. Simulating such systems implies different challenges with respect to the ones faced when simulating AMPs on membranes [Frederix et al., 2018; Orsi, 2018].

In theory, the set up of the system is quite straightforward: only the solvent characteristics and optionally the experimental salt concentration need to be matched, then a random initial configuration of the molecules - in the desired concentration - would allow the simulation of the process of interest. In reality, reproducing the experimental conditions often implies working with very large systems: with the level of dilution of the solutions employed in the experiments, a considerable volume needs to be simulated to host enough copies of the peptide to observe the assembly of large enough oligomers. However with this approach the time scale useful to witness a spontaneous assembly would greatly exceed the computational time available. For that reason, two main strategies have been adopted: coarse-grain simulations and the use of pre-assembled structures. Other routes include the choice of an implicit solvent model, or the use of other techniques such as Monte Carlo (a probabilistic exploration of the phase space rather than a dynamical algorithm) which can sometimes be less time consuming. Here, as we focus on MD simulations, we give a few examples of the first two strategies mentioned, which can be adopted in this framework.

Many studies have been performed with the coarse-grain MARTINI force field to witness assembly of surfactants [Wu and Yang, 2012], polymers [Wang et al., 2012a; Bochicchio and Pavan, 2017] and lipids [Lee and Pastor, 2011; Brocos et al., 2012], and a few focussed on peptides as well [Guo et al., 2012; Seo et al., 2012]. For example, the assembly in water of peptide amphiphiles (PAs) into cylindrical fibers has been simulated at the coarse-grain level [Lee et al., 2012], showing a transition from small micelles to longer fibres (Figure 2.6). This example of minimal PA structures is particularly interesting for the study of AMPs as well, as it shares with them the amphiphatic character, so that having a general knowledge on how similar sequences assemble together would help in tuning their aggregation properties in water prior to the delivery to the membrane. In the work mentioned, pre-assembled fibres have been simulated as well at the atomistic level, to confirm their stability in solution.

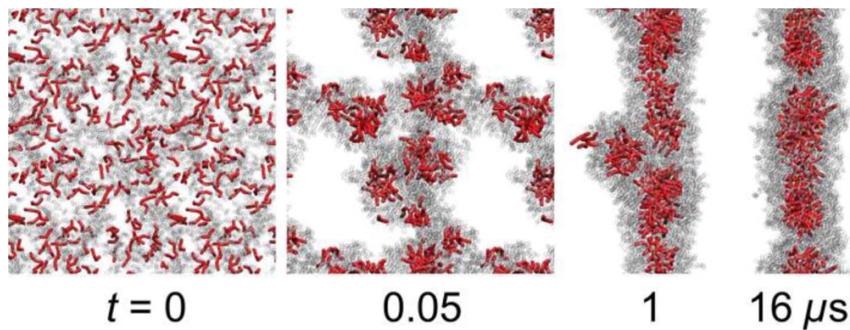


Figure 2.6: Process of peptide amphiphiles (PA) fiber formation assessed through MARTINI simulations. Hydrophobic tails in red, peptides in gray. Reproduced from Lee et al. [2012].

The second approach consists in preparing the system in a pre assemble state that is somewhat suggested by experimental evidences and using MD simulations to verify whether the conformation is kept or it is disrupted, and which one out of many is the one most energetically favoured. It has been widely employed in cases where the final assembly was hypothesised to have a high degree of order, achievable only with a long sampling. This approach has been used to prove that a branched peptide can self-assemble in bilayers first, and then that a larger hypothetical structure assembled in the shape of a capsule was stable in the run time of the coarse-grain simulations employed. Specifically, the capsule has been build to match the peptide density on the self-assembled double layer and to respect the constraints derived from the presence of the curvature [Gudlur et al., 2012].

Similar approaches have been crucial in elucidating the assembly process of viral capsids: capsids are very large systems and the assembly of their protein subunits is mediated by energy barriers. For such reasons, already pre-assembled systems have been simulated to understand the interaction between the components and thus the first mechanisms of the assembly. This has been done recurring to ultra coarse-grain or elastic network models [Grime et al., 2016] first, and only in the most recent advances to atomistic simulations [Perilla et al., 2016; Hadden and Perilla, 2018]. Additionally, smaller portions of a capsid can be simulated, to obtain a minimal information on the cohesion of its blocks [Abi Mansour and Ortoleva, 2014].

The examples above show how Molecular Dynamics simulations have been employed for the investigation of many different systems during the years,

adapting the resolution, set up and the techniques related to better query the systems of interest. Such overview suggests then that simulations would be a suitable tool to investigate the system of interest of this work, namely the self-assembling antimicrobial peptide capzip. The two aspects of its behaviour will be studied separately, adopting the necessary approximations and strategies to make the simulations efficient and to query the related questions at each time.

The details of the systems simulated and the specific parameters used for each of the system simulated in the following chapters can be found in the relative sections, together with an extensive explanation of the motivation of the choices made.

Chapter 3

Capzip simulations

In Chapter 1 we introduced the molecule capzip and its properties, highlighting the unknowns of its mechanisms of action. In Chapter 2 we presented a review on Molecular Dynamics simulations, proving their successes in elucidating the behaviour of self-assembling and antimicrobial peptides in the past. Now, we employ this technique to understand better our system of interest. Given the exoticism of the unit, and the little atomistic information at disposal, modelling such peptide must proceed in a stepwise manner.

The first aim is to elucidate which structures it forms in solution and what interactions are keeping the molecules together. To understand the latter ones it is important to retain the highest level of detail possible, and for this reason we resorted to atomistic simulations first. This description has an high computational cost though, preventing the simulation of very large systems for a very long time, as it would be required to reproduce the natural assembly from a dispersed solution. Thus, we simulated increasingly complex pre-assembled blocks, verifying each time their behaviour in solution and inferring whether they are suitable to form a stable supramolecular structure. This approach, fully explained in Section 3.1, lead to the model of a minimal capsule, which has been subsequently investigated at coarser levels to explore its behaviour on longer time scales.

Of the pre-built structures a few selected ones were simulated in contact with model membranes to understand the determinants of their antimicrobial activity. Details on these simulations and the specific techniques employed to enhance the sampling are given in Section 3.2.

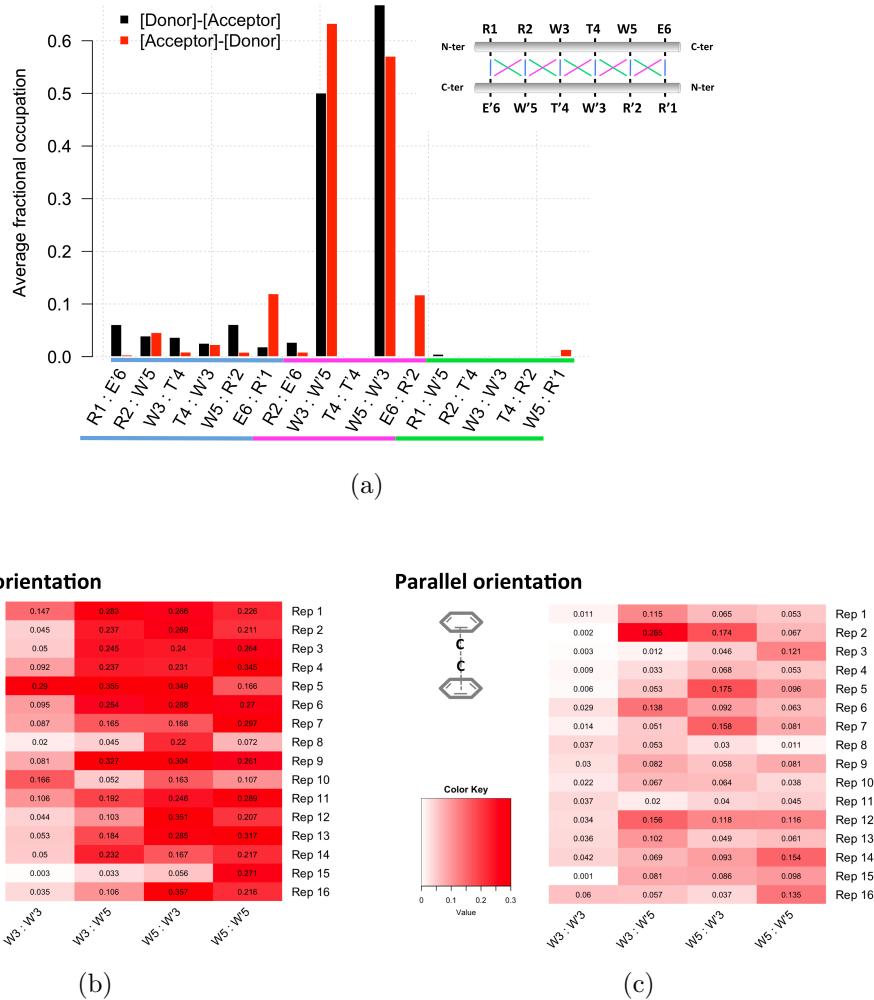


Figure 3.1: (a) Presence of backbone hydrogen bonds between amino acids in two facing antiparallel RRWTWE chains. The top right inset shows a scheme of the initial configuration. All the pairs highlighted in blue, green and pink are reported in the histogram; the same color code appears in the bar labels. Occupancy is averaged over 16 simulations of 20 ns. (b, c) For each replica and possible pair of facing Tryptophan residues in the β -sheet, the map gives the fraction of time for which a parallel or perpendicular π -stacking interaction has been observed. The second and fourth column correspond arrangement with the most populated hydrogen bonds.

3.1 Modelling the assembly

As previously mentioned, the antimicrobial sequence of capzip is designed with opposite charges at its extremes to favour an antiparallel β -sheets pairing with other copies of itself. MD simulations of two RRWTWE sequences paired in

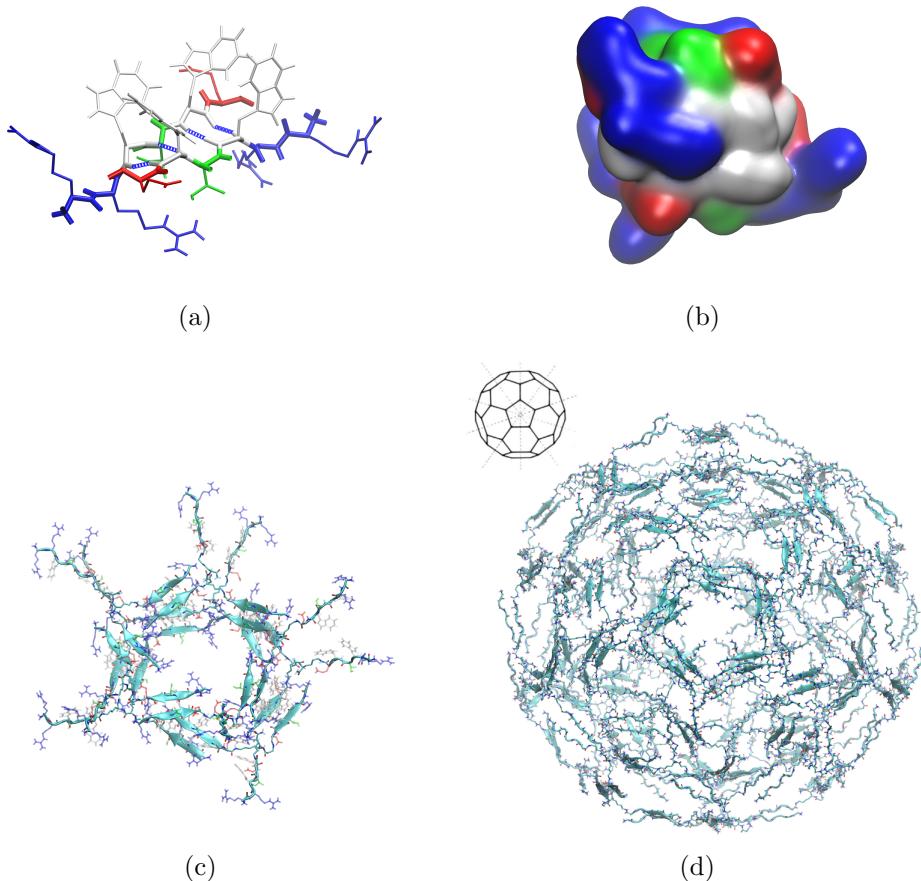


Figure 3.2: (a) Detail of β -sheet pairing with facing Tryptophan residues forming hydrogen bonds between their backbone atoms (bonds representation coloured by residue type, and hydrogen bond representation). (b) Two stacking β -sheets in surface representation, coloured by residue type. In white the partially buried hydrophobic patch. (c) A pentagonal subunit: ten antimicrobial molecules arranged in two stacking pentagons. Chains are paired in antiparallel β -sheets within each pentagon, and the two are interfacing with their Tryptophan residues in contact. (d) Atomistic structure of the buckyball simulated (bonds and cartoon representation) and geometrical model for comparison.

this fashion confirm that the assembly is stabilised by opposite charge interactions (with statistics gather over 16 replicas, each run for 20 ns). Moreover, backbone hydrogen bonds form between Tryptophan residues of facing strands, after a rearrangement of the mutual position of the backbones (Figure 3.1(a)). Finally, π -stacking contributes to the interaction as well, albeit in minor measure (Figure 3.1(b), (c)).

The favourable hydrophobic interactions between Tryptophan residues re-

sult in the creation of a hydrophobic patch which includes four of them (in white in Figure 3.2(a) on one side of the β -sheet plane. This creates an amphiphilic structure where the hydrophobic core is segregated from the remaining charged residues distributed at the other positions. The combination of two stacking β -sheets, paired to match their hydrophobic patches, constitutes an effective supramolecular assemblies to reduce solvent exposure of such residues.

This pairing strategy, however, needs to be applied in the context of full molecules assembly. The quasi three-fold symmetry of capzip suggests a regular geometric arrangement. The best examples of organised protein structures can be found in viral capsids, which are composed by the regular repetition of highly symmetric protein subunits. Inspired by this, we tested whether a geometrical organisation can represent a stable capsule, choosing as representative geometry a truncated icosahedron (buckyball).

Preliminary atomistic simulations (100 ns) were run on a pentagonal subunit formed by ten molecules arranged in two stacking pentagons (as in Figure 3.2(c)), proving the cohesion between molecules belonging to the subunit. Specifically, the number of contacts between backbone C_α s did not decrease in time but augmented slightly at the beginning, due to the compaction of the unpaired arms toward the core of the structure (Figure 3.3(a)). Moreover, for each pair of facing chains, it is computed the distance between their centres of mass. Figure 3.3(b) reports the variance of this distance over its average value, as a measure of the cohesion of the subunit, showing that in the majority of the cases less than 2% of variability is observed.

The pentagonal subunit respects the building principles of a) β -sheet pairing between antimicrobial sequences and b) double layer structure to screen the hydrophobic patches, and will constitute a face of the icosahedron. Each capzip molecule is centred in one vertex of the polygon, with the branches laying alongside the edges departing from it. On each edge two branches coming from opposite sides meet in an antiparallel fashion. When possible, a β -sheets with paired Tryptophan residues is organised (as in Figure 3.2(a)). The stacking pentagons interact through the hydrophobic patches of their β -sheet which are arranged in stacked positions.

The full truncated icosahedron was assembled from twelve pentagonal subunits (Figure 3.2(d)). As each subunit is formed by two stacked pentagons, the resulting structure has two concentric layers, for a total of 120 molecules,

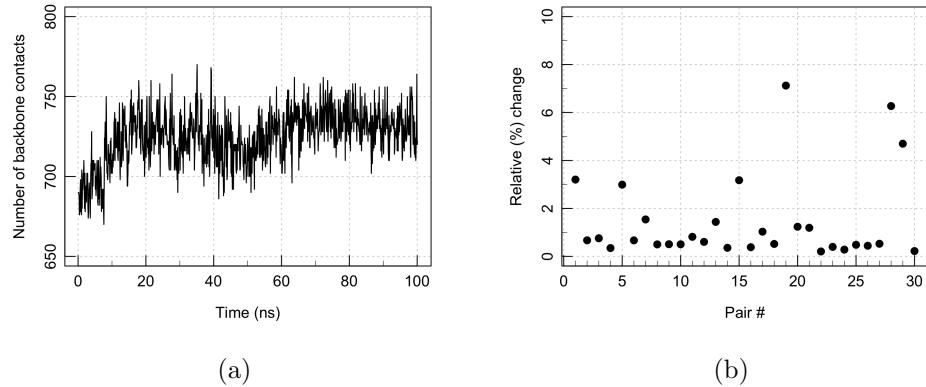


Figure 3.3: (a) Number of backbone contacts during a simulation of a pentagonal subunit. (b) Variability of the inter chain average distance between facing chains. The 30 pairs defined as facing are the chains belonging to the same β -sheet (5 for each of the two stacking pentagons), and for each stacking β -sheet the 4 possible inter-pentagon (inter-layer) pairs of chains.

and initial radius of 7.7 nm. This geometry represents a minimal model of the possible structures of capzip assembly in solution as, given the flexibility of the molecule, different geometries are possible, though proceeding from analogous interactions between the components. The final structure was simulated at atomistic and coarse-grain levels, respectively with the GROMOS 53a6 [Oostenbrink et al., 2004], SIRAH [Machado et al., 2018] and MARTINI [Marrink et al., 2007; Monticelli et al., 2008] force fields (with both standard and polar water [Yesylevskyy et al., 2010] to compare the two models). From the final configurations of the MARTINI coarse-grain model (standard water), atomistic coordinates were obtained and simulated, to be compared with the original atomistic dynamics. Moreover, additional simulations were run at all the coarse-grain levels on a structure made of one layer only (i.e. build from pentagonal block make by one pentagon only), to prove whether the bilayer structure was more energetically favoured.

A multiscale analysis is needed also to investigate the antimicrobial activity. Being highly costing to simulate a full truncated icosahedron on a bilayer patch at the atomistic level, the pentagonal subunit employed to build the complete structure (Figure 3.2(c)) was taken as representative of the latter. It was simulated close to the membrane plane, parallel to it, to avoid spending time in sampling non bound conformations (Figure 3.4(a)). This, together

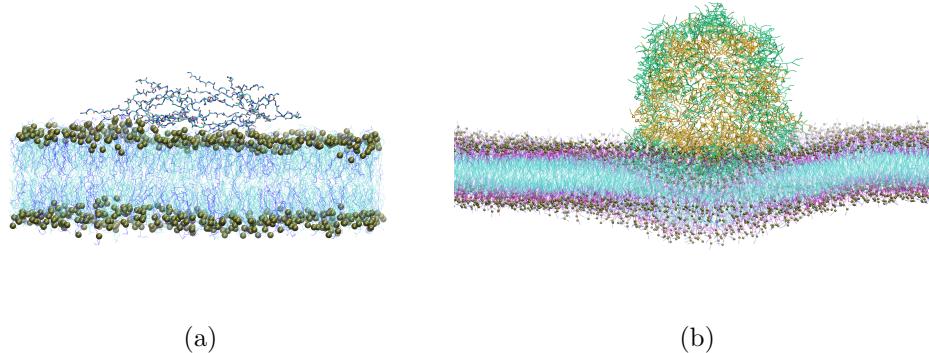


Figure 3.4: (a) Atomistic structure of a pentagonal subunit on a 740 lipid bacterial model membrane of composition DLPC:DLPG 3:1 (initial configuration). Peptide backbone in line and cartoon representation; lipid in cyan lines (DLPC) and blue ones (DLPG), all lipids phosphate in golden van der Waals beads. (b) coarse-grain (MARTINI) representation of the buckyball on a 2880 lipids bacterial model membrane (final configuration of the trajectory). Protein in bonds representation: green outer buckyball layer, yellow inner one. Lipid in line representation, coloured by bead type, and lipids phosphate in golden van der Waals beads.

with a tailored use of an applied electric field (see Section 3.2), will speed up simulations considerable.

To observe the natural binding of the peptide to the membrane, the process of the full buckyball approaching a model membrane was simulated with a MARTINI coarse-grain description (Figure 3.4(b)).

Two membrane patches were simulated for both resolutions, a model bacterial and a model mammalian membrane, to identify the different interactions with the peptide. The first one presents 25% of anionic lipids (DLPG), and the rest are zwitterionic (DLPC), while the second has only DLPC lipids. The choice of the bacterial model was dictated by the experiments performed on the capzip system, and the mammalian one employed the same zwitterionic lipid chosen in the bacterial to simplify the comparison.

3.2 Simulations details

Atomistic simulations All simulations were performed with the GROMACS software, version 5.5 and 2016 [Berendsen et al., 1995; Abraham et al., 2015, 2018].

The atomistic coordinates for the peptidic supramolecular assemblies described in Section 3.1 were built combining GROMACS tools and the MOE software [ULC, 2018]. Simulations were run with the GROMOS 53a6 force field [Oostenbrink et al., 2004]. Parameters for the central residue connected with the peptidic chains are computed with the ATB software [Malde et al., 2011; Koziara et al., 2014]; the ones for the joining bonds are derived from tabulated values of analogous peptide bonds.

The systems were solvated with SPC water [Berendsen et al., 1981] and counter ions were added (Na^+ or Cl^-); further ions are introduced to reach the concentration of 150 mM (for the species with opposite charge with respect to the counter ions), to reproduce the experimental conditions. For simulations of a single β -sheets and of the pentagonal subunit, the systems were energy minimised with a steepest descent algorithm, then equilibrated in the NVT ensemble with decreasing positional restraints at increasing temperatures (100 K, 200 K, 250 K, 300 K and respectively 1000, 1000, 500, 250 $\text{kJ/mol}\cdot\text{nm}^2$ restraints, 100 ps each); then in the NPT ensemble, without restraints, at the same temperatures steps and for the same time. Production was followed for 100 ns.

For the truncated icosahedron structure the above equilibration was still insufficient. Due to the construction procedure, two thirds of the branches are not properly paired along the edges. Therefore, after an NVT equilibration as above, strong flat-bottom restraints (1000 $\text{kJ/mol}\cdot\text{nm}^2$) were placed between the center of mass of imperfectly aligned branches throughout the NPT heating, to penalise their mutual separation with respect to their initial distance (100 ps runs at 100 K, 200 K and 250 K and 35 ns at 300 K). This was followed by a series of 10 ns runs at 300 K with decreasing restraints strength (750, 500 and 250 $\text{kJ/mol}\cdot\text{nm}^2$) and by a free production run (100 ns). Three different replicas were run, generated from the final configuration of the 300 K NPT run with 1000 $\text{kJ/mol}\cdot\text{nm}^2$ restraints.

Throughout all the simulations, the temperature was maintained by independently coupling the protein and the solvent (plus ions) to two external temperature baths using a velocity rescale thermostat [Bussi et al., 2007] with coupling constant τ_T of 0.1 ps. The pressure was kept at 1 bar by Berendsen [Berendsen et al., 1984] or Parrinello-Rahman barostat [Parrinello and Rahman, 1981] (for the equilibration phases and the production run respectively)

Capzip in solution simulations

System (Nr peptides)	Model	Time (ns)	Replicas
β -sheet (Fraction)	GR	100	32
Pentagonal subunit (10)	GR	100	1
Buckyball bilayer (120)	GR	100	3
Buckyball bilayer & monolayer (120)	SI	1000	3 & 1
Buckyball bilayer & monolayer (120)	MA	1000	3 & 2
Buckyball bilayer & monolayer (120)	MA_P	1000	2 & 1

Table 3.1: Table of simulations of capzip assembly in water. Model specifies the force field: GR = united atom GROMOS 53a6 [Oostenbrink et al., 2004], SI = coarse-grain SIRAH [Machado et al., 2018], MA = coarse-grain MARTINI [Marrink et al., 2007; Monticelli et al., 2008], MA_P = coarse-grain MARTINI with polar water [Yesylevskyy et al., 2010].

using an isotropic coupling, with isothermal compressibility $4.5 \times 10^{-5} \text{ bar}^{-1}$ and coupling constant τ_P of 1 ps. Electrostatic interactions were treated using the smooth Particle Mesh Ewald (PME) algorithm [Essmann et al., 1995], with a short-range cutoff of 0.9 nm. The van der Waals interactions were treated with a plain 0.9 nm cutoff. A SPC water model [Berendsen et al., 1981] is used. All atomistic dynamic runs were performed using a 2 fs time step. An overview of simulations of peptide assembly in solution is given in Table 3.1.

The atomistic coordinates for the bacterial model membrane patch were built with the PACKMOL software [Martínez et al., 2009], from pdb files of a single DLPC [Poger et al., 2010] and DLPG [Kukol, 2009] molecule. Two patches were built, made respectively of 512 and 740 lipids, with composition DLPC:DLPG (3:1). The initial area per lipid was set to 7 nm^2 , above the values found experimentally for either lipid species ($0.608 \pm 0.012 \text{ nm}^2$ for DLPC [Kučerka et al., 2011] and $0.656 \pm 0.012 \text{ nm}^2$ for DLPG [Pan et al., 2012]). The correct area per lipid of the mixture was reached during a 400 ns equilibration, the final configuration of which was used for simulations with the peptide on the membrane. A similar procedure was held for DLPC, producing a patch of 748 lipids with an initial area per lipid of $0.656 \pm 0.012 \text{ nm}^2$. This patch was used for control simulations against the 740 lipids bacterial one.

The initial configuration of atomistic simulations of the peptide on the membrane was generated from the equilibrated lipid bilayer (after 400 ns run) and the equilibrated pentagonal subunit (after 100 ns run with positional restraints on the C_α), placing the pentagon plane parallel to the membrane one and close to it (Figure 3.4). The inflategro [Kandt et al., 2007] script was used to solve the partial overlap of the peptide side chains with the lipid molecules, removing the ones overlapping. The sizes of the two patches fit the pentagonal subunit with respectively 3.5 nm and 5.4 nm distance between its periodic boundary images (along both x and y).

For simulations involving membranes, the version 54a7 of the GROMOS force field [Schmid et al., 2011] was initially chosen, and it is thus the one selected for simulations involving the 512 lipid membrane patch, but upon further research the 54a8 [Oostenbrink et al., 2005; Reif et al., 2013] was deemed more suitable ad thus selected for the following tuns on the 740 lipids patches. Lipid parameters were taken from [Poger et al., 2010] for DLPC, while for DLPG they were built from the ones available in the literature for POPG [Kukol, 2009].

The simulations set-up parameters are as above, except for the use of three thermal coupling groups (peptide, membrane, water plus ions), a semi-isotropic pressure coupling, and a larger cut off radius for both Coulomb and van der Waals interactions (1.2 nm). Additionally, for the 512 lipids membrane, a Reaction Field [Tironi et al., 1995] was used instead of PME long range electrostatic treatment. Control simulations on membrane patches without peptide showed that the results in terms of area per lipid are compatible with the ones obtained from PME.

Each membrane patch was first equilibrated for 50 ps in NPT conditions at 50 K, then the temperature was gradually increased up to 300 K in 500 ps, and finally a 400 ns production was run. The final configuration was used as initial structure for the peptide-membrane simulations. A similar equilibration procedure was followed for peptide-membrane systems.

Additional simulations were performed applying an external electric field to the membrane, from the side hosting the peptide to the opposite one, to mimic the membrane potential and verify how the presence of the peptide affects the membrane response to external stimuli. In a first test performed on the 512 lipids bacterial patch, the field was increased by 20 mV/nm steps

every 200 ns, until reaching the electroporation value, which resulted to be 130 mV/nm. For the bacterial membrane patches of both sizes, other simulations were performed starting from the unperturbed membrane configuration and the threshold field, in three replicates each. Three control runs were performed on the 512 patch without peptide, at the threshold value of the field, and one at the higher value of 140 mV/nm.

To be noticed that membrane field is estimated around 20 mV/nm across mammalian membranes and 35 mV/nm (from, respectively, a -70/-90 mV and -130/-150 mV potential [Wilson et al., 2011] and an estimate membrane thickness of 4 nm), but previous computational work explored the effects of fields up to 500 mV/nm [Tieleman, 2004; Böckmann et al., 2008; Piggot et al., 2011], to witness poration within the simulations time, according to the resources available at the time.

An overview of simulations of peptide-membrane systems is given in Table 3.2 (control simulations on pure membrane are listed in Table ??).

SIRAH coarse-grain simulations SIRAH coarse-grain simulations were run with the SIRAH force field [Machado et al., 2018]. Peptide coordinates for the buckyball geometry were obtained from the atomistic ones using the converter distributed with force field. Parameters for the central residue were built from comparison with similar chemical moieties. All simulations were run adding Cl^- counter ions to balance the positive charges of the peptide and additional Na^+ and Cl^- ones to reach a 150 mM concentration (of the cationic species).

While for simulations of peptide assembly in solution we resorted to a multiscale procedure, comparing up to four force fields, for simulations on membrane we focussed only on atomistic and MARTINI simulations only. This has been performed in the interest of time, and following some concerns arisen in preliminary SIRAH runs of the capsule on a membrane, which showed an unstable behaviour (contraction and expansion of the membrane). Given that there are little benchmarks so far on the performances of SIRAH for lipids, and none for systems as large as the one studied here, we are continuing pursuing the investigation at an exploratory level, working in contact with the developers of the force field to help elucidating its limitations and work out the best equilibration procedure to treat large membrane patches.

Capzip on membrane simulations					
	Peptides	Lipids	Model	Time (μ s)	Replicas
Bacterial	10	512	GR	500	2
	10	740	GR	500	1
	120	2880	MA	10000	2
	120	2880	MA_P	10000	1
Mammalian	10	740	GR	500	1
	120	2880	MA	10000	1
	120	2880	MA_P	10000	1
Electroporation simulations					
	Peptides	Lipids	Model	Time (ns)	Rep. E (mV/nm)
Bacterial	10	512	GR	$75^P, 20^P, 71^P$	3 130
	10	740	GR	$60^P, 50^P, 70^P$	3 130
	120	2880	MA_P	500	1 20
	120	2880	MA_P	168^P	1 40
Mamm.	10	740	GR	$20^P, 28^P, 39^P$	3 130
	10	2888	MA_P	500	1 20
	10	2888	MA_P	??	1 40

Table 3.2: Table of simulations of peptide-membrane complexes. All the mentioned ones run at 150 mM concentration of NaCl. Models: GR = united atom GROMOS, MA = coarse-grain MARTINI, MA_P = coarse-grain MARTINI with polar water. Superscript P denotes poration. For electroporation simulations on pure membranes, see SI Table 3.6. A number of 10 peptides denotes the pentagonal subunit, 120 the buckyball bilayer.

For SIRAH simulations, the temperature coupling was performed with a velocity rescale thermostat [Bussi et al., 2007] and coupling constant τ_T of 0.1 ps, and the pressure coupling at 1 bar pressure, with 4.5×10^{-5} bar $^{-1}$ isothermal compressibility, using a Parrinello-Rahman barostat [Parrinello and Rahman, 1981] with a τ_P of 6 ps. Electrostatic interactions were treated using the PME algorithm [Essmann et al., 1995], with a short-range cutoff of 1.2

nm and relative dielectric constant of 1. The van der Waals interactions are treated with a plain 1.2 nm cutoff.

After energy minimization, a 4 ns NVT equilibration was run at 300 K, followed by a 10 ns NPT run, both with positional restraints ($1000 \text{ kJ/mol}\cdot\text{nm}^2$) on the solute. Two 10 ns run (NPT ensemble, 300 K) were then performed with backbone restraints of 1000 and $100 \text{ kJ/mol}\cdot\text{nm}^2$, respectively. Similar to the procedure adopted for the atomistic simulations, during the latter, flat bottom positional restraints ($1000 \text{ kJ/mol}\cdot\text{nm}^2$) were enforced on the unpaired branches during the NPT. Three additional 10 ns equilibrations were run at 300 K, with no backbone restraints and decreasing flat bottom ones (respectively 750, 500 and $250 \text{ kJ/mol}\cdot\text{nm}^2$). Finally the production run was carried on for 1 μs . All runs were performed with a 20 fs time step.

MARTINI coarse-grain simulations For MARTINI [Marrink et al., 2007; Monticelli et al., 2008] coarse-grain simulations, peptide coordinates and parameters were obtained from the atomistic ones using `martinize.py` [de Jong et al., 2013], and `pycgtool.py` [Graham et al., 2017] for the central residue. Parameters for the joining bonds were derived from tabulated values of analogous ones.

The bacterial and mammalian model membranes, hosting 2880 lipids each, were built with `insane.py` [Wassenaar et al., 2015], with composition DLPC:DLPG 3:1 and pure DLPC respectively. The simulations parameters used for lipids are consistent with Ref. [Marrink et al., 2003]. The peptide-membrane systems were built placing the buckyball at a minimum distance of 1 nm from the membrane surface.

For either the simulations in solution and with the membrane, counter ions were added for the standard MARTINI simulations, while for the ones run with Polar MARTINI, additional ones to reach a 150 mM concentration were inserted.

For simulations performed with the standard water model, the temperature coupling was performed with a velocity rescale thermostat [Bussi et al., 2007] with a coupling constant τ_T of 1 ps. An isotropic or semi-isotropic pressure coupling was applied (for peptide in solution or peptide on membrane simulations) at 1 bar pressure, with $4.6 \times 10^{-5} \text{ bar}^{-1}$ isothermal compressibility, using a Berendsen [Berendsen et al., 1984] or Parrinello-Rahman barostat [Par-

rinello and Rahman, 1981] (equilibration and production phase respectively) with a τ_P of 12 ps. An isotropic or semi-isotropic coupling has been used for simulation without and with membranes respectively. Coulomb interactions were treated with a Reaction Field scheme [Tironi et al., 1995] and cut off radius of 1.1 nm, van der Waals interaction with a cut off scheme and the same cut off radius. The relative dielectric constant is set to 15. Simulations performed with the polar water model were run with the parameters above, except the relative dielectric constant set to 2.5, and the choice of a PME scheme for the long range Coulomb interaction (1.2 nm cut off radius).

For simulations of the capsule in solution, for both water models, after energy minimization four 10 ns equilibration runs (NPT ensemble, 300 K, 10 fs time step) were performed with flat bottom positional restraints on the unpaired branches (respectively at 750, 500 and 250 kJ/mol·nm²) - as done for the other force fields. Finally the production run was carried on for 1 μ s. All runs were performed with a 20 fs time step.

It is interesting to notice that the MARTINI force field with the standard water model produce very similar results even without such refined equilibration procedure. However, we choose to follow it to have results more consistent with the other parametrisations and be certain that the differences observed are due only to the force field and not the equilibration procedure.

The membranes used in the simulations are equilibrated for 1 μ s and the final configuration used to build the peptide-membrane system, together with the equilibrated capsule structure (before production). The full system is then after energy minimised and equilibrated for 500 ps. Production is followed for 10 μ s for simulations with the standard water, and 1 μ s with polar water.

The adoption of the polar water model allows to perform electroporation experiments also with the MARTINI force field (while the standard water, not bearing a dipole, is unable to screen the externally applied electric field, resulting in unphysical effects). We thus resorted to a procedure similar to the one employed for atomistic simulations, testing an external electric field of magnitude 20 mV/nm and 40 mV/nm, finding that the latter is sufficient to trigger electroporation in model bacterial membranes in less than 500 ns (while the former is not). It is likely that longer simulations would allow to observe this behaviour even with lower values of the force field, however, in the interest of time, we selected this value for further investigation.

Specifically, we selected a configuration from the early stages of poration (where the pore showed a diameter of 2 nm) and continue the simulations switching to an isotropic pressure coupling. Indeed,

3.3 Analysis

Simulations in solution Several structural analysis were performed on the outcome of the simulations of the buckyball in solution. The Radius of gyration (R_g), the Root Mean Square Deviation with respect to the initial configuration (RMSD) were computed with the GROMACS software. To get the average distribution of the mass of the capsule around its center, the Radial Distribution Function (RDF) of the protein masses around their center of mass was computed (with GROMACS). The profile could be fitted with a Gaussian function (with the R [?] software), so that the position of its maximum can be taken as an average radius of the capsule, and its Full Width at Half Maximum (FWHM) as an estimate of the thickness of the bilayer.

The dynamical character of the structure was assessed computing the correlation of motion between the molecules. The central atom (or bead) from which the branches depart was taken as reference. For all the pairs i and j of such reference positions, the covariance of motion $\sigma^2(i, j)$ was computed (with GROMACS), where the covariance is the sum of the components along each axis: $\sigma^2(i, j) = \sigma^2(i, j)_x + \sigma^2(i, j)_y + \sigma^2(i, j)_z$. Then for each pair, this measure was normalised as:

$$\text{corr}(i, j) = \frac{\sigma^2(i, j)}{\sqrt{\sigma^2(i, i) \cdot \sigma^2(j, j)}}. \quad (3.1)$$

Another structural measure performed concerns the pairing of the branches. Two branches are defined as paired if their center of mass is closer than a cut off distance of 1.2 nm. This simple measure discards any more precise information on the orientation of the chains with respect to each other, and aims at checking whether the network of molecules present in an ideal buckyball structure is maintained. In the ideal buckyball, contacts within the same layer sum up to 90 for each layer (for a total of 180 in a bilayer). This measure can be easily applied to any description (atomistic or coarse-grain) without disagreement in the interpretation. The computational pipeline combines GROMACS tools

and a post-processing in R language.

To characterise the chemical determinants that promote the assembly, we investigate the interactions between amino acids of different types, computing the number of contacts between backbone and side chains of single amino acids, filtered for the ones present at least 50% of the simulation time, and classified them by amino acid type. We define contacts between amino acids backbones if the C α of two residues (or the corresponding coarse-grain bead) are closer than a cutoff distance of 0.6 nm; and between side chains if selected reference atoms in the side chain are closer than the same cut off; finally mixed ones if the proximity is between a C α and the side chain reference atom. As side chains reference, we took the heavy atom or bead farthest away from the backbone (respectively for GROMOS, SIRAH and MARTINI: CZ/SC2/BCZ for Arg, CZ2/SC4/BNE for Trp, OG1/SC1/BPG for Thr and CD/SC1/BCD for Glu). The functions to perform the analysis were built on the ones implemented in the MDAnalysis software (see Appendix –).

Further investigation has been carried on for atomistic simulations computing the hydrogen bonds between amino acids and grouping them by amino acid type and by region of occurrence (e.g. between two backbones, side chains or connecting a backbone atom and a side chain one).

Finally, the Solvent Accessible Surface Area was computed for each amino acid (with GROMACS). Its value was averaged in time and over all the residues of the same type. It was normalised over its reference value for each residue type X, obtained as the measure of the SASA of X from a Gly-X-Gly tripeptide. The resulting measure (named Q_{SASA}) takes into account the size of the side chain of each amino acid, giving a measure of exposure which can be compared between different residues. This normalisation is somewhat inappropriate for coarse-grain models, however we employed it as it provides nevertheless a coarse regularisation for size effects.

Simulations on a membrane To assess the influence of the peptide or of the presence of an electric field on the membrane, the area per lipid (ApL) is monitored, i.e. how much space each lipid molecule has on the membrane plane. In the case of approximately flat membranes aligned to the xy plane, this can be assessed from the product of the lateral dimensions of the simulation box divided by the number of lipids in one leaflet. For highly curved membranes,

as the ones obtained under the effect of a strong electric field, this measure does not reflect the true spacing between lipids. In that case the true ApL was computed through the algorithm developed by Braun et al. [Braun et al., 2011]: it first identifies the undulating reference surface $u(\mathbf{r})$ of the membrane fitting a Fourier series to a reference atom per lipid (presently the Phosphorus), then it computes the true ApL based on the interpolated surface $u(\mathbf{r})$ as:

$$ApL = \frac{1}{N} \int_{x_{box}} \int_{y_{box}} \left[\sqrt{1 + (\nabla u(\mathbf{r}))^2} \right] dx dy \quad (3.2)$$

where the gradient $\nabla u(\mathbf{r})$ is zero for a perfectly flat surface.

For atomistic simulations, the average order of lipid tails has been assessed computing the deuterium order parameters S_{CD} of the acyl chains for each lipid bilayer leaflet separately. S_{CD} measures the orientation with respect to the bilayer normal of a carbon-hydrogen bond in a given position along the chain for each lipid in the bilayer. Their spread is evaluated according to the ensemble average:

$$S_{CD} = \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle. \quad (3.3)$$

As the GROMOS force field employs a united-atom representation, the tetrahedral positions of the hydrogens are constructed based on the neighbouring carbons positions.

To evaluate the mobility of lipids on the membrane plane, for each simulation we extracted the trajectory of the phosphorus atom of every lipid, in the top and bottom leaflet separately, removing the collective motion of the leaflet with respect to the simulation box frame. Each trajectory was used to compute the Mean Square Displacement (MSD) of the lipid as a function of time. These curves were averaged over all the lipids of a given species belonging to the leaflet of interest, and, for a given interval of time, on all the possible time windows of that length fitting within the simulation time analysed. The diffusion coefficient D was obtained from a linear fit of the average MSD profile, following Einstein equation in two dimensions:

$$\langle (x - x_0)^2 \rangle = 4Dt. \quad (3.4)$$

The fit was performed in the regions which showed a linear dependence. This implied discarding the first interval of the profile, where the behaviour is not

linear, and the last 50 ns, where the poorer statistics leads to more noisy data. The exact amount of the discarded initial region depends on the system studied, e.g. whether the protein is present on the membrane or not.

For atomistic simulations of the protein on the membrane, the network of hydrogen bonds between the protein and lipids has been computed, classifying the interactions by the protein residue type and the lipid species they occur between. Finally, another useful information was given by the number of lipids of a species within a threshold distance from the protein, in time. The aforementioned diffusion constant can also be computed selectively on the lipids which, at the initial time, are within this threshold distance or, conversely, which are further away from it.

To quantify whether a lipid patch is undergoing transition to a gel phase, the regular packing of the lipids' chains was quantified by the hexagonal order parameter S_6 [Uppulury et al., 2015]. Specifically, a chain was represented by its position on the xy plane (parallel to the membrane surface), computed as the average x and y position of its carbon atoms. For each chain j it was computed the set of neighbouring chains $\{n\}$ as the ones within a 0.65 nm radius from j . Then S_6 is defined as:

$$S_{6,j} = \frac{1}{6} \left| \sum_{k \in \{n\}} e^{6i\theta_{jk}} \right| \quad (3.5)$$

with θ_{jk} being the angle between the vector connecting j and k , and the x axis (and i is the imaginary unit). A chain is in gel phase if it has an hexagonal order parameter larger than 0.72 [Uppulury et al., 2015].

3.4 Results: capsule in solution

We list here the results for the capsule in solution first: starting from the atomistic simulations, we then proceed to compare the outcome with the ones obtained by different coarse-grain models.

3.4.1 Atomistic simulations

We discuss first the results from atomistic simulations, color coded as black or in shades of grey in the following sections. However, all the plots present back

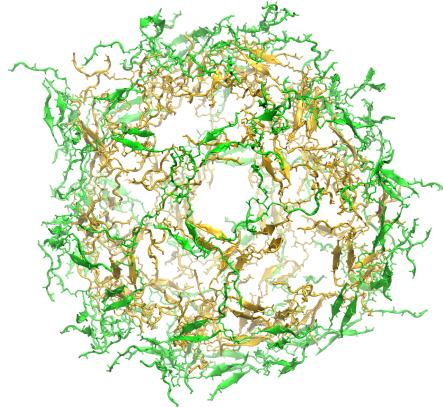


Figure 3.5: Final snapshots (100 ns) from the simulation of a buckyball in solution (replica 3): bonds and cartoon representation, backbone only, green external layer and yellow internal one.

to back the results from coarse-grain runs as well, anticipating the discussion of Section 3.4.2, as this would make easier the comparison later on.

Global capsule structure Atomistic simulations of the buckyball in solution show a consistently equilibrated structure across the three replicas (Figure 3.5 and SI movie –). This is proven by both the stable value of the protein R_g and the almost plateauing backbone RMSD (Figure 3.6(a) and (b)). It is interesting to notice that previous simulations performed with a shorter equilibration, without the phase employing flat bottom restraints, resulted in the immediate disruption of several connections in the buckyball network, with resulting larger R_g . This suggests that the structural pairing present in the structure can form only when the chains are in close contact. For this to happen, many conformations need to be sampled, and this is compatible with the long time of assembly observed experimentally (up to 7-10 hours). The RDF of the protein masses around the buckyball centre of mass shows a Gaussian profile (Figure 3.6(c)). The fact that no masses are observed nearby the origin means that the molecules do not collapse to the center and the central cavity is maintained. To be noticed that, given the way RDF is computed and normalised, a uniformly full object would display a flat distribution. A fit of the RDF to a Gaussian curve returns a mean value of 5.1 nm and a FWHM of 2.2 nm, which gives an estimate of the bilayer thickness. A similar computation is repeated for the inner and outer layer separately, providing 1.1 nm of distance

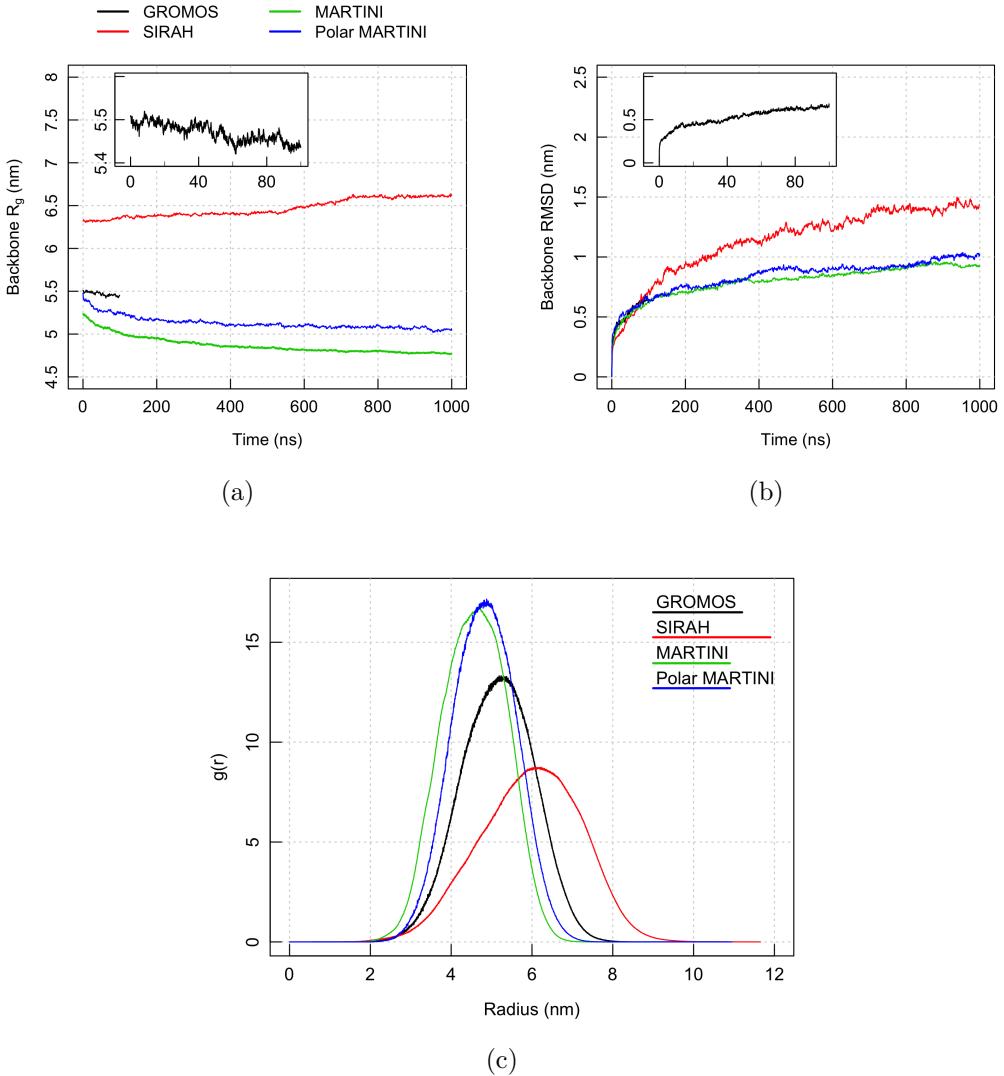


Figure 3.6: (a) R_g and (b) RMSD computed respectively on the Protein and its backbone. Results are displayed for simulations performed in GROMOS (100 ns), SIRAH, MARTINI and MARTINI with polar water (all 1 μ s). Inset: zoom on the GROMOS values. (c) RDF of Protein masses around their center of mass, computed on both layers, displayed for the same simulation set up as in (a,b). For each label of the legend, the bar has length of the respective RDF FWHM (thickness estimate). All results are showed for Replica 1 of each simulation setup.

between the two distributions means. This interlayer distance is compatible with the distance between the backbones of stacking β -sheet in structures like densely packed amyloids with cross- β sheet quaternary structure (1.0 nm [Sunde et al., 1997]).

This thickness value hints at the fact that the two layers are closely packed. This is confirmed by the analysis of the pairwise correlations of motion: molecules at the same polar coordinates (i.e. stacking radially one on the other) have a positive correlation and so move coherently, while the ones at opposite poles, as well the ensemble of all possible pairs, do not show particular correlation (Figure 3.7).

Contacts between molecules The measure of paired branches in the buckyball network, shown in Figure 3.8, reports an average (in time) of 240 pairings between arms belonging to the same layer (summing over inner and outer), four pairing per molecule, and around 60 only for inter-layer ones. This first value is larger than the one predicted for a perfect buckyball arrangement (which would be three), likely due to the fact that the structure contracts slightly with respect to the initial size and some branches deviate from the original position, locating themselves in the proximity of multiple neighbours. On the contrary, few inter layer contacts are observed within 1.2 nm distance cut off because of the steric hindrance of the side chains which are located in between the two layers, as they keep the average position of inter-layer backbone stretches at a distance greater than the cut off chosen.

To look more into the details of the interactions between arms of the molecules, we computed the contacts between backbone and/or side chains of amino acids, normalised by the number of molecules present, and classified the ones which survive more than 50% of the simulation time by residue type (Figure 3.9). The number of backbone contacts per molecule is around 3 for Threonine, Glutamic acid, 4 for Arginine, and between 6 and 7 for Tryptophan residues. As in each molecule there are 3 Threonine and 3 Glutamic acid but 6 Arginine and 6 Tryptophan, this proves that on average each residue is well paired with another one, except for Arginines: only two thirds of them are paired on average, likely because of their terminal position. At the same time however the bar plot shows also that there is no rigid arrangement between branches. Indeed, for example, Tryptophan residues are not chiefly paired to

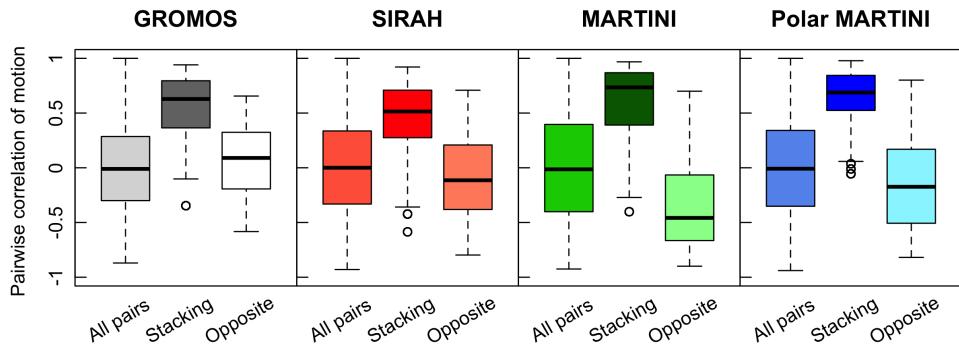


Figure 3.7: Distribution of the correlation of motion between different molecules in the buckyball simulations. Black band: median of the distribution; box: first and third quartiles; whiskers: maximum and minimum, outliers excluded (hollow dots). Results are shown for Replica 1 of each simulation setup.

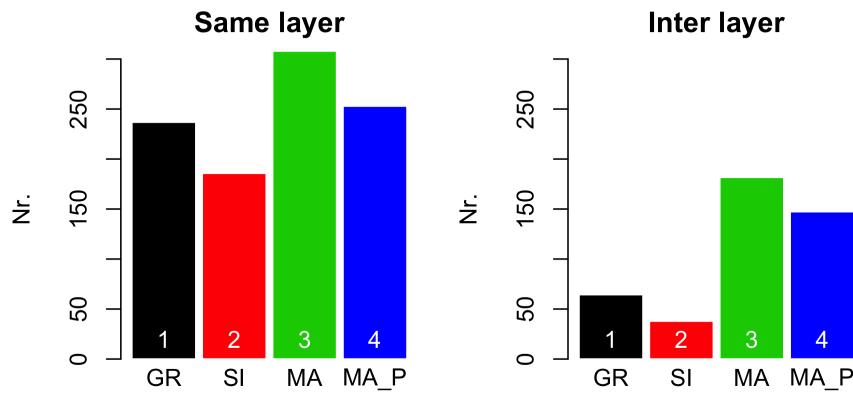


Figure 3.8: Number of paired branches within the same layer and between layers, as defined in the main text. Results are shown for Replica 1 of each simulation setup.

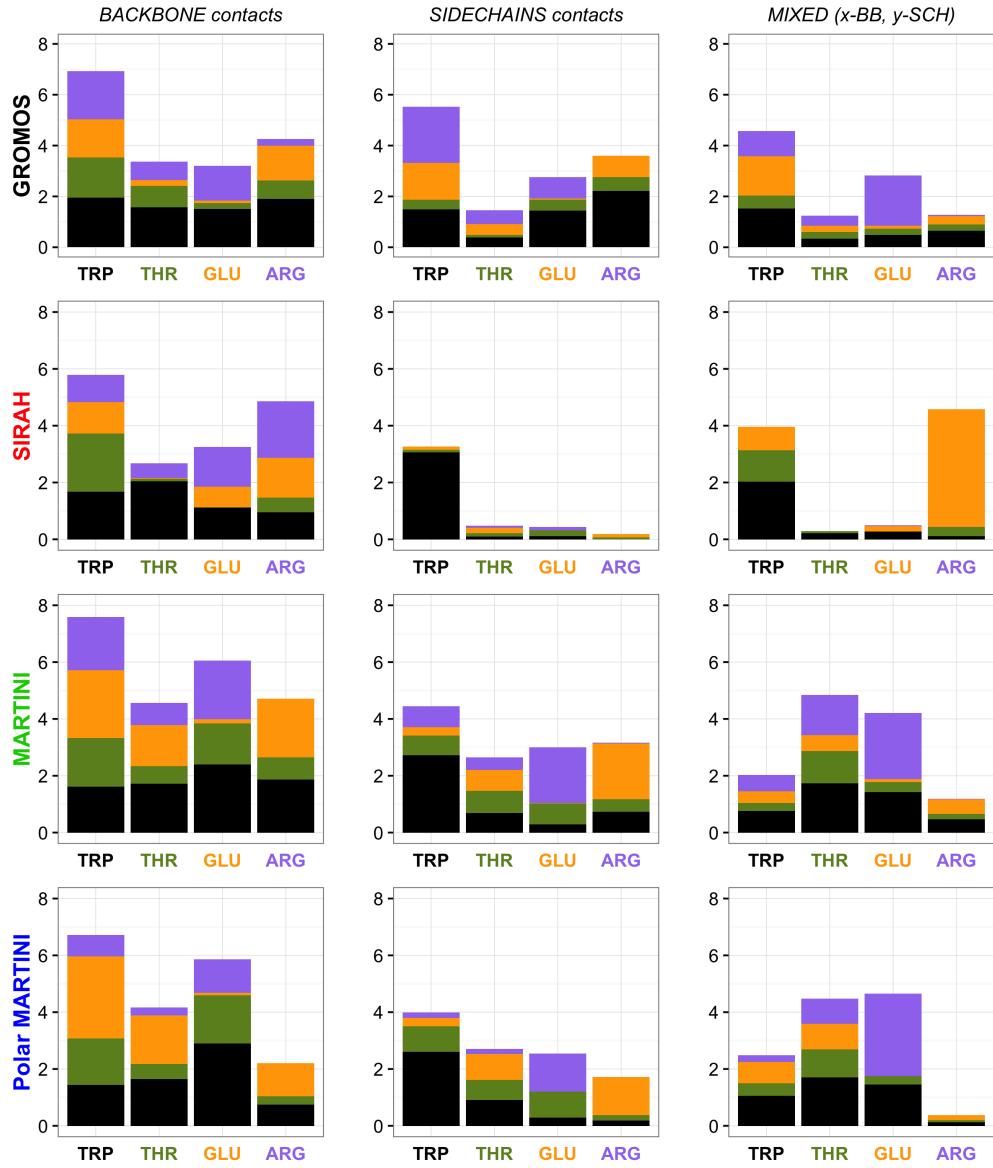


Figure 3.9: Contacts with persistence greater than 50% between C_{α} s (Backbone contacts), between side chains (Sidechain contacts) and mixed (in which case, for each bar, the residue on the x -axis contributes to the contact with its backbone, and the bar is split by the identity of the partner engaging its side chain). The simulation setup is reported along the y -axis. Results are shown for Replica 1 of each simulation setup.

Tryptophan ones, as the optimal arrangement would be, suggesting flexibility in the structure. Nevertheless, the analysis highlights that Trp residues are key to form contacts with the neighbours and (from the central column) that their cation- π interaction with Arginine through their side chains is an important element of the structure.

Hydrogen bonds interaction Some of these contacts are mediated by an hydrogen bond interaction: the number of them present during the simulations is computed, and divided by the number of molecules. We classify the hydrogen bonds occurring between backbones and/or side chains of the amino acids, breaking them by amino acid type. Tryptophan contributes to a large number of backbone hydrogen bonds (Figure 3.10 top), especially with other Tryptophan residues, consistently with what found in the analysis of contacts carried on above. Arginin side chains are the most prone to establish H-bonds as a donor with many different amino acid side chains, but especially with Glutamic acid as expected from the facing positions they occupy in the molecules arrangement.

Chemical characteristics of the surface: SASA Finally, it is important to understand what residues are exposed at the surface of the structure, especially in view of future applications: in order to make the peptide co-assemble with other products, the two must have a compatible chemical character. To understand what surface the peptide exposes to the solution, we compute the average Solvent Accessible Surface Area (SASA) per molecule and break it down by amino acid type. Figure 3.11 shows that half of the accessible surface is represented by the charged residues Arginine, while Tryptophan contribute to less than one quarter to it, despite having bulky side chains. For atomistic simulations, we can compare the SASA per each residue X, with the reference SASA computed on simulations of a Gly-X-Gly tripeptide (computation performed with the software POPS [?]). The resulting ratio Q_{SASA} is greater than one for both the Arginines: while for the terminal one (ARG1) it is suspected, the fact that also the second has $Q_{SASA} > 1$ proves that these residues are highly exposed in solution. On the contrary, all the other residues have values around 0.5: for Tryptophan this is due to their propensity to be buried inside the structure, while for Glutamic acid and especially Threonine, despite

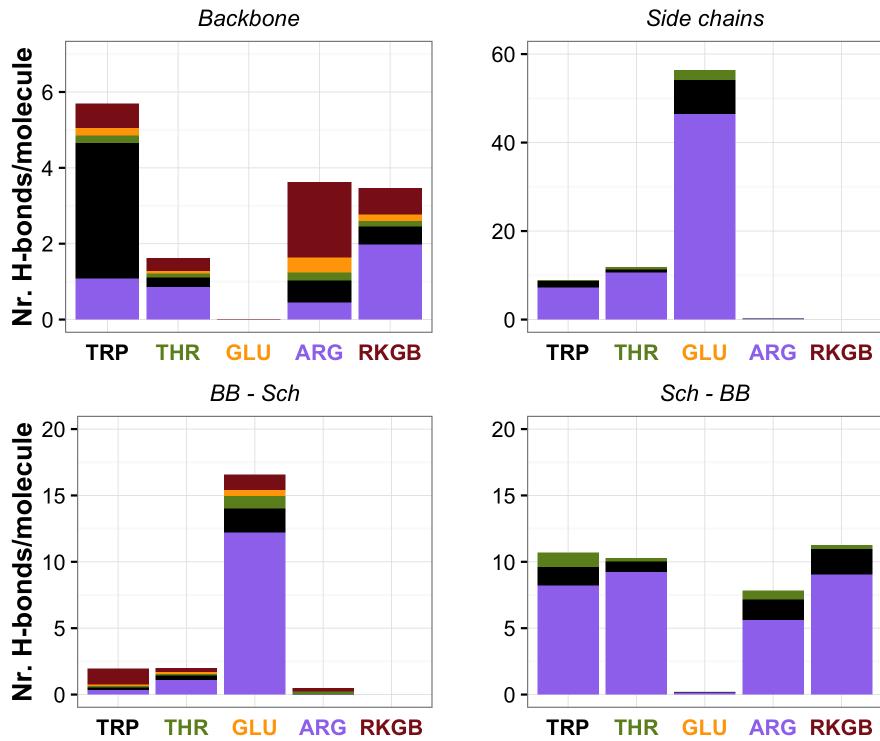


Figure 3.10: Average number of hydrogen bonds per molecule occurring between amino acids, including the central scaffold RKGB, in a 100 ns atomistic simulation of the buckyball in solution. Result are shown for Replica 1. For each bar, the residue on the *x*-axis is the acceptor, and the bar is split by the identity of the donors. In the case of Backbone - Side chain and Side chain - backbone, the first mentioned correspond to the acceptor (and thus the residue on the *x*-axis).

having charged or polar side chains, they are shielded from the solvent by the large Trp side chains nearby them.

3.4.2 Multiscale comparison of model capsule

We performed a multiscale analysis of the capsule structure simulating it with different coarse-grain force fields, with a twofold aim: first we wanted to simulate the assembly for a longer time, to observe how its structural properties are maintained on the medium time scale (of the order of the microsecond). Second, we believe that proving the stability of the capsule with different descriptions strengthens the evidence that the assembly proposed is indeed a favourable arrangement of the molecules in solution.

As mentioned before, to this aim we compared simulations run with the

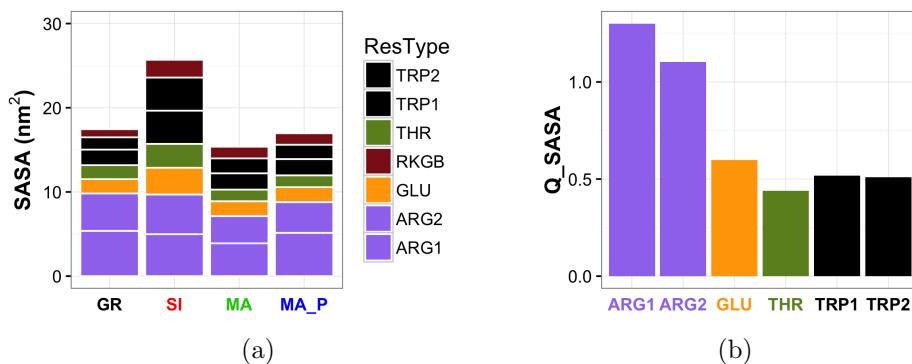


Figure 3.11: (a) Break down of the Solvent Accessible Surface Area (SASA) per residue type. Results are shown for Replica 1 of each simulation setup. (b) Normalised SASA over the reference SASA computed for each amino acid type X as the value in a Gly-X-Gly tripeptide.

SIRAH, MARTINI force fields and MARTINI used in conjunction with polar water. The investigation is also useful to compare how the different descriptions differ and to infer the advantages of each model. We first comment on the quantities already analysed at the atomistic level (if applicable), and then extend the analysis to simulations of a monolayer capsule, which has been modelled to prove the greater stability of the bilayer one.

Force field comparison As foreseeable, the structures obtained with coarse-grain force fields are slightly different with respect to the atomistic one and among themselves. The SIRAH simulated bilayer capsule has a more open structure with respect to the atomistic one, with a skewed and broader radial density profile, while MARTINI with the standard water model provides a more compact configuration (respectively 6.0 nm and 4.6 nm average radius, with 2.9 nm and 1.9 nm average thickness - see Figure 3.6(a) and (c)). MARTINI with polar water provides a slightly larger structure with respect to standard MARTINI (4.8 nm average radius), with a comparable thickness. This is likely due to the poorer properties of solvation of standard MARTINI water, which cause the protein beads to preferentially interact between themselves rather than with the solvent.

The correlation of motion between molecules for all the coarse-grain force field is similar to the results of the atomistic one, with a slight anticorrelation between molecules at the opposite poles (Figure 3.7). This is due to the

contraction or expansion happening at the beginning at the simulation, when the capsule adjusts to the equilibrium size, which depend to some extent on the force field. These effects though are slightly more pronounced for the MARTINI force fields, probably due to the greater cohesion between the beads which makes them moving coherently.

The number of chain paired in the SIRAH simulations is fewer than in the atomistic ones (Figure 3.8), in line with a more expanded structure, while MARTINI simulations propose a higher number, consistently with the reduced size of the capsule. In particular, the inter layer contacts are significantly higher (more than the double), as would be expected when the two layers are closer and more strongly bound, as suggested by the values of the stacking molecules correlation. Finally, MARTINI suggest a more compact structure than Polar MARTINI, privileging contacts between residues rather than contacts with water molecules.

Breaking down the contact analysis by residue type (Figure 3.9), coarse-grain simulations shows necessarily a different organisation, due to the fact that different parametrisation results in different proportions between the side chains volumes of different amino acids. Never the less, in all the representations, Tryptophan has a prominent role in establishing contacts with its neighbours at the backbone level, while different force field suggest different role of the side chains. Quite surprisingly, the SIRAH force field does not promote interactions between side chains a part from the Tryptophan ones with themselves. This seems due to the more expanded structure of the capsule and the larger solvation of the amino acids, which are more exposed in solution (see next paragraph discussing SASA values). Moreover, none of the coarse-grain descriptions seem to capture the preferential cation- π interaction between Arginine and Tryptophan.

Finally, the values of SASA cannot be compared between force fields, because of the different dimensions and number of atoms/beads employed. However, it is interesting to notice that consistently across force fields, Arginine constitute around half of the exposed surface, with the exception of SIRAH, where the more open structure results in other residues to be more exposed as well (Figure 3.11(a)).

Overall, these results suggest that every coarse-grain force field has a particular propensity for an equilibrium distance between peptidic components

due to the different property of solvation of the water model chosen. To better understand this, we computed the Coulomb and Lennard-Jones contribution to the energy due to peptide-peptide interactions or peptide-solvent ones. Figure 3.12(a) plots these value for the second half of the simulated time for each coarse-grain force field (to be noticed that for the SIRAH description the Protein-Protein terms include both the short range interactions and the 1-4 interactions, i.e. the ones computed between atoms separated by three bonds. These interactions are not part of the force field for MARTINI models, so only the short range ones are accounted in those cases).

A great difference can be seen among force field, as expected. Looking at the Coulomb component, it is clear that the mean field approach adopted by MARTINI, which consists in adopting a high relative dielectric constant ($\epsilon = 15$) to compensate the absence of water dipoles, decreases massively the contribution of the Protein-Protein electrostatics to the overall energy with respect to the two other models. Indeed this approach decreases the interactions also on the short range, while in reality the screening effect due to water can be seen only on long distances (of about 1 nm in a 0.1 M salt solution), and, due to the $1/r$ behaviour of the Coulomb interaction, these short scale contributions are clearly important. The partial reversion to $\epsilon = 2.5$, together with an introduction of the water dipole, increases the amount of the Coulomb contribution by 10-fold (in absolute value). This is higher than the ratio of the ϵ between the two models (equal to 6), likely due to the fact that the different screening allows a rearrangement of the charges which brings opposite ones closer to each other.

SIRAH simulations instead, run at $\epsilon = 1$, present Coulomb components 2.5 times larger than the ones of Polar MARTINI, suggesting that the two water models, both representing the separation of water charges in an approximate way, give a similar energy contribution despite being different.

This difference in Coulomb energies affects both the Protein-Protein and Protein-Water interactions consistently, however it can have consequences on the dynamics because it changes their proportion with respect to the Lennard-Jones contributions. These interactions are not changed between the two MARTINI models, making their role predominant in standard MARTINI, while competing with the Coulomb contribution in Polar MARTINI (still being larger overall). On the contrary, SIRAH parametrisation opts for a smaller role

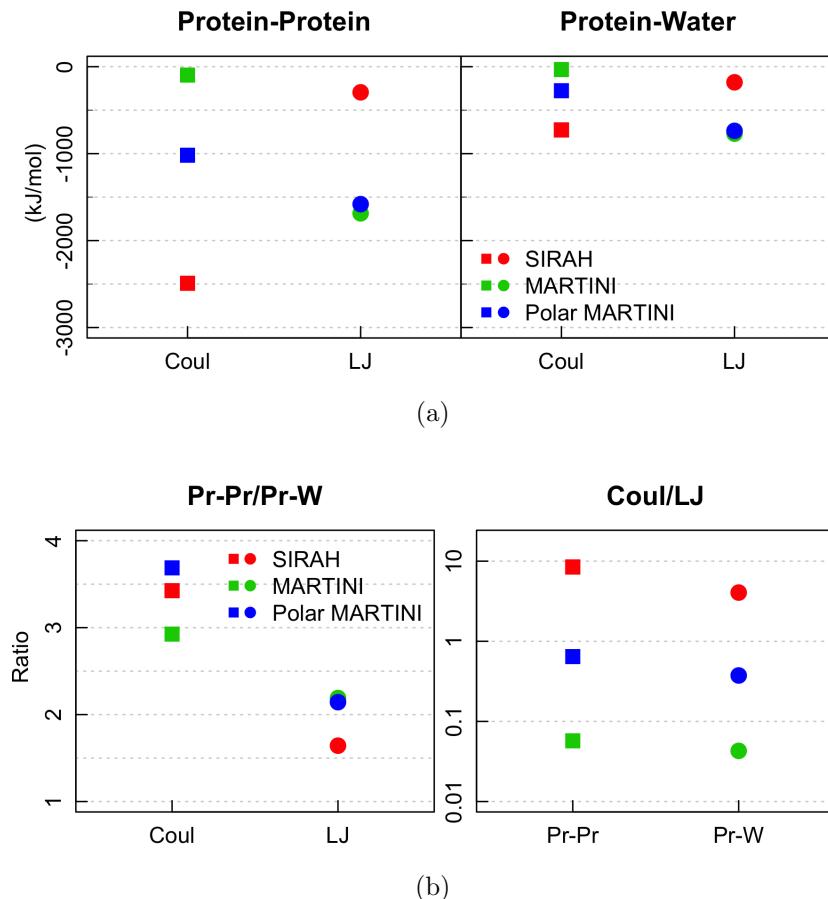


Figure 3.12: (a) Protein-Protein and Protein-Water non-bonded interactions, normalised per molecule. Values obtained as average on the second half of the trajectory of Replica 1 for each setup. (b) Ratio between the Protein-Protein and Protein-Water interactions for each force field, for Coulomb and Lennard-Jones respectively; or between Coulomb and Lennard/Jones, for Protein-Protein and Protein-Water interactions separately (note the log scale on y). Values computed as for plot a).

of Lennard-Jones with respect to the Coulomb contribution. Therefore, protein electrostatics have an increasing contribution in MARTINI, Polar MARTINI and SIRAH respectively, both in absolute terms and with respect to the Lennard-Jones contribution.

This can partially explain the differences in sizes observed across the models: despite the Protein-Protein Coulomb energy is more negative in SIRAH, because of the way it is computed, with a smaller dielectric constant these contributions are less screened, thus the many positive amino acids composing (which are more abundant than the negative ones) the capsule can repel each other more effectively.

Finally, to understand the ratio between the Protein-Protein and Protein-Water components, for both Coulomb and Lennard-Jones, is computed (Figure 3.12(b)). All the three force fields present values in the same range. The ratio of Lennard-Jones interactions is slightly lower in SIRAH than in the two MARTINI models, and the ratio of Coulomb interactions is between 3 and 4 for all the models. Interestingly, Polar MARTINI has a greater Protein-Protein component with respect to standard MARTINI (in absolute value, which means a more negative contribution). From this can be deduced that the slightest more open structure observed in Polar MARTINI is likely not due to the more favourable solvation properties of the polar water model, but from a different balance of electrostatic versus Lennard-Jones, as explained before.

An interesting follow up on this topic would be investigating whether a tuning of the dielectric constant used in either cases can produce more consistent results between the two water model. However, the choice of the constant was optimised to reproduce at best the properties of bulk water and solvation free energies of ions for both cases. This then raises the question whether the protein parametrisation performs equally good in conjunction with both model, or which one is more compatible.

Bilayer versus monolayer capsules To prove also on large scale objects that the bilayer structure is indeed essential to grant a structure which does not disassemble or change shape, we perform simulations of a monolayer capsule (specifically taking the external layer of the capsule already simulated) in the three coarse-grain force fields employed so far. We focus here on the results obtained with the SIRAH and Polar MARTINI models, as the ones observed

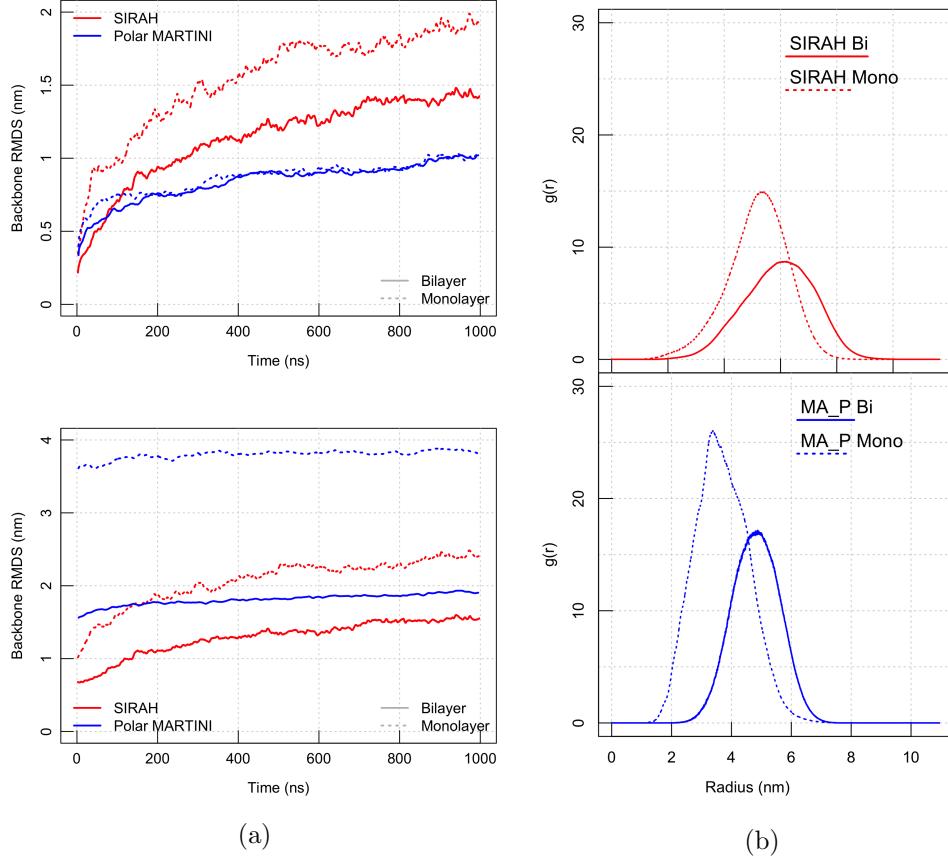


Figure 3.13: (a) RMSD of the monolayer and bilayer structures for SIRAH and Polar MARTINI force field, with respect to the initial structure of the production (top), and initial geometrical configuration (external layer of Figure 3.2(d)). (b) RDF of Protein masses around their center of mass. For each label of the legend, the bar has length of the respective RDF FWHM (thickness estimate). Results are shown for Replica 1 of each simulation setup.

with standard MARTINI resemble closely the Polar version.

The RMSD with respect to the initial configuration of the production (Figure 3.13(a), top) shows that the monolayer undergoes a larger conformational change than the bilayer in the SIRAH force field. This effect is not traceable in Polar MARTINI within the same plot. However, if we consider the RMSD with respect to the initial structure built (the geometrically regular polyhedra as in Figure 3.2(d)), it is clear that for Polar MARTINI the discrepancy between monolayer and bilayer is even more pronounced (Figure 3.13(a), bottom). This is likely due to the fact that rearrangements are more favoured in

short time during the more coarser simulations.

This larger change of the monolayer is due to a larger contraction of the structure, which collapses within its center (Figure 3.13(b)). This is evident especially for Polar MARTINI, while for the $1\ \mu\text{s}$ SIRAH simulations, together with the contraction it is observed that Tryptophan side chains have a tendency to become closer, modifying locally the structure and causing a local puckering (see Figure 3.14). Likely this mechanism is present also in the MARTINI ones, but is it overshadowed by the general contraction, which brings the chains close enough to screen the hydrophobic components without a puckering of the arms as the one observed in SIRAH.

Finally, the SASA of each residue type computed on the initial configuration is higher for the monolayer than the bilayer, while this difference partially levels out during the simulations, due to the rearrangements mentioned above (Figure 3.15).

3.4.3 Discussion

All the findings above strengthen the hypothesis that a favourable packing must allow for the screening of hydrophobic residues, arranging multiple copies of them close enough to interact together. For this reason, the bilayer structure is more prone to maintain its original configuration.

Based on this, we deemed that the capsules observed experimentally must have a non-monolayer structure: a single layer would not provide enough structural stability and, moreover, is not compatible with the thickness observed in some of the TEM images collected experimentally.

The multiscale investigation of this system provided both information on how it is structured, and a useful comparison between force fields, proving how the same system can be described in different ways from different models. This limited comparison might still be useful especially in the light of the comparatively new SIRAH force field, which has been less assessed than the MARTINI description.

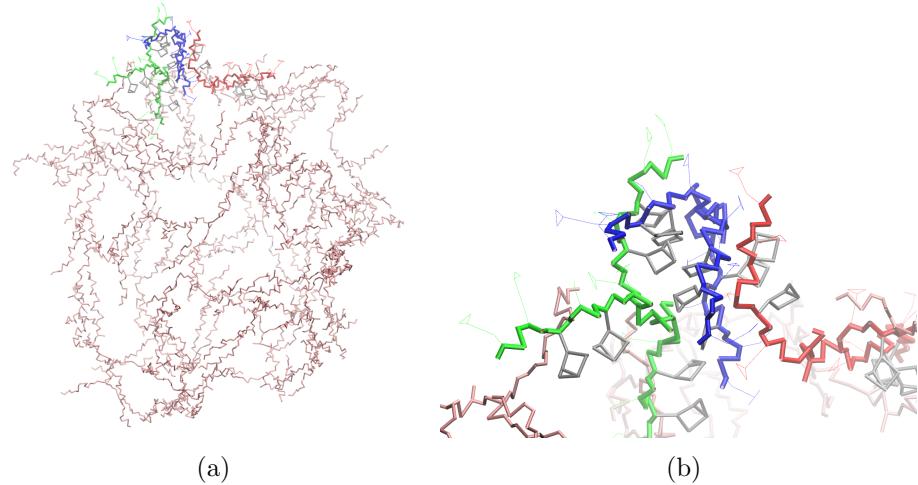


Figure 3.14: [REDO] (a) Final configuration of a SIRAH monolayer simulations (Replica 1), with highlighted a protruding puckered structure recurrently occurring several positions of the capsule. (b) Detail of the puckered structure. Representations: backbone of the full structure in pink bonds; backbone of highlighted triskelion molecules in green, blue and red bonds; side chains of selected molecules in green, blue and red lines, excluded Tryptophan, in silver bonds.

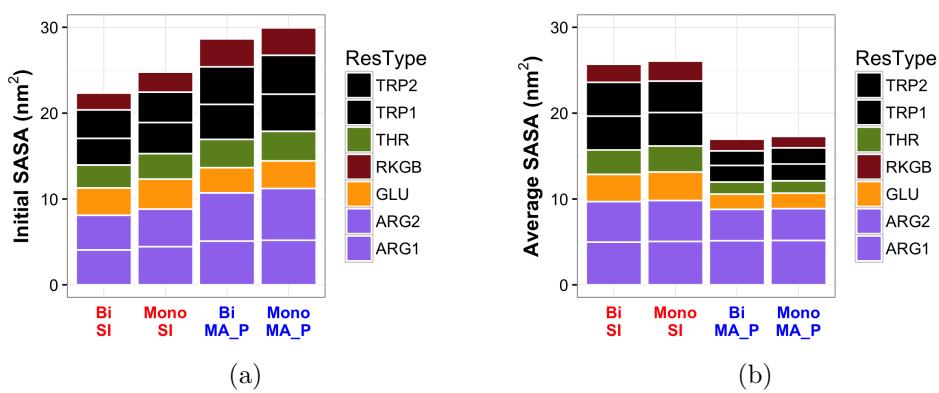


Figure 3.15: Break down of the Solvent Accessible Surface Area (SASA) per molecule by residue types for simulations of the bilayer and monolayer structure. Results are shown for Replica 1 of each simulation setup. (a) SASA compute from the initial configuration; (b) from the average over the production run.

3.5 Results: peptide-membrane interactions

We now discuss the results obtained from atomistic simulations of the pentagonal peptidic subunit in contact with a model bacterial or mammal membrane, both under standard simulation conditions and with an applied external electric field. In this we elucidate the local effect of the peptide on the lipid organisation within the bilayer. We then assess the process of membrane binding through coarse grain simulations, without and with electric field, comparing the MARTINI and Polar MARTINI force fields for the simulation without an external potential, and necessarily resorting to Polar MARTINI for the ones involving the electric contribution.

3.5.1 Atomistic simulations of the bacterial model membrane

Influence of the peptide on the 512-lipids patch For the model of bacterial membrane chosen (composition DLPC/DLPG 3:1), the ApL obtained from simulations was, according to the setup, $0.5796 \pm 0.0052 \text{ nm}^2$ or $0.5686 \pm 0.0036 \text{ nm}^2$ (respectively 512 lipids with GROMOS 54a7 force field or 740 lipids with GROMOS 54a8 force field - Figure 3.16(a), lines 1 and 5). The differences can be attributed to the force field, however extensive tests on many phosphocholine lipids (see Chapter 4) suggests that the trend is generally the opposite, with 54a8 giving larger values than 54a7. Therefore we attribute the discrepancy partly to reproducibility issues, partly to the different sizes of the patches which has a great effect, as it will be shown in the following.

Regarding the smaller patch, the presence of the pentagonal peptide sub-unit (made of ten antimicrobial molecules, bearing a $+60e$ charge) made the ApL decrease by 9% with respect to the apo value (Figure 3.16(a), line 3). A similar effect (qualitatively and quantitatively) was observed applying an electric field of 70 mV/nm intensity or combining the presence of the peptide with a weak field of 20 mV/nm, comparable with physiological values (Figure 3.16(a), lines 2 and 4). Likely, they give comparable perturbations as the high protein charge produces a similar field which interacts strongly with the anionic lipid species.

The fact that all the different perturbations lead to a similar reduction of ApL also from the quantitative point of view suggests that the membrane is approaching the maximum packing allowed. Consistently, the diffusion coef-

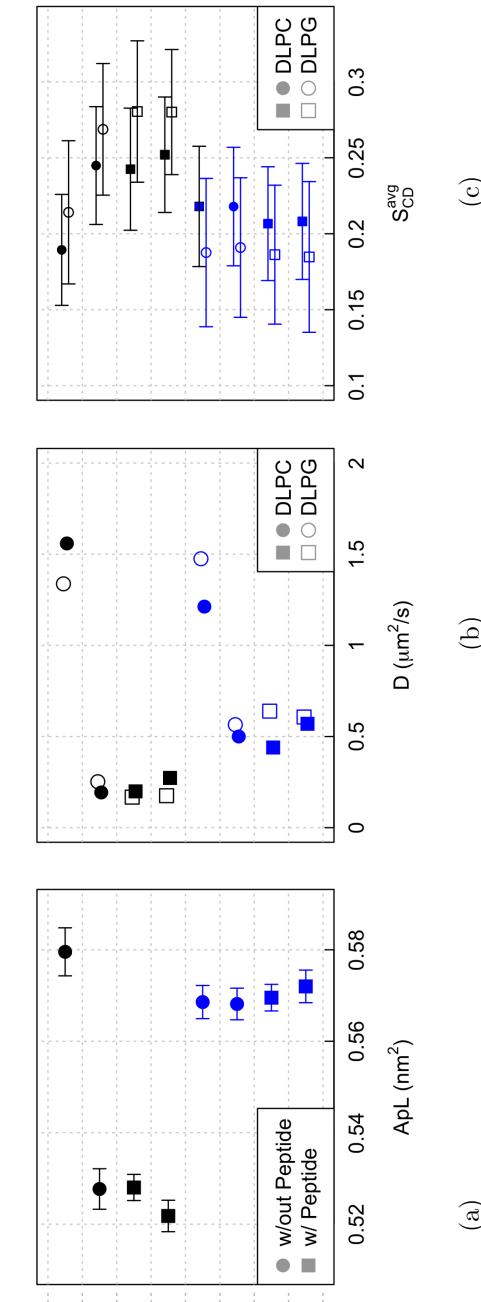


Figure 3.16: (a) Area per lipid (A_{PL}); (b) lateral lipid diffusion coefficient (D), (c) average tail order S_{CD} from simulations of model bacterial membrane patches simulated under different conditions. On the left of (a) and (c) a schematic of the system setup shows the number of lipids L , number of peptides P , and electric field E applied in the z direction (in mV/mm) in each case. Black points refer to simulations on 512 lipids patches, blue ones on 740 lipids patches; squares denotes presence of peptide molecules; in (b) full symbol are for DLPC and hollow for DLPG. Bars denote the standard error; for the diffusion coefficient, these are smaller than the point size. When more replicas are run, results are shown from replica 1. Analysis performed discarding the first 200 ns of simulations.

ficient D had a 6-fold reduction for both species of lipids involved for any of the perturbations mentioned and the order parameter S_{CD}^{avg} increased slightly (Figure 3.16(b)) and (c)). The variability on this last quantity is high, showing that some regions reached a S_{CD}^{avg} value larger than 0.3, which is often associated to transition to the gel phase [Pluhackova et al., 2016]. Computing the hexagonal order parameter S_6 for the bottom leaflet of the system i presence of the peptidic subunit shows that only a few lipids were in the gel phase ($S_6 \geq 0.72$), without the appearance of ordered clusters (Figure 3.17), however none of them has this property in the apo simulations.

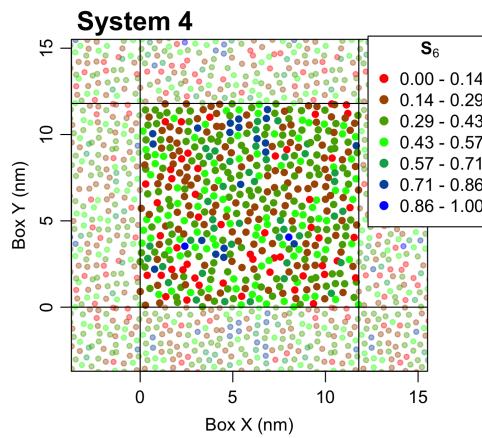


Figure 3.17: Hexagonal order parameter S_6 for the bottom leaflet lipid acyl chains, computed on the last frame of a simulation of a 512 bacterial patch with 10 peptide molecules (line 3 in 3.16). Chain plotted by the average xy position of their carbon atoms, and colour coded by the S_6 value; chains of the periodic images shown faded out; boundaries of the simulation box in solid black lines.

Influence of the peptide on the 740-lipids patch It is remarkable that when the same perturbations were applied to a larger membrane, ApL and tail order parameter S_{CD}^{avg} were equal within the error between the different conditions tested, while the diffusion coefficient D was still significantly affected, being reduced by 2.5 times (Figure 3.16(a)-(c), lines 5 to 8).

This suggests that the simulation conditions can affect remarkably the outcome and the possible interpretation: we reckon that the larger impact on the smaller membrane is due to the fact that local perturbations are effectively communicated to other parts of the membrane through periodic boundary

conditions, while this is less effective in a larger patch. Alternatively, this can also hint at the fact that the 54a8 force field proposes a more rigid structure of the membrane with respect to GROMOS 54a7. However, results from simulations of the membrane alone presents similar values of both ApL and D (with a slight reduction in the ApL when going to the larger membrane). We reserve further tests on this subject for simulations under an applied electric field, as it will speed up the appearance of any differences between the two setups.

Locally reduced mobility The common effect observed on both membrane sizes is a diminished mobility of the lipids. To understand the local determinants of this outcome, we computed the diffusion coefficient for lipids based on their distance from the peptide at the initial frame of the production run (Table 3.3). All lipids are slowed down in their diffusion, with the ones closer to the protein being more affected. To be noticed that values of the diffusion coefficient smaller than $1 \mu\text{m}^2/\text{s}$ (as the ones observed for the lipids within 1 nm from the peptide) are usually associated to the gel phase [Scomparin et al., 2009].

This effect is due to electrostatic and hydrogen bond interactions between the lipids and the peptide. We computed their number in the large membrane simulations, monitoring the ones which persisted more than 40% of the time (Table 3.4), singling out the ones involving Arginine and Tryptophan residues. Arginine promotes long life bonds, and the number of them is equally divided between DLPC and DLPG residues, despite these last are way less in number. Tryptophan promotes a few persistent bonds, but in less measure (about 20-25% of the ones involving Arginine residues). The remaining ones persisting more than 40% of the time involve Threonine residues. These interactions, together with the electrostatic coupling between positive residues and DLPG, bond some lipids to remain close to the peptide, slowing down their dynamics which in turn influence the one of the molecule nearby.

We think that this reduced “fluidity” is an important perturbation of the membrane structure, as it allows for less rearrangements in case of further external stimuli.

Electroporation results As it is not possible to witness penetration of the peptide through the membrane within the available simulation time, we opted

DLPC			DLPG		Any		
	Region	Nr.	D ($\mu\text{m}^2/\text{s}$)	Nr.	D ($\mu\text{m}^2/\text{s}$)	Nr.	D ($\mu\text{m}^2/\text{s}$)
Peptide	$d < 1$	93	0.248(3)	38	0.212(3)	131	0.239(3)
	$d < 2$	152	0.278(3)	52	0.288(5)	204	0.280(3)
	$d < 3$	239	0.353(3)	70	0.318(3)	309	0.344(3)
	$d > 3$	329	0.408(3)	98	0.638(3)	427	0.465(3)
Apo	All	568	1.213(5)	169	1.475(8)	737	1.278(6)

Table 3.3: Diffusion coefficients of lipids in a 740 lipids bacterial patch, simulated in presence of the pentagonal peptide subunit. The values are computed for groups of lipids which, at the initial time, were within 1 nm, 2 nm, 3 nm from the peptide or further away. Values are split for the two lipid species involved. The values for the patch simulated without the peptide are reported as reference. Error from linear fit in parenthesis.

E=0	Total	$\tau \geq 40\%$	E=20	Total	$\tau \geq 40\%$
All	2041	62	All	2238	75
	R W	R W		R W	R W
DLPC	668 73	25 8	DLPC	798 90	26 8
DLPG	928 114	13 3	DLPG	980 104	19 5

Table 3.4: Number of hydrogen bonds between the pentagonal peptide sub-unit and lipids in simulations of the 740 lipid bacterial patch with the sub-unit and no electric field (left) or a field with value 20 mV/nm (right). Values are divided by the total number counted and the number of those present for more than 40% of the simulated time. The total peptide-lipid hydrogen bonds and the one between each lipid species and either Arginine or Tryptophan residues are given.

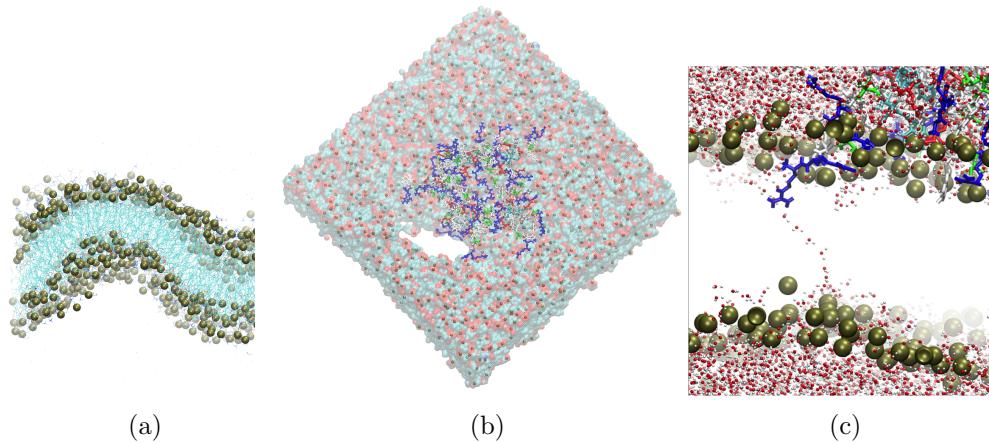


Figure 3.18: Panel a: Membrane deformation due to an electric field applied to the membrane (512 lipids, $E = -130 \text{ mV/nm}$). Panel b: pore formation due to the action of the peptide. Panel c: pore precursor due to Arginine insertion and water penetration.

to perform electroporation simulations. An electric field of increasing intensity was applied along the negative z direction perpendicular to the membrane, with the peptide on the positive z side, to model the field generated by the transmembrane potential.

The field was increased of 20 mV/nm every 200 ns (or 10 mV/nm when approaching the poration threshold), showing that the critical value of 130 mV/nm triggered poration in presence of the peptide. This was confirmed by three replicas run from the initial unperturbed membrane configuration with a 130 mV/nm field, all resulting in poration after 20 ns, 75 ns and 71 ns respectively (Figure 3.18(b)). Similarly, for simulations with the same field on a larger patch (740 lipids) poration occurs at 60 ns, 50 ns and 70 ns respectively. As a control, a 512 lipids bacterial membrane is simulated under the same conditions, for 600 ns in three replicates: despite a reduced diffusion coefficient and the appearance of curved regions, no poration was observed (Figure 3.18(a)). The appearance of a curvature made necessary to compute the area per lipid taking it into account as explained in Section 3.3. The three replicas run at $E = 130 \text{ mV/nm}$ provided values of 0.520 nm^2 , 0.514 nm^2 and 0.550 nm^2 , and their discrepancy is due to the different shapes the membrane adopts during the run (because of that the values were computed over the last 10 ns only). Indeed, once a deformation appears it can be quickly enhanced by the

electric field and small casual variations in its initial phases can bring to very different outcomes.

The electroporation threshold for the membrane alone was set at the higher value of 140 mV/nm: out of three simulations run with such value of the electric field, two resulted in disruption around time equal to 160 ns and –, while another presented a curved but still intact membrane after 200 ns.

This procedure of verifying the conditions promoting poration was performed on the 512 lipids patch. The poration mechanism, both with and without peptide, is preceded by a high deformation (curvature) of the membrane. However this does not happen with the large patch, for which simulations in presence of the peptide and an electric field of 130 mV/nm lead to poration, but this was initiated in a flatter membrane conformation (Figure 3.18(b)). To investigate whether this is due to the different size, or rather the force field used (54a7 versus 54a8), we run an additional control simulation of the 512 lipids patch with the 54a8 force field and the electric field set at the electroporation threshold of 140 mV/nm, as this would promote a quick insurgence of the curvature, if any. The patch developed a curved shape similar to the outcome from the corresponding 54a7 simulations. The membrane was not electroporated in the 200 ns run, as was observed in one out of the three replicas of the 54a7 simulations. Remarkably, the ApL values of these two simulations are quite similar (0.554 versus 0.556 nm² for 54a7 and 54a8 respectively, computed taking into account the curvature), despite the differences outlined generally between the two force fields (see Chapter 4). However, from the other electroporation simulations it emerges that there is a high variability in them, so that the consistency of the two values might be fortuitous.

This proves that the size of the patch has the largest influence in the outcome of the simulations.

MEASURE CURVATURE?

Regardless the shape the membrane assumes before disruption, the peptide speeds up the collapse mechanism as the charged Arginine residues insert into the membrane core interacting with the negatively charged phosphate of the lipids, and promote the penetration of water molecule (Figure 3.18(c)). Additionally, the peptide enhances the invagination of the membrane and these two aspects together (stronger invagination and Arginine insertion) allow pore formation at a value of the electric field lower than the one necessary for elec-

troporation.

This findings, together with the previous results, suggests that the rigidification process (measured by the reduction of the diffusion coefficient), regardless the membrane compactness (measured by the area per lipid), is a proxy for poration as it makes the lipids less able to accommodate perturbations and to seal a forming water channel.

3.5.2 Atomistic simulations of the mammalian model membrane

After the investigation of a model bacterial membrane, we focussed on a mammalian one, modelled as a DLPC patch. The investigation proceeded directly with simulations of a large patch

say 130 rupture apl (larger, error in dlpg) hbond with protein curvature in peptide one?

hb prot (590, 28) hb arg (461, 17) hb trp (57, 5)
0.495(1) diffusion

3.5.3 Coarse grain simulations of the buckyball on model membranes

Standard MARTINI simulations Coarse grain simulations allow to model the behaviour of the full capsule interacting with the membrane. At first, simulations with the standard MARTINI model were run, due to their higher computational speed (the full system comprising a 2880 lipids patch and the capsule measures approximately 30 nm along each side of the simulation box).

Coarse-grain simulations of the full buckyball on a DLPC:DLPG membrane (3:1 ratio) confirms the binding of the peptide on the latter, driven from charge-charge recognition: in both the replicas run, the peptide approaches the membrane after about 2 μ s, remaining bound up to the 10 μ s simulated. Post-binding, the capsule diffuses on the membrane and produces an increasingly high curvature on it, in a process which tends to maximize the contact area (see SI Movie –). No poration is observed, probably due to the force field characteristics which stabilise the structure of both the membrane and the peptide assembly. Additionally, longer time scales might be needed to observe poration. Instead, we opted for an electroporation experiment, in line with what done in the atomistic case (see next paragraph).

A quantitative analysis on the membrane patch shows that the lipids diffusion coefficient is decreased after the binding, by 9% for DLPC and by 15% for DLPG (see SI Table 3.5 and SI Figure ??). Moreover, as DLPG is negatively charged, it is recruited around the peptide and remain bound to it, diffusing much slower than the surrounding lipids.

	MARTINI		Polar MARTINI		
	DLPC	DLPG	DLPC	DLPG	
Bacterial	Pre	310.0(4)	302.6(5)	E 0	307(2)
	Post	267.6(2)	245.5(2)	E-20	264.7(7)
	Protein, post				
Mamm.	DLPC		DLPC		
	Any	308.0(3)	E 0	349.5(3)	
			E-20	278.4(7)	

Table 3.5: Analysis of lipid diffusion in a coarse-grain simulation of the buckyball binding to a model bacterial membrane (DLPC:DLPG 3:1). Data for Replica 1, binding time 2.14 μ s. Protein lateral diffusion after binding to the membrane 16.50(2) $\mu\text{m}^2/\text{s}$. ^a Number of lipids within 2 nm distance from the buckyball in the last 5 μ s of simulations. ^b In a simplistic model in which the dumping of lipid lateral diffusion coefficient is due to some lipids bound to the buckyball, and thus they diffuse at its pace, while the other are free to move as before, it is valid the following equation: $D_{post} = D_{pre} \cdot (N_{lipid} - N_{bound}) + D_{protein} \cdot N_{bound}$.

prot diffusion on bact membrane after binding martini non polar counterions 28.4(2) martini polar mix 177(2) martini polar E20 53.3(6)

A control simulations with a pure DLPC membrane shows no binding of the peptide: the distance of the icosahedron from the membrane is on average 3 nm and never smaller than 1 nm. This does not exclude a binding on longer time scales, but hints at the fact that, not being driven by the opposite charge recognition, the peptide-membrane interaction is weaker and less likely to happen.

3.6 Rest

Electroporation simulations						
	Lipids	Model	Time (ns)	Rep.	E (mV/nm)	
Bacterial	512	GR (54a7)	400	1	0, 70, 120	
	512	GR (54a7)	600	3	130	
	512	GR (54a7)	150 ^{EP} , 154 ^{EP} , 200	3	140	
	512	GR (54a8)	200	1	140	
	740	GR (54a8)	400 (350,finish & redo prop)	1	0	
	740	GR (54a8)	400	1	20	
	720	MA	1000	1	0	
Mamm.	2880	MA_P	500	1	20, 40	
	748	GR (54a8)	400	1	0	
	748	GR (54a8)	20 ^{EP} , 28 ^{EP} , 39 ^{EP}	3	130	
	722 (54a8)	MA	1000	1	0	
	2888	MA_P	500	1	20, 40	

Table 3.6: Table of control simulations of membranes complexes. All run at 150 mM concentration of NaCl. Models: GR = united atom GROMOS, MA = coarse-grain MARTINI, MA_P = coarse-grain MARTINI with polar water. Superscript EP refers to the rupture of the membrane.

3.7 Additional material

No peptide		Bacterial	Mamm.
E = 0	512	54a7 (1)	54a7 no salt (1)
	740	54a8 (1)	54a8 (1R)
E = 130	512	54a7 (3)	–
	740	–	54a8 (1)
+ Bact/512/E-140: 54a8 (1R)			
Peptide		Bacterial	Mamm.
E = 0	512	54a7 (1)	–
	740	54a8 (1)	54a8 (1R)
E = 130	512	54a7 (3)	–
	740	54a8 (3)	54a8 (3)

Table 3.7: Table of the simulations run for different systems. Electric field expressed in mV/nm. For each cell, the syntax reports: Nr. of lipids *Force field* (Nr. of replicas).

Chapter 4

Lipid parametrisation

SIMULATIONS of lipids can be a challenging task, especially when the membrane simulated has not been tested experimentally, and thus no comparison can be performed on standard properties such as the area per lipid. Moreover, the parameters describing each lipid species must be carefully chosen, to aim at an accurate reproduction of the natural properties of the membrane.

The work performed in Chapter 3 involved the simulation of a mixed membrane which, to our knowledge, has not been experimentally tested. Additionally, one of the lipids involved (DLPG) had not been parametrised before for the GROMOS force field. The procedure of DLPG parametrisation brought to our attention the discrepancy existing between the GROMOS parametrisation of protein and the one employed for lipids, as the two of them are obtained through different procedures: fit to hydration properties for proteins, and quantum mechanics computations for lipids.

The set of lipid parameters derived originally from Chiu et al. [1995] and updated multiple times up to the work of Poger et al. [2010], is still regarded as a standard for phosphocholines, and thus used to validate the most recent parametrisation of the force field (GROMOS 54a8 [Reif et al., 2012]). However, it does not take into account recent evolutions of the force field, aimed at reparametrising some constitutive moieties which happen to be also part of lipids. For this reason, we performed a reparametrisation of the lipid head group which includes these new descriptions, aiming at an improved consistency with the rest of the force field and monitoring whether they would also improve the agreement with the experimental quantities.

The resulting parametrisation was successful in reproducing key proper-

ties such as the area per lipid, and proposed a different interaction with a test peptide with respect to the one outlines by the previous parameter set. Specifically, the new parameters promote a weaker protein-membrane interaction allowing simulations to be less biased from the initial conditions chosen. This is of crucial importance for all simulations in general, and in particular when the sampling is reduced due to computational resources issues.

This chapter consists of the paper “Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description”, published in the *Journal of Chemical Theory and Computation*, together with its Supporting Information. The project was conceived by Prof. Franca Fraternali, Dr. Christian Margreitter and myself. I carried out all the work, under their guidance; specifically I implemented the new parameter sets, performed and analysed the simulations and selected the best performing parameters to be released. References for the paper and its supplementary material are included separately after each of them (with a numeric notation). All references feature in the bibliography of this thesis as well.

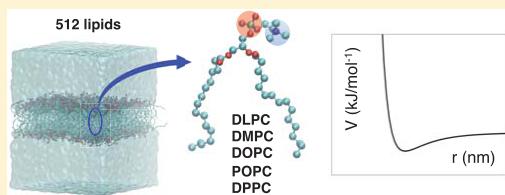
Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description

Irene Marzuoli,¹ Christian Margreitter,¹ and Franca Fraternali^{*1,2}

¹ Randall Centre for Cell and Molecular Biology, King's College London, London SE1 1UL, U.K.

² Supporting Information

ABSTRACT: Membranes are a crucial component of both bacterial and mammalian cells, being involved in signaling, transport, and compartmentalization. This versatility requires a variety of lipid species to tailor the membrane's behavior as needed, increasing the complexity of the system. Molecular dynamics simulations have been successfully applied to study model membranes and their interactions with proteins, elucidating some crucial mechanisms at the atomistic detail and thus complementing experimental techniques. An accurate description of the functional interplay of the diverse membrane components crucially depends on the selected parameters that define the adopted force field. A coherent parameterization for lipids and proteins is therefore needed. In this work, we propose and validate new lipid head group parameters for the GROMOS 54A8 force field, making use of recently published parametrizations for key chemical moieties present in lipids. We make use additionally of a new canonical set of partial charges for lipids, chosen to be consistent with the parameterization of soluble molecules such as proteins. We test the derived parameters on five phosphocholine model bilayers, composed of lipid patches four times larger than the ones used in previous studies, and run 500 ns long simulations of each system. Reproduction of experimental data like area per lipid and deuterium order parameters is good and comparable with previous parameterizations, as well as the description of liquid crystal to gel-phase transition. On the other hand, the orientational behavior of the head groups is more realistic for this new parameter set, and this can be crucial in the description of interactions with other polar molecules. For that reason, we tested the interaction of the antimicrobial peptide lactoferricin with two model membranes showing that the new parameters lead to a weaker peptide–membrane binding and give a more realistic outcome in comparing binding to antimicrobial versus mammal membranes.



1. INTRODUCTION

Cellular membranes are key promoters and regulators of many biological processes due to their crucial role in segregating the external world from the organism. Small molecule transport, drug permeation, intracellular signaling, and antibody response are all regulated by the cell membrane or by membrane-related components.^{1–8} To fully comprehend and ultimately influence the bespoke processes, it is paramount to understand membranes and their constituting lipids in atomistic detail. However, due to the complexity of those systems, researchers have resorted to the use of simplified model membranes, which can be synthesized and characterized *in vitro*. This enables the individual contributions of the components involved to be disentangled. Indeed, for the cellular membrane to be able to perform different functions, its composition is necessarily complex. Lipids are one of the main components and can be present in up to hundreds of different species.⁹ In addition, many transmembrane proteins tessellate the cell surface, promoting signaling pathways and influencing the membrane's structural and mechanical properties.^{10,11} Phospholipid bilayers and micelles have been investigated, in particular, as these lipids represent the main components of the eukaryotic and the inner bacterial membranes. Both have been modeled selecting

specific phospholipids to emulate the appropriate surface charge or to reproduce the human cell membrane fluidity by introducing, for example, cholesterol.^{12,13} As these simplified membranes retain the core characteristics of their different biological templates,¹⁴ they can be used to test the membrane interaction with proteins, peptides, antimicrobial molecules, or drugs.

Experiments can provide global properties of membranes and, despite the great accuracy of techniques like NMR and X-ray scattering in measuring the average position of atoms in rigid structures, they face challenges when characterizing the biologically relevant fluid phase, as opposed to the gel one that emerges at lower temperatures.^{15–18} Alongside experimental characterization, molecular dynamic (MD) simulations have played a central role in the investigation of the behavior of lipids, due to the atomistic spatial resolution they provide. Therefore, MD simulations complement our understanding of membranes' behavior and are also important for the study of lipid systems in combination with proteins, providing detailed insights into the mechanisms of their interactions. In the past,

Received: May 25, 2019

Published: August 21, 2019

MD simulations have been successfully employed to reproduce typical phenomena in membranes, such as lipids' flip-flop,^{19,20} vesicle formation,^{21,22} aggregation into bilayers,^{23–25} and stress-induced^{26–29} and peptide-induced pore formations.^{30–32} Moreover, the implementation of more realistic models of bacterial membranes, by including a more diverse set of components into the simulated systems, has been pursued^{33,34} to test specific interactions with antimicrobial peptides and understand their selectivity.^{35,36}

The reliability of such simulations depends on the accurate parameterizations of lipids and proteins, which need to be validated against experimental data. Moreover, the two descriptions must be consistently integrated into the force fields used, i.e., be derived with the same parameterization procedure. Different approaches to the problem are possible, which resulted in the development of multiple force fields suitable for simulations of biomolecules: for example, the CHARMM^{37–39} and AMBER⁴⁰ force fields are parameterized from quantum mechanics calculations, while GROMOS96⁴¹ is calibrated to match global properties like the hydration free energy of chemical moieties. All of them have been constantly updated to meet the new experimental values available and more faithfully reproduce the different species involved.

However, it is a very difficult task to parameterize the constituents of a complex system so that all parameters are consistent with the rest of the force field and reproduce both the single-molecule observables and the collective behavior. In the present work, we consider the parameterization of phospholipids in the context of the GROMOS96 force field,⁴¹ addressing some of the inconsistencies in the lipid head group parameters commonly used so far, particularly in consideration that these contribute to the description of recognition processes at the interface.

In the past, lipid simulations using the GROMOS96 force field suffered from difficulties involved in transferring the pre-existing parameters, calibrated mainly for peptides in an aqueous environment, to the amphiphilic environment of the lipid assembly. This resulted in the failure to reproduce the membranes' behavior properly^{42–44} and therefore a series of modifications were adopted, particularly in the choice of lipid-specific Lennard-Jones interactions^{44–46} and partial charges.⁴²

In the light of recent reparameterizations of a set of choline moieties⁴⁷ and of phosphate-containing species,⁴⁸ we undertake the task of updating the parameters used for lipids, in particular, phosphocholines, as they contain both these chemical moieties. Within this work, we show that it is possible to integrate the recently computed partial charges within simulations while maintaining good agreement with the available experimental data. We also test the transferability of the new phosphate charges onto lipids without a choline head group, namely, phosphoethanolamine (POPE) and phosphatidylglycerol (POPG).

Most importantly, the new description of phosphocholine is consistent with the GROMOS96 parameterization philosophy, based on the decomposition of large molecules into smaller compounds and subsequently fitting their parameters to experimental hydration free energies. Together with adjustments to specific van der Waals potentials, we believe that the parameters presented here will contribute to improving the accuracy of the description of membrane–solvent and membrane–protein interactions. To this aim, we compared the available parameters with the one proposed in this work, simulating the interaction of an antimicrobial peptide with two

model membranes, highlighting the differences in the mechanisms observed, and comparing them with the available experimental evidence.

2. METHODS

2.1. Background to Lipid Force Fields. The most recent iteration of the lipids' parameters commonly used in simulations with the GROMOS force field is the one by Poger and Mark.⁴⁴ They employed partial charges derived quantum-mechanically by Chiu et al.,⁴² combined with a modified repulsion between the choline methyl groups and the OM oxygen atoms in the phosphate with respect to the standard choline–OM one.

The original set of Chiu charges⁴² was derived from ab initio Hartree–Fock self-consistent field calculations⁵² and Mulliken population analysis.⁵³ Slight modifications were applied to make each individual charge group sum up to an integer value, following the GROMOS96 philosophy. Despite the resulting charges that differ substantially from the ones used for the same chemical groups in different chemical contexts, the GROMOS community employed this set as it gave results in closer agreement with the available experimental data.

The refinement of van der Waals parameters for aliphatic alkanes, together with the bond, bond angle, and torsional parameters for the ester groups,^{49,50} prompted the reparameterization of the lipids' head group description; in particular, Chandrasekhar et al.⁴⁵ recomputed the head group torsional parameters from ab initio quantum-mechanical torsional profiles of each of the fragments composing the head group (Figure 1).

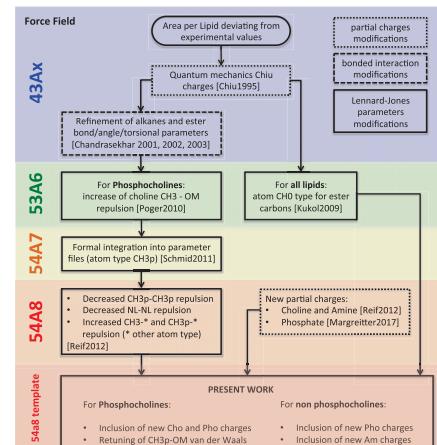


Figure 1. Evolution of the lipid parameters in the GROMOS force field. References—Chiu1995: ref 42; Chandrasekhar2001, 2002, 2003: refs 45, 49, 50; Poger2010: ref 44; Kukol2009: ref 46; Schmid2011: ref 51; Reif2012: ref 47; Margreitter2017: ref 48.

The modifications above were included in the 53A6 version of the GROMOS force field. However, additional change was necessary to match the experimental results. Therefore, Poger and Mark⁴⁴ introduced a change in the CH3 choline and OM repulsion. The C12 parameter (related to the Pauli repulsion) between the newly introduced atom-type CH3p (to represent

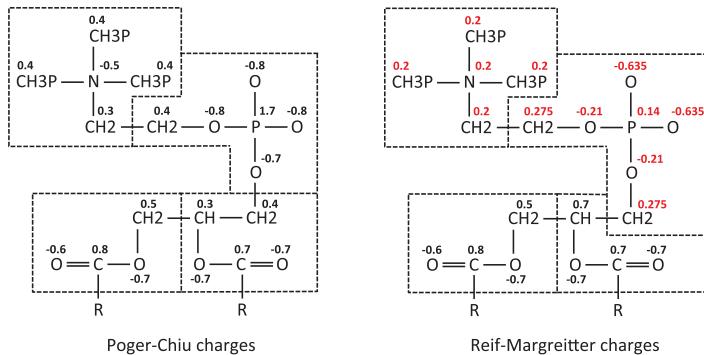


Figure 2. Partial charges for the phosphocholine head groups and the glycerol and ester moieties in the Chiu⁴² scheme (left) and the one tested in the current work (right). Red font denotes values that have been changed between the two. Atoms belonging to the same charge groups are enclosed by the same dashed polygon.

Table 1. Table of Simulations for Phosphocholine Bilayers^a

simulations of phosphocholine lipids				
sim	charges ^b	FF	CH3p-OM C12 ^c (kJ mol ⁻¹ nm ¹²)	CH3p-CH3p C12 ^d (kJ mol ⁻¹ nm ¹²)
1	Chiu	S4A7	1.58×10^{-5}	2.66×10^{-5}
2	Chiu	S4A8	6.93×10^{-6}	6.48×10^{-5}
3	RM	S4A8_v1	1.10×10^{-5}	6.48×10^{-5}
4	RM	S4A8_v2	1.58×10^{-5}	6.48×10^{-5}
5	RM	S4A8_v3	4.50×10^{-5}	6.48×10^{-5}

^aAll are run for 500 ns and systems consisting of 500 lipid molecules (256 per layer), using a particle mesh Ewald (PME) long-range electrostatic scheme. ^bCharge set: Chiu from Ref 42, Reif–Margreitter (RM) as illustrated in the present work. ^cAs a reference, the standard C12 parameter in S4A7/S4A8 for CH3-OM is 4.44×10^{-6} kJ mol⁻¹ nm¹². ^dThe CH3-CH3 C12 parameters are 2.66×10^{-5} for each parameter set.

the united methyl atoms in the choline head group of lipids) and the oxygen-type OM, present in the phosphate group, was increased by a factor of 3.5. This modification was optimized and tested against experimental values, increases the spacing between individual lipids, and thus leads to the appropriate area per lipid (ApL).^{51,54} The new atom-type CH3p has all of the characteristics of CH3, except for the bespoke parameter, i.e., all Lennard-Jones interactions involving OM. These Poger–Chiu parameters have been successful in reproducing membrane behavior and were used in many MD applications.^{33,55}

Later, Reif et al.⁴⁷ enhanced the methyl–methyl repulsion for both CH3 and CH3p in the S4A8 parameter set, which allowed for a decrease in the large repulsion value between CH3p and OM previously introduced⁴⁴ while still reproducing experimental values. The S4A8 parameter set contains two additional, nonlipid-specific, modifications important for this work: the choline Lennard-Jones parameters and partial charges, and the phosphate partial charges. The C12 Lennard-Jones repulsion term for the NL nitrogen atom type (present in the choline moiety) was increased to successfully prevent oversolvation.^{47,54} To the same end, the +1 e total charge was evenly distributed over all five atoms, which resulted in a better approximation of the experimentally obtained hydration free energy in comparison to the S4A7 parameter set. Similarly, Margreitter et al.⁴⁸ calibrated the partial charges of four phosphate species and enhanced the reproduction of experimental data. The relevant phosphate-containing species for this work is dimethyl-phosphate, a

compound not directly present in force field versions prior to S4A8.

Another approach to lipid parameterization was proposed by Kukol,⁴⁶ namely, the use of the already available CH0 atom type for the ester carbons in place of the standard C atom type, in conjunction with the Chiu charges. This atom type, designed to describe a bare sp³ carbon bound to four heavy atoms, has a repulsion energy term 10–40 times larger than a bare carbon bound to other atom types, enforcing a greater spacing between lipid molecules and thereby increasing the ApL. As this modification is also applicable in the absence of a choline head group and does not require the introduction of another atom type, this method can be used to parameterize POPC and POPG.

2.2. Parameterization Strategy. In an effort to enhance the consistency of the force field, we integrated the new partial charges for the choline and phosphate moieties [Reif–Margreitter (RM) charge set] into the lipid building blocks of GROMOS 54A8 so that the entire phosphocholine head group now follows the common GROMOS-like modeling approach (Figure 2). Only the partial charges of the ester groups remain as described in the Chiu set, a deviation from the canonical parameterization strategy necessary to match the experimental area per lipid values: Chandrasekhar et al. showed in ref 56 that the replacement of the ester charges with the standard ones for the ester moiety (parameterized to reproduce the experimental free energies of hydration of a series of alkane esters⁵⁷) resulted in a much smaller area per lipid, not compatible with the experimental values.

The introduction of the new head group charges required a refinement of the CH3p-OM Lennard-Jones repulsion, as the 54A8 value was set considering the original Chiu charges. Ideally, one would always try to keep the force field terms as much transferable as possible. Nevertheless, the complexity and anisotropic nature of some biological environments can be difficult to parametrize with single chemical groups, as the same chemical group can behave differently according to the context it is inserted in. Lipid systems are one of such examples, and to maintain the correct physical behavior of the system, we used specific C12 parameters for the CH3p-OM repulsion in phosphocholine lipid atoms. This allows for a more balanced description of the physicochemical properties of the lipid bilayer and a better match with the available experimental observables.

Aiming at this, and to disentangle the effect of charge parameterization versus the CH3p-OM repulsion, we tested three different values of such Lennard-Jones parameter with the new charges while control simulations were run using the Chiu partial charges and the GROMOS 54A7 or 54A8 parameter set for each lipid (Table 1). The phosphocholines tested are 1,2-lauroyl-*sn*-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC), and 2-oleoyl-1-palmitoyl-*sn*-glycero-3-phosphocholine (POPC), which have different tail lengths and numbers of unsaturated bonds as in previous works^{44,58} (Table 2).

Table 2. Details of the Systems Simulated: Lipid Name, Tail Composition, Initial ApL (and Reference from Which the Initial Coordinates Are Taken), Simulation Temperature, and Gel–Liquid Phase Experimental Transition Temperature

lipid bilayer systems				
lipid ^a	tails ^b	ApL ₀ (nm ²)	T _{MD} (K)	T _C (K)
DLPC	12:0/12:0	0.632 ⁴⁴	303	276.4 ^{53–66}
DMPC	14:0/14:0	0.616 ⁵⁸	303	296.9 ^{53–66}
DOPC	18:1c9/18:1c9	0.649 ⁵⁸	303	255.7 ^{53–66}
POPC	16:0/18:1c9	0.638 ⁵⁸	303	270.5 ^{53–66}
DPPC	16:0/16:0	0.631 ⁵⁸	323	314.2 ^{53–66}
POPE	16:0/18:1c9	0.568 ³³	313	299.3 ⁵⁷
POPG	16:0/18:1c9	0.602 ³³	303	268.1 ⁶⁸

^aDLPC: 1,2-lauroyl-*sn*-glycero-3-phosphocholine, DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine, DOPC: 1,2-dioleoyl-*sn*-glycero-3-phosphocholine, POPC: 2-oleoyl-1-palmitoyl-*sn*-glycero-3-phosphocholine, DPPC: 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine, POPE: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine, POPG: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-rac-glycerol). ^bExample: 16:0/18:1c9 indicates that tail 1 has 16 carbons with no unsaturated bonds and tail 2 has 18 carbons with one unsaturated bond between carbons 9 and 10—ester carbon counts as number 1.

To prove the transferability of the new phosphate charges to other lipid species, which do not contain a choline head group (and thus an enhanced repulsion, which has an impact on the ApL), test simulations of a phosphoethanolamine (POPE) and a phosphoglycerol (POPG, Table 2) bilayer have been performed. These lipids have amine and glycerol head groups, respectively. The parameterization of both takes advantage of the Kukol approach⁴⁶ employing a CH0 atom for the ester moieties to enhance the repulsion between lipids. For POPE

and POPG, simulations were run with the standard parameters from ref 33 (denoted as Piggot–Chiu in the present work) or with the updated phosphate partial charges (Supporting Information (SI) Table 2).

The evolution of simulation techniques seen in the recent years suggested two other changes in the simulation setup: first, the original set of parameters was designed to be used with a twin-range cutoff scheme and a reaction field long-range electrostatic contribution,⁵⁹ but the twin-range cutoff is no longer supported in the latest versions of the GROMACS software used for the present work.⁶⁰ Additionally, the PME algorithm⁶¹ for long-range electrostatic treatment is currently the predominant method used for protein dynamics. In the context of unifying the two fields of protein and lipid simulations, we therefore opted for a PME long-range treatment, running a control simulation (on the DPPC bilayer) with a reaction field scheme to assess the impact of such a change (SI Table 1).

The other change we adopted in comparison to the earlier work was a larger system size. Due to computational limitations, the original parameterization was performed on a 128-lipid bilayer,⁴⁴ but recent advances allow for larger systems to be simulated and we therefore used membranes four times as large (512 lipids). This larger size allows to track larger undulations of the membrane, as the effect of periodic boundary conditions (PBCs) is less restrictive. Again, a control simulation on a 128 DPPC membrane has been run to test the relevance and the effect of this change (SI Table 1).

Finally, the improvements reached with the adoption of the new parameters are monitored through the comparison with experimental values, but it is useful to have benchmarks derived from other simulation experiments. For that purpose, we compare some key properties with the values obtained by the all-atom CHARMM36 force field.^{37,62} Despite a thorough comparison beyond the present work, it is relevant to observe whether the changes introduced by the new parameters are going in the direction of the outcomes proposed by other descriptions.

2.3. Simulation Systems. Seven pure lipid bilayers have been simulated, five of which contain phosphocholines, one phosphoethanolamine (POPE), and one phosphoglycerol (POPG), as described in Table 2. Every bilayer is formed by 512 lipids (256 per leaflet), generated by replicating an equilibrated 128-lipid system from the literature two times in the *x* and *y* directions (see Table 2).

Water molecules were added to reach a minimum distance of 7.5 nm between periodic copies of the membrane along the *z*-direction, with a ratio of 85–120 H₂O per lipid. This distance is larger than the one used in the previous parameterization publications because we observed an enhanced undulatory behavior for larger membranes and therefore a higher distance is necessary to avoid interactions between periodic replicas in the *z*-direction.

2.4. Simulation Parameters. All simulations were run using the GROMACS software version 2016.3,^{60,69,70} under periodic boundary conditions in a rectangular box. The temperature was maintained by coupling the membrane and the solvent independently to an external bath using the Berendsen thermostat⁷¹ with a coupling time τ_T of 0.1 ps, at the reference temperatures indicated in Table 2, which are above the gel–liquid phase transition temperature for each lipid. The pressure was kept at 1 bar with a semi-isotropic coupling using a Berendsen barostat,⁷¹ applying isothermal

325 compressibility of 4.5×10^{-5} bar $^{-1}$ and a coupling constant τ_p
 326 of 0.5 ps. Covalent bond lengths of the lipids were constrained
 327 using the LINCS algorithm.⁷² The geometry of the simple
 328 point charge water molecules was constrained using
 329 SETTLE.⁷³ A 2 fs time step was used, with a Verlet integration
 330 scheme. The PME⁶¹ long-range treatment was applied to the
 331 electrostatic interactions beyond a 1.4 nm cutoff, and the
 332 reaction field scheme⁵⁹ control simulation was run with the
 333 same cutoff radius. A plain cutoff was used for van der Waals
 334 interactions, with a cutoff radius of 1.4 nm.

335 Each system was initially energy-minimized and then
 336 simulated at 50 K for 10 ps. Subsequently, the temperature
 337 was increased gradually over 500 ps until the final simulation
 338 temperature. The system was then simulated for 500 ns. The
 339 equilibration of the systems was monitored by examining the
 340 time evolution of the potential energy and the area per lipid:
 341 200 ns is found to be sufficient to reach equilibration for all of
 342 the bilayers (SI Figure 2) so that the analysis has been
 343 performed over the last 300 ns of the production run, with
 344 frames stored every 100 ps. An overview of the simulations
 345 performed is given in Table 1 and SI Tables 1 and 2.

346 **2.5. Analysis.** To calibrate the lipid parameters, we used
 347 the observables listed below, as common practice in standard
 348 parameterization procedures.^{58,74}

349 **2.5.1. Area per Lipid.** For systems where the membrane is
 350 aligned to the xy plane, the area per lipid (A_{pL}) can be
 351 computed from the product of the lateral dimensions of the
 352 simulation box divided by the number of lipids in one leaflet.
 353 As shown in SI Figure 2 for DPPC, after 100 ns of simulation,
 354 the A_{pL} oscillates around a value with fluctuations of the same
 355 magnitude, indicating equilibration. To allow further time for
 356 local rearrangements, we restrict our analyses to the last 300 ns
 357 of the simulations.

358 The equilibration protocol was verified on the DPPC
 359 bilayer, repeating the computation of the A_{pL} on two
 360 nonoverlapping time windows, specifically between 200 and
 361 350 ns and between 350 and 500 ns. For all of the parameter
 362 sets, the two windows gave compatible values of the A_{pL} ,
 363 confirming the convergence of the simulations (SI Figure 3).
 364 The above procedure is valid if the membrane is flat or has
 365 minor undulations only. To test this and verify that deviations
 366 from planarity are not influencing the results, the A_{pL} was
 367 recomputed taking into account membrane undulations
 368 according to the procedure outlined in ref 75. The differences
 369 with the values computed from the simulation box dimensions
 370 were between 0.20 and 0.46%, which is lower than the error
 371 derived from the standard deviation across the simulation for
 372 any of the area per lipid computed.

373 As such, our computations are of value in rating the results
 374 against experimental outcomes and/or to compare parameter
 375 sets, as a local measure would not significantly improve the
 376 comparison.

377 **2.5.2. Isothermal Area Compressibility Module.** Following
 378 the protocol in ref 58, we computed the isothermal area
 379 compressibility module (K_A) from the fluctuations of the A_{pL}
 380 values according to

$$K_A = \frac{k_B \langle T \rangle \langle A_{pL} \rangle}{n_L \sigma_{A_{pL}}^2} \quad (1)$$

382 where k_B is the Boltzmann constant, $\langle T \rangle$ and $\langle A_{pL} \rangle$ are the
 383 ensemble averages of the temperature and the area per lipid,

384 respectively, n_L is the number of lipids in one leaflet, and $\sigma_{A_{pL}}^2$
 385 is the variance of A_{pL} .

386 **2.5.3. Bilayer Thickness.** From the electron density profiles,
 387 the bilayer thickness can be evaluated in several ways and
 388 compared to the values from X-ray scattering experiments: the
 389 hydrophobic thickness (D_{HH}) is measured as the distance
 390 between the phosphorus peaks in the two layers, as these
 391 atoms have the highest electron density, while the Luzzati
 392 thickness (D_B)⁵⁸ is defined as

$$D_B = b_z - \int_{-b_z/2}^{+b_z/2} \rho_W(z) dz \quad (2)$$

393 where b_z is the z -dimension of the simulation box and $\rho_W(z)$ is
 394 the volume fraction of water (vs other components) along z
 395 and normalized to 1 in the bulk water region

$$\rho_W(z) = \frac{n_W(z)V_W}{dV} \quad (3)$$

396 where $n_W(z)$ is the time-averaged number of water molecules
 397 in a bin of width dz , V_W is the specific volume of the water
 398 model used (taken from ref 76), and dV is the time-averaged
 399 volume of a slice.

400 **2.5.4. Dipole Potential.** The dipole potential along the z -
 401 direction (perpendicular to the membrane plane) can be
 402 computed from the charge density along z ($\rho(z)$) via a double
 403 integration⁷⁷

$$\psi(z) = -\frac{1}{\epsilon_0} \int_{z_0}^z \int_{z_0}^{z'} \rho(z'') dz'' dz' \quad (4)$$

404 Several choices are possible for the two integration constants,⁷⁸
 405 and for the present work, they are selected to set the dipole
 406 potential to zero in the middle of the bulk water region, at both
 407 sides of the membrane.

408 **2.5.5. Deuterium Order Parameter of Lipid Chains.** The
 409 deuterium order parameters S_{CD} of the acyl chains for each
 410 lipid bilayer were calculated and compared between the
 411 different sets studied. S_{CD} evaluates the average order of the
 412 lipid tails by measuring the orientation with respect to the
 413 bilayer normal of a carbon–hydrogen bond in a given position
 414 along the chain for each lipid in the bilayer. Their spread is
 415 evaluated according to the ensemble average

$$S_{CD} = \frac{1}{2} (3 \cos^2 \theta - 1) \quad (5)$$

416 As the GROMOS force field employs a united-atom
 417 representation, the tetrahedral positions of the hydrogens are
 418 constructed based on the neighboring carbons' positions.

419 **2.5.6. Hydration of Head Groups.** To estimate and
 420 compare the hydration of lipid molecules, we computed the
 421 distribution of the distances between the oxygen of water and
 422 the nearest lipid atom. For each individual chemical group, the
 423 distance between the water oxygen and the nearest atom
 424 within that group was calculated. A quantitative measure for
 425 hydration was obtained by integrating the distribution up to
 426 the first peak or second one (for phosphate and glycerol).

427 **2.5.7. Orientation of Head Groups.** We computed the
 428 orientation of the lipids' head groups as the distribution of the
 429 angle between the P–N vector (joining the phosphorus atom
 430 and the choline nitrogen) and the outward normal to the
 431 membrane. The orientation of the *sn*-1 and -2 carbonyl dipoles
 432 with respect to the bilayer normal has also been calculated.

Table 3. Average Area per Lipid (in nm²) over the Last 300 ns of Simulations for Phosphocholine Bilayers^a

ID	charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	0.608(4)	0.591(4)	0.600(5)	0.604(4)	0.616(4)
2	Chiu/54A8	0.626(5)	0.612(4)	0.623(4)	0.623(4)	0.635(5)
3	RM/54A8_v1	0.631(4)	0.616(5)	0.625(6)	0.629(5)	0.638(5)
4	RM/54A8_v2	0.652(5)	0.643(6)	0.649(5)	0.650(5)	0.657(5)
5	RM/54A8_v3	0.690(6)	0.684(6)	0.690(6)	0.687(6)	0.693(5)
0	RM/54A7					0.603(5)
RF	Chiu/54A7					0.603(4)
small	Chiu/54A7					0.594(11)
experimental ^b		0.608–0.632	0.589–0.660	0.674–0.725	0.643–0.683	0.570–0.717
CHARMM36 ³⁷		0.644(4)	0.608(2)	0.690(3)	0.647(2)	0.629(3)

^aThe number in parentheses is the standard deviation of the last digit. All simulations are run at 303 K, except for DPPC (run at 323 K). Analogous values for POPE and POPG are reported in SI Tables 6 and 8. ^bWe report the maximum and minimum values of a review of experimental results given in Table 1 of ref 74. Only values referring to the temperature simulated are considered.

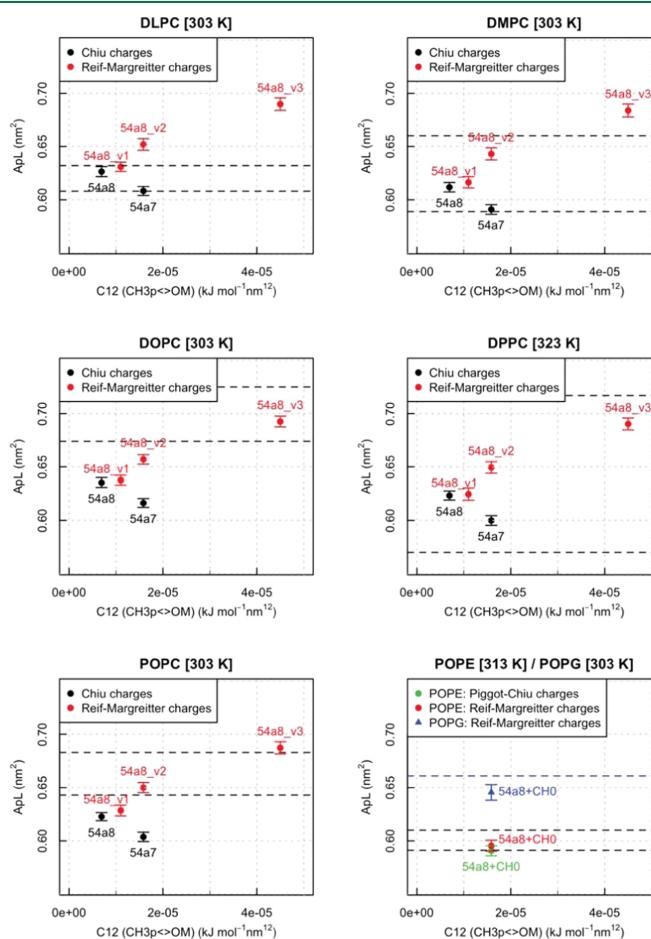


Figure 3. Area per lipid obtained for the five sets of parameters and seven lipid species. Error bars are the standard deviation over the 300 ns analyzed. Dashed lines indicate the range of experimental values from Table 1 in refs 74 and 33. For the plot reporting POPE and POPG values, the black dashed lines refer to POPE and the blue one to POPG.

2.5.8. Lateral Diffusion. For each simulation, we extracted the trajectory of the phosphorus atom of every lipid in the top and bottom leaflets separately, removing the collective motion of the leaflet. These trajectories were used to compute the mean-square displacement (MSD) for each lipid as a function of time, discarding the first 200 ns of production. This figure was averaged over all of the lipids in the leaflet and, for a given interval of time, on all of the possible time windows of that length-fitting within the simulation time analyzed. The diffusion coefficient D was obtained from a linear fit of the average MSD profile, following the Einstein equation⁷⁹ in two dimensions

$$\langle (x - x_0)^2 \rangle = 4Dt \quad (6)$$

The fit was performed discarding the first 50 ns of the profile, where the behavior is not linear, and the last 100 ns, where the poorer statistics leads to more noisy data. Coefficients obtained for the two leaflets were averaged to give the value reported.

2.5.9. Tilt Modulus. We computed the tilt modulus following the theoretical framework explained in ref 80. According to this, the angle θ a lipid forms with the local normal to the membrane follows the distribution

$$P(\theta) = C \sin(\theta) \exp\left(-\frac{\kappa_t^1 \theta^2}{2k_B T}\right) \quad (7)$$

where C is a normalization constant, k_B is the Boltzmann constant, T is the temperature, and κ_t^1 is the tilt modulus. This can be extracted from a fit of the distribution or, for computational reasons, from a fit of $\ln(P(\theta)(\sin \theta)^{-1})$. The direction in which a lipid points is taken as the vector joining the center of mass of the terminal atoms of the tails and the center of mass of selected atoms in the head group. Specifically, the last three carbons of each tail are taken as the reference for the first group, and the phosphorus and the carbon from which the two tails divert for the second. The computation was performed using a dedicated python module⁸⁰ available on the openStructure platform.⁸¹

2.6. Phase Transition. The set of new parameters performing best according to the previous observables was tested for sensitivity to temperature variations. A DPPC bilayer was chosen as the reference system and simulated at two additional temperatures: 303 and 333 K (SI Table 1), the first of which is below the experimentally determined liquid to gel-phase transition temperature.^{63–66} As DPPC has also been used to perform the other control simulations, we opted for this model membrane for consistency reasons.

Besides the standard analysis described before, the local area per lipid was computed using a Dirichlet tessellation⁸² of the lipid head positions projected onto the horizontal plane parallel to the membrane (one leaflet at the time). The tessellation divides the plane into polygons, each enclosing one head position. Every polygon comprises the locations on the plane, which are closer to the position of the head enclosed by that polygon than to any other heads.

Moreover, to quantify whether and how many lipids undergo face transition during the simulations, the regular packing of each of their chains was quantified by the hexagonal order parameter S_6 , as previously reported in the literature.⁸³ Specifically, a chain was represented by its position on the xy plane (parallel to the membrane surface), computed as the average x and y positions of its carbon atoms. For each chain, j ,

the set of neighboring chains was defined as the ones within a 0.65 nm radius from j . Then, S_6 is defined as

$$S_{6,j} = \frac{1}{6} \left| \sum_k e^{6i\theta_{jk}} \right| \quad (8)$$

with θ_{jk} being the angle between the vector connecting j and k and the x axis (i is the imaginary unit). A chain is in the gel phase if it has an hexagonal order parameter larger than 0.72.⁸³

3. RESULTS AND DISCUSSION

In general, the parameters described in this work are shown to reproduce the available experimental target values well while, at the same time, are likely to allow for a better description of lipid–protein interactions, since the head groups are updated to the recent GROMOS force field.

3.1. Area per Lipid and Isothermal Area Compressibility Module. We report in Table 3 and Figure 3 the values of ApL for the simulation run. From such computations, it emerges that the increase of the CH3p–OM repulsion has a nonlinear effect on the area per lipid, as reported in ref 54. On the contrary, the comparison between simulation ID1 and the control ID0 for DPPC, which differ only in their partial charges, shows an almost identical ApL value (SI Figure 4). This suggests that the charge redistribution in the head group affects the global structure of the bilayer and the ApL less dramatically than the adopted value for the Lennard-Jones repulsion.

The comparison with the control simulation using a smaller membrane shows that larger systems allow for the evaluation of the ApL with a smaller error, as local fluctuations are averaged over a larger number of lipids. The standard deviation computed for the 128-lipid system is compatible with those reported in both the original⁵⁸ and a more recent publication,⁸⁴ in which the same system size was used.

The ApL from the simulation with a reaction field treatment for the long-range electrostatic term does not differ significantly from the one obtained with a PME treatment, being only slightly higher, which is in consistence with what was found in ref 85.

Finally, the values found using the Chiu charge set and the S4A7 force field (ID1 in Table 1) are systematically lower than those obtained in the original publications,^{44,58} despite employing the same charge set and force field, while a better agreement is shown with those obtained more recently by Reif et al. for DPPC.⁵⁴ We attribute this to the different versions of GROMACS used, as the integration algorithm has recently been updated, affecting the calculated properties. Moreover, the double-cutoff scheme is no longer supported, preventing a faithful reproduction of the simulation setup used in ref 58.⁵³⁹ The variability caused by these changes has been extensively investigated by Reif et al.⁸⁴ and reflects the observed discrepancy between the present and previous results.⁵⁴²

From the considerations above, we suggest parameter set RM/S4A8_v1 as the one that best reproduces all of the tested lipids at once. For DOPC and POPC bilayers, however, parameter set RM/S4A8_v2 performs slightly better: it must be noticed that these two species present unsaturated bonds along the tails, whose influence might not be fully represented by any of the parameter sets. Indeed, it has been suggested that only a polarizable force field would be able to correctly capture the dynamics of the hydrophobic region of the membranes,⁸⁶

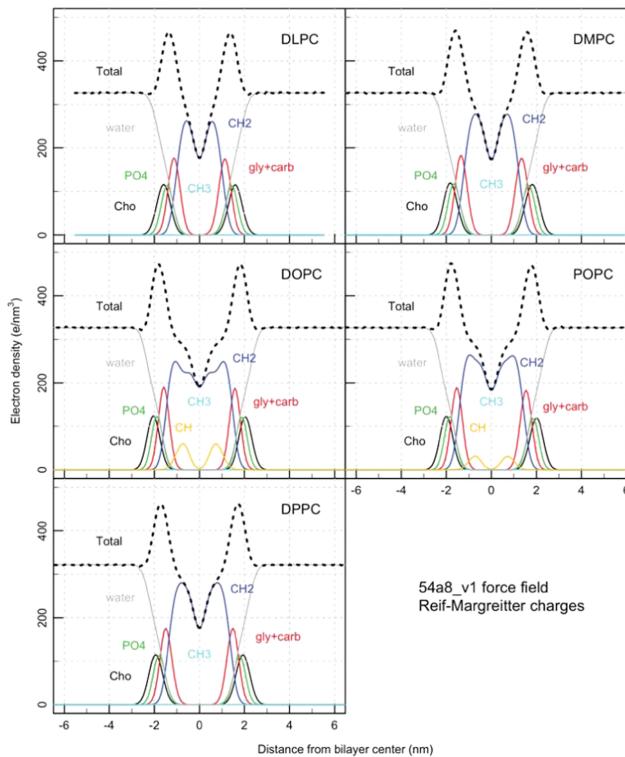


Figure 4. Electron density profiles of the hydrated DLPC, DMPC, DOPC, DPPC, and POPC bilayers (total) and of their individual components (Cho: choline, PO4: phosphate, gly + carb: glycerol and carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID3 (54A8_v1 force field, Reif–Margreiter charges).

Table 4. Bilayer Thickness for Phosphocholine Bilayers, Derived from the Electron Density Profiles (Example in Figure 4) According to the Phosphate or Luzzati Methods^a

ID	charges/FF	hydrophobic thickness D_{HH} (nm)				
		DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	2.83	3.59	3.05	3.30	4.30
3	RM/54A8_v1	2.72	3.48	2.89	3.22	4.06
experiment ^b		3.08	3.44–3.60	3.53–3.71	3.70	3.42–3.83
Luzzati thickness D_B (nm)						
ID	charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	3.11	3.54	4.13	4.00	3.93
3	RM/54A8_v1	3.04	3.48	3.94	3.88	3.75
experiment ^b		3.14	3.63–3.96	3.59–3.87	3.68	3.50–3.83

^aAll simulations were run at 303 K, except for DPPC (323 K). ^bValues from ref 44 and Table 2 in ref 58.

taking in proper account the difference between saturated and unsaturated bonds.

For POPE and POPG, we resorted to the modification proposed by Kukol,⁴⁶ i.e., the use of the CH0 atom type for the ester carbons (see Section 2). For both lipids, a good agreement with experimental ApL values could be achieved using the new partial charge parameters (Figure 3).

Along the same lines, when comparing the results with the ones obtained with the CHARMM36 force field in its original publication,⁶² we find DOPC, presenting an unsaturated bond

in each tail, to be the most diverging. In particular, CHARMM36 better captures the spacing between the lipids, enhanced due to the presence of the double bond, and we suspect that this is due to its all-atom description.

Results of the isothermal area compressibility calculations confirm the finding of refs 58 and 62 that K_A values obtained from simulation are about 1.5–3 times larger than those measured experimentally (SI Table 3). This holds for all parameter sets tested. Set RM/54A8_v1 performs better than

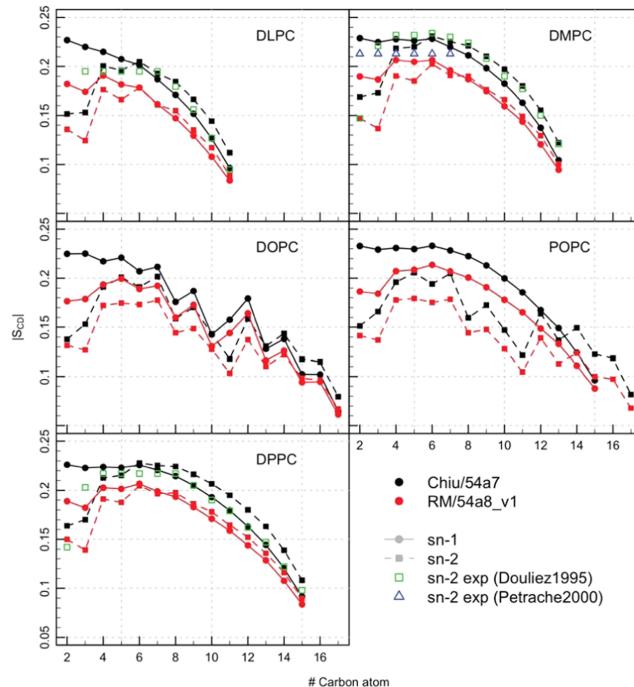


Figure 5. Deuterium order parameter S_{CD} profiles of the *sn*-1 (solid curves) and *sn*-2 (dashed curves) fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC, and POPC bilayers calculated from simulations ID1 (54A7 force field, Chiu charges) and ID3 (54A8_v1 force field, Reif-Margreitter charges). The S_{CD} values are averaged over all of the lipid *sn*-1 and -2 acyl chains in the systems. Experimental values Douliez1995 from ref 89 and Petrache2000 ones from ref 90.

571 Chiu/54A7 for all of the lipids tested but DLPC, for which the
572 results are equivalent.

573 The overestimation of the compressibility is likely due to the
574 underestimation of ApL fluctuations during dynamic simu-
575 lations. The K_A value computed for the small, 128 lipids,
576 DPPC bilayer patch is smaller than the one computed for the
577 512 lipid ones (342 and 499 mN m⁻¹, respectively), as the
578 small patch exhibits higher fluctuations of the ApL (see Section
579 3.1). It is thus evident that the size of the system plays a pivotal
580 role in obtaining correct fluctuations and global properties.

581 **3.2. Electron and Charge Density Profile.** Across
582 simulations with different parameters, the electron density
583 qualitatively maintains the same profile for each phosphocho-
584 line lipid. In Figure 4, the density for parameter set 54A8_v1
585 and all of the lipids is displayed (SI Figures 10–13 show the
586 same plot for the other parameter set), while in SI Figures 5–
587 9, panel (b), the total and the phosphate group electron
588 densities are shown for the five parameter sets tested, for one
589 lipid at the time. The peak broadness shows a direct
590 relationship with the packing density of the bilayer: larger
591 ApL values correspond to shallower profile of the density,
592 due to fluctuations of the membrane along the z axis and to
593 deeper penetration of water molecules into the bilayer.

594 The bilayer thickness was evaluated from the electron
595 density profiles, as explained in Section 2. Our parameters are
596 overall in better agreement with the Luzzati estimate of the
597 thickness rather than the hydrophobic one, but altogether,
598 these measurements (phosphate and Luzzati thickness) do not

strongly discriminate between sets. In Table 4, the values for 599 t4
the Chiu/54A7 and RM/54A8_v1 sets are shown (see SI 600
Table 4 for the complete results).

601 Further comparison of the dipole potential profiles, obtained 602
from the charge density, shows how the RM/54A8_v1 charge 603
set gives results closer to the ones obtained in all-atom 604
simulations^{77,87} (see SI Section 1 for a complete discussion).

605 **3.3. Order Parameter of the Acyl Chains.** For all of the 606
lipids and parameter sets, S_{CD} is lower than 0.25, which 607
indicates that the tails are generally disordered and the 608
membrane has not transitioned to a gel-like state,⁸⁸ even for 609
the simulation with the lowest ApL. Figure 5 and SI Figures 610 f5
14–17 display the computed values for specific parameter sets, 611
and SI Figures 5–9, panel (c), show a cross-parameter 612
comparison for each lipid. Comparing these different sets, 613
simulations denoted by ID from 1 to 5 show a consistently 614
decreasing S_{CD} , in line with the increased area per lipid and 615
decreased bilayer thickness. This indicates that when the lipid 616
molecules are constrained in space, their tails tend to be 617
stretched and ordered. The presence of unsaturated bonds in 618
the DOPC and POPC lipids is captured, by all parameter sets, 619
as a decrease in S_{CD} at the positions related to those bonds.⁶²⁰
The main difference due to the introduction of the new 621
charges is in the decreased order observed for the first and 622
second carbon bonds of the *sn*-1 tail, which show S_{CD} values 623
smaller than the ones for the third carbon bond, while with the 624
Chiu charges, a constant increase is observed with decreasing 625
carbon index for tail *sn*-1.

Overall, the RM/54A8_v1 set is within the range of experimental values^{89,90} (Figure 5); in particular, it captures the low order of the first *sn*-2 carbon atom (numbered 2) well, while the Chiu/54A7 set presents closer values in the central region of the tails. However, it must be noticed that variability is found within the experimental data (see the different experimental values reported in Figure 5). Therefore, without aiming at a perfect fit to such a small pool of experimental data, we consider set RM/54A8_v1 as sufficiently accurate in representing the experimental findings, in particular, in better reproducing the regions in the vicinity of the head group, while the description of the hydrophobic core remains less accurate and subject to improvement.

3.4. Hydration of Head Groups and Glycerol/Carbonyl Moieties. The hydration of functional groups of lipids is a key characteristic for both their dynamics and potential interactions with other molecules, such as proteins. From the distribution of distances between water oxygens and the nearest atom of various lipid groups, it emerges that the new partial charges modify the hydration profile of the lipid head group (Figure 6 shows the comparison between parameter sets for DPPC and SI Figures 17–20 for the other lipid bilayers).

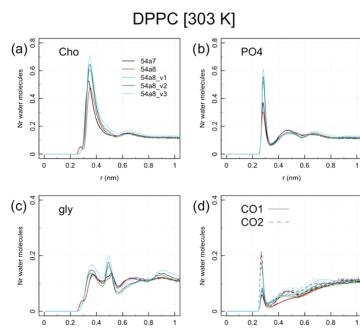


Figure 6. Distribution of the distance between the water oxygen and the nearest lipid head group atom for simulation DPPC. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the *sn*-1 and *sn*-2 positions.

The choline major peak at 0.38 nm and the phosphate one at 0.30 nm are higher and sharper when employing the RM charges rather than the Chiu ones, reflecting an increased average hydration of these two moieties. Additionally, for the simulations run with the RM charges, the choline profile does not display the first, low intensity, peak obtained with the Chiu set at 0.28 nm: indeed, the charges of choline and the modification of the C12 Lennard-Jones repulsion for the NL

atom type introduced in parameter set 54A8 were optimized to successfully prevent oversolvation, repelling water from its core.^{47,54} The profiles of the other components are partially influenced, as well. For the RM charges, the second peak for glycerol increases its value and the two ester peaks have more similar values between them (Figure 6, panel (c), and SI Figures 17–20), which is consistent with deeper water penetration.

To quantify the observed differences, the hydration profiles were integrated up to the first peak or the second one in the case of phosphate and glycerol (SI Table 5). The results show that the average number of water molecules around the choline group is higher for the RM/54A8_v1 set than for the Chiu/54A7 one by one water molecule. This seems to contrast with the increased hydrophobicity of the newly parameterized choline moiety; however, this might partially be explained by the changed orientation of the head groups (see Section 3.5) and by the new parameterization of phosphate,⁴⁸ which accounted for the hydrogen bond potential of the most solvent-accessible atoms, leading to a better solvation of the head group in comparison to the Chiu/54A7 and Chiu/54A8 sets.

The integration up to the second peak of the distribution of distances between the water oxygens and any lipid head group atom gives values between 12 and 17 water molecules per lipid, which is in agreement with the experimental range of 10–20.^{91–94} Again, parameter set RM/54A8_v1 results in more hydrated head groups (about one water molecule more for each lipid) with respect to Chiu/54A7. Notably, the average number of water molecules increases, as expected, for the simulations resulting in a larger ApL (RM/54A8_v2 and RM/54A8_v3). The trends above are confirmed by solvent-accessible surface area values, which are higher for the choline head groups described by the RM charge set with respect to the Chiu one, while the values are closer between parameter sets for the phosphate and glycerol moieties, which are more deeply buried (SI Figure 21).

The increased hydration might be of relevance when simulating interactions with peptides and proteins. Moreover, as shown in a recent comparison between different lipid force fields,⁶² the Chiu/54A7 parameter set results in a slightly less hydrated head group with respect to the CHARMM36 and Lipid14 force fields; therefore, the new set of parameters achieves values closer to them.

3.5. Orientation of the Head Groups and Carbonyl Moieties. The orientation of the head groups, defined by the angle of the P–N vector with the outward bilayer normal, is similar for all of the lipids within the same parameter set (see SI Figure 22, top row). This indicates that the nature of the tails does not strongly affect the behavior of the head group, which is to be expected. Comparing different sets for DPPC

Table 5. Binding Time of Lactoferricin (LFC) Peptide to the Model Membranes in Examination^a

	binding time (ns)							
	DLPC/DLPG 3:1				POPC			
	OI	OII	OIII	OIV	OI	OII	OIII	OIV
Chiu/54A8	0.5	3.8	6.3	2.8	13.9	21.9	(65.0)	13.9
RM/54A8_v1	2.8	75.2	62.0	9.6	1.0	NA	NA	NA

^aNA denotes no binding observed in the time simulated (100 ns). For POPC/OIII simulated with parameter sets Chiu/54A8, the binding time is in parentheses as LFC approaches the membrane but maintains a 0.4 nm distance (± 0.02), which is higher than the threshold chosen to define a binding event.

708 (Figure 7 and SI Figure 23), it emerges that with the Chiu
709 charges, the distribution of P–N angles is bimodal, with

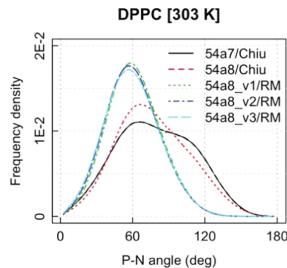


Figure 7. Distribution of the P–N, CO1, and CO2 angles with respect to the outward normal to the bilayer.

710 preferred values around 60 and 90°, while the new charges
711 restrict the motion to the 60° configuration. Recent
712 experimental data support a value around 60° (see refs 95
713 and 96), as opposed to 90° as reported previously.⁹⁷

714 It is noteworthy that this property was not part of the
715 calibration process, i.e., the agreement with the experimental
716 observables in ref 96 is most likely due to a more accurate
717 description of the solvation of the choline and phosphate
718 moieties. Simulations performed by Botan et al.⁹⁸ confirm that
719 smaller angles with respect to the membrane normal are caused
720 by a higher level of head group hydration, which is in line with
721 conclusions from the previous section. This difference in the
722 predominant configuration of the lipids' head group will most
723 probably influence the interaction with proteins or peptides
724 approaching the interfacial region, providing a different
725 binding recognition landscape.

726 The orientation of the *sn*-1 and *sn*-2 carbonyl dipoles with
727 respect to the bilayer normal is again similar across different
728 lipids (SI Figure 22, middle and bottom rows). The
729 introduction of the RM charges has a small effect on these
730 dipoles, as a result of the spatial rearrangement of the nearby
731 head group. The most probable value for CO1 is shifted from
732 110 to 120° (Chiu vs RM charges), while the one for CO2
733 from 135 to 150°.

734 **3.6. Lipid Lateral Diffusion.** To correctly reproduce the
735 membrane and its functions, its dynamical characteristics are as
736 important as its structural ones. To address this, lipid lateral
737 diffusion can be measured and compared against experimental
738 data. Lateral diffusion is influenced by the area per lipid, with a
739 tighter packing preventing larger displacements but is not
740 solely determined by it.

741 Lateral diffusion coefficients (D) measured from simulations
742 are shown in Figure 8. As anticipated, the set with largest ApL
743 (S4A8_v3) presents the highest values; however, parameter set
744 S4A8_v1 gives significantly higher diffusion coefficients than
745 those obtained with the Chiu/S4A7 set, despite the values of
746 ApL being similar. SI Figure 24 depicts a comparison of the
747 diffusion coefficient of DPPC between ID0 and ID1, which
748 differ only in the partial charges of the head groups. This
749 confirms that the RM charges (ID0) allow for more mobility of
750 the lipids with respect to Chiu ones (ID1), independent from
751 all other modifications to the force field.

752 Regarding the simulation conditions, the use of a reaction
753 field scheme increases the mobility by 34%, whereas the size of
754 the patch decreases it by a small but significant amount (19%;

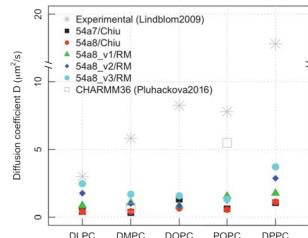


Figure 8. Lateral diffusion coefficient of DLPC, DMPC, DOPC, DPPC, and POPC bilayers for different parameter sets.

755 see SI Figure 24). It is known that periodic boundary 755
756 conditions affect the evaluation of lipid diffusion;^{99,100} therefore, the larger the system simulated, the more accurate 757 the reproduction of the experimental values. However, the 758 change in the D due to the electrostatic treatment and the 759 patch size, taken in absolute terms (i.e., a difference of about 760 0.4 and 0.2 $\mu\text{m}^2 \text{s}^{-1}$, respectively), are small in comparison with 761 the effect due to the adoption of the new parameters (between 762 2 and 6 $\mu\text{m}^2 \text{s}^{-1}$).⁷⁶³

764 The comparison with experimental values is challenging due 764
765 to the fact that different experimental techniques report values, 765
766 which are an order of magnitude apart. Pogore et al. gave an 766 overview of this variability for DPPC bilayers in Table 2 of ref 767
74 and observed that values span from 0.5 to 50 $\mu\text{m}^2 \text{s}^{-1}$. In 768 this view, the values obtained in the present work for DPPC 769 are well within the range, regardless of the parameter set 770 chosen. However, we report experimental values from ref 101 771 obtained through pulsed-field gradient nuclear magnetic 772 resonance as a guide. Additionally, we report the values 773 obtained with CHARMM36 in ref 62. The CHARMM36 774 benchmarks are present only for two of the phosphocholines 775 analyzed in this work and show that the values obtained with 776 this force field span a broader range. The consistently low 777 values of D computed with the different GROMOS parameter 778 sets in this work are in agreement with what was found in the 779 literature.^{102,103}⁷⁸⁰

781 **3.7. Tilt Modulus.** We report in Figure 9 the values of κ_t^1 781 obtained for each of the phosphocholines considered and each 782 parameter set tested. For comparison, we plot the experimental 783 values obtained by Nagle et al. in ref 104 and the results from 784 simulations using the CHARMM36 force field.¹⁰⁵ Given that 785 the data show quite a large spread in their values depending on 786

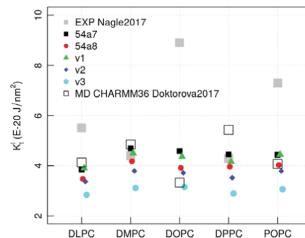


Figure 9. Tilt modulus κ_t^1 computed from the distribution of lipid tilt angles along the last 300 ns of the trajectories. The results are compared with the experimental values from ref 104 and the ones obtained (with the same procedure as employed here) from CHARMM36 simulations in ref 105.

787 the actual experimental setup used, for the comparison, we
 788 selected values, which were all obtained under the same
 789 conditions, for both the experiments and the computational
 790 results.

791 The plot shows that the tilt modulus κ_t^l varies between 3 and
 792 $5 \times 10^{-20} \text{ J nm}^{-2}$. In simulations resulting in larger ApL (e.g.,
 793 parameter set S4A8_v3), the lipids are in a less dense
 794 environment and can better accommodate changes in their
 795 orientations resulting in a lower tilt modulus (the tilt modulus
 796 gives the energy necessary for tilting the lipids per unit area).

797 The comparison with the experimental values is very good
 798 for DMPC and DPPC, while it is poorer for DLPC and very
 799 poor for DOPC and POPC, which harbor unsaturated bonds
 800 in the tails. Results from the CHARMM36 simulations show
 801 more variability between different phosphocholines, but a
 802 similar if not lower agreement with the experiment (for
 803 example, for DOPC). In general, comparing the results
 804 together, we think that we achieved a sufficiently qualitative
 805 agreement with the previous computational literature.

806 The discrepancy with experiments (both for our results and
 807 for the ones from ref 105) is likely due to computational
 808 limitations: as the tilt is retrieved from an ensemble
 809 distribution, larger and longer simulations are more likely to
 810 give a better result. Moreover, as briefly mentioned at various
 811 points in the manuscript, artifacts arising in the hydrophobic
 812 regions of lipids (such as the suboptimal modeling of
 813 unsaturated tails) will probably only be resolved using a
 814 polarizable description.

815 **3.8. Phase Transition Behavior.** The previous analysis
 816 points to parameter set S4A8_v1 as the one that best
 817 reproduces the experimental properties for each of the lipids
 818 simulated. Therefore, we test this set further to assess its ability
 819 to reproduce the change in lipid behavior under different
 820 temperature conditions. As mentioned in Section 2.6, we use
 821 as test system the DPPC bilayer patch.

822 The comparison between simulations at 323 and 333 K
 823 shows that the global area per lipid increases with temperature,
 824 consistently with what was expected. Parameter set S4A8_v1
 825 captures the increase in lipid spacing, with a slight under-
 826 estimation of ApL at 333 K with respect to the experiments
 827 (for S4A8_v1 and experiments, respectively: 0.624(6) vs
 828 0.631(13) nm² at 323 K and 0.634(5) vs 0.650(13) nm² at 333
 829 K). The experimental ApL is measured at a different
 830 temperature with the same experimental setup.¹⁶

831 The simulation at 303 K shows the formation of a patch of
 832 ordered lipids, suggesting that the parameters can reproduce
 833 different phases: two nucleation sites for the gelification
 834 process are observed in both the upper leaflet and lower leaflet,
 835 in nonmatching positions, and the gel front extends over time.
 836 To classify the phase a lipid belongs to, we computed the
 837 hexagonal order parameter S_6 for each chain from the last
 838 frame of the simulation (time point 400 ns). Figure 10 shows
 839 the position of the chains on the xy plane, for each leaflet
 840 separately, color-coded by S_6 : the regions where S_6 is larger
 841 correspond to a densely packed area with a quasi-hexagonal
 842 lattice. In particular, the center of the ordered patches has S_6
 843 values larger than 0.72 (last two colors of the scale), i.e., it can
 844 be classified as a gel. Overall, 20% of chains have undergone
 845 this transition within the time simulated.

846 Averaging over all of the lipids, S_6 at 303 K is 0.45. As a
 847 comparison, we computed this average quantity on the last
 848 frame of the simulations performed at 323 and 333 K, finding

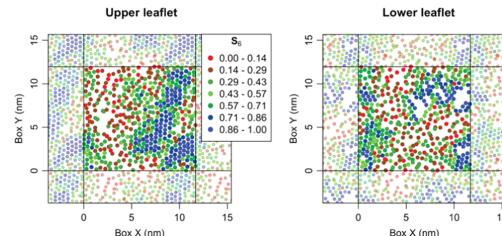


Figure 10. Hexagonal order parameter S_6 for lipid acyl chains computed on the last frame of a DPPC bilayer simulation using S4A8_v1 parameters. Each point corresponds to the average position of the carbon atoms of the respective chain on the xy plane. Thus, every lipid is represented by two points in these plots. Solid black lines denote the boundaries of the simulation box and chains of the periodic images (used for the computation of S_6 boundaries) are shown faded out. Colors from red to blue denote an increasing S_6 value: the last two indicate gel-phase lipids.

0.28 and 0.27, respectively (with only six and three chains above the gel threshold of 0.72).

A hexagonal order can be obtained when the tails are well ordered and parallel to each other, standing in a vertical straight conformation (SI Figure 25 shows a detail of a well-ordered gel patch). We thus compute the S_{CD} order parameter of the acyl chains averaging separately over the lipids for which at least one chain has an S_6 value larger than the 0.72 threshold (168 lipids overall) and for the others. The last 100 ns of the simulation time was used. These values are compared to the average S_{CD} from the simulations at 323 and 333 K. Figure 11

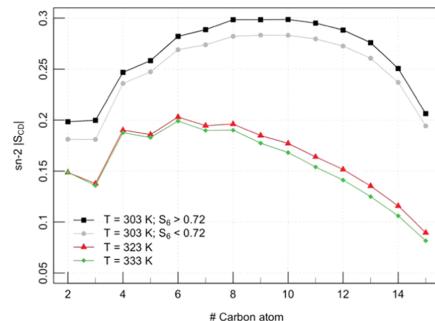


Figure 11. Order parameter for the acyl chain sn-2 for a DPPC bilayer simulated at a different temperature. The average is performed including both leaflets. For the simulation at 303 K, the lipids were split in two groups according to their hexagonal order parameter S_6 and the acyl chain order parameter S_{CD} computed for each of them.

shows highly ordered tails for the membrane simulated below the transition temperature, for both the gel and nongel lipids (classified according to the S_6 threshold). This suggests that the full patch is undergoing a phase transition, but the completion of the process is not seen due to the short simulation time scale. As a comparison, tails at 323 and 333 K are much more disordered, with a slight decrease in order with increasing temperature.

Finally, we computed the local area per lipid chain from a Dirichlet tessellation of the same set of points used to calculate

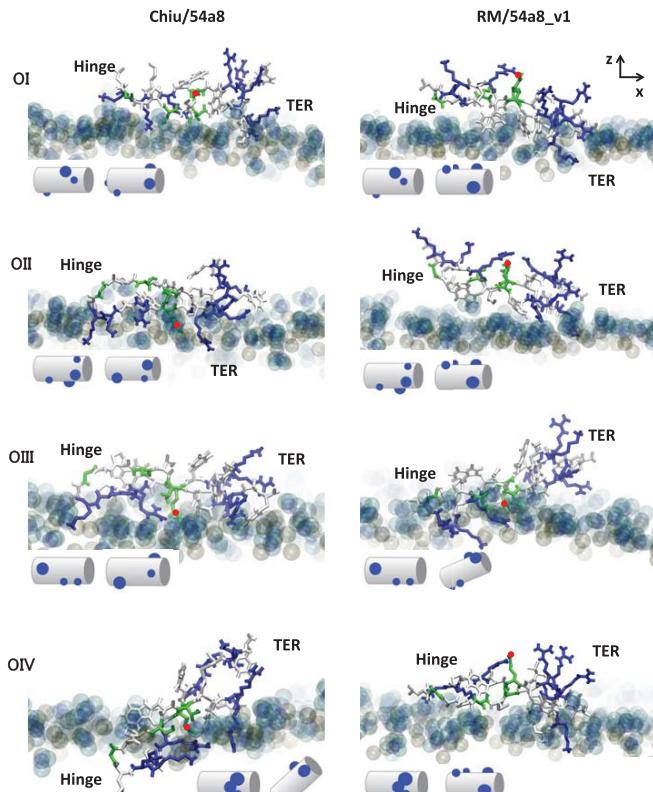


Figure 12. Final configurations of the simulations of LFC on a DLPC/DLPG 3:1 membrane, starting from four different initial orientations OI–OIV. OII, OIII, and OIV are obtained from OI with an anticlockwise rotation of, respectively, 90, 180, and 270° along the main axis (x axis in the top right panel). LFC is colored by residue type (blue charged, green hydrophilic, white hydrophobic); phosphorus atoms are shown in golden beads. The terminal and hinge regions of 1LFC are indicated (TER, hinge), together with GLN7 as a red dot to help the visualization. The insets show a cartoon representation of the initial and final configurations, highlighting the positive patches as blue dots.

870 the hexagonal order parameter. The average values over the
 871 gel-phase tails (multiplied by 2) give an A_pL of 0.438 ± 0.038
 872 nm 2 , while the remaining of the chains have widely spread
 873 values, correlated to their S_6 parameter, giving an average of
 874 0.57 ± 0.18 nm 2 . The values found for the gel patches (at 303
 875 K) are close to the experimental outcomes by Nagle et al. of
 876 0.473 nm 2 at 293 K¹⁰⁶ and $0.479(2)$ at 297 K.¹⁰⁷ The value
 877 computed from simulations is smaller likely because it is
 878 computed only over the tails perfectly packed in a hexagonal
 879 lattice.

880 Altogether, these results prove that the newly developed
 881 parameters can successfully reproduce the gel phase when a
 882 lipid patch is simulated below the phase transition temper-
 883 ature.

884 **3.9. Transferability to POPE and POPG.** As mentioned
 885 above, the areas per lipid values of POPE and POPG
 886 simulations, where the phosphate partial charges have been
 887 replaced with the RM values and, in the case of POPE, the
 888 amine partial charges have been updated according to the
 889 54A8 force field, are in good agreement with the available
 890 experimental data.

891 For POPE, a slightly enhanced hydration is obtained from
 892 the update of the phosphate charges (from 5.7 to 6.4 water
 893 molecules per lipid with experimental values between 4 and
 894 7;^{108,109} see SI Table 7) with similar results in terms of
 895 thickness D_B (3.89 nm for Chiu and 3.92 for RM set, with an
 896 experimental value of 4.13 nm;¹¹⁰ SI Table 6 and SI Figures 26
 897 and 27). Overall, these results confirm the transferability of the
 898 new phosphate charges to different types of phospholipids.
 899

4. INTERACTION WITH PROTEINS

899 The adoption of the updated parameters enhances the
 900 consistency with the GROMOS parameters for protein
 901 simulations. To test how this affects the simulations of
 902 peptides interacting with a membrane, we performed addi-
 903 tional simulations of a small antimicrobial peptide on the
 904 surface of two different model membranes.
 905

906 The peptide selected is bovine lactoferricin (PDB code
 907 1LFC). It has a length of 25 amino acids and adopts a β -
 908 hairpin conformation in solution, with many aromatic
 909 hydrophobic residues on one side and charged amino acids
 910 distributed all over.¹¹¹ This peptide is antimicrobial and
 911 therefore found to preferentially bind bacterial membranes.
 912

versus mammal ones, as shown by NMR experiments on LFC subsequences (namely, LFC_{4–9}¹¹² and LFC_{4–14}¹¹³).

One can model the bacterial membrane by a mixture of zwitterionic and anionic lipids. The latter are characteristic of the cell wall of both Gram-positive and -negative bacteria.^{114,115} In this study, we selected the mixture DLPC/DLPG with a 3:1 ratio that has been used to elucidate the antimicrobial activity of lactoferricin-derived peptides.¹¹⁶ As for the mammal membrane description, we used POPC as it has been often used in molecular dynamics simulations with this purpose.^{117–119} Despite being rather simple, these or similar model membranes have often been used in experiments to test, among others, the effects of antimicrobial peptides upon binding.¹⁴

Molecular simulations can shed light on the differences in the binding process of LFC to antimicrobial and mammal model membranes. Our parameterization should then reflect a sensible difference in the binding behavior for these two cases. The exact binding mechanism of LFC to a membrane is not fully understood, and a number of experimental papers have hypothesized binding modes for the interactions of this peptide with model membranes. A mutation study in LFC_{1–15} suggests that Trp residues anchor the peptide to the membrane as the antimicrobial activity of the peptide was retained only when Trp was mutated in equally hydrophobic amino acids.¹²⁰ However, the role of Trp seems to be different in other antimicrobial peptides, where they reside at the lipid–water interface and form hydrogen bonds with the moieties nearby.¹²¹ As experiments on the full-length peptide (25 amino acids) have not yet been reported and the full sequence contains additional charged and hydrophobic residues, it remains unclear whether this additional region would change the aforementioned binding mechanism. We therefore decided to use molecular dynamics simulations to elucidate molecular determinants in discriminating the binding of the peptide to mammal and bacterial membranes.

In order to not be biased by the initial configuration adopted in the simulation, we performed multiple simulations with different initial orientations of the peptide relative to the membrane. The hairpin main axis was aligned to the membrane plane and the peptide rotated around this axis in steps of 90°, leading to four different starting orientations named OI, OII, OIII, and OIV (SI Figure 29). This allows different segments of the sequence (and thus amino acids with different chemical characteristics) to face the membrane in the initial positioning. The initial minimum distance between the peptide and the lipids was set to 2 nm. The simulation length was 100 ns each, sufficient to see the binding process in all of the control cases.

The simulations have been performed for the proposed RM/54A8_v1 parameter set and the Chiu/54A8 one (control cases) to compare with the most recent set available in GROMOS and highlight the difference of the newly parameterized lipid head groups.

To quantify the outcome of the simulations, we monitored the time at which the peptide binds (always irreversibly) to the membrane as the time at which the minimum distance between the peptide and the membrane is below 0.3 nm (Table S). The cutoff was chosen, analyzing the configurations after LFC bound to the membrane, which resulted generally to stabilize around a minimum distance of 0.25 nm. The minimum distance was computed every 100 ps, and a running average was applied with a 10 frame window. Additionally,

insertion depth of each amino acid in the membrane has been calculated as the difference between the *z* position of the lowest atom of the amino acid and the average of the maximum *z* coordinate of the five lipids closest to it.

Table S shows the different binding times for LFC against a mixed DLPC/DLPG or pure POPC membrane patch. For the mixed, anionic membrane, the new parameters favor a slower and weaker binding process. Indeed, with parameter set Chiu/54A8, the peptide is quickly sequestered by the lipids due to the opposite charge interaction. This favors an unspecific binding, dependent on the sequence facing the membrane in the initial configuration (Figure 12 and SI Movie 1). In Figure 13, the average insertion in the membrane after the binding is

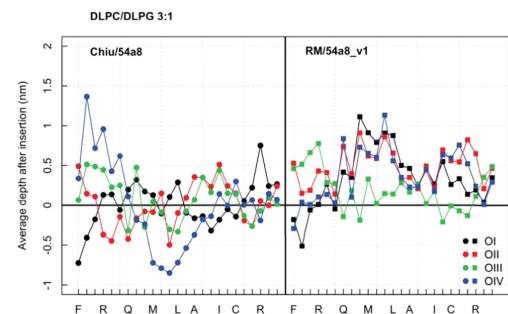


Figure 13. Average insertion depth of each amino acid after binding of the LFC peptide as per Table S. The zero value is the top of the membrane plane so that a negative depth means insertion into the membrane. Some reference amino acids are displayed at the bottom.

plotted for each amino acid: Chiu/54A8 favors a deep insertion of differently charged residues for different runs. The RM/54A8_v1 simulations produce a less inserted configuration of LFC and a more consistent protrusion of the hinge region out of the membrane, i.e., the central stretch of amino acids between Met and Leu (Figure 13), as three out of the four simulations (all but OIII) show this behavior.

The angular orientation of the peptide around its axis has been computed as the angle formed with the *z* axis by the backbone carbon and nitrogen bonded via a hydrogen bond (amino acids 7 with 19 and 9 with 17), confirming that the new set of parameters allows for more freedom in the reorientation of the initial configurations, while the previous one tends to keep them close to the original configuration (SI Figure 30). Additionally, the new set of parameters seems to favor the reorientation of the peptide as to face the Trp residues toward the membrane surface in three out of the four simulations, in contrast with the results from the previous parameterization. This preference for the interfacial region is a known mechanism in the membrane binding of Trp- and Arg-rich peptides.¹²¹

When simulating a pure POPC membrane (here considered as a mammal membrane model), the resulting binding poses obtained with the Chiu/54A8 and different initial conditions are consistent among each other; in particular, the three amino acids Lys12–Leu13–Gly14 located at the hinge of the hairpin promote the insertion, while the terminal region stays exposed in solution. Therefore, parameter set Chiu/54A8 discriminates between the two membranes as it suggests a weaker binding to the mammal one. However, with the parameter set RM/

1017 S4A8_v1, three out of the four simulations result in no binding
1018 at all, in agreement with experimental findings.¹¹² The
1019 remaining simulation (O1) shows a quick binding event,
1020 promoted by the terminal regions as observed for three out of
1021 four simulations with the DLPC/DLPG 3:1 membrane.

1022 The results above highlight that the new parameters show a
1023 membrane-binding process less dependent on the initial
1024 conditions, allowing for a dynamical rearrangement of the
1025 protein at the membrane interface. This comes at the expenses
1026 of a longer sampling time needed to observe binding events for
1027 most of the configurations chosen. Future work will focus on
1028 systematic comparisons of available peptide–membrane
1029 simulations with other parameterizations and on longer
1030 simulated times.

1031 The difference between the behavior on a model bacterial or
1032 mammal membrane is more pronounced for the new
1033 parameters, and this is consistent with the selective
1034 antimicrobial action of the peptide and its low hemolytic
1035 activity.^{112,113} Overall, we think that these new parameters
1036 show promising characteristics for the simulation of mem-
1037 brane–peptide interactions within the GROMOS force field,
1038 particularly for the study of interfacial absorption.

5. CONCLUSIONS

1039 In this work, we present a reparameterization of a range of
1040 phospholipids in the context of the GROMOS force field,
1041 taking advantage of recent optimizations reported for key
1042 chemical groups in these molecules. The effect of the newly
1043 adopted head group partial charges has been tested extensively
1044 to ensure that they match experimentally observable character-
1045 istics of lipid bilayers. In parallel, we tested the effect of the van
1046 der Waals repulsion between the choline methyl groups and
1047 the phosphate oxygens, as it was modified by Poger et al. to
1048 reproduce the experimental area per lipid values while using
1049 the partial charges derived by Chiu et al. A summary of the
1050 updated parameters and simulation conditions is available in SI
1051 Table 10.

1052 The work proves that the new charges are suitable to
1053 describe all of the phosphocholine bilayers tested, matching
1054 the experimental values as successfully as the previous
1055 parameter set. The major advantage of the Reif–Margreitter
1056 set lies in the partial charges of the head group, which are
1057 derived by applying the GROMOS parameterization philoso-
1058 phy rather than quantum mechanics calculations, thereby
1059 providing a description, which is more consistent with the
1060 parameters adopted for other biomolecules such as proteins
1061 within this force field. By using the updated partial charges for
1062 the choline (more hydrophobic) and phosphate (more
1063 hydrophilic) groups, the parameters also show a better
1064 reproduction of the average head group orientation, which
1065 was recently reassessed by experiments. The value of the
1066 Lennard-Jones repulsion term found to best reproduce the
1067 experimental values is the one in set S4A8_v1, which is set to a
1068 value in between that of the S4A7 and S4A8 parameter sets.
1069 In the Reif–Margreitter parameter set, only the partial
1070 charges of the ester groups remain as described in the Chiu
1071 charge set. Preliminary work has been started to test the
1072 influence of the ester charges in combination with the new
1073 ones for the head group but, in accordance to what was
1074 previously found by Chandrasekhar et al. (ref 56), the
1075 replacement of the ester charges with the standard ones for
1076 the ester moiety resulted in values of area per lipid too low
1077 with respect to the experimental findings. As mentioned

1078 previously, it is possible that this discrepancy with the rest of
1079 the force field can only be avoided by adopting a polarizable
1080 force field.⁸⁶ However, in the absence of further sophisticated
1081 changes to the force field parameterization, we are confident
1082 that the proposed parameters are a major step forward in the
1083 description of lipid head groups and they should enable
1084 improved modeling of the interaction of lipids with water and
1085 other soluble molecules.
1085

1086 The new phosphate partial charges have been proved to
1087 transfer well to other phospholipids not presenting a choline
1088 head. For those lipids, the Kukol modification, which takes
1089 advantage of a different atom type for the ester carbon, is
1090 adopted to obtain the correct area per lipid.
1090

1091 Finally, the performance in reproducing some specific
1092 peptide–membrane interactions was tested. In this respect,
1093 the new parameter set shows significant differences with
1094 respect to the latest Chiu/S4A8 set: it better discriminates the
1095 binding of an antimicrobial sequence on a bacterial versus a
1096 mammal membrane. Additionally, it favors a weaker and more
1097 dynamic binding, which is less biased from the initial
1098 conditions of the simulations.
1098

1099 In conclusion, we believe that the new Reif–Margreitter
1100 charge set together with the GROMOS S4A8_v1 parameter set
1101 is a major improvement on the previous iteration of the
1102 GROMOS lipid force field and should be particularly suited for
1103 protein–membrane systems, such as studies including small
1104 antimicrobial peptides, which rely on an accurate peptide–
1105 membrane recognition.
1105

■ ASSOCIATED CONTENT

S Supporting Information

1106 The Supporting Information is available free of charge on the
1107 ACS Publications website at DOI: 10.1021/acs.jctc.9b00509.
1109

1110 Additional control simulations; simulations for POPE
1111 bilayer; compressibility values for PC lipids; thickness
1112 values for PC lipids; lipid moiety for PC lipids, POPE
1113 and POPG bilayers; properties of POPE and POPG
1114 bilayers; previous parameters and simulation setup (SI
1115 Tables 1–10); dipole potential; time series of the ApL
1116 for a DPPC bilayer; ApL for different time windows of
1117 the same simulation and control simulations; compar-
1118 ison of key properties across parameter sets for DLPC,
1119 DMPC, DOPC, POPC, and DPPC bilayers; electron
1120 density profiles of bilayers for parameter sets Chiu/
1121 S4A7, Chiu/S4A8, RM/S4A8_v2, and RM/S4A8_v3;
1122 deuterium order parameter S_{CD} of bilayers for parameter
1123 sets Chiu/S4A8, RM/S4A8_v2, and RM/S4A8_v3;
1124 hydration profiles for DLPC, DMPC, DOPC, and
1125 POPC bilayers across parameter sets; phosphocholine
1126 moieties SASA; head group orientation, lipid tail
1127 comparison and parameter set comparison; diffusion
1128 coefficient of DPPC in control simulations; snapshot of
1129 DPPC patch partially in gel phase; key properties of
1130 POPE bilayer for parameter sets Chiu/S4A8/CH0 and
1131 RM/S4A8/CH0; key properties of POPG bilayer for
1132 parameter set RM/S4A8/CH0; scheme of the 1LFC
1133 orientations; 1LFC orientation with respect to mem-
1134 brane plane along simulations (SI Figures 1–30) (PDF)
1135 Simulation of LFC on a bacterial membrane for both
1136 parameter sets Chiu/S4A8 (left half) and RM/S4A8_v1
1137 (right half), starting from orientation OII (Movie 1)
1138 (MP4)

1139 ■ AUTHOR INFORMATION

1140 Corresponding Author

1141 *E-mail: franca.fraternali@kcl.ac.uk.

1142 ORCID

1143 Irene Marzuoli: 0000-0001-7536-6144

1144 Franca Fraternali: 0000-0002-3143-6574

1145 Notes

1146 The authors declare no competing financial interest.

1147 The topology and the parameter files in GROMACS format for
 1148 the lipid molecules and the atomic coordinates of equilibrated
 1149 bilayers of DLPC, DMPC, DOPC, POPC, DPPC, POPE, and
 1150 POPG are available at <http://fraternalilab.kcl.ac.uk/>
 1151 [wordpress/biomembrane-simulations](#). Note that these param-
 1152 eter files include the post-translationally modified residues
 1153 specified for the GROMOS 54A8 force field earlier.¹²²

1154 ■ ACKNOWLEDGMENTS

1155 The authors thank Chris Oostenbrink and Maria M. Reif for
 1156 careful reading of the manuscript and feedback. I.M.
 1157 acknowledges funding by the Engineering and Physical
 1158 Sciences Research Council (EPSRC) through the Centre for
 1159 Doctoral Training “Cross Disciplinary Approaches to Non-
 1160 Equilibrium Systems” (CANES, Grant No. EP/L015854/1).
 1161 C.M. acknowledges funding by the MRC/BBSRC Systems
 1162 Immunology of the Lifecourse programme grant: MABRA
 1163 “Multiscale analysis of B cell responses in ageing” (MR/
 1164 L01257X/1). This work was performed using resources
 1165 provided by the Cambridge Service for Data Driven Discovery
 1166 (CSD3) operated by the University of Cambridge Research
 1167 Computing Service (<http://www.cs3.cam.ac.uk/>), provided
 1168 by Dell EMC and Intel using Tier-2 funding from the
 1169 Engineering and Physical Sciences Research Council (capital
 1170 grant EP/P020259/1), and DiRAC funding from the Science
 1171 and Technology Facilities Council (www.dirac.ac.uk).

1172 ■ REFERENCES

- (1) Cooper, G. M. *Transport of Small Molecules*; Sinauer Associates, 2000.
- (2) Cocucci, E.; Kim, J.; Bai, Y.; Pabla, N. Role of Passive Diffusion, Transporters, and Membrane Trafficking-Mediated Processes in Cellular Drug Transport. *Clin. Pharmacol. Ther.* **2017**, *101*, 121–129.
- (3) Mälkiä, A.; Murtomäki, L.; Urtti, A.; Kontturi, K. Drug permeation in biomembranes: In vitro and in silico prediction and influence of physicochemical properties. *Eur. J. Pharm. Sci.* **2004**, *23*, 13–47.
- (4) Scott, D.; Ghosh, A.; Di, L.; Maurer, T. Passive drug permeation through membranes and cellular distribution. *Pharmacol. Res.* **2017**, *117*, 94–102.
- (5) Zhao, R.; Babani, S.; Gao, F.; Liu, L.; Goldman, I. D. The mechanism of transport of the multitargeted antifolate (MTA) and its cross-resistance pattern in cells with markedly impaired transport of methotrexate. *Clin. Cancer Res.* **2000**, *6*, 3687–3695.
- (6) Shi, Y.; Massagué, J. Mechanisms of TGF-β Signaling from Cell Membrane to the Nucleus. *Cell* **2003**, *113*, 685–700.
- (7) Jean-Louis, S.; Akare, S.; Ali, M. A.; Mash, E. A.; Meuillet, E.; Martinez, J. D. Deoxycholic acid induces intracellular signaling through membrane perturbations. *J. Biol. Chem.* **2006**, *281*, 14948–14960.
- (8) Hodder, A. N.; Crewther, P. E.; Anders, R. F. Specificity of the protective antibody response to apical membrane antigen 1. *Infect. Immun.* **2001**, *69*, 3286–3294.
- (9) Fahy, E.; Subramaniam, S.; Brown, H. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Seyama, Y.; Shaw, W.; Shimizu, T.; Spener, F.; van Meer, G.; VanNieuwenhze, M. S.; White, S. H.; Witztum, J. L.; Dennis, E. A. A comprehensive classification system for lipids. *J. Lipid Res.* **2005**, *46*, 839–862.
- (10) Sharpe, H. J.; Stevens, T. J.; Munro, S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **2010**, *142*, 158–169.
- (11) Entova, S.; Billod, J.-M.; Swiecki, J.-M.; Martín-Santamaría, S.; Imperiali, B. Insights into the key determinants of membrane protein topology enable the identification of new monotopic folds. *eLife* **2018**, *7*, e40889.
- (12) Zhang, L.; Rozek, A.; Hancock, R. E. Interaction of cationic antimicrobial peptides with model membranes. *J. Biol. Chem.* **2001**, *276*, 35714–35722.
- (13) Risselada, H. J.; Marrink, S. J. The molecular face of lipid rafts in model membranes. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17367–17372.
- (14) Chan, Y.-H. M.; Boxer, S. G. Model membrane systems and their applications. *Curr. Opin. Chem. Biol.* **2007**, *11*, 581–587.
- (15) Nagle, J. F.; Tristram-Nagle, S. Structure of lipid bilayers. *Biochim. Biophys. Acta* **2000**, *1469*, 159–195.
- (16) Kučerka, N.; Heberle, F.; Pan, J.; Katsaras, J. Structural Significance of Lipid Diversity as Studied by Small Angle Neutron and X-ray Scattering. *Membranes* **2015**, *5*, 454–472.
- (17) Pan, J.; Heberle, F. A.; Tristram-Nagle, S.; Szymanski, M.; Koepfinger, M.; Katsaras, J.; Kučerka, N. Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 2135–2148.
- (18) Kučerka, N.; Nieh, M.-P.; Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808*, 2761–2771.
- (19) Gurtovenko, A. A.; Vattulainen, I. Molecular Mechanism for Lipid Flip-Flops. *J. Phys. Chem. B* **2007**, *111*, 13554–13559.
- (20) Tielemans, D.; Marrink, S.-J. Lipids Out of Equilibrium: Energetics of Desorption and Pore Mediated Flip-Flop. *J. Am. Chem. Soc.* **2006**, *128*, 12462–12467.
- (21) Marrink, S.-J.; Mark, A. E. The Mechanism of Vesicle Fusion as Revealed by Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2003**, *11144*.
- (22) de Vries, A. H.; Mark, A. E.; Marrink, S.-J. Molecular Dynamics Simulation of the Spontaneous Formation of a Small DPPC Vesicle in Water in Atomistic Detail. *J. Am. Chem. Soc.* **2004**, *126*, 4488–4489.
- (23) Marrink, S.-J.; Risselada, J.; Mark, A. E. Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chem. Phys. Lipids* **2005**, *135*, 223–244.
- (24) Marrink, S.-J.; Lindahl, E.; Edholm, O.; Mark, A. E. Simulation of the Spontaneous Aggregation of Phospholipids into Bilayers. *J. Am. Chem. Soc.* **2001**, *123*, 8638–8639.
- (25) Skjevik, Å. A.; Madej, B. D.; Dickson, C. J.; Teigen, K.; Walker, R. C.; Gould, I. R. All-atom lipid bilayer self-assembly with the AMBER and CHARMM lipid force fields. *Chem. Commun.* **2015**, *51*, 4402–4405.
- (26) Vernier, P. T.; Ziegler, M. J.; Sun, Y.; Chang, W. V.; Gundersen, M. A.; Tielemans, D. P. Nanopore Formation and Phosphatidylserine Externalization in a Phospholipid Bilayer at High Transmembrane Potential. *J. Am. Chem. Soc.* **2006**, *128*, 6288–6289.
- (27) Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Ion transport across transmembrane pores. *Biophys. J.* **2007**, *92*, 4209–4215.
- (28) Marrink, S.-J.; Jähnig, F.; Berendsen, H. J. Proton transport across transient single-file water pores in a lipid membrane studied by molecular dynamics simulations. *Biophys. J.* **1996**, *71*, 632–647.
- (29) Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Molecular dynamics simulations of hydrophilic pores in lipid bilayers. *Biophys. J.* **2004**, *86*, 2156–2164.
- (30) Sengupta, D.; Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochim. Biophys. Acta, Biomembr.* **2008**, *1778*, 2308–2317.
- (31) Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Antimicrobial Peptides in Action. *J. Am. Chem. Soc.* **2006**, *128*, 12156–12161.

- 1270 (32) Lipkin, R.; Pino-Angeles, A.; Lazaridis, T. Transmembrane
1271 Pore Structures of β -Hairpin Antimicrobial Peptides by All-Atom
1272 Simulations. *J. Phys. Chem. B* **2017**, *121*, 9126–9140.
- 1273 (33) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Electroporation of
1274 the *E. coli* and *S. aureus* Membranes: Molecular Dynamics Simulations
1275 of Complex Bacterial Membranes. *J. Phys. Chem. B* **2011**, *115*,
1276 13381–13388.
- 1277 (34) Khalid, S.; Berglund, N. A.; Holdbrook, D. A.; Leung, Y. M.;
1278 Parkin, J. The membranes of Gram-negative bacteria: progress in
1279 molecular modelling and simulation. *Biochem. Soc. Trans.* **2015**, *43*,
1280 162–167.
- 1281 (35) Pino-Angeles, A.; Leveritt, J. M.; Lazaridis, T. Pore Structure
1282 and Synergy in Antimicrobial Peptides of the Magainin Family. *PLOS*
1283 *Comput. Biol.* **2016**, *12*, No. e1004570.
- 1284 (36) Lopes, D.; Jakobtorweihen, S.; Nunes, C.; Sarmento, B.; Reis, S.
1285 Sheding light on the puzzle of drug-membrane interactions:
1286 Experimental techniques and molecular dynamics simulations. *Prog.*
1287 *Lipid Res.* **2017**, *58*, 24–44.
- 1288 (37) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.;
1289 Tobias, D. J.; Mondragón-Ramírez, C.; Vorobyov, I.; MacKerell, A.
1290 D.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force
1291 Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**,
1292 *114*, 7830–7843.
- 1293 (38) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.;
1294 MacKerell, A. D. Optimization of the CHARMM Additive Force
1295 Field for DNA: Improved Treatment of the B1/BII Conformational
1296 Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348–362.
- 1297 (39) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de
1298 Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an
1299 improved force field for folded and intrinsically disordered proteins.
1300 *Nat. Methods* **2017**, *14*, 71–73.
- 1301 (40) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.;
1302 Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of
1303 Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem.*
1304 *Theory Comput.* **2015**, *11*, 3696–3713.
- 1305 (41) Scott, W.; Hünenberger, P.; Tironi, I.; Mark, A.; Billeter, S.;
1306 Fennem, J.; Torda, A.; Huber, T.; Krüger, P.; van Gunsteren, W. F.
1307 The GROMOS Biomolecular Simulation Program Package. *J. Phys.*
1308 *Chem. A* **1999**, *103*, 3596–3607.
- 1309 (42) Chiu, S.-W.; Clark, M.; Balaji, V.; Subramaniam, S.; Scott, H.
1310 L.; Jakobsson, E. Incorporation of surface tension into molecular
1311 dynamics simulation of an interface: a fluid phase lipid bilayer
1312 membrane. *Biophys. J.* **1995**, *69*, 1230–1245.
- 1313 (43) Tieleman, D. P.; Berendsen, H. J. C. Molecular dynamics
1314 simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer
1315 with different macroscopic boundary conditions and parameters. *J.*
1316 *Chem. Phys.* **1996**, *105*, 4871.
- 1317 (44) Poger, D.; Van Gunsteren, W. F.; Mark, A. E. A new force field
1318 for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **2010**,
1319 *31*, 1117–1125.
- 1320 (45) Chandrasekhar, I.; Kastenholz, M.; Lins, R. D.; Oostenbrink,
1321 C.; Schuler, L. D.; Tieleman, D. P.; van Gunsteren, W. F. A consistent
1322 potential energy parameter set for lipids: dipalmitoylphosphatidylcholine
1323 line as a benchmark of the GROMOS96 45A3 force field. *Eur.*
1324 *Biophys. J.* **2003**, *32*, 67–77.
- 1325 (46) Kukol, A. Lipid Models for United-Atom Molecular Dynamics
1326 Simulations of Proteins. *J. Chem. Theory Comput.* **2009**, *5*, 615–626.
- 1327 (47) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New
1328 Interaction Parameters for Charged Amino Acid Side Chains in the
1329 GROMOS Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 3705–3723.
- 1330 (48) Margreitter, C.; Reif, M. M.; Oostenbrink, C. Update on
1331 phosphate and charged post-translationally modified amino acid
1332 parameters in the GROMOS force field. *J. Comput. Chem.* **2017**, *38*,
1333 714–720.
- 1334 (49) Chandrasekhar, I.; van Gunsteren, W. F. Sensitivity of
1335 molecular dynamics simulations of lipids to the size of the ester
1336 carbon. *Curr. Sci.* **2001**, *1325–1327*.
- 1337 (50) Chandrasekhar, I.; van Gunsteren, W. F. A comparison of the
1338 potential energy parameters of aliphatic alkanes: molecular dynamics
simulations of triacylglycerols in the alpha phase. *Eur. Biophys. J.* **2002**, *31*, 89–101.
- (51) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843–856.
- (52) Bransden, B. H.; Joachain, C. *J. Physics of Atoms and Molecules*; Prentice Hall, 2003; p 1114.
- (53) Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. *I. J. Chem. Phys.* **1955**, *23*, 1833–1840.
- (54) Reif, M. M.; Winger, M.; Oostenbrink, C. Testing of the GROMOS Force-Field Parameter Set 54A8: Structural Properties of Electrolyte Solutions, Lipid Bilayers, and Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 1247–1264.
- (55) Berglund, N. A.; Piggot, T. J.; Jefferies, D.; Sessions, R. B.; Bond, P. J.; Khalid, S. Interaction of the Antimicrobial Peptide Polymyxin B1 with Both Membranes of *E. coli*: A Molecular Dynamics Study. *PLOS Comput. Biol.* **2015**, *11*, No. e1004180.
- (56) Chandrasekhar, I.; Oostenbrink, C.; van Gunsteren, W. F. Simulating the Physiological Phase of Hydrated DPPC Bilayers: The Ester Moiety. *Soft Mater.* **2004**, *2*, 27–45.
- (57) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- (58) Poger, D.; Mark, A. E. On the Validation of Molecular Dynamics Simulations of Saturated and cis-Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *J. Chem. Theory Comput.* **2010**, *6*, 325–336.
- (59) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (60) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B. *GROMACS User Manual*, version 2016; GROMACS Development Team, 2018.
- (61) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (62) Pluhackova, K.; Kirsch, S. A.; Han, J.; Sun, L.; Jiang, Z.; Unruh, T.; Böckmann, R. A. A Critical Comparison of Biomembrane Force Fields: Structure and Dynamics of Model DMPC, POPC, and POPE Bilayers. *J. Phys. Chem. B* **2016**, *120*, 3888–3903.
- (63) Lewis, R. N. A. H.; Mak, N.; McElhaney, R. N. A differential scanning calorimetric study of the thermotropic phase behavior of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry* **1987**, *26*, 6118–6126.
- (64) Huang, C. H.; Lapides, J. R.; Levin, I. W. Phase-transition behavior of saturated, symmetric chain phospholipid bilayer dispersions determined by Raman spectroscopy: correlation between spectral and thermodynamic parameters. *J. Am. Chem. Soc.* **1982**, *104*, 5926–5930.
- (65) Mabrey, S.; Sturtevant, J. M. Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 3862–3866.
- (66) Davis, P. J.; Fleming, B. D.; Coolbear, K. P.; Keough, K. M. W. Gel to liquid-crystalline transition temperatures of water dispersions of two pairs of positional isomers of unsaturated mixed-acid phosphatidylcholines. *Biochemistry* **1981**, *20*, 3633–3636.
- (67) Wang, Z. Q.; Lin, H. N.; Li, S.; Huang, C. H. Calorimetric studies and molecular mechanics simulations of monounsaturated phosphatidylethanolamine bilayers. *J. Biol. Chem.* **1994**, *269*, 23491–23499.
- (68) Borle, F.; Seelig, J. Ca^{2+} binding to phosphatidylglycerol bilayers as studied by differential scanning calorimetry and ^3H - and ^{31}P -nuclear magnetic resonance. *Chem. Phys. Lipids* **1985**, *36*, 263–283.
- (69) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

- 1408 (70) Abraham, M. J.; Murtola, T.; Schulz, R.; Päll, S.; Smith, J. C.;
1409 Hess, B.; Lindahl, E. GROMACS: High performance molecular
1410 simulations through multi-level parallelism from laptops to super-
1411 computers. *SoftwareX* **2015**, *1–2*, 19–25.
- 1412 (71) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.;
1413 DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an
1414 external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- 1415 (72) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M.
1416 LINCS: A linear constraint solver for molecular simulations. *J.
1417 Comput. Chem.* **1997**, *18*, 1463–1472.
- 1418 (73) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of
1419 the SHAKE and RATTLE algorithm for rigid water models. *J.
1420 Comput. Chem.* **1992**, *13*, 952–962.
- 1421 (74) Poger, D.; Caron, B.; Mark, A. E. Validating lipid force fields
1422 against experimental data: Progress, challenges and perspectives.
1423 *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 1556–1565.
- 1424 (75) Braun, A. R.; Brandt, E. G.; Edholm, O.; Nagle, J. F.; Sachs, J.
1425 N. Determination of electron density profiles and area from
1426 simulations of undulating membranes. *Biophys. J.* **2011**, *100*, 2112–
1427 2120.
- 1428 (76) Chiu, S.-W.; Pandit, S. A.; Scott, H. L.; Jakobsson, E. An
1429 Improved United Atom Force Field for Simulation of Mixed Lipid
1430 Bilayers. *J. Phys. Chem. B* **2009**, *113*, 2748–2763.
- 1431 (77) Ding, W.; Palaiokostas, M.; Wang, W.; Orsi, M. Effects of Lipid
1432 Composition on Bilayer Membranes Quantified by All-Atom
1433 Molecular Dynamics. *J. Phys. Chem. B* **2015**, *119*, 15263–15274.
- 1434 (78) Gurtovenko, A. A.; Vattulainen, I. Calculation of the
1435 electrostatic potential of lipid bilayers from molecular dynamics
1436 simulations: Methodological issues. *J. Chem. Phys.* **2009**, *130*,
1437 No. 215107.
- 1438 (79) Einstein, A.; Fürt, R. *Investigations on the Theory of Brownian
1439 Movement*; Dover Publications, 1956; p 119.
- 1440 (80) Johner, N.; Harries, D.; Khelashvili, G. Implementation of a
1441 methodology for determining elastic properties of lipid assemblies
1442 from molecular dynamics simulations. *BMC Bioinf.* **2016**, *17*, No. 161.
- 1443 (81) Biasini, M.; Mariani, V.; Haas, J.; Scheuber, S.; Schenk, A. D.;
1444 Schwede, T.; Philippsen, A. OpenStructure: a flexible software
1445 framework for computational structural biology. *Bioinformatics* **2010**,
1446 *26*, 2626–2628.
- 1447 (82) Aurenhammer, F. Franz, Voronoi diagrams – a survey of a
1448 fundamental geometric data structure. *ACM Comput. Surv.* **1991**, *23*,
1449 345–405.
- 1450 (83) Uppulury, K.; Coppock, P. S.; Kindt, J. T. Molecular Simulation
1451 of the DPPE Lipid Bilayer Gel Phase: Coupling between Molecular
1452 Packing Order and Tail Tilt Angle. *J. Phys. Chem. B* **2015**, *119*, 8725.
- 1453 (84) Reißer, S.; Poger, D.; Stroet, M.; Mark, A. E. Real Cost of
1454 Speed: The Effect of a Time-Saving Multiple-Time-Stepping
1455 Algorithm on the Accuracy of Molecular Dynamics Simulations. *J.
1456 Chem. Theory Comput.* **2017**, *13*, 2367–2372.
- 1457 (85) Anézo, C.; de Vries, A. H.; Höltje, H.-D.; Tieleman, D. P.;
1458 Marrink, S.-J. Methodological Issues in Lipid Bilayer Simulations. *J.
1459 Phys. Chem. B* **2003**, *107*, 9424–9433.
- 1460 (86) Chowdhary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.;
1461 MacKerell, A. D.; Roux, B. A Polarizable Force Field of
1462 Dipalmitoylphosphatidylcholine Based on the Classical Drude
1463 Model for Molecular Dynamics Simulations of Lipids. *J. Phys.
1464 Chem. B* **2013**, *117*, 9142–9160.
- 1465 (87) Warshavik, D. T.; Muellner, M. J.; Chachisvilis, M. Effect of
1466 membrane tension on the electric field and dipole potential of lipid
1467 bilayer membrane. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808*,
1468 2608–2617.
- 1469 (88) Vermeer, L. S.; de Groot, B. L.; Réat, V.; Milon, A.; Czaplicki, J.
1470 Acyl chain order parameter profiles in phospholipid bilayers:
1471 computation from molecular dynamics simulations and comparison
1472 with 2H NMR experiments. *Eur. Biophys. J.* **2007**, *36*, 919–931.
- 1473 (89) Douliez, J. P.; Léonard, A.; Dufourc, E. J. Restatement of order
1474 parameters in biomembranes: calculation of C-C bond order
1475 parameters from C-D quadrupolar splittings. *Biophys. J.* **1995**, *68*,
1476 1727–1739.
- 1477 (90) Petrache, H. I.; Dodd, S. W.; Brown, M. F. Area per lipid and
1478 acyl length distributions in fluid phosphatidylcholines determined by
1479 (2)H NMR spectroscopy. *Biophys. J.* **2000**, *79*, 3172–3192.
- 1480 (91) Borle, F.; Seelig, J. Hydration of Escherichia coli lipids:
1481 Deuterium T1 relaxation time studies of phosphatidylglycerol,
1482 phosphatidylethanolamine and phosphatidylcholine. *Biochim. Biophys.
1483 Acta, Biomembr.* **1983**, *735*, 131–136.
- 1484 (92) Ulrich, A. S.; Volke, F.; Watts, A. The dependence of
1485 phospholipid head-group mobility on hydration as studied by
1486 deuterium-NMR spin-lattice relaxation time measurements. *Chem.
1487 Phys. Lipids* **1990**, *55*, 61–66.
- 1488 (93) Ulrich, A.; Watts, A. Molecular response of the lipid headgroup
1489 to bilayer hydration monitored by 2H-NMR. *Biophys. J.* **1994**, *66*, 1489
1490 1441–1449.
- 1491 (94) Lairion, F.; Filler, R.; Disalvo, E. Reversed micelles as model
1492 systems to study interfacial properties of lipid bilayers. *Colloids Surf., B
1493 2002*, *25*, 369–371.
- 1494 (95) Semchyshyn, D. J.; Macdonald, P. M. Conformational
1495 response of the phosphatidylcholine headgroup to bilayer surface
1496 charge: torsion angle constraints from dipolar and quadrupolar
1497 couplings in biccles. *Magn. Reson. Chem.* **2004**, *42*, 89–104.
- 1498 (96) Doux, J. P. F.; Hall, B. A.; Killian, J. A. How lipid headgroups
1499 sense the membrane environment: an application of ¹⁴N NMR.
1500 *Biophys. J.* **2012**, *103*, 1245–1253.
- 1501 (97) Bildt, G.; Gally, H. U.; Seelig, A.; Seelig, J.; Zaccai, G. Neutron
1502 diffraction studies on selectively deuterated phospholipid bilayers.
1503 *Nature* **1978**, *271*, 182–184.
- 1504 (98) Botan, A.; Favela-Rosales, F.; Fuchs, P. F. J.; Javanainen, M.;
1505 Kanduc, M.; Kulig, W.; Lamberg, A.; Loison, C.; Lyubartsev, A.;
1506 Miettinen, M. S.; Monticelli, L.; Määttä, J.; Ollila, O. H. S.; Retegan,
1507 M.; Rög, T.; Santuz, H.; Tynkkynen, J. Toward Atomistic Resolution
1508 Structure of Phosphatidylcholine Headgroup and Glycerol Backbone
1509 at Different Ambient Conditions. *J. Phys. Chem. B* **2015**, *119*, 15075–
1510 15088.
- 1511 (99) Venable, R. M.; Ingólfsson, H. I.; Lerner, M. G.; Perrin, B. S.;
1512 Camley, B. A.; Marrink, S. J.; Brown, F. L. H.; Pastor, R. W. Lipid and
1513 Peptide Diffusion in Bilayers: The Saffman-Delbrück Model and
1514 Periodic Boundary Conditions. *J. Phys. Chem. B* **2017**, *121*, 3443–
1515 3457.
- 1516 (100) Vögele, M.; Köfinger, J.; Hummer, G. Hydrodynamics of
1517 Diffusion in Lipid Membrane Simulations. *Phys. Rev. Lett.* **2018**, *120*,
1518 No. 268104.
- 1519 (101) Lindblom, G.; Orädd, G. Lipid lateral diffusion and membrane
1520 heterogeneity. *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788*, 234–
1521 244.
- 1522 (102) Poger, D.; Mark, A. E. Lipid Bilayers: The Effect of Force
1523 Field on Ordering and Dynamics. *J. Chem. Theory Comput.* **2012**, *8*,
1524 4807–4817.
- 1525 (103) Piggot, T. J.; Piñeiro, Á.; Khalid, S. Molecular Dynamics
1526 Simulations of Phosphatidylcholine Membranes: A Comparative
1527 Force Field Study. *J. Chem. Theory Comput.* **2012**, *8*, 4593–4609.
- 1528 (104) Nagle, J. F. Experimentally determined tilt and bending
1529 moduli of single-component lipid bilayers. *Chem. Phys. Lipids* **2017**,
1530 205, 18–24.
- 1531 (105) Doktorova, M.; Harries, D.; Khelashvili, G. Determination of
1532 bending rigidity and tilt modulus of lipid membranes from real-space
1533 fluctuation analysis of molecular dynamics simulations. *Phys. Chem.
1534 Chem. Phys.* **2017**, *19*, 16806–16818.
- 1535 (106) Nagle, J. F.; Cognet, P.; Dupuy, F. G.; Tristram-Nagle, S.
1536 Structure of gel phase DPPC determined by X-ray diffraction. *Chem.
1537 Phys. Lipids* **2019**, *218*, 168–177.
- 1538 (107) Sun, W.-J.; Suter, R. M.; Knewton, M. A.; Worthington, C.
1539 R.; Tristram-Nagle, S.; Zhang, R.; Nagle, J. F. Order and disorder in
1540 fully hydrated unoriented bilayers of gel-phase dipalmitoylphospho-
1541 tidylycholine. *Phys. Rev. E* **1994**, *49*, 4665–4676.
- 1542 (108) McIntosh, T. J.; Simon, S. A. Area per molecule and
1543 distribution of water in fully hydrated dilauroylphosphatidylethanol-
1544 amine bilayers. *Biochemistry* **1986**, *25*, 4948–4952.

- 1545 (109) Marinov, R.; Dufourc, E. J. Thermotropicism and hydration
1546 properties of POPE and POPE-cholesterol systems as revealed by
1547 solid state ^{2}H and ^{3}P -NMR. *Eur. Biophys. J.* **1996**, *24*, 423–431.
1548 (110) Rand, R. P.; Fuller, N.; Parsegian, V. A.; Rau, D. C. Variation
1549 in hydration forces between neutral phospholipid bilayers: evidence
1550 for hydration attraction. *Biochemistry* **1988**, *27*, 7711–7722.
1551 (111) Hwang, P. M.; Zhou, N.; Shan, X.; Arrowsmith, C. H.; Vogel,
1552 H. J. Three-Dimensional Solution Structure of Lactoferricin B, an
1553 Antimicrobial Peptide Derived from Bovine Lactoferrin [†]. *Biochemistry*
1554 **1998**, *37*, 4288–4298.
1555 (112) Nguyen, L. T.; Schibli, D. J.; Vogel, H. J. Structural studies
1556 and model membrane interactions of two peptides derived from
1557 bovine lactoferricin. *J. Pept. Sci. A* **2005**, *11*, 379–389.
1558 (113) Schibli, D. J.; Hwang, P. M.; Vogel, H. J. The structure of the
1559 antimicrobial active center of lactoferricin B bound to sodium dodecyl
1560 sulfate micelles. *FEBS Lett.* **1999**, *446*, 213–217.
1561 (114) Sohlenkamp, C.; Geiger, O. Bacterial membrane lipids:
1562 diversity in structures and pathways. *FEMS Microbiol. Rev.* **2016**, *40*,
1563 133–159.
1564 (115) Ma, H.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan,
1565 A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S. Modeling
1566 Diversity in Structures of Bacterial Outer Membrane Lipids. *J. Chem.*
1567 *Theory Comput.* **2017**, *13*, 811–824.
1568 (116) Castelletto, V.; de Santis, E.; Alkassem, H.; Lamarre, B.;
1569 Noble, J. E.; Ray, S.; Bella, A.; Burns, J. R.; Hoogenboom, B. W.;
1570 Ryadnov, M. G. Structurally plastic peptide capsules for synthetic
1571 antimicrobial viruses. *Chem. Sci.* **2016**, *7*, 1707–1711.
1572 (117) Kandasamy, S. K.; Larson, R. G. Binding and insertion of α -
1573 helical anti-microbial peptides in POPC bilayers studied by molecular
1574 dynamics simulations. *Chem. Phys. Lipids* **2004**, *132*, 113–132.
1575 (118) Vivcharuk, V.; Tomberli, B.; Tolokh, I. S.; Gray, C. G.
1576 Prediction of binding free energy for adsorption of antimicrobial
1577 peptide lactoferricin B on a POPC membrane. *Phys. Rev. E* **2008**, *77*,
1578 No. 031913.
1579 (119) Wang, Y.; Schlamadinger, D. E.; Kim, J. E.; McCammon, J. A.
1580 Comparative molecular dynamics simulations of the antimicrobial
1581 peptide CM15 in model lipid bilayers. *Biochim. Biophys. Acta,*
1582 *Biomembr.* **2012**, *1818*, 1402–1409.
1583 (120) Strom, M. B.; Haug, B. E.; Rekdal, O.; Skar, M. L.; Stensen,
1584 W.; Svendsen, J. S. Important structural features of 15-residue
1585 lactoferricin derivatives and methods for improvement of antimicro-
1586 bial activity. *Biochem. Cell Biol.* **2002**, *80*, 65–74.
1587 (121) Chan, D. I.; Prenner, E. J.; Vogel, H. J. Tryptophan- and
1588 arginine-rich antimicrobial peptides: Structures and mechanisms of
1589 action. *Biochim. Biophys. Acta, Biomembr.* **2006**, *1758*, 1184–1202.
1590 (122) Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.;
1591 Zagrovic, B. A Systematic Framework for Molecular Dynamics
1592 Simulations of Protein Post-Translational Modifications. *PLoS*
1593 *Comput. Biol.* **2013**, *9*, No. e1003154.

Supporting Information:

Lipid head group parameterisation for GROMOS 54a8: a consistent approach with
protein force field description

Irene Marzuoli, Christian Margreitter and Franca Fraternali

Randall Centre for Cell and Molecular Biology, King's College London, London

List of Tables

1	Table of additional control simulations	7
2	Table of simulations for POPE bilayer	7
3	Table of compressibility values for PC lipids	8
4	Table of thickness values for PC lipids	8
5	Table of water molecules per lipid moiety for PC lipids	9
6	Table of computed properties of POPE bilayer	10
7	Table of water molecules per lipid moiety for POPE bilayer	10
8	Table of computed properties of POPG bilayer	10
9	Table of water molecules per lipid moiety for POPG bilayer	10
10	Comparison with previous parameters and simulations setup	11

List of Figures

1	Dipole potential	4
2	Time series of the area per lipid for a DPPC bilayer	12
3	Comparison of Area per Lipid for different time windows of the same simulation	13
4	Area per lipid for control simulations	14
5	Comparison of key properties across parameter sets for a DLPC bilayer	15
6	Comparison of key properties across parameter sets for a DMPC bilayer . . .	16
7	Comparison of key properties across parameter sets for a DOPC bilayer . . .	17
8	Comparison of key properties across parameter sets for a POPC bilayer . . .	18
9	Comparison of key properties across parameter sets for a DPPC bilayer . . .	19
10	Electron density profiles of bilayers for parameter set Chiu/54A7	20
11	Electron density profiles of bilayers for parameter set Chiu/54A8	21
12	Electron density profiles of bilayers for parameter set RM/54A8_v2	22
13	Electron density profiles of bilayers for parameter set RM/54A8_v3	23
14	Deuterium order parameter S_{CD} of bilayers for parameter set Chiu/54A8 . .	24
15	Deuterium order parameter S_{CD} of bilayers for parameter set RM/54A8_v2 .	25
16	Deuterium order parameter S_{CD} of bilayers for parameter set RM/54A8_v3 .	26
17	Hydration profiles for DLPC bilayer across parameter sets	27
18	Hydration profiles for DMPC bilayer across parameter sets	28
19	Hydration profiles for DOPC bilayer across parameter sets	29
20	Hydration profiles for POPC bilayer across parameter sets	30
21	Phosphocholine moieties SASA	31
22	Head group orientation, lipid tails comparison	32
23	Head group orientation, parameter sets comparison	33
24	Diffusion coefficient of DPPC in control simulations	34
25	Snapshot of DPPC patch partially in gel phase	35
26	Key properties of POPE bilayer for parameter set Chiu/54A8/CH0	36
27	Key properties of POPE bilayer for parameter set RM/54A8/CH0	37

28	Key properties of POPG bilayer for parameter set RM/54A8/CH0	38
29	Scheme of the 1LFC orientations chosen for the peptide-membrane simulations	39
30	1LFC orientation with respect to membrane plane along simulations	40

1 Supplementary results: dipole potential

For the five phospholipids and the five different parameters tested, we computed the dipole potential across the membrane, according to the procedure outlined in Section 2.5 of the paper.

The results are outlined in Figure 1 and show a striking difference between the profiles obtained with the Chiu partial charges set and the Reif-Margreitter one: in particular, while the former suggests local maxima of the dipole potential in the head regions, the latter results in a potential rising to its higher value in the center of the tails region.

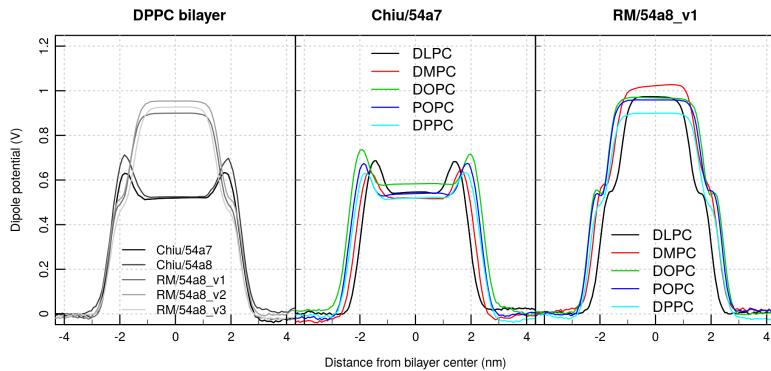


Figure 1: Dipole potential for respectively: (left) a DPPC bilayer, simulated with the five parameter sets explored in the manuscript; (center) the five different phospholipid considered and parameter set Chiu/54a7; (right) the five different phospholipid considered and parameter set RM/54a8_v1.

The results for the Chiu set of partial charges (and either the 54a7 or 54a8 GROMOS force field) are compatible with what obtained previously by Gurtenko et al. in Ref. [7] using the same charges and the Berger force field (a slight modification of GROMOS tailored for lipids simulations). On the contrary, the new results are qualitatively more similar to what obtained with the all atom force field CHARMM36 in Ref. [4] and [23], as they both suggest a peak of the potential in the lipid tail regions. However, with respect to the CHARMM36 profiles, the ones obtained by the RM/54a8_v1 parameters give a flatter trend in the center

of the bilayer. Moreover, the peaks relative to the phosphate regions are shallower for the RM/54a8_v1 simulations with respect to the CHARMM ones.

To discriminate between the two behaviours it is necessary a comparison with experimental results. Few works compute the profile of the dipole potential along the bilayer normal, focussing instead on measuring the difference between the dipole potential on the two sides of the bilayer [20, 19, 13]. Never the less, Wang et al. [22] measured its profile for ester-DPhPC lipid vesicles, suggesting a peak in the central region of the bilayer. This lipid is derived from DPPC by adding an ester linkage: as such, its profile will not be identical to the one produced by DPPC, but being the modification in the ester region only, we can assume that the qualitative trend remains the same and thus can be used for the present comparison.

Additionally to experimental data, theoretical profiles obtained from ab initio calculations and geometrical assumptions on the structure of lipids bilayers suggest indeed a plateau in the tail region [3] - or at the very least do not support the presence of peaks in the head regions, as showed when the Chiu charge set is used in the simulations.

For all the above reasons we once more consider the new set of parameter (RM/54a8_v1) as a better compromise than the previous sets Chiu/54a7 or Chiu/54a8.

Supplementary Movie 1: file LFC_54a8_v1_0.mp4

The movie displays the simulation of LFC on a bacterial membrane (composition DLPC:DLPG with 3:1 ratio) for both parameter set Chiu/54a8 (left half) and RM/54a8_v1 (right half), starting from orientation OII. The lipid choline and phosphate atoms are shown as van der Waals beads, the protein as cyan cartoon and bonds coloured by residue type (blue positive, green polar, white hydrophobic) except for the two Tryptophan residues which are coloured in orange.

Control simulations

Sim	Charges ^a	FF	Long range ES ^b	Nr lipids	T (K)
0	RM	54A7	PME	512	323
RF	Chiu	54A7	RF	512	323
Small	Chiu	54A7	PME	128	323
303	RM	54A8_v1	PME	512	303
333	RM	54A8_v1	PME	512	333

Table 1: Table of additional control simulations, run on the DPPC bilayer. ^a Charge set: Chiu from Ref [2], RM Reif-Margreitter as illustrated in the present work. ^b RF reaction field [21], PME Particle Mesh Ewald summation [6].

Simulations of POPE

Sim	Charges ^a	FF	Ester carbon type
1CH0	PC	54a8	CH0
CH0	RM	54a8	CH0

Table 2: Table of simulations for POPE. All are run for 500 ns and systems consisting of 512 lipid molecules (256 per layer), using a PME long range electrostatic scheme. ^a Charge set: PC Piggot-Chiu from Ref [14], RM Reif-Margreitter as illustrated in the present work.

 K_A (mN m^{-1})

ID	Charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	553	477	607	515	499
2	Chiu/54a8	471	500	471	679	612
3	RM/54a8_v1	562	354	485	397	329
4	RM/54a8_v2	365	318	518	466	406
5	RM/54a8_v3	332	300	426	345	376
Experiment ^a	234(23)	188-265(18)	231(20)	180-330	248(20)	

Table 3: Average isothermal area compressibility module K_A over the last 300 ns of simulations for phosphocholine bilayers. All simulations run at 303 K, except for DPPC (323 K). ^a Values from Ref. [9] and Table 1 in Ref. [15], in parenthesis the standard error on the last digit is shown.

Hydrophobic thickness D_{HH} (nm)

ID	Charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	2.83	3.59	3.05	3.30	4.30
2	Chiu/54a8	2.78	3.52	2.90	3.14	4.21
3	RM/54a8_v1	2.72	3.48	2.89	3.22	4.06
4	RM/54a8_v2	2.61	3.33	2.82	3.08	4.06
5	RM/54a8_v3	2.38	3.13	2.65	2.86	3.77
Experiment ^a	3.08	3.44-3.60	3.53-3.71	3.70	3.42-3.83	

Luzzati thickness D_B (nm)

ID	Charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	3.11	3.54	4.13	4.00	3.93
2	Chiu/54a8	3.14	3.61	3.97	3.96	3.96
3	RM/54a8_v1	3.04	3.48	3.94	3.88	3.75
4	RM/54a8_v2	3.00	3.34	3.87	3.77	3.74
5	RM/54a8_v3	2.80	3.21	3.62	3.57	3.50
Experiment ^a	3.14	3.63-3.96	3.59-3.87	3.68	3.50-3.83	

Table 4: Bilayer thickness, according to the Phosphate or Luzzati methods, derived from the electron density profiles of phosphocholine bilayers. All simulations run at 303 K, except for DPPC (323 K). ^a Values from Ref. [16] and Table 1 in Ref. [15]

Nr water molecules per group

Group	ID	DLPC	DMPC	DOPC	DPPC	POPC
All	1	3.9 (13.2)	3.8 (12.9)	3.9 (13.3)	3.7 (12.7)	3.8 (13.1)
	2	3.6 (12.7)	3.5 (12.4)	3.6 (12.9)	3.4 (12.3)	3.6 (12.6)
	3	4.3 (14.1)	4.3 (13.8)	4.4 (14.3)	4.2 (13.7)	4.3 (14.1)
	4	4.6 (15.1)	4.6 (15.0)	4.6 (15.1)	4.5 (14.7)	4.6 (15.1)
	5	5.1 (16.9)	5.1 (16.8)	5.0 (16.9)	4.9 (16.6)	5.1 (16.8)
Cho	1	13.1	12.8	13.3	12.6	13.1
	2	12.6	12.3	12.7	12.2	12.5
	3	14.3	14.1	14.5	13.9	14.3
	4	15.3	15.2	15.4	14.9	15.3
	5	17.0	16.9	17.0	16.7	17.0
PO4	1	3.1 (9.8)	3.0 (9.6)	3.1 (9.9)	3.1 (9.6)	3.1 (9.8)
	2	2.6 (9.3)	2.5 (9.0)	2.6 (9.4)	2.6 (9.2)	2.5 (9.2)
	3	4.0 (9.8)	4.0 (9.6)	4.0 (9.9)	3.9 (9.8)	4.0 (9.8)
	4	4.3 (10.3)	4.3 (10.3)	4.3 (10.4)	4.2 (10.3)	4.3 (10.4)
	5	4.8 (11.5)	4.7 (11.5)	4.8 (11.5)	4.7 (11.5)	4.8 (11.5)
Gly	1	0.3 (3.5)	0.3 (3.4)	0.3 (3.5)	0.3 (3.4)	0.3 (3.5)
	2	0.3 (3.4)	0.2 (3.2)	0.3 (3.4)	0.2 (3.3)	0.2 (3.3)
	3	0 (3.5)	0 (3.5)	0 (3.6)	0 (3.4)	0 (3.5)
	4	0 (3.8)	0 (3.8)	0 (3.8)	0 (3.7)	0 (3.8)
	5	0 (4.3)	0 (4.3)	0 (4.2)	0 (4.2)	0 (4.3)
CO1	1	1.4	1.4	1.4	1.4	1.4
	2	1.5	1.5	1.5	1.5	1.5
	3	1.2	1.2	1.2	1.1	1.2
	4	1.3	1.2	1.3	1.1	1.3
	5	1.3	1.3	1.3	1.3	1.3
CO2	1	0.5	0.5	0.5	0.5	0.5
	2	0.6	0.6	0.6	0.6	0.6
	3	0.7	0.6	0.7	0.6	0.7
	4	0.7	0.7	0.7	0.7	0.7
	5	0.8	0.8	0.8	0.8	0.8

Table 5: Integrated value of water molecules around each lipid and for their constituting chemical groups (Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions). The integration is performed up to the first peak and, if applicable, to the second (the latter of which is given in brackets). Peak positions are the same throughout the simulations, except for glycerol in ID3 to ID5, where the first peak is not shown (thus the value of 0, see main text for interpretation).

9

POPE (313 K) structural properties

ID	Charges/FF/ester	ApL (nm ²)	K _A (mN m ⁻¹)	D _{HH} (nm)	D _B (nm)
1CH0	Chiu/54a8/CHO	0.591(5)	437	2.48	3.92
CH0	RM/54a8/CHO	0.595(5)	361	2.52	3.89
Experiment ^a		0.566	—	—	4.13

Table 6: Area per lipid, lateral compressibility and bilayer thickness, according to the phosphate or Luzzati methods, derived from the electron density profiles for POPE bilayer. ^a From [17].

POPE (313 K) Nr water molecules per group

ID	Charges/FF/ester	All	Am	PO4	Gly	CO1	CO2
1CH0	Chiu/54a8/CHO	1.6 (5.7)	2.2	3.3 (10.6)	0 (3.4)	0.8	0.5
CH0	RM/54a8/CHO	2.1 (6.4)	2.5	4.2 (11.3)	0 (3.6)	0.6	0.6

Table 7: Number of water molecules per lipid hydrating the POPE bilayer and its different polar moieties (Am: amine, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions). Integration was performed up to the first peak and when relevant to the second one (value in bracket). Peak positions are the same for both simulations.

POPG (303 K) structural properties

ID	Charges/FF/ester	ApL (nm ²)	K _A (mN m ⁻¹)	D _{HH} (nm)	D _B (nm)
CH0	RM/54a8/CHO	0.646(7)	198	3.76	3.98
Experiment ^a		0.661(13)	—	3.74(7)	—

Table 8: Area per lipid, lateral compressibility and bilayer thickness, according to the phosphate or Luzzati methods, derived from the electron density profiles. ^a From [11].

POPG (303 K) Nr water molecules per group

ID	Charges/FF/ester	All	Etam	PO4	Gly	CO1	CO2
H0	RM/54a8/CHO	0.2 (3.7)	0.5 (7.8)	3.9	3.2 (6.2)	0.5	0.4

Table 9: Number of water molecules per lipid hydrating the POPG bilayer and its different polar moieties (Etam: etamine, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions). Integration was performed up to the first peak, and when relevant to the second one (value in bracket). Peak positions are the same for both simulations.

10

Comparison with the previous parameterization work

	Poger et al. [16, 15]	Present work	
Lipid parameters			
Choline partial charges	Chiу [2]	Reif [18]	
Phosphate partial charges	Chiу [2]	Margreitter [8]	
CH3p-OM C12 [$\text{kJ mol}^{-1}\text{nm}^{12}$]	$1.58E-5$ [16]	$1.10E-5$ Present work	
CH3p-CH3p C12 [$\text{kJ mol}^{-1}\text{nm}^{12}$]	$2.66E-5$ [10]	$6.48E-5$ [18]	
Simulations details			
GROMACS version	3.2.1	2016.3	[1]
Long range electrostatics	Reaction Field [21]	PME [6]	
Cut-off ^a	Twin-range	Single	
Short range cut-off ^a	0.8 nm	NA	
Long range cut-off	1.4 nm	1.4 nm	
System size (# of lipids)	128	512	
Sampling time	120 ns	300 ns	
# replicas	2	1	

Table 10: Table of the changes occurred between simulations run in Poger et al. [16, 15] and the present work, in terms of parameters used and simulation conditions. ^a Settings not reproducible with GROMACS version 2016.3 used in the present work.

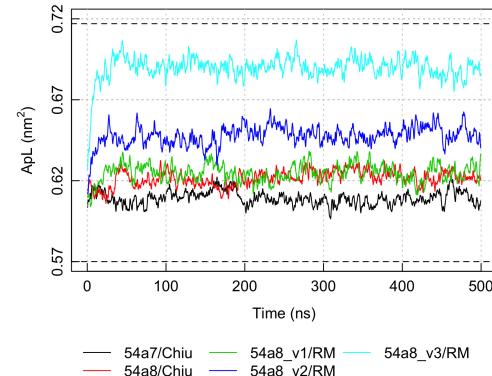


Figure 2: Time series of the area per lipid for a DPPC bilayer, for simulation ID 1 to 5 (see Table 1 for details).

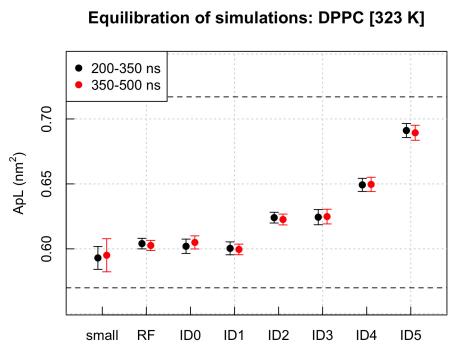


Figure 3: Area per lipid (computed from the simulation box sides) of a DPPC bilayer, for the five test simulation (ID 1 to 5 in Table 1) and the control ones (SI Table 1), computed for two non overlapping time intervals.

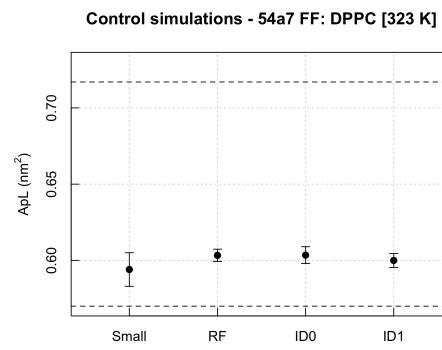


Figure 4: Area per lipid (computed from the simulation box sides) for a DPPC bilayer, for the control simulations as described in SI Table 1). Simulation ID1 (see Table 1) has been included for comparison, as each control simulation (small, RF, ID0) differs from ID1 by only one property (either the system size, long-range electrostatic treatment or the head group partial charges).

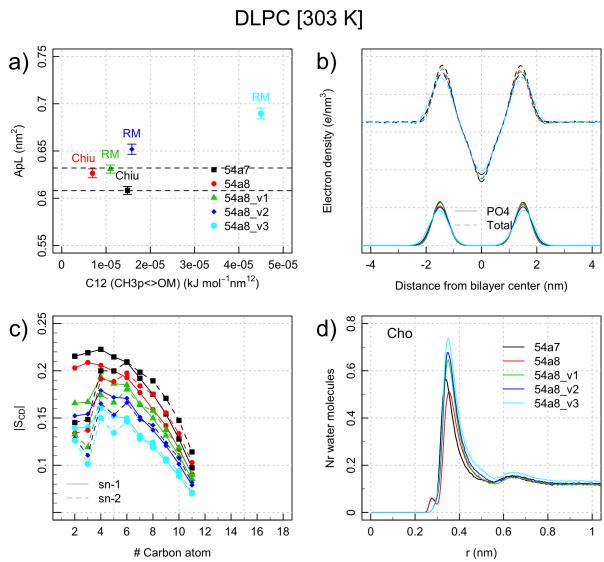


Figure 5: Area per lipid (a), electron density of CH₂ tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline atom for a DLPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

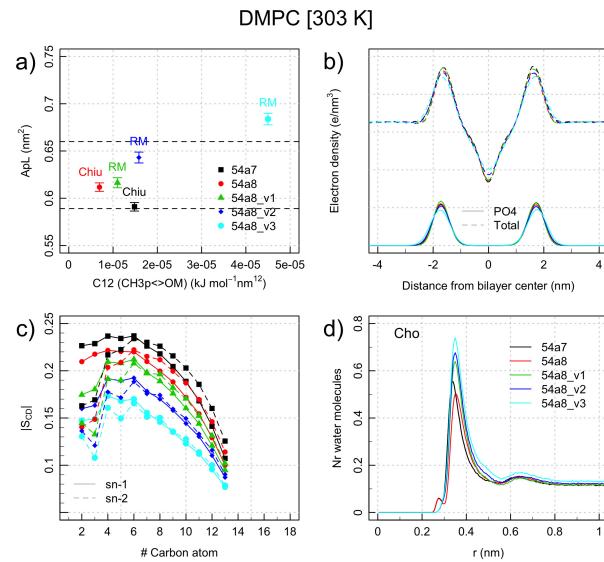


Figure 6: Area per lipid (a), electron density of CH₂ tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a DMPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

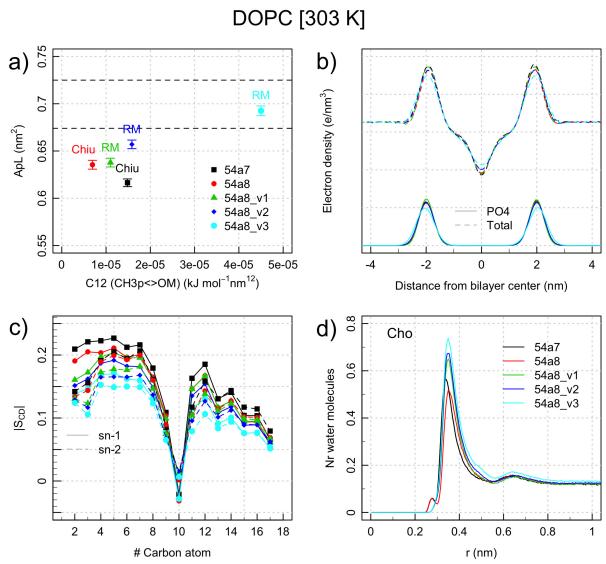


Figure 7: Area per lipid (a), electron density of CH₂ tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a DOPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

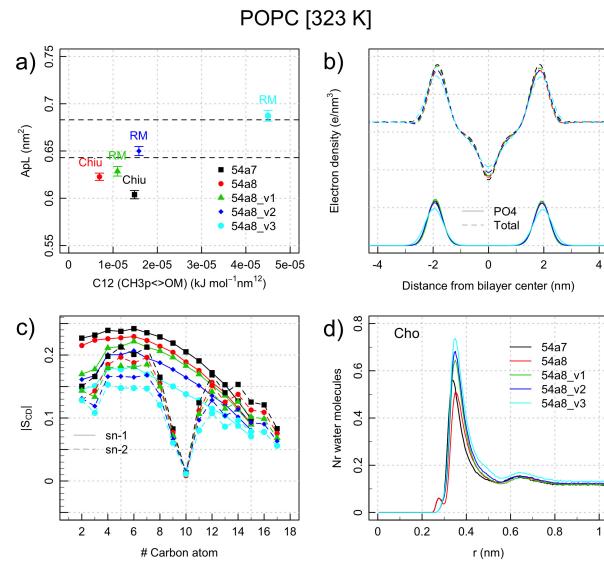


Figure 8: Area per lipid (a), electron density of CH₂ tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a POPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

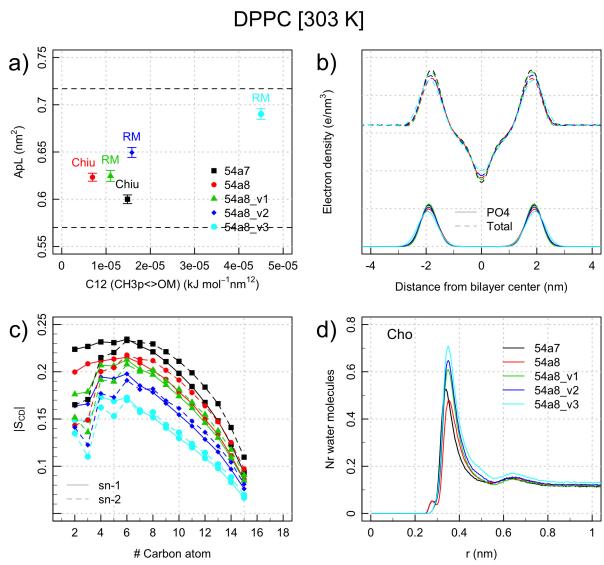


Figure 9: Area per lipid (a), electron density of CH₂ tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a DPPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

19

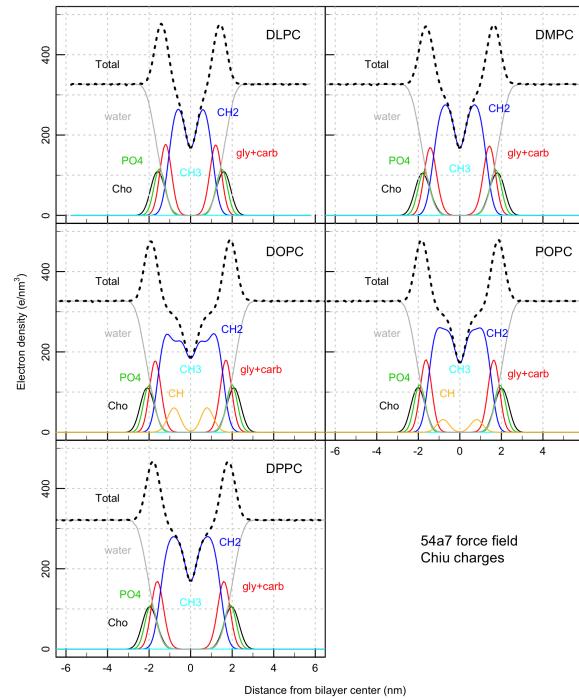


Figure 10: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID1 (54A7 force field, Chiu charges).

20

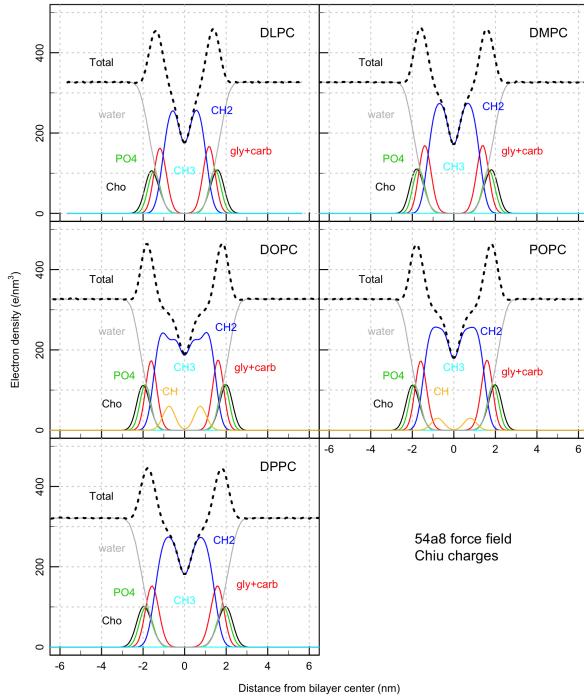


Figure 11: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID2 (54A8 force field, Chiu charges).

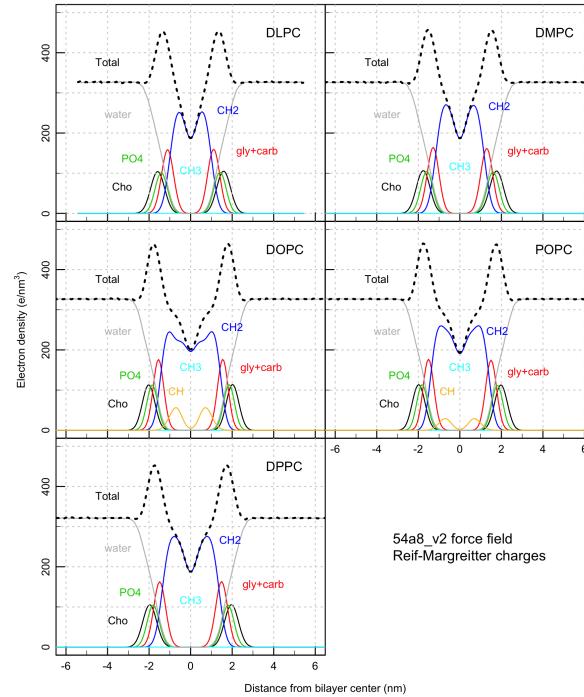


Figure 12: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID4 (54A8_v2 force field, Reif-Margreitter charges).

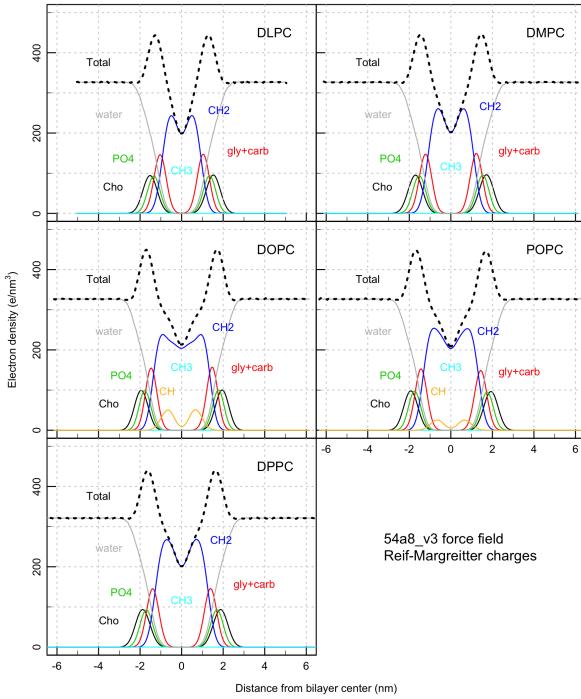


Figure 13: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID5 (54A8_v3 force field, Reif-Margreitter charges).

23

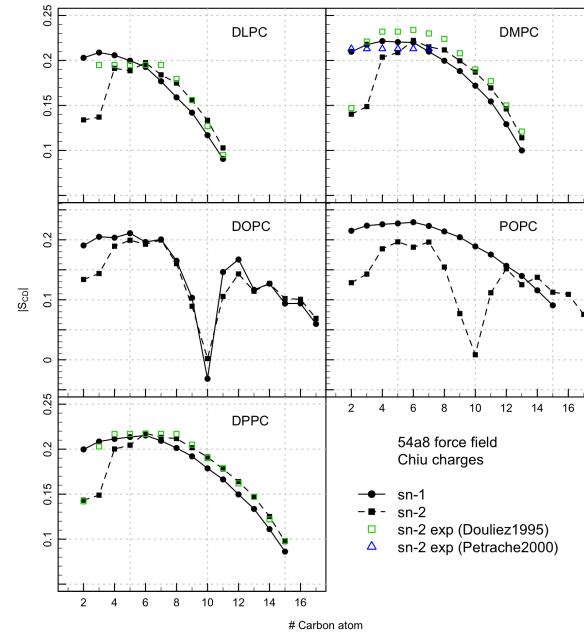


Figure 14: Deuterium order parameter S_{CD} profiles of the sn-1 and sn-2 fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers calculated from simulations ID2 (54A8 force field, Chiu charges). The S_{CD} values are averaged over all the lipid sn-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values: Douliez1995 from Ref. [5], Petrache2000 from Ref. [12].

24

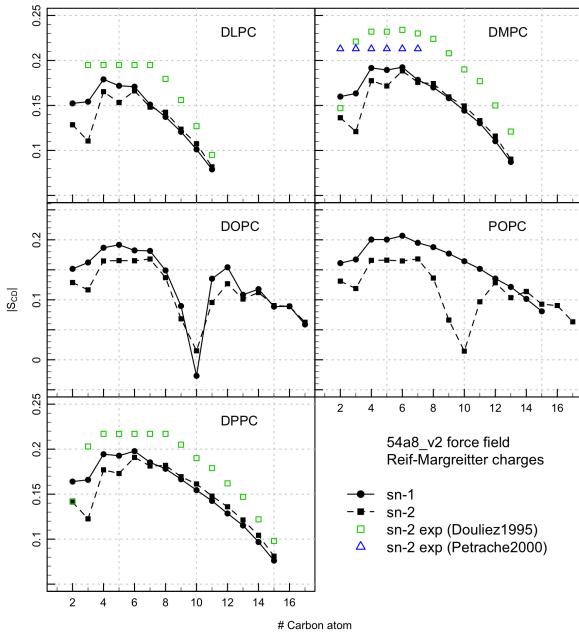


Figure 15: Deuterium order parameter S_{CD} profiles of the sn-1 and sn-2 fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers calculated from simulations ID4 (54A₈_v2 force field, Reif-Margreitter charges). The S_{CD} values are averaged over all the lipid sn-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values: Douliez1995 from Ref. [5]. Petrache2000 from Ref. [12].

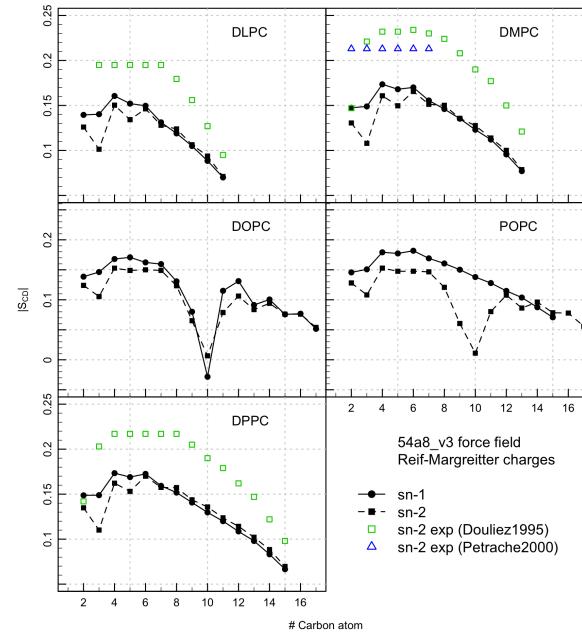


Figure 16: Deuterium order parameter S_{CD} profiles of the sn-1 and sn-2 fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers calculated from simulations ID5 (54A₈_v3 force field, Reif-Margreitter charges). The S_{CD} values are averaged over all the lipid sn-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values: Douliez1995 from Ref. [5]. Petrache2000 from Ref. [12].

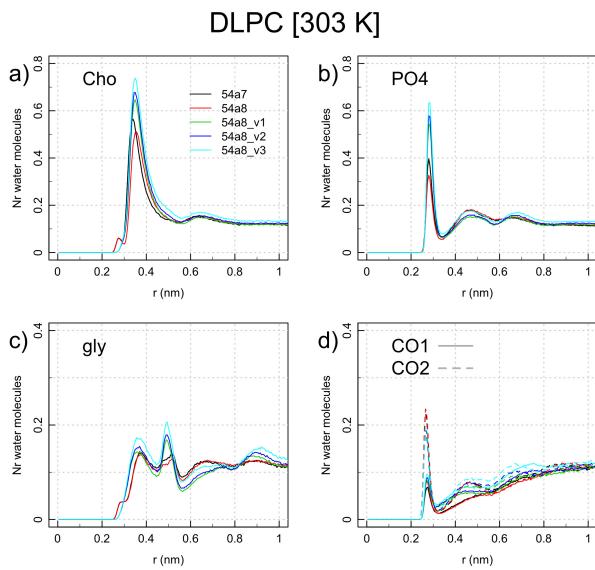


Figure 17: Distribution of the distance between the water oxygen and the nearest lipid head-group atom in given moieties for simulations of a DLPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

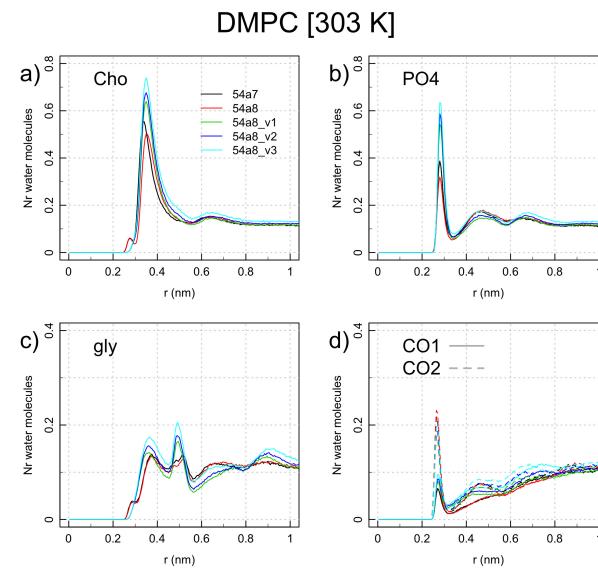


Figure 18: Distribution of the distance between the water oxygen and the nearest lipid head-group atom in given moieties for simulations of a DMPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

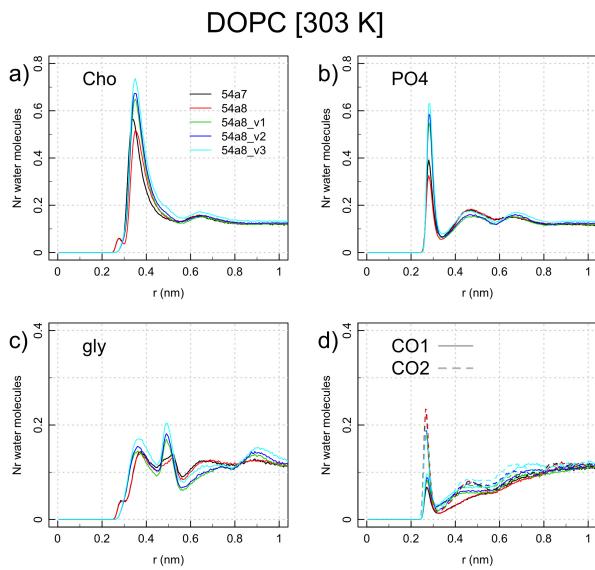


Figure 19: Distribution of the distance between the water oxygen and the nearest lipid head-group atom in given moieties for simulations of a DOPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

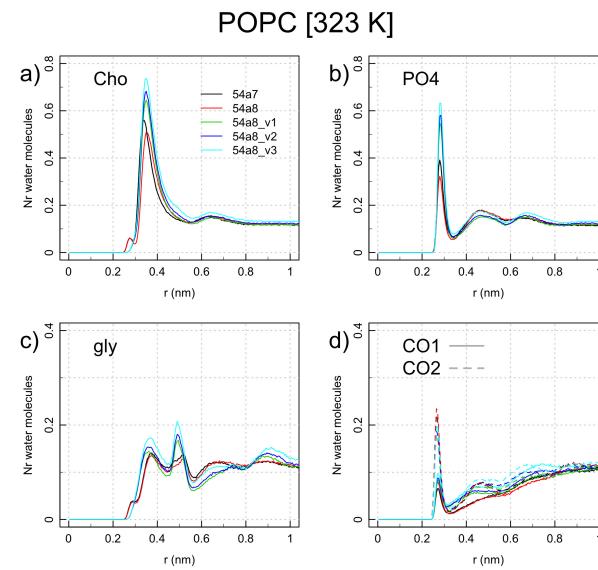


Figure 20: Distribution of the distance between the water oxygen and the nearest lipid head-group atom in given moieties for simulations of a POPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

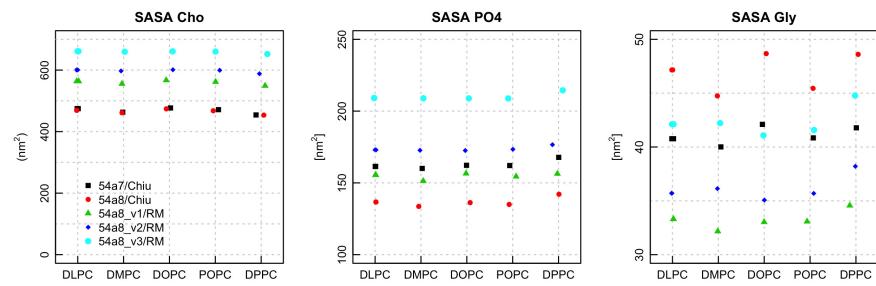


Figure 21: Values of SASA for Choline, Phosphate and Glycerol moieties for all the five phosphocholine and each of the parameter sets simulated.

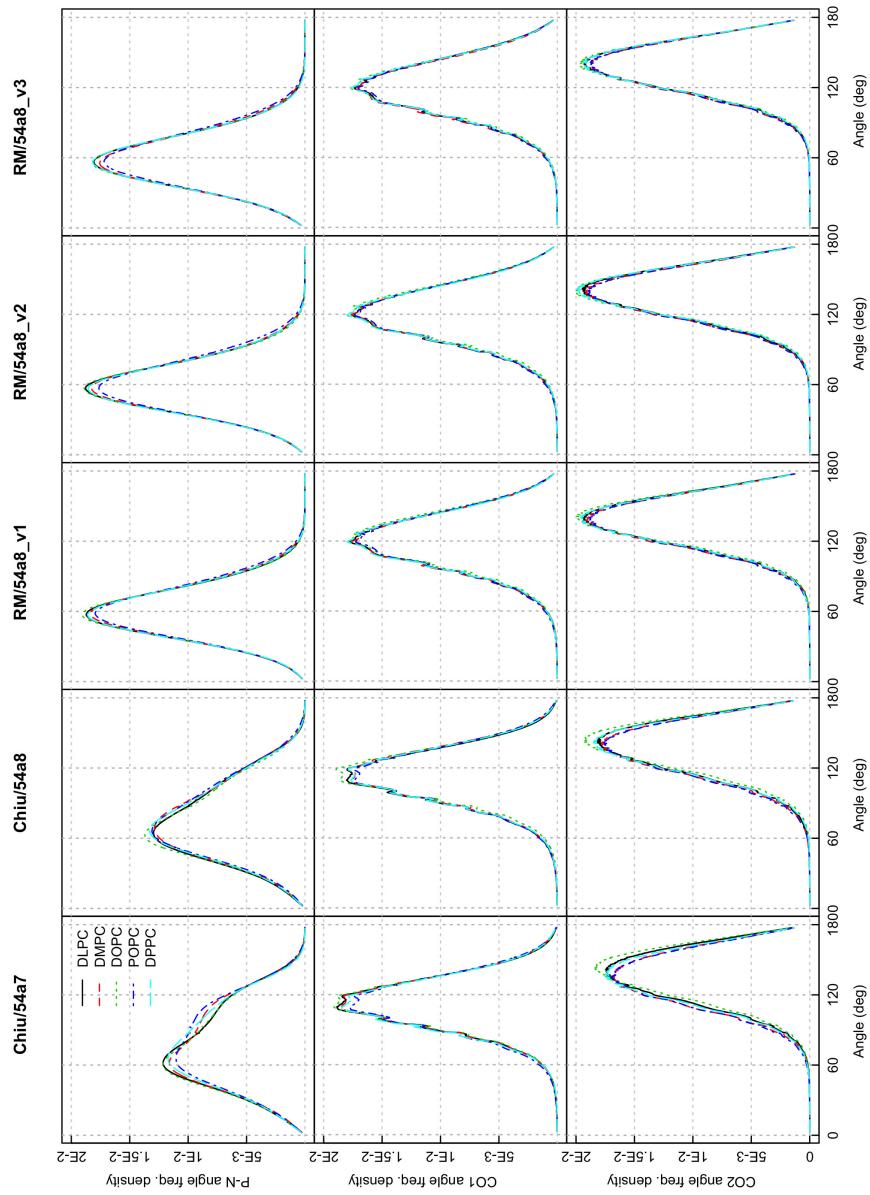


Figure 22: Distribution of the P-N, CO1 and CO2 angles with respect to the outward bilayer normal across different lipids for simulations ID 1 to 5.

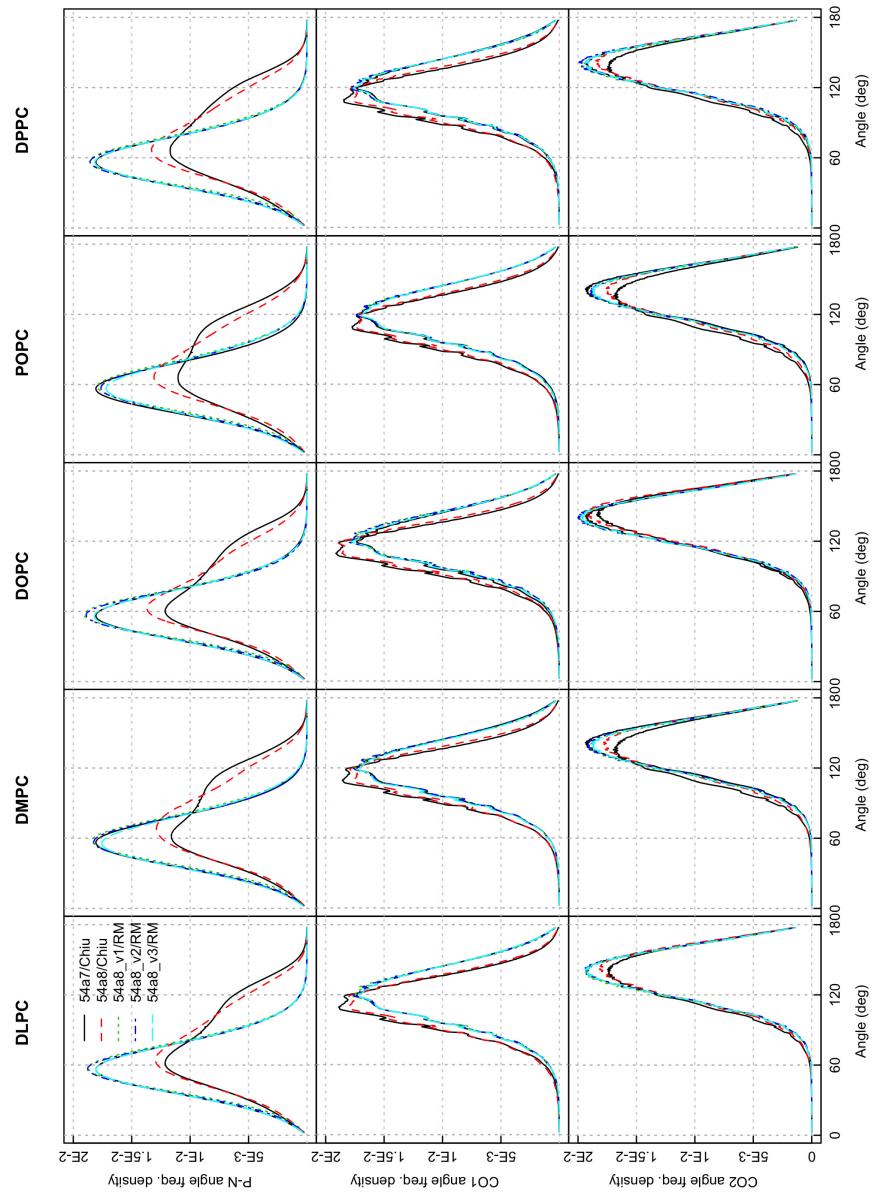


Figure 23: Distribution of the P-N, CO1 and CO2 angles with respect to the outward bilayer normal across different simulations for each lipid.

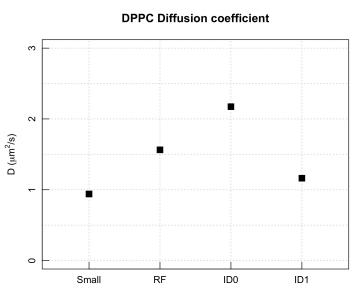


Figure 24: Lateral diffusion coefficient for DPPC in the control simulations as per SI Table 1.

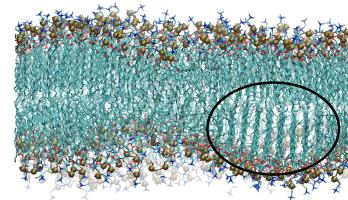


Figure 25: Snapshot after 400 ns simulation of a DPPC patch at 303 K with the 54a8_v1 parameter set. In the circle a detail of the ordered lipid tails which have already transitioned in a gel phase.

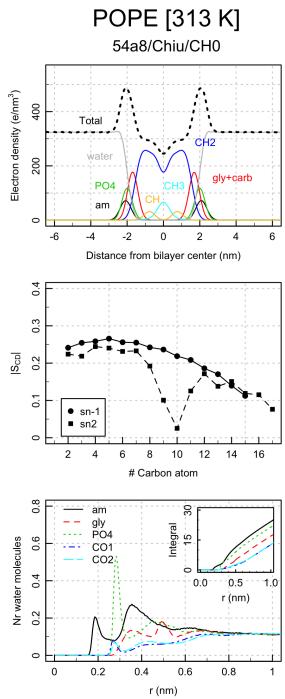


Figure 26: Properties of POPE bilayer simulated with the 54A8 force field, Chiu charges and CH0 esters. Top: electron density profiles (Total: whole bilayer, am: amide, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of acyl chains, CH: CHdCH groups of oleyl chains, CH3: terminal methyls); middle: deuterium order parameter S_{CD} of the sn-1 and sn-2 fatty acyl chains (proS hydrogen only); bottom: distribution of the distance between the water oxygen and the nearest lipid headgroup atom - insets: integral of the distribution (am: amine, PO4: phosphate, gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions).

36

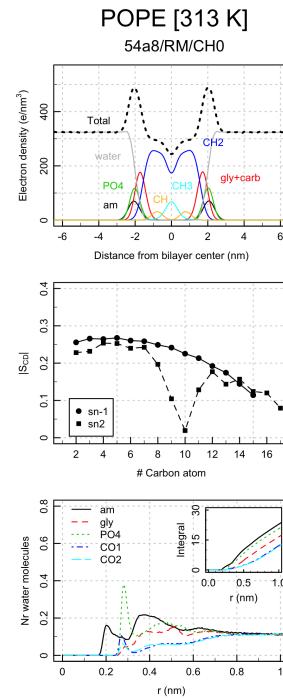


Figure 27: Properties of POPE bilayer simulated with the 54A8 force field, RM charges and CH0 esters. Top: electron density profiles (Total whole bilayer, am: amide, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2 methylenes of acyl chains, CH CHdCH groups of oleyl chains, CH3 terminal methyls); middle: deuterium order parameter S_{CD} of the sn-1 and sn-2 fatty acyl chains (proS hydrogen only); bottom: distribution of the distance between the water oxygen and the nearest lipid headgroup atom - insets: integral of the distribution (am: amine, PO4: phosphate, gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions).

37

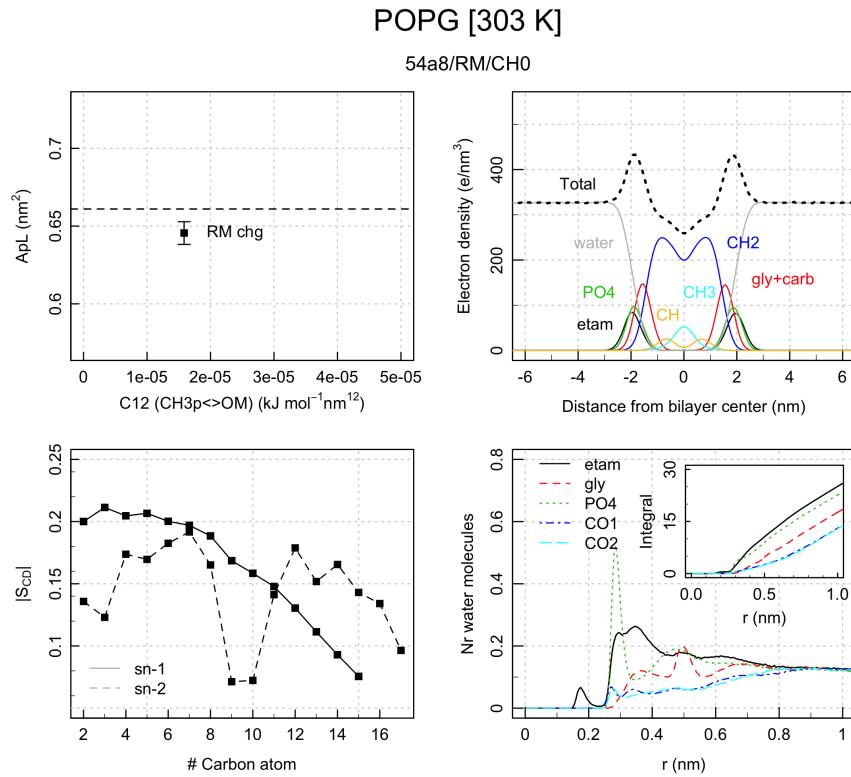


Figure 28: Properties of POPG bilayer simulated with the 54A8 force field, RM charges and CH0 esters. Top left: area per lipid and experimental value (dashed line); top right: electron density profiles (Total: whole bilayer, etam: etamine, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of acyl chains, CH: CHdCH groups of oleoyl chains, CH3: terminal methyls); bottom left: deuterium order parameter S_{CD} of the sn-1 and sn-2 fatty acyl chains (proS hydrogen only); bottom right: distribution of the distance between the water oxygen and the nearest lipid headgroup atom - insets: integral of the distribution (etam: etamine, PO4: phosphate, gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions).

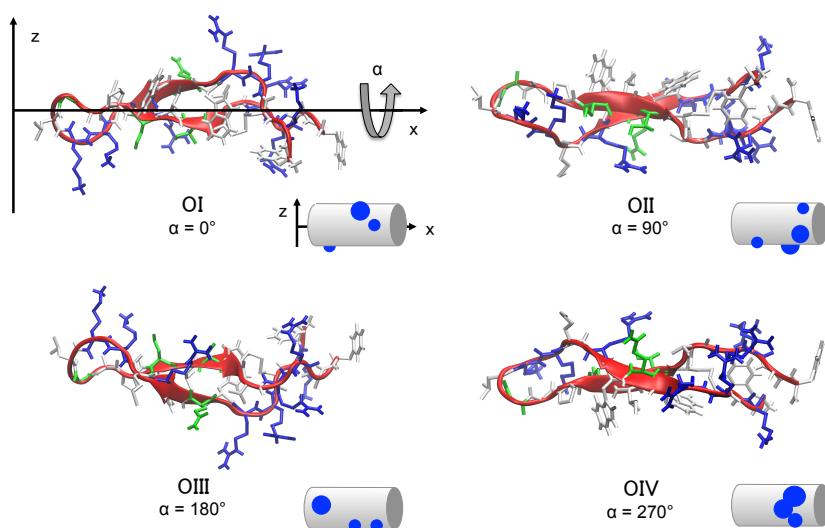


Figure 29: Four different orientations of the peptide 1LFC (bovine lactoferricin) used to sample different initial conditions for the peptide-membrane simulations. OII, OIII and OIV are obtained from OI with an anticlockwise rotation along the main axis (shown as x-axis in OI) of respectively 90° , 180° , 270° . The membrane plane is placed parallel to the main axis of 1LFC; the residues displayed in the bottom are the ones facing the lipids. Backbone red cartoon representation and whole peptide bonds representation coloured by residue type. The insets shows a cartoon representation of the portion of LFC facing the membrane, with blue dots showing the positive patches.

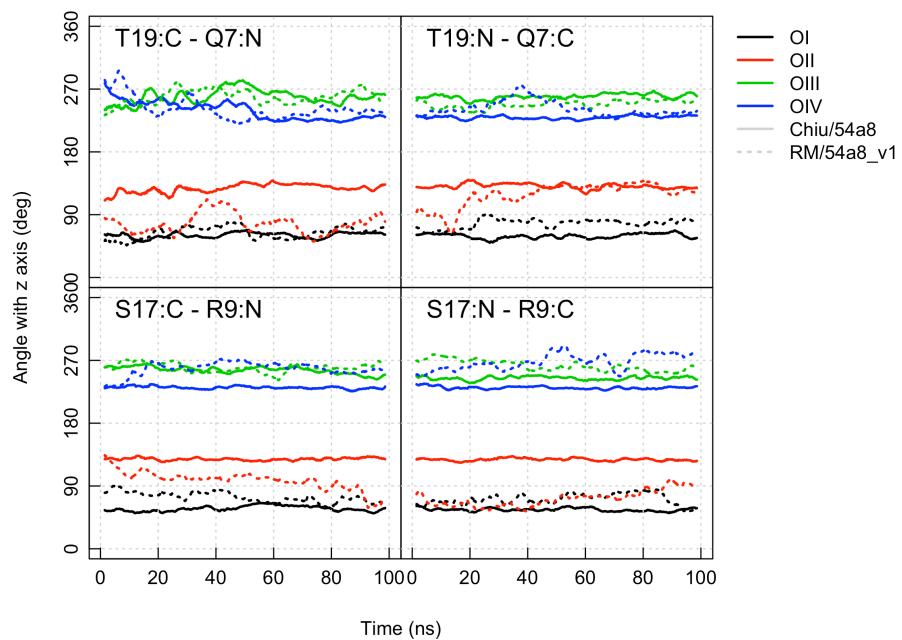


Figure 30: Angle with respect to the z-axis formed by the β -sheet plane. The backbone of amino acids 19 and 17 form hydrogen bonds with the backbone of amino acids 7 and 9 respectively. The vectors connecting the heavy atoms bonded to the O-H pairs forming them are taken as reference to compute the orientation. Solid lines refer to parameter set Chiu/54a8 and dashed ones to set RM/54a8_v1.

References

- [1] M J Abraham et al. *GROMACS User Manual version 2016*. www.gromacs.org, 2018.
- [2] S.-W. Chiu et al. “Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane.” In: *Biophysical journal* 69.4 (1995), pp. 1230–45. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(95)80005-6.
- [3] R J Clarke. “The dipole potential of phospholipid membranes and methods for its detection”. In: *Advances in Colloid and Interface Science* 89-90 (2001), pp. 263–281. ISSN: 0001-8686. DOI: 10.1016/S0001-8686(00)00061-0.
- [4] W Ding et al. “Effects of Lipid Composition on Bilayer Membranes Quantified by All-Atom Molecular Dynamics”. In: *The Journal of Physical Chemistry B* 119.49 (2015), pp. 15263–15274. ISSN: 1520-6106. DOI: 10.1021/acs.jpcb.5b06604.
- [5] J P Douliez, A Léonard, and E J Dufourc. “Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splittings.” In: *Biophysical journal* 68.5 (1995), pp. 1727–39. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(95)80350-4.
- [6] U Essmann et al. “A smooth particle mesh Ewald method”. In: *The Journal of Chemical Physics* 103.19 (1995), pp. 8577–8593. ISSN: 0021-9606. DOI: 10.1063/1.470117.
- [7] A A Gurtovenko and I Vattulainen. “Calculation of the electrostatic potential of lipid bilayers from molecular dynamics simulations: Methodological issues”. In: *The Journal of Chemical Physics* 130.21 (2009), p. 215107. ISSN: 0021-9606. DOI: 10.1063/1.3148885.
- [8] C Margreitter, M M Reif, and C Oostenbrink. “Update on phosphate and charged post-translationally modified amino acid parameters in the GROMOS force field”. In: *Journal of Computational Chemistry* 38.10 (2017), pp. 714–720. ISSN: 01928651. DOI: 10.1002/jcc.24733.
- [9] J F Nagle and S Tristram-Nagle. “Structure of lipid bilayers.” In: *Biochimica et biophysica acta* 1469.3 (2000), pp. 159–95. ISSN: 0006-3002.

- [10] C Oostenbrink et al. “A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6”. In: *Journal of Computational Chemistry* 25.13 (2004), pp. 1656–1676. ISSN: 1096-987X. DOI: 10.1002/jcc.20090. URL: <http://dx.doi.org/10.1002/jcc.20090>.
- [11] J Pan et al. “Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering”. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1818.9 (2012), pp. 2135–2148. ISSN: 00052736. DOI: 10.1016/j.bbamem.2012.05.007.
- [12] H I Petrache, S W Dodd, and M F Brown. “Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by (2)H NMR spectroscopy.” In: *Biophysical journal* 79.6 (2000), pp. 3172–92. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(00)76551-9.
- [13] A D Pickar and R Benz. “Transport of oppositely charged lipophilic probe ions in lipid bilayer membranes having various structures”. In: *The Journal of Membrane Biology* 44.3-4 (1978), pp. 353–376. ISSN: 0022-2631. DOI: 10.1007/BF01944229.
- [14] T J Piggot, D A Holdbrook, and S Khalid. “Electroporation of the E. coli and S. Aureus Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes”. In: *The Journal of Physical Chemistry B* 115.45 (2011), pp. 13381–13388. ISSN: 1520-6106. DOI: 10.1021/jp207013v.
- [15] D Poger and A E Mark. “On the Validation of Molecular Dynamics Simulations of Saturated and cis -Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment”. In: *Journal of Chemical Theory and Computation* 6.1 (2010), pp. 325–336. ISSN: 1549-9618. DOI: 10.1021/ct900487a.
- [16] D Poger, W F Van Gunsteren, and A E Mark. “A new force field for simulating phosphatidylcholine bilayers”. In: *Journal of Computational Chemistry* 31.6 (2010), pp. 1117–1125. ISSN: 01928651. DOI: 10.1002/jcc.21396.

- [17] R P Rand et al. “Variation in hydration forces between neutral phospholipid bilayers: evidence for hydration attraction”. In: *Biochemistry* 27.20 (1988), pp. 7711–7722. ISSN: 0006-2960. DOI: [10.1021/bi00420a021](https://doi.org/10.1021/bi00420a021).
- [18] M M Reif, P H Hünenberger, and C Oostenbrink. “New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field”. In: *Journal of Chemical Theory and Computation* 8.10 (2012), pp. 3705–3723. ISSN: 1549-9618. DOI: [10.1021/ct300156h](https://doi.org/10.1021/ct300156h).
- [19] J Schamberger and R J Clarke. “Hydrophobic Ion Hydration and the Magnitude of the Dipole Potential”. In: *Biophysical Journal* 82.6 (2002), pp. 3081–3088. ISSN: 0006-3495. DOI: [10.1016/S0006-3495\(02\)75649-X](https://doi.org/10.1016/S0006-3495(02)75649-X).
- [20] T Starke-Peterkovic and R J Clarke. “Effect of headgroup on the dipole potential of phospholipid vesicles”. In: *European Biophysics Journal* 39.1 (2009), pp. 103–110. ISSN: 0175-7571. DOI: [10.1007/s00249-008-0392-y](https://doi.org/10.1007/s00249-008-0392-y).
- [21] I G Tironi et al. “A generalized reaction field method for molecular dynamics simulations”. In: 102.13 (1995), pp. 5451–5459. ISSN: 0021-9606. DOI: [10.1063/1.469273](https://doi.org/10.1063/1.469273).
- [22] L Wang, P S Bose, and F J Sigworth. “Using cryo-EM to measure the dipole potential of a lipid membrane”. In: *Proceedings of the National Academy of Sciences* 103.49 (2006), pp. 18528–18533. ISSN: 0027-8424. DOI: [10.1073/PNAS.0608714103](https://doi.org/10.1073/PNAS.0608714103).
- [23] D T Warshaviak, M J Muellner, and M Chachisvilis. “Effect of membrane tension on the electric field and dipole potential of lipid bilayer membrane”. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1808.10 (2011), pp. 2608–2617. ISSN: 00052736. DOI: [10.1016/j.bbamem.2011.06.010](https://doi.org/10.1016/j.bbamem.2011.06.010).

4.1 Additional material: effects of membrane undulation

During the revision process of the paper which constitute this Chapter, it was brought to our attention that membrane undulations might be incompatible with the computation of ApL performed from the sides of the simulation box. The more a membrane patch is deviating from a flat geometry, the more approximate would be such method. Moreover, as the electron density profile is computed as an average over all the x and y positions, undulations of the membrane would result in a broader profile, with more uncertainty on the exact positions of the electron density peaks.

A brief comment on this appears on Section 2.5.1 of the paper, granting that the procedure chosen was suitable for the analysis of the simulations presented, and we would like to provide additional evidence to strengthen the case. This analysis supports as well the choice of using the box sides approach in Chapter 3 for simulation without or with low electric field, switching to a computation which takes into account undulations only for the highly curved membranes obtained with a high electric field.

The computation of ApL accounting for the membrane curvature is performed using a Fourier approach according to the work of Braun et al. [2011] and using the software made available from the same publication. This allows to obtain the ratio between the “true” area per lipid (i.e. computed taking into account the undulations of the membrane) and the projected one (i.e. from the box sides approach). Figure 4.1 reports the ratio $\text{ApL}_{\text{und}}/\text{ApL}_{\text{proj}}$ for each lipid and each parameter set: the values span between 1.0020 and 1.0046, which translates into an ApL correction between 0.20 and 0.46%, or, in absolute value, 0.001-0.003 nm². These discrepancies are lower than the standard deviation for any of the area per lipid computed, showing that the more accurate computation has no consequence for the purpose of evaluating the ApL and comparing parameter sets. The parameter set showing less discrepancy is Chiu/54a7, compatible with the fact that on average it produces smaller area per lipid: a more packed bilayer results in less freedom for the lipids to undulate.

Regarding the electron density, to check the influence of the averaging procedure over different locations of the patch, we computed the electron density for the phosphorus atoms on the full patch (256 lipid per leaflet), on a

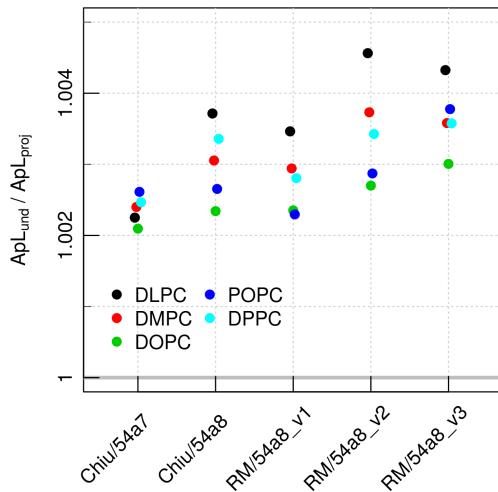


Figure 4.1: Ratio between area per lipid computed taking into account the undulations of the membrane (according to the procedure devised in Braun et al. [2011]) and the one computed from the projection on the plane parallel to the membrane. A value of 1 would be obtained for a perfectly flat membrane.

medium patch of roughly one forth in dimension (60 lipids per leaflet, see Figure 4.2(b) for a scheme), and finally a small 12 lipids per leaflet patch. It must be noticed that the GROMACS software, while computing the density perpendicular to the bilayer plane (presently along the z axis), centres the patch analysed around its center of mass, at each time step. This holds for GROMACS versions 5 onward, an version 2016.3 was used for all the analyses presented.

The resulting densities from the three different patches described above are normalised by the number of lipids present in each and compared. Figure 4.2(a) shows that the profiles for the full and medium patches are almost identical, while the smaller patch produces narrower peaks. Indeed, computing the full width at half maximum (FWHM) of these peaks, the FWHM of the small patch peaks is 16% smaller than the FWHM of the large patch peaks. However the distance between phosphorus peaks differ between patches by around 0.1 Å only, showing that the undulations have an effect mainly (or only) on the peaks broadness. As such, the hydrophobic thickness (measured as the distance between the phosphorous peaks) is virtually unvaried, and the difference imputable to the computing procedure is smaller than both the experimental thickness error and the differences arising from the use of different parameter

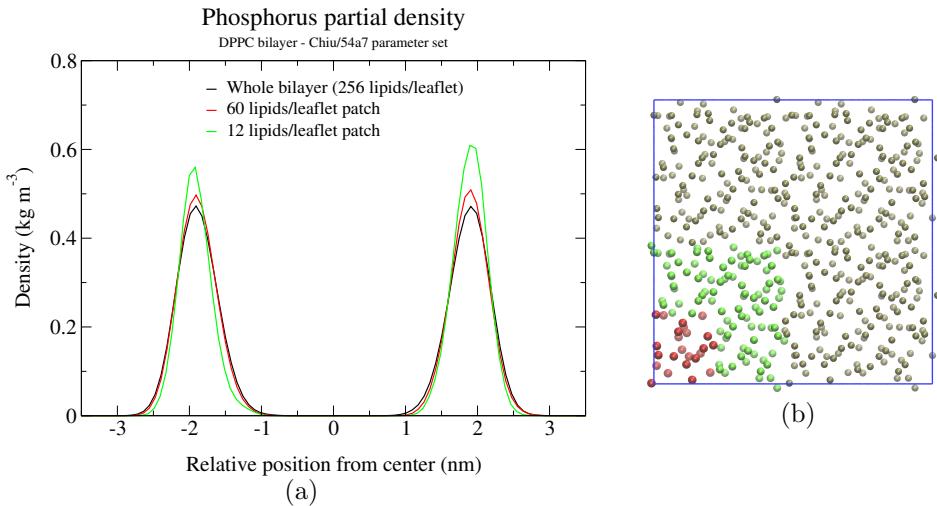


Figure 4.2: (a) Electron density of the phosphorus atom computed on the full membrane, a medium and a small patch. (b) van der Waals representation of the phosphorus atoms of the DPPC lipid bilayer (initial configuration) used for simulations. The simulation box is highlighted in blue. The green plus red regions constitute the medium patch mentioned in the analysis; the red one constitute the small patch.

sets. Further analysis must be done on the Luzzati thickness estimate, which depends on the broadness of the water profile. However, it must be noticed that also for this measure, the range of experimental results is quite broad and thus, as for many properties assessed in the paper, likely to be larger than the differences which can derive from the analysis procedure.

Appendices

A.1 On the derivation of the GP predictive distribution

This appendix gives a sketch of the procedure by which Eq. (??) is obtained, which substantially relies on the properties of multivariate Gaussian distributions. For full details on this one can consult the excellent Refs. [?] and [?].

Bibliography

- (2019). Carbon Nanotube. https://www.123rf.com/photo_30155287_carbon-nanotube-isolated-on-white-background.html.
- Abi Mansour, A. and Ortoleva, P. J. (2014). Multiscale Factorization Method for Simulating Mesoscopic Systems with Atomic Precision. *Journal of Chemical Theory and Computation*, 10(2):518–523.
- Abouzeed, Y. M., Baucheron, S., and Cloeckaert, A. (2008). ramR mutations involved in efflux-mediated multidrug resistance in *Salmonella enterica* serovar Typhimurium. *Antimicrobial agents and chemotherapy*, 52(7):2428–34.
- Abraham, E. P. and Chain, E. (1988). An enzyme from bacteria able to destroy penicillin. 1940. *Reviews of infectious diseases*, 10(4):677–8.
- Abraham, M. J., Murtola, T., Schulz, R., et al. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25.
- Abraham, M. J., van der Spoel, D., Lindahl, E., et al. (2018). *GROMACS User Manual version 2016*. www.gromacs.org.
- Aguilera, O., Ostolaza, H., Quirós, L., et al. (1999). Permeabilizing action of an antimicrobial lactoferricin-derived peptide on bacterial and artificial membranes. *FEBS Letters*, 462(3):273–277.
- Ahmadi, S., Barrios Herrera, L., Chehelamirani, M., et al. (2018). Multiscale modeling of enzymes: QM-cluster, QM/MM, and QM/MM/MD: A tutorial review. *International Journal of Quantum Chemistry*, 118(9):e25558.

- Alom, M. Z., Taha, T. M., Yakopcic, C., et al. (2019). A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics*, 8(3):292.
- Andersen, H. C. (1980). Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393.
- Anézo, C., de Vries, A. H., Höltje, H.-D., et al. (2003). Methodological Issues in Lipid Bilayer Simulations. *The Journal of Physical Chemistry B*, 107(35):9424–9433.
- Anguela, X. M. and High, K. A. (2019). Entering the Modern Era of Gene Therapy. *Annual Review of Medicine*, 70(1):273–288.
- Anisimov, V. M., Lamoureux, G., Vorobyov, I. V., et al. (2005). Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *Journal of Chemical Theory and Computation*, 1(1):153–168.
- Arnold, R., Cole, M., and McGhee, J. R. (1977). A bactericidal effect for human lactoferrin. *Science*, 197(4300):263–265.
- Arnold, R. R., Brewer, M., and Gauthier, J. J. (1980). Bactericidal activity of human lactoferrin: sensitivity of a variety of microorganisms. *Infection and immunity*, 28(3):893–8.
- Arseneault, M., Bédard, S., Boulet-Audet, M., et al. (2010). Study of the Interaction of Lactoferricin B with Phospholipid Monolayers and Bilayers. *Langmuir*, 26(5):3468–3478.
- Asmar, A. T. and Collet, J.-F. (2018). Lpp, the Braun lipoprotein, turns 50 major achievements and remaining issues. *FEMS Microbiology Letters*, 365(18).
- Aurenhammer, F. and Franz (1991). Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405.
- Baker, C. M., Ferrari, M. J., and Shea, K. (2018). Beyond dose: Pulsed antibiotic treatment schedules can maintain individual benefit while reducing resistance. *Scientific Reports*, 8(1):5866.

- Balatti, G. E., Ambroggio, E. E., Fidelio, G. D., et al. (2017). Differential Interaction of Antimicrobial Peptides with Lipid Structures Studied by Coarse-Grained Molecular Dynamics Simulations. *Molecules (Basel, Switzerland)*, 22(10).
- Band, V. and Weiss, D. (2014). Mechanisms of Antimicrobial Peptide Resistance in Gram-Negative Bacteria. *Antibiotics*, 4(1):18–41.
- Barducci, A., Bonomi, M., and Parrinello, M. (2010). Linking Well-Tempered Metadynamics Simulations with Experiments. *Biophysical Journal*, 98(9):L44–L46.
- Barducci, A., Bonomi, M., and Parrinello, M. (2011). Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5):826–843.
- Barrangou, R. (2015). The roles of CRISPRCas systems in adaptive immunity and beyond. *Current Opinion in Immunology*, 32:36–41.
- Barrera, E. E., Machado, M. R., and Pantano, S. (2019). Fat SIRAH: Coarse-grained phospholipids to explore membrane-protein dynamics. *bioRxiv*, page 627570.
- Bellamy, W., Takase, M., Yamauchi, K., et al. (1992). Identification of the bactericidal domain of lactoferrin. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1121(1-2):130–136.
- Bennett, W. F. D., Hong, C. K., Wang, Y., et al. (2016). Antimicrobial Peptide Simulations and the Influence of Force Field on the Free Energy for Pore Formation in Lipid Bilayers. *Journal of Chemical Theory and Computation*, 12(9):4524–4533.
- Berendsen, H., Postma, J., Van Gunsteren, W., et al. (1981). Interaction models for water in relation to protein hydration. In Pullman, B., editor, *Intermolecular Forces*, pages 331–342. D. Reidel Publishing Company, Dordrecht.
- Berendsen, H., van der Spoel, D., and van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1-3):43–56.

- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., et al. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690.
- Berglund, N. A., Piggot, T. J., Jefferies, D., et al. (2015). Interaction of the Antimicrobial Peptide Polymyxin B1 with Both Membranes of *E. coli*: A Molecular Dynamics Study. *PLOS Computational Biology*, 11(4):e1004180.
- Berman, H. M., Westbrook, J., Feng, Z., et al. (2000). The protein data bank. <http://www.rcsb.org/>.
- Bernardi, R. C., Melo, M. C. R., and Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et biophysica acta*, 1850(5):872–7.
- Best, R. B., Clarke, J., and Karplus, M. (2005). What Contributions to Protein Side-chain Dynamics are Probed by NMR Experiments? A Molecular Dynamics Simulation Analysis. *Journal of Molecular Biology*, 349(1):185–203.
- Bhadra, P., Yan, J., Li, J., et al. (2018). AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. *Scientific Reports*, 8(1):1697.
- Biasini, M., Mariani, V., Haas, J., et al. (2010). OpenStructure: a flexible software framework for computational structural biology. *Bioinformatics*, 26(20):2626–2628.
- Billal, D. S., Feng, J., Leprohon, P., et al. (2011). Whole genome analysis of linezolid resistance in *Streptococcus pneumoniae* reveals resistance and compensatory mutations. *BMC Genomics*, 12(1):512.
- Biocomputing Group (2016). Image gallery.
- Birkegård, A. C., Halasa, T., Toft, N., et al. (2018). Send more data: a systematic review of mathematical models of antimicrobial resistance. *Antimicrobial resistance and infection control*, 7:117.
- Blair, J. M. A., Webber, M. A., Baylay, A. J., et al. (2014). Molecular mechanisms of antibiotic resistance. *Nature Publishing Group*, 13.

- Bochicchio, D. and Pavan, G. M. (2017). From Cooperative Self-Assembly to Water-Soluble Supramolecular Polymers Using Coarse-Grained Simulations. *ACS Nano*, 11(1):1000–1011.
- Böckmann, R. A., de Groot, B. L., Kakorin, S., et al. (2008). Kinetics, Statistics, and Energetics of Lipid Membrane Electroporation Studied by Molecular Dynamics Simulations. *Biophysical Journal*, 95(4):1837–1850.
- Boisselier, E. and Astruc, D. (2009). Gold nanoparticles in nanomedicine: preparations, imaging, diagnostics, therapies and toxicity. *Chemical Society Reviews*, 38(6):1759.
- Bonomi, M., Heller, G. T., Camilloni, C., et al. (2017). Principles of protein structural ensemble determination. *Current Opinion in Structural Biology*, 42:106–116.
- Borle, F. and Seelig, J. (1983). Hydration of Escherichia coli lipids: Deuterium T1 relaxation time studies of phosphatidylglycerol, phosphatidylethanolamine and phosphatidylcholine. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 735(1):131–136.
- Borle, F. and Seelig, J. (1985). Ca²⁺ binding to phosphatidylglycerol bilayers as studied by differential scanning calorimetry and 2H- and 31P-nuclear magnetic resonance. *Chemistry and Physics of Lipids*, 36(3):263–283.
- Botan, A., Favela-Rosales, F., Fuchs, P. F. J., et al. (2015). Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions. *The Journal of Physical Chemistry B*, 119(49):15075–15088.
- Bransden, B. H. and Joachain, C. J. (2003). *Physics of atoms and molecules*. Prentice Hall.
- Braun, A. R., Brandt, E. G., Edholm, O., et al. (2011). Determination of electron density profiles and area from simulations of undulating membranes. *Biophysical journal*, 100(9):2112–20.
- Brocos, P., Mendoza-Espinosa, P., Castillo, R., et al. (2012). Multiscale molecular dynamics simulations of micelles: coarse-grain for self-assembly and atomic resolution for finer details. *Soft Matter*, 8(34):9005.

- Brogden, K. A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nature Reviews Microbiology*, 3(3):238–250.
- Bugg, T. D. H., Wright, G. D., Dutka-Malen, S., et al. (1991). Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. *Biochemistry*, 30(43):10408–10415.
- Büldt, G., Gally, H. U., Seelig, A., et al. (1978). Neutron diffraction studies on selectively deuterated phospholipid bilayers. *Nature*, 271(5641):182–184.
- Bunker, A., Magarkar, A., and Viitala, T. (2016). Rational design of liposomal drug delivery systems, a review: Combined experimental and computational studies of lipid membranes, liposomes and their PEGylation. *BBA - Biomembranes*, 1858(10):2334–2352.
- Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101.
- Carpenter, T. S., Parkin, J., and Khalid, S. (2016). The Free Energy of Small Solute Permeation through the *Escherichia coli* Outer Membrane Has a Distinctly Asymmetric Profile. *The Journal of Physical Chemistry Letters*, 7(17):3446–3451.
- Castelletto, V., de Santis, E., Alkassem, H., et al. (2016). Structurally plastic peptide capsules for synthetic antimicrobial viruses. *Chem. Sci.*, 7(3):1707–1711.
- Chan, D. I., Prenner, E. J., and Vogel, H. J. (2006). Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1758(9):1184–1202.
- Chan, Y.-H. M. and Boxer, S. G. (2007). Model membrane systems and their applications. *Current opinion in chemical biology*, 11(6):581–7.
- Chandrasekhar, I., Kastenholz, M., Lins, R. D., et al. (2003). A consistent potential energy parameter set for lipids: dipalmitoylphosphatidylcholine as a benchmark of the GROMOS96 45A3 force field. *European biophysics journal : EBJ*, 32(1):67–77.

- Chandrasekhar, I., Oostenbrink, C., and van Gunsteren, W. F. (2004). Simulating the Physiological Phase of Hydrated DPPC Bilayers: The Ester Moiety. *Soft Materials*, 2(1):27–45.
- Chandrasekhar, I. and van Gunsteren, W. F. (2001). Sensitivity of molecular dynamics simulations of lipids to the size of the ester carbon.
- Chandrasekhar, Wilfred F. van Gunst, I. and van Gunsteren, W. F. (2002). A comparison of the potential energy parameters of aliphatic alkanes: molecular dynamics simulations of triacylglycerols in the alpha phase. *European Biophysics Journal*, 31(2):89–101.
- Chen, C., Starr, C. G., Troendle, E. P., et al. (2019). Simulation-guided rational *de novo* design of a small pore-forming antimicrobial peptide. *Journal of the American Chemical Society*, page jacs.8b11939.
- Cherkasov, A., Hilpert, K., Jenssen, H., et al. (2009). Use of Artificial Intelligence in the Design of Small Peptide Antibiotics Effective against a Broad Spectrum of Highly Antibiotic-Resistant Superbugs. *ACS Chemical Biology*, 4(1):65–74.
- Chiu, S.-W., Clark, M., Balaji, V., et al. (1995). Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. *Biophysical journal*, 69(4):1230–45.
- Chiu, S.-W., Pandit, S. A., Scott, H. L., et al. (2009). An Improved United Atom Force Field for Simulation of Mixed Lipid Bilayers. *The Journal of Physical Chemistry B*, 113(9):2748–2763.
- Chou, K.-C. (2015). Impacts of Bioinformatics to Medicinal Chemistry. *Medicinal Chemistry*, 11:218–234.
- Chowdhary, J., Harder, E., Lopes, P. E. M., et al. (2013). A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *The Journal of Physical Chemistry B*, 117(31):9142–9160.
- Cieplak, P., Caldwell, J., and Kollman, P. (2001). Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol

- and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *Journal of Computational Chemistry*, 22(10):1048–1057.
- Cipcigan, F., Carrieri, A. P., Pyzer-Knapp, E. O., et al. (2018). Accelerating molecular discovery through data and physical sciences: Applications to peptide-membrane interactions. *The Journal of Chemical Physics*, 148(24):241744.
- City of Vienna Climate Protection Programme (2012). The Viennese Database for Disinfectants (WIDES Database). <https://www.wien.gv.at/english/environment/protection/oekokauf/disinfectants/>.
- Clarke, R. J. (2001). The dipole potential of phospholipid membranes and methods for its detection. *Advances in Colloid and Interface Science*, 89-90:263–281.
- Cochran, A. G., Skelton, N. J., and Starovasnik, M. A. (2001). Tryptophan zippers: Stable, monomeric β -hairpins. *Proceedings of the National Academy of Sciences*, 98(10):5578–5583.
- Cocucci, E., Kim, J., Bai, Y., et al. (2017). Role of Passive Diffusion, Transporters, and Membrane Trafficking-Mediated Processes in Cellular Drug Transport. *Clinical Pharmacology & Therapeutics*, 101(1):121–129.
- Collu, F., Spiga, E., Lorenz, C. D., et al. (2015). Assembly of Influenza Hemagglutinin Fusion Peptides in a Phospholipid Bilayer by Coarse-grained Computer Simulations. *Frontiers in molecular biosciences*, 2:66.
- Cooper, G. M. (2000). *Transport of Small Molecules*. Sinauer Associates.
- Costerton, J. W., Stewart, P. S., and Greenberg, E. P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science (New York, N.Y.)*, 284(5418):1318–22.
- Crombez, L., Aldrian-Herrada, G., Konate, K., et al. (2009). A New Potent Secondary Amphiphatic Cellpenetrating Peptide for siRNA Delivery Into Mammalian Cells. *Molecular Therapy*, 17(1):95–103.

- Davis, P. J., Fleming, B. D., Coolbear, K. P., et al. (1981). Gel to liquid-crystalline transition temperatures of water dispersions of two pairs of positional isomers of unsaturated mixed-acid phosphatidylcholines. *Biochemistry*, 20(12):3633–3636.
- Daya, S. and Berns, K. I. (2008). Gene therapy using adeno-associated virus vectors. *Clinical microbiology reviews*, 21(4):583–93.
- de Jong, D. H., Singh, G., Bennett, W. F. D., et al. (2013). Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation*, 9(1):687–697.
- de Vries, A. H., Mark, A. E., and Marrink, S.-J. (2004). Molecular Dynamics Simulation of the Spontaneous Formation of a Small DPPC Vesicle in Water in Atomistic Detail. *Journal of the American Chemical Society*, 126(14):4488–4489.
- Delcour, A. H. (2009). Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1794(5):808–816.
- Depan, D., Shah, J., and Misra, R. (2011). Controlled release of drug from folate-decorated and graphene mediated drug delivery system: Synthesis, loading efficiency, and drug release response. *Materials Science and Engineering: C*, 31(7):1305–1312.
- Deslouches, B., Phadke, S. M., Lazarevic, V., et al. (2005). De Novo Generation of Cationic Antimicrobial Peptides: Influence of Length and Tryptophan Substitution on Antimicrobial Activity. *Antimicrobial Agents and Chemotherapy*, 49(1):316–322.
- Dickson, C. J., Madej, B. D., Skjevik, Å. A., et al. (2014). Lipid14: The Amber Lipid Force Field. *Journal of Chemical Theory and Computation*, 10(2):865–879.
- Ding, W., Palaiokostas, M., Wang, W., et al. (2015). Effects of Lipid Composition on Bilayer Membranes Quantified by All-Atom Molecular Dynamics. *The Journal of Physical Chemistry B*, 119(49):15263–15274.

- Doktorova, M., Harries, D., and Khelashvili, G. (2017). Determination of bending rigidity and tilt modulus of lipid membranes from real-space fluctuation analysis of molecular dynamics simulations. *Physical Chemistry Chemical Physics*, 19(25):16806–16818.
- Dolejska, M., Villa, L., Poirel, L., et al. (2013). Complete sequencing of an IncHI1 plasmid encoding the carbapenemase NDM-1, the ArmA 16S RNA methylase and a resistance-nodulation-cell division/multidrug efflux pump. *Journal of Antimicrobial Chemotherapy*, 68(1):34–39.
- Douglas, S. M., Bachelet, I., and Church, G. M. (2012). A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science*, 335(6070):831–834.
- Douliez, J.-P., Ferrarini, A., and Dufourc, E.-J. (1998). On the relationship between C-C and C-D order parameters and its use for studying the conformation of lipid acyl chains in biomembranes. *The Journal of Chemical Physics*, 109(6):2513–2518.
- Douliez, J. P., Léonard, A., and Dufourc, E. J. (1995). Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splittings. *Biophysical journal*, 68(5):1727–39.
- Doux, J. P. F., Hall, B. A., and Killian, J. A. (2012). How lipid headgroups sense the membrane environment: an application of N NMR. *Biophysical journal*, 103(6):1245–53.
- Dror, R. O., Dirks, R. M., Grossman, J., et al. (2012). Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of Biophysics*, 41(1):429–452.
- Ebbengaard, A., Mordhorst, H., Overgaard, M. T., et al. (2015). Comparative Evaluation of the Antimicrobial Activity of Different Antimicrobial Peptides against a Range of Pathogenic Bacteria. *PLOS ONE*, 10(12):e0144611.
- Ebenhan, T., Gheysens, O., Kruger, H. G., et al. (2014). Antimicrobial peptides: their role as infection-selective tracers for molecular imaging. *BioMed research international*, 2014:867381.

- Editor BBC (2016). Artificial intelligence: Go master Lee Se-dol wins against AlphaGo program - BBC News. <https://www.bbc.co.uk/news/technology-35797102>.
- Einstein, A. and Furt, R. (1956). *Investigations on the theory of Brownian movement*. Dover Publications.
- Entova, S., Billod, J.-M., Swiecicki, J.-M., et al. (2018). Insights into the key determinants of membrane protein topology enable the identification of new monotopic folds. *eLife*, 7.
- Erol, O., Uyan, I., Hatip, M., et al. (2017). Recent Advances in Bioactive 1D and 2D Carbon Nanomaterials for Biomedical Applications. *Nanomedicine: Nanotechnology, Biology and Medicine*.
- Essmann, U., Perera, L., Berkowitz, M. L., et al. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19):8577–8593.
- Fan, T., Yu, X., Shen, B., et al. (2017). Peptide Self-Assembled Nanostructures for Drug Delivery Applications. *Journal of Nanomaterials*, 2017:1–16.
- Farnaud, S. and Evans, R. W. (2003). Lactoferrin - a multifunctional protein with antimicrobial properties. *Molecular immunology*, 40(7):395–405.
- Farrotti, A., Conflitti, P., Srivastava, S., et al. (2017). Molecular Dynamics Simulations of the Host Defense Peptide Temporin L and Its Q3K Derivative: An Atomic Level View from Aggregation in Water to Bilayer Perturbation. *Molecules*, 22(7):1235.
- Fjell, C. D., Hiss, J. A., Hancock, R. E. W., et al. (2011). Designing antimicrobial peptides: form follows function. *Nature Reviews Drug Discovery*, 11(1):37.
- Floyd, J. L., Smith, K. P., Kumar, S. H., et al. (2010). LmrS Is a Multidrug Efflux Pump of the Major Facilitator Superfamily from *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 54(12):5406–5412.
- Fornili, A., Autore, F., Chakroun, N., et al. (2012). ProteinWater Interactions in MD Simulations: POPS/POPSCOMP Solvent Accessibility Analysis,

- Solvation Forces and Hydration Sites. pages 375–392. Springer, New York, NY.
- Fraternali, F. and van Gunsteren, W. (1996). An Efficient Mean Solvation Force Model for Use in Molecular Dynamics Simulations of Proteins in Aqueous Solution. *Journal of Molecular Biology*, 256(5):939–948.
- Frederix, P. W. J. M., Patmanidis, I., and Marrink, S. J. (2018). Molecular simulations of self-assembling bio-inspired supramolecular systems and their connection to experiments. *Chemical Society Reviews*, 47(10):3470–3489.
- Gahr, M., Speer, C. P., Damerau, B., et al. (1991). Influence of Lactoferrin on the Function of Human Polymorphonuclear Leukocytes and Monocytes. *Journal of Leukocyte Biology*, 49(5):427–433.
- Gan, L., Chen, S., and Jensen, G. J. (2008). Molecular organization of Gram-negative peptidoglycan. *Proceedings of the National Academy of Sciences*, 105(48):18953–18957.
- Gifford, J. L., Hunter, H. N., and Vogel, H. J. (2005). Lactoferricin. *Cellular and Molecular Life Sciences*, 62(22):2588–2598.
- Glukhov, E., Stark, M., Burrows, L. L., et al. (2005). Basis for selectivity of cationic antimicrobial peptides for bacterial versus mammalian membranes. *The Journal of biological chemistry*, 280(40):33960–7.
- Gonçalves, M. B., Dreyer, J., Lupieri, P., et al. (2013). Structural prediction of a rhodamine-based biosensor and comparison with biophysical data. *Phys. Chem. Chem. Phys.*, 15(6):2177–2183.
- Google (2016). Google AI Blog. <https://ai.googleblog.com/2019/>.
- Gordon, E., Mouz, N., Duée, E., et al. (2000). The crystal structure of the penicillin-binding protein 2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance. *Journal of Molecular Biology*, 299(2):477–485.
- Graham, J. A., Essex, J. W., and Khalid, S. (2017). PyCGTOOL: Automated Generation of Coarse-Grained Molecular Dynamics Models from

- Atomistic Trajectories. *Journal of Chemical Information and Modeling*, 57(4):650–656.
- Grime, J. M. A., Dama, J. F., Ganser-Pornillos, B. K., et al. (2016). Coarse-grained simulation reveals key features of HIV-1 capsid self-assembly. *Nature communications*, 7:11568.
- Group of BioMolecular Simulations (2019). Sirah force field for amino acids, solvent and nucleotides.
- Gudlur, S., Sukthankar, P., Gao, J., et al. (2012). Peptide Nanovesicles Formed by the Self-Assembly of Branched Amphiphilic Peptides. *PLoS ONE*, 7(9).
- Guo, C., Luo, Y., Zhou, R., et al. (2012). Probing the Self-Assembly Mechanism of Diphenylalanine-Based Peptide Nanovesicles and Nanotubes. *ACS Nano*, 6(5):3907–3918.
- Guo, L., Lim, K. B., Poduje, C. M., et al. (1998). Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. *Cell*, 95(2):189–98.
- Gurtovenko, A. A. and Vattulainen, I. (2007). Molecular Mechanism for Lipid Flip-Flops. *The Journal of Physical Chemistry B*, 111(48):13554–13559.
- Gurtovenko, A. A. and Vattulainen, I. (2009). Calculation of the electrostatic potential of lipid bilayers from molecular dynamics simulations: Methodological issues. *The Journal of Chemical Physics*, 130(21):215107.
- Hadden, J. A. and Perilla, J. R. (2018). All-atom virus simulations. *Current Opinion in Virology*, 31:82–91.
- Hallock, K. J., Lee, D.-K., and Ramamoorthy, A. (2003). MSI-78, an analogue of the magainin antimicrobial peptides, disrupts lipid bilayer structure via positive curvature strain. *Biophysical journal*, 84(5):3052–60.
- Hancock, R. E. W. and Sahl, H.-G. (2006). Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nature Biotechnology*, 24(12):1551–1557.

- Haria, N. R., Monticelli, L., Fraternali, F., et al. (2014). Plasticity and conformational equilibria of influenza fusion peptides in model lipid bilayers. *BBA - Biomembranes*, 1838(4):1169–1179.
- Hart, K., Foloppe, N., Baker, C. M., et al. (2012). Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *Journal of Chemical Theory and Computation*, 8(1):348–362.
- Heo, L. and Feig, M. (2018). Experimental accuracy in protein structure refinement via molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 115(52):13276–13281.
- Hess, B., Bekker, H., Berendsen, H. J. C., et al. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472.
- Hilpert, K., Volkmer-Engert, R., Walter, T., et al. (2005). High-throughput generation of small antibacterial peptides with improved activity. *Nature Biotechnology*, 23(8):1008–1012.
- Hodder, A. N., Crewther, P. E., and Anders, R. F. (2001). Specificity of the protective antibody response to apical membrane antigen 1. *Infection and immunity*, 69(5):3286–94.
- Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697.
- Huang, C. H., Lapides, J. R., and Levin, I. W. (1982). Phase-transition behavior of saturated, symmetric chain phospholipid bilayer dispersions determined by Raman spectroscopy: correlation between spectral and thermodynamic parameters. *Journal of the American Chemical Society*, 104(22):5926–5930.
- Huang, J. and MacKerell, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25):2135–2145.

- Huang, J., Rauscher, S., Nawrocki, G., et al. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, 14(1):71–73.
- Huber, T., Torda, A. E., and van Gunsteren, W. F. (1994). Local elevation: A method for improving the searching properties of molecular dynamics simulation. *The Journal of Physical Chemistry A*, 101(33):5926–5930.
- Hughes, G. A. (2005). Nanostructure-mediated drug delivery. *Nanomedicine: Nanotechnology, Biology and Medicine*, 1(1):22–30.
- Hwang, P. M., Zhou, N., Shan, X., et al. (1998). Three-Dimensional Solution Structure of Lactoferricin B, an Antimicrobial Peptide Derived from Bovine Lactoferrin. *Biochemistry*, 37(12):4288–4298.
- Jahani, S., Shakiba, A., and Jahani, L. (2015). The Antimicrobial Effect of Lactoferrin on Gram-Negative and Gram-Positive Bacteria. *International Journal of Infection*, 2(3).
- Jain, S. and Pillai, J. (2017). Bacterial membrane vesicles as novel nanosystems for drug delivery. *International journal of nanomedicine*, 12:6329–6341.
- Jämbeck, J. P. M. and Lyubartsev, A. P. (2012). Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids. *The Journal of Physical Chemistry B*, 116(10):3164–3179.
- Jean-Louis, S., Akare, S., Ali, M. A., et al. (2006). Deoxycholic acid induces intracellular signaling through membrane perturbations. *The Journal of biological chemistry*, 281(21):14948–60.
- Jhong, J.-H., Chi, Y.-H., Li, W.-C., et al. (2019). dbAMP: an integrated resource for exploring antimicrobial peptides with functional activities and physicochemical properties on transcriptome and proteome data. *Nucleic Acids Research*, 47(D1):D285–D297.
- Jia, Z., O'Mara, M. L., Zuegg, J., et al. (2011). The Effect of Environment on the Recognition and Binding of Vancomycin to Native and Resistant Forms of Lipid II. *Biophysical Journal*, 101(11):2684–2692.

- Jiang, Q., Song, C., Nangreave, J., et al. (2012). DNA Origami as a Carrier for Circumvention of Drug Resistance. *Journal of the American Chemical Society*, 134(32):13396–13403.
- Jiang, Z., Vasil, A. I., Gera, L., et al. (2011). Rational Design of α -Helical Antimicrobial Peptides to Target Gram-negative Pathogens, *Acinetobacter baumannii* and *Pseudomonas aeruginosa*: Utilization of Charge, ‘Specificity Determinants’, Total Hydrophobicity, Hydrophobe Type and Location as Design Parameters to Improve the Therapeutic Ratio. *Chemical Biology & Drug Design*, 77(4):225–240.
- Johner, N., Harries, D., and Khelashvili, G. (2016). Implementation of a methodology for determining elastic properties of lipid assemblies from molecular dynamics simulations. *BMC Bioinformatics*, 17(1):161.
- Jolivet-Gougeon, A. and Bonnaure-Mallet, M. (2014). Biofilms as a mechanism of bacterial resistance. *Drug Discovery Today: Technologies*, 11:49–56.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., et al. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935.
- Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energies and Properties of Organic Liquids.
- Jorgensen, W. L. and Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666.
- Joseleau-Petit, D., Liébart, J.-C., Ayala, J. A., et al. (2007). Unstable Escherichia coli L Forms Revisited: Growth Requires Peptidoglycan Synthesis. *Journal of Bacteriology*, 189(18):6512–6520.
- Kagan, B. L., Selsted, M. E., Ganz, T., et al. (1990). Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. *Proceedings of the National Academy of Sciences*, 87(1):210–214.

- Kandasamy, S. K. and Larson, R. G. (2004). Binding and insertion of α -helical anti-microbial peptides in POPC bilayers studied by molecular dynamics simulations. *Chemistry and Physics of Lipids*, 132(1):113–132.
- Kandt, C., Ash, W., and Peter Tieleman, D. (2007). Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41(4):475–488.
- Kapoor, G., Saigal, S., and Elongavan, A. (2017). Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of anaesthesiology, clinical pharmacology*, 33(3):300–305.
- Khalid, S., Berglund, N. A., Holdbrook, D. A., et al. (2015). The membranes of Gram-negative bacteria: progress in molecular modelling and simulation. *Biochemical Society Transactions*, 43(2):162–167.
- Khalid, S., Piggot, T. J., and Samsudin, F. (2019). Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned? *Accounts of Chemical Research*, 52(1):180–188.
- Khandelia, H. and Kaznessis, Y. N. (2005). Molecular dynamics simulations of helical antimicrobial peptides in SDS micelles: What do point mutations achieve? *Peptides*, 26(11):2037–2049.
- Kikhney, A. G. and Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters*, 589(19PartA):2570–2577.
- Kim, J. (2017). Phage as a therapeutic agent.
- Kirkpatrick, C. H., Green, I., Rich, R. R., et al. (1971). Inhibition of Growth of *Candida albicans* by Iron-Unsaturated Lactoferrin: Relation to Host-Defense Mechanisms in Chronic Mucocutaneous Candidiasis. *Journal of Infectious Diseases*, 124(6):539–544.
- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598):671–680.
- Klauda, J. B., Venable, R. M., Freites, J. A., et al. (2010). Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B*, 114(23):7830–7843.

- Kleandrova, V. V., Russo, J. M., Speck-Planche, A., et al. (2016). Enabling the Discovery and Virtual Screening of Potent and Safe Antimicrobial Peptides. Simultaneous Prediction of Antibacterial Activity and Cytotoxicity. *ACS Combinatorial Science*, 18(8):490–498.
- Kleckner, I. R. and Foster, M. P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1814(8):942–968.
- Kleinjung, J. and Fraternali, F. (2014). Design and application of implicit solvent models in biomolecular simulations. *Current Opinion in Structural Biology*, 25:126–134.
- Kleinjung, J., Fraternali, F., Martin, S. R., et al. (2003). Thermal unfolding simulations of apo-calmodulin using leap-dynamics. *Proteins: Structure, Function, and Bioinformatics*, 50(4):648–656.
- Kleinjung, J., Scott, W. R. P., Allison, J. R., et al. (2012). Implicit Solvation Parameters Derived from Explicit Water Forces in Large-Scale Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 8(7):2391–2403.
- Kojima, S. and Nikaido, H. (2013). Permeation rates of penicillins indicate that Escherichia coli porins function principally as nonspecific channels. *Proceedings of the National Academy of Sciences*, 110(28):E2629–E2634.
- Koziara, K., Stroet, M., Malde, A., et al. (2014). Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies. *Journal of Computer-Aided Molecular Design*, 28(3):221–233.
- Krol, S. (2012). Challenges in drug delivery to the brain: Nature is against us. *Journal of Controlled Release*, 164(2):145–155.
- Kučerka, N., Heberle, F., Pan, J., et al. (2015). Structural Significance of Lipid Diversity as Studied by Small Angle Neutron and X-ray Scattering. *Membranes*, 5(3):454–472.
- Kučerka, N., Nieh, M.-P., and Katsaras, J. (2011). Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function

- of temperature. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(11):2761–2771.
- Kukol, A. (2009). Lipid Models for United-Atom Molecular Dynamics Simulations of Proteins. *Journal of Chemical Theory and Computation*, 5(3):615–626.
- Kumariya, R., Sood, S. K., Rajput, Y. S., et al. (2015). Increased membrane surface positive charge and altered membrane fluidity leads to cationic antimicrobial peptide resistance in Enterococcus faecalis. *Biochimica et biophysica acta*, 1848(6):1367–75.
- Ladokhin, A. S. and White, S. H. (2001). ‘Detergent-like’ permeabilization of anionic lipid vesicles by melittin. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1514(2):253–260.
- Lai, Y. and Gallo, R. L. (2009). AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense. *Trends in Immunology*, 30(3):131–141.
- Laio, A. and Parrinello, M. (2002). Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12562–6.
- Lairion, F., Filler, R., and Disalvo, E. (2002). Reversed micelles as model systems to study interfacial properties of lipid bilayers. *Colloids and Surfaces B: Biointerfaces*, 25(4):369–371.
- Lammers, T., Subr, V., Ulbrich, K., et al. (2009). Simultaneous delivery of doxorubicin and gemcitabine to tumors in vivo using prototypic polymeric drug carriers. *Biomaterials*, 30(20):3466–3475.
- Lata, S., Mishra, N. K., and Raghava, G. P. (2010). AntiBP2: improved version of antibacterial peptide prediction. *BMC Bioinformatics*, 11(Suppl 1):S19.
- Lauer, K. B., Borrow, R., and Blanchard, T. J. (2017). Multivalent and Multipathogen Viral Vector Vaccines. *Clinical and Vaccine Immunology*, 24(1):e00298–16.

- Lavigne, J.-P., Sotto, A., Nicolas-Chanoine, M.-H., et al. (2013). An adaptive response of *Enterobacter aerogenes* to imipenem: regulation of porin balance in clinical isolates. *International Journal of Antimicrobial Agents*, 41(2):130–136.
- Lee, E. H., Hsin, J., Sotomayor, M., et al. (2009). Discovery Through the Computational Microscope. *Structure*, 17(10):1295–1306.
- Lee, H. and Pastor, R. W. (2011). Coarse-Grained Model for PEGylated Lipids: Effect of PEGylation on the Size and Shape of Self-Assembled Structures. *The Journal of Physical Chemistry B*, 115(24):7830–7837.
- Lee, O.-S., Cho, V., and Schatz, G. C. (2012). Modeling the Self-Assembly of Peptide Amphiphiles into Fibers Using Coarse-Grained Molecular Dynamics. *Nano Letters*, 12(9):4907–4913.
- Lehrer, R. I. (2004). Primate defensins. *Nature Reviews Microbiology*, 2(9):727–738.
- Leipzig, J. (2016). A review of bioinformatic pipeline frameworks. *Briefings in Bioinformatics*, 18(3):bbw020.
- Leontiadou, H., Mark, A. E., and Marrink, S.-J. (2004). Molecular dynamics simulations of hydrophilic pores in lipid bilayers. *Biophysical journal*, 86(4):2156–64.
- Leontiadou, H., Mark, A. E., and Marrink, S. J. (2006a). Antimicrobial Peptides in Action. *Journal of the American Chemical Society*, 128(37):12156–12161.
- Leontiadou, H., Mark, A. E., and Marrink, S.-J. (2006b). Antimicrobial Peptides in Action. *Journal of the American Chemical Society*, 128(37):12156–12161.
- Leontiadou, H., Mark, A. E., and Marrink, S.-J. (2007). Ion transport across transmembrane pores. *Biophysical journal*, 92(12):4209–15.
- Lewis, R. N. A. H., Mak, N., and McElhaney, R. N. (1987). A differential scanning calorimetric study of the thermotropic phase behavior of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry*, 26(19):6118–6126.

- Liechty, W. B., Kryscio, D. R., Slaughter, B. V., et al. (2010). Polymers for drug delivery systems. *Annual review of chemical and biomolecular engineering*, 1:149–73.
- Lin, T.-Y. and Weibel, D. B. (2016). Organization and function of anionic phospholipids in bacteria. *Applied Microbiology and Biotechnology*, 100(10):4255–4267.
- Lindblom, G. and Orädd, G. (2009). Lipid lateral diffusion and membrane heterogeneity. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1788(1):234–244.
- Linko, V., Ora, A., and Kostiainen, M. A. (2015). DNA Nanostructures as Smart Drug-Delivery Vehicles and Molecular Devices. *Trends in Biotechnology*, 33(10):586–594.
- Lipkin, R., Pino-Angeles, A., and Lazaridis, T. (2017). Transmembrane Pore Structures of β -Hairpin Antimicrobial Peptides by All-Atom Simulations. *The Journal of Physical Chemistry B*, 121(39):9126–9140.
- Liu, S., Bao, J., Lao, X., et al. (2018). Novel 3D Structure Based Model for Activity Prediction and Design of Antimicrobial Peptides. *Scientific Reports*, 8(1):11189.
- Livermore, D. (2008). Defining an extended-spectrum β -lactamase. *Clinical Microbiology and Infection*, 14:3–10.
- Lobanovska, M. and Pilla, G. (2017). Penicillin's Discovery and Antibiotic Resistance: Lessons for the Future? *The Yale journal of biology and medicine*, 90(1):135–145.
- Lobo, F. P., Mota, B. E. F., Pena, S. D. J., et al. (2009). Virus-Host Coevolution: Common Patterns of Nucleotide Motif Usage in Flaviviridae and Their Hosts. *PLoS ONE*, 4(7):e6282.
- Long, K. S., Poehlsgaard, J., Kehrenberg, C., et al. (2006). The Cfr rRNA Methyltransferase Confers Resistance to Phenicols, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A Antibiotics. *Antimicrobial Agents and Chemotherapy*, 50(7):2500–2505.

- Loose, C., Jensen, K., Rigoutsos, I., et al. (2006). A linguistic model for the rational design of antimicrobial peptides. *Nature*, 443(7113):867–869.
- Lopes, D., Jakobtorweihen, S., Nunes, C., et al. (2017). Shedding light on the puzzle of drug-membrane interactions: Experimental techniques and molecular dynamics simulations. *Progress in Lipid Research*, 65:24–44.
- Ma, H., Cummins, D. D., Edelstein, N. B., et al. (2017). Modeling Diversity in Structures of Bacterial Outer Membrane Lipids. *Journal of Chemical Theory and Computation*, 13(2):811–824.
- Ma, Y., Nolte, R. J., and Cornelissen, J. J. (2012). Virus-based nanocarriers for drug delivery. *Advanced Drug Delivery Reviews*, 64(9):811–825.
- Mabrey, S. and Sturtevant, J. M. (1976). Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proceedings of the National Academy of Sciences of the United States of America*, 73(11):3862–6.
- Machado, M. R., González, H. C., and Pantano, S. (2017). MD Simulations of Viruslike Particles with Supra CG Solvation Affordable to Desktop Computers. *Journal of Chemical Theory and Computation*, 13(10):5106–5116.
- Machado, M. R., Guisasola, E. E. B., Klein, F., et al. (2018). The SIRAH force field 2.0: Altius, Fortius, Citius. *bioRxiv*, page 436774.
- MacKerell, A. D., Bashford, D., Bellott, M., et al. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins . *The Journal of Physical Chemistry B*, 102(18):3586–3616.
- Macpherson, J. A., Theisen, A., Masino, L., et al. (2019). Functional cross-talk between allosteric effects of activating and inhibiting ligands underlies PKM2 regulation. *eLife*, 8.
- Mahlapuu, M., Håkansson, J., Ringstad, L., et al. (2016). Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. *Frontiers in Cellular and Infection Microbiology*, 6:194.

- Maier, J. A., Martinez, C., Kasavajhala, K., et al. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation*, 11(8):3696–3713.
- Malay, A. D., Miyazaki, N., Biela, A., et al. (2019). An ultra-stable gold-coordinated protein cage displaying reversible assembly. *Nature*, 569(7756):438–442.
- Malde, A., Zuo, L., Breeze, M., et al. (2011). An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *Journal of Chemical Theory and Computation*, 7(12):4026–4037.
- Mälkiä, A., Murtomäki, L., Urtti, A., et al. (2004). Drug permeation in biomembranes: In vitro and in silico prediction and influence of physicochemical properties. *European Journal of Pharmaceutical Sciences*, 23(1):13–47.
- Mantravadi, P., Kalesh, K., Dobson, R., et al. (2019). The Quest for Novel Antimicrobial Compounds: Emerging Trends in Research, Development, and Technologies. *Antibiotics*, 8(1):8.
- Margreitter, C., Reif, M. M., and Oostenbrink, C. (2017). Update on phosphate and charged post-translationally modified amino acid parameters in the GROMOS force field. *Journal of Computational Chemistry*, 38(10):714–720.
- Marinov, R. and Dufourc, E. J. (1996). Thermotropism and hydration properties of POPE and POPE-cholesterol systems as revealed by solid state²H and³¹P-NMR. *European Biophysics Journal*, 24(6):423–431.
- Marrink, S. J., de Vries, A. H., and Mark, A. E. (2003). Coarse Grained Model for Semiquantitative Lipid Simulations.
- Marrink, S.-J., Lindahl, E., Edholm, O., et al. (2001). Simulation of the Spontaneous Aggregation of Phospholipids into Bilayers. *Journal of the American Chemical Society*, 123(35):8638–8639.
- Marrink, S.-J. and Mark, A. E. (2003). The Mechanism of Vesicle Fusion as Revealed by Molecular Dynamics Simulations. *Journal of the American Chemical Society*.

- Marrink, S. J., Risselada, H. J., Yefimov, S., et al. (2007). The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B*, 111(27):7812–7824.
- Marrink, S.-J., Risselada, J., and Mark, A. E. (2005). Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chemistry and Physics of Lipids*, 135(2):223–244.
- Marrink, S. J. and Tieleman, D. P. (2013). Perspective on the Martini model. *Chemical Society reviews*, 42(16):6801–22.
- Martínez, L., Andrade, R., Birgin, E. G., et al. (2009). PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13):2157–2164.
- Martyna, G. J., Tuckerman, M. E., Tobias, D. J., et al. (1996). Explicit reversible integrators for extended systems dynamics. *Molecular Physics*, 87(5):1117–1157.
- Masaoka, Y., Tanaka, Y., Kataoka, M., et al. (2006). Site of drug absorption after oral administration: Assessment of membrane permeability and luminal concentration of drugs in each segment of gastrointestinal tract. *European Journal of Pharmaceutical Sciences*, 29(3-4):240–250.
- Matsuzaki, K. (2009). Control of cell selectivity of antimicrobial peptides. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1788(8):1687–1692.
- Matsuzaki, K., Murase, O., Fujii, N., et al. (1996). An Antimicrobial Peptide, Magainin 2, Induced Rapid Flip-Flop of Phospholipids Coupled with Pore Formation and Peptide Translocation. *Biochemistry*, 35(35):11361–11368.
- Matsuzaki, K., Sugishita, K.-i., Harada, M., et al. (1997). Interactions of an antimicrobial peptide, magainin 2, with outer and inner membranes of Gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1327(1):119–130.
- McIntosh, T. J. and Simon, S. A. (1986). Area per molecule and distribution of water in fully hydrated dilrauroylphosphatidylethanolamine bilayers. *Biochemistry*, 25(17):4948–4952.

- Meißner, R. H., Schneider, J., Schiffels, P., et al. (2014). Computational Prediction of Circular Dichroism Spectra and Quantification of Helicity Loss upon Peptide Adsorption on Silica. *Langmuir*, 30(12):3487–3494.
- Mettler Toledo (2018). Polymerization reactions.
- Migoń, D., Jaśkiewicz, M., Neubauer, D., et al. (2018). Alanine Scanning Studies of the Antimicrobial Peptide Aurein 1.2. *Probiotics and Antimicrobial Proteins*, pages 1–13.
- Milano, G. and Kawakatsu, T. (2009). Hybrid particle-field molecular dynamics simulations for dense polymer systems. *The Journal of Chemical Physics*, 130(21):214106.
- Mills, M. and Andricioaei, I. (2008). An experimentally guided umbrella sampling protocol for biomolecules. *The Journal of chemical physics*, 129(11):114101.
- Mingozzi, F. and High, K. A. (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nature Reviews Genetics*, 12(5):341–355.
- Mitragotri, S., Burke, P. A., and Langer, R. (2014). Overcoming the challenges in administering biopharmaceuticals: formulation and delivery strategies. *Nature Reviews Drug Discovery*, 13(9):655–672.
- Miyamoto, S. and Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8):952–962.
- Monticelli, L., Kandasamy, S. K., Periole, X., et al. (2008). The MARTINI Coarse-Grained Force Field: Extension to Proteins. *Journal of chemical theory and computation*, 4(5):819–34.
- Moskowitz, S. M., Ernst, R. K., and Miller, S. I. (2004). PmrAB, a two-component regulatory system of *Pseudomonas aeruginosa* that modulates resistance to cationic antimicrobial peptides and addition of aminoarabinose to lipid A. *Journal of bacteriology*, 186(2):575–9.
- Mulliken, R. S. (1955). Electronic Population Analysis on LCAOMO Molecular Wave Functions. I. *The Journal of Chemical Physics*, 23(10):1833–1840.

- Naafs, M. A. B. (2018). The Antimicrobial Peptides: Ready for Clinical Trials? *Biomedical Journal of Scientific & Technical Research*, 7(4):001–005.
- Nagle, J. F. (2017). Experimentally determined tilt and bending moduli of single-component lipid bilayers. *Chemistry and Physics of Lipids*, 205:18–24.
- Nagle, J. F., Cognet, P., Dupuy, F. G., et al. (2019). Structure of gel phase DPPC determined by X-ray diffraction. *Chemistry and Physics of Lipids*, 218:168–177.
- Nagle, J. F. and Tristram-Nagle, S. (2000). Structure of lipid bilayers. *Biochimica et biophysica acta*, 1469(3):159–95.
- Naldini, L. (2011). Ex vivo gene transfer and correction for cell-based therapies. *Nature Reviews Genetics*, 12(5):301–315.
- Nelson, D. C., Garbe, J., and Collin, M. (2011). Cysteine proteinase SpeB from Streptococcus pyogenes - a potent modifier of immunologically important host and bacterial proteins. *Biological Chemistry*, 392(12):1077–88.
- Nguyen, L. T., Haney, E. F., and Vogel, H. J. (2011). The expanding scope of antimicrobial peptide structures and their modes of action. *Trends in Biotechnology*, 29(9):464–472.
- Nguyen, L. T., Schibli, D. J., and Vogel, H. J. (2005). Structural studies and model membrane interactions of two peptides derived from bovine lactoferricin. *Journal of Peptide Science*, 11(7):379–389.
- Nicolas, J., Mura, S., Brambilla, D., et al. (2013). Design, functionalization strategies and biomedical applications of targeted biodegradable/biocompatible polymer-based nanocarriers for drug delivery. *Chem. Soc. Rev.*, 42(3):1147–1235.
- Niewiadomska, A. M., Jayabalasingham, B., Seidman, J. C., et al. (2019). Population-level mathematical modeling of antimicrobial resistance: a systematic review. *BMC Medicine*, 17(1):81.
- Nikaido, E., Shirozaka, I., Yamaguchi, A., et al. (2011). Regulation of the AcrAB multidrug efflux pump in *Salmonella enterica* serovar Ty-

- phimurium in response to indole and paraquat. *Microbiology*, 157(3):648–655.
- Nordmann, P., Poirel, L., Walsh, T. R., et al. (2011). The emerging NDM carbapenemases. *Trends in Microbiology*, 19(12):588–595.
- Nosé, S. and Klein, M. (1983). Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076.
- Ogawa, W., Onishi, M., Ni, R., et al. (2012). Functional study of the novel multidrug efflux pump KexD from Klebsiella pneumoniae. *Gene*, 498(2):177–182.
- Okamoto, Y. (2004). Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *Journal of Molecular Graphics and Modelling*, 22(5):425–439.
- O’neill, J. (2016). TACKLING DRUG-RESISTANT INFECTIONS GLOBALLY: FINAL REPORT AND RECOMMENDATIONS THE REVIEW ON ANTIMICROBIAL RESISTANCE.
- Oostenbrink, C., Soares, T. A., van der Vegt, N. F. A., et al. (2005). Validation of the 53A6 GROMOS force field. *European Biophysics Journal*, 34(4):273–284.
- Oostenbrink, C., Villa, A., Mark, A. E., et al. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25(13):1656–1676.
- Orsi, M. (2018). Molecular simulation of self-assembly. *Self-assembling Biomaterials*, pages 305–318.
- Orsi, M. and Essex, J. W. (2011). The ELBA Force Field for Coarse-Grain Modeling of Lipid Membranes. *PLoS ONE*, 6(12):e28637.
- Pan, J., Heberle, F. A., Tristram-Nagle, S., et al. (2012). Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1818(9):2135–2148.

- Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190.
- Pattni, B. S., Chupin, V. V., and Torchilin, V. P. (2015). New Developments in Liposomal Drug Delivery. *Chemical Reviews*, 115(19):10938–10966.
- Pattni, B. S. and Torchilin, V. P. (2015). Targeted Drug Delivery Systems: Strategies and Challenges. pages 3–38. Springer, Cham.
- Perilla, J. R., Hadden, J. A., Goh, B. C., et al. (2016). All-Atom Molecular Dynamics of Virus Capsids as Drug Targets. *The journal of physical chemistry letters*, 7(10):1836–44.
- Periole, X., Cavalli, M., Marrink, S.-J., et al. (2009). Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *Journal of Chemical Theory and Computation*, 5(9):2531–2543.
- Peschel, A., Jack, R. W., Otto, M., et al. (2001). *Staphylococcus aureus* Resistance to Human Defensins and Evasion of Neutrophil Killing via the Novel Virulence Factor Mprf Is Based on Modification of Membrane Lipids with l-Lysine. *The Journal of Experimental Medicine*, 193(9):1067–1076.
- Peschel, A. and Sahl, H.-G. (2006). The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature Reviews Microbiology*, 4(7):529–536.
- Petrache, H. I., Dodd, S. W., and Brown, M. F. (2000). Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by (2)H NMR spectroscopy. *Biophysical journal*, 79(6):3172–92.
- Petrov, D., Margreitter, C., Grandits, M., et al. (2013). A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. *PLoS Computational Biology*, 9(7):e1003154.
- Phillips, J. C., Braun, R., Wang, W., et al. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802.

- Pickar, A. D. and Benz, R. (1978). Transport of oppositely charged lipophilic probe ions in lipid bilayer membranes having various structures. *The Journal of Membrane Biology*, 44(3-4):353–376.
- Piggot, T. J., Allison, J. R., Sessions, R. B., et al. (2017). On the Calculation of Acyl Chain Order Parameters from Lipid Simulations. *Journal of Chemical Theory and Computation*, 13(11):5683–5696.
- Piggot, T. J., Holdbrook, D. A., and Khalid, S. (2011). Electroporation of the *E. coli* and *S. Aureus* Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes. *The Journal of Physical Chemistry B*, 115(45):13381–13388.
- Piggot, T. J., Piñeiro, Á., and Khalid, S. (2012). Molecular Dynamics Simulations of Phosphatidylcholine Membranes: A Comparative Force Field Study. *Journal of Chemical Theory and Computation*, 8(11):4593–4609.
- Pino-Angeles, A., Leveritt, J. M., and Lazaridis, T. (2016). Pore Structure and Synergy in Antimicrobial Peptides of the Magainin Family. *PLOS Computational Biology*, 12(1):e1004570.
- Pirtskhalava, M., Gabrielian, A., Cruz, P., et al. (2016). DBAASP v.2: an enhanced database of structure and antimicrobial/cytotoxic activity of natural and synthetic peptides. *Nucleic Acids Research*, 44(13):6503–6503.
- Pluhackova, K., Kirsch, S. A., Han, J., et al. (2016). A Critical Comparison of Biomembrane Force Fields: Structure and Dynamics of Model DMPC, POPC, and POPE Bilayers. *The Journal of Physical Chemistry B*, 120(16):3888–3903.
- Poger, D., Caron, B., and Mark, A. E. (2016). Validating lipid force fields against experimental data: Progress, challenges and perspectives. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(7):1556–1565.
- Poger, D. and Mark, A. E. (2010). On the Validation of Molecular Dynamics Simulations of Saturated and cis -Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *Journal of Chemical Theory and Computation*, 6(1):325–336.

- Poger, D., Van Gunsteren, W. F., and Mark, A. E. (2010). A new force field for simulating phosphatidylcholine bilayers. *Journal of Computational Chemistry*, 31(6):1117–1125.
- Pokhrel, N. and Maibaum, L. (2018). Free Energy Calculations of Membrane Permeation: Challenges Due to Strong HeadgroupSolute Interactions. *Journal of Chemical Theory and Computation*, 14(3):1762–1771.
- Ponder, J. W., Wu, C., Ren, P., et al. (2010). Current Status of the AMOEBA Polarizable Force Field. *The Journal of Physical Chemistry B*, 114(8):2549–2564.
- Poulou, A., Voulgari, E., Vrioni, G., et al. (2013). Outbreak Caused by an Ertapenem-Resistant, CTX-M-15-Producing Klebsiella pneumoniae Sequence Type 101 Clone Carrying an OmpK36 Porin Variant. *Journal of Clinical Microbiology*, 51(10):3176–3182.
- Raetz, C. R. H. and Whitfield, C. (2002). Lipopolysaccharide Endotoxins. *Annual Review of Biochemistry*, 71(1):635–700.
- Rand, R. P., Fuller, N., Parsegian, V. A., et al. (1988). Variation in hydration forces between neutral phospholipid bilayers: evidence for hydration attraction. *Biochemistry*, 27(20):7711–7722.
- Reif, M. M., Hünenberger, P. H., and Oostenbrink, C. (2012). New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *Journal of Chemical Theory and Computation*, 8(10):3705–3723.
- Reif, M. M., Winger, M., and Oostenbrink, C. (2013). Testing of the GROMOS Force-Field Parameter Set 54A8: Structural Properties of Electrolyte Solutions, Lipid Bilayers, and Proteins. *Journal of Chemical Theory and Computation*, 9(2):1247–1264.
- Reißen, S., Poger, D., Stroet, M., et al. (2017). Real Cost of Speed: The Effect of a Time-Saving Multiple-Time-Stepping Algorithm on the Accuracy of Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 13(6):2367–2372.

- Ren, P. and Ponder, J. W. (2003). Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *Jurnal of Physical Chemistry B*, 24(107):5933–5947.
- Risselada, H. J. and Marrink, S. J. (2008). The molecular face of lipid rafts in model membranes. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45):17367–72.
- Rossi, G., Manfrin, A., and Lutolf, M. P. (2018). Progress and potential in organoid research. *Nature Reviews Genetics*, 19(11):671–687.
- Saar-Dover, R., Bitler, A., Nezer, R., et al. (2012). D-Alanylation of Lipoteichoic Acids Confers Resistance to Cationic Peptides in Group B Streptococcus by Increasing the Cell Wall Density. *PLoS Pathogens*, 8(9):e1002891.
- Samsudin, F., Boags, A., Piggot, T. J., et al. (2017). Braun’s Lipoprotein Facilitates OmpA Interaction with the Escherichia coli Cell Wall. *Bio-physical Journal*, 113(7):1496–1504.
- Sánchez, L., Calvo, M., and Brock, J. H. (1992). Biological role of lactoferrin. *Archives of disease in childhood*, 67(5):657–61.
- Sandoval-Perez, A., Pluhackova, K., and Böckmann, R. A. (2017). Critical Comparison of Biomembrane Force Fields: ProteinLipid Interactions at the Membrane Interface. *Journal of Chemical Theory and Computation*, 13(5):2310–2321.
- Santos, R., Ursu, O., Gaulton, A., et al. (2017). A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery*, 16(1):19–34.
- Schamberger, J. and Clarke, R. J. (2002). Hydrophobic Ion Hydration and the Magnitude of the Dipole Potential. *Biophysical Journal*, 82(6):3081–3088.
- Schibli, D. J., Hwang, P. M., and Vogel, H. J. (1999). The structure of the antimicrobial active center of lactoferricin B bound to sodium dodecyl sulfate micelles. *FEBS Letters*, 446(2-3):213–217.
- Schmid, N., Eichenberger, A. P., Choutko, A., et al. (2011). Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal*, 40(7):843–856.

- Schoonen, L. and van Hest, J. C. M. (2014). Functionalization of protein-based nanocages for drug delivery applications. *Nanoscale*, 6(13):7124–7141.
- Scomparin, C., Lecuyer, S., Ferreira, M., et al. (2009). Diffusion in supported lipid bilayers: Influence of substrate and preparation technique on the internal dynamics. *The European Physical Journal E*, 28(2):211–220.
- Scott, D., Ghosh, A., Di, L., et al. (2017). Passive drug permeation through membranes and cellular distribution. *Pharmacological Research*, 117:94–102.
- Scott, W., Hünenberger, P., Tironi, I., et al. (1999). The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, 103(19):3596–3607.
- Semchyschyn, D. J. and Macdonald, P. M. (2004). Conformational response of the phosphatidylcholine headgroup to bilayer surface charge: torsion angle constraints from dipolar and quadrupolar couplings in bicelles. *Magnetic Resonance in Chemistry*, 42(2):89–104.
- Sengupta, D., Leontiadou, H., Mark, A. E., et al. (2008). Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(10):2308–2317.
- Seo, M., Rauscher, S., Pomès, R., et al. (2012). Improving Internal Peptide Dynamics in the Coarse-Grained MARTINI Model: Toward Large-Scale Simulations of Amyloid- and Elastin-like Peptides. *Journal of Chemical Theory and Computation*, 8(5):1774–1785.
- Shao, Y., Gan, Z., Epifanovsky, E., et al. (2015). Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Molecular Physics*, 113(2):184–215.
- Sharpe, H. J., Stevens, T. J., and Munro, S. (2010). A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell*, 142(1):158–69.
- Shau, H., Kim, A., and Golub, S. H. (1992). Modulation of natural killer and lymphokine-activated killer cell cytotoxicity by lactoferrin. *Journal of Leukocyte Biology*, 51(4):343–349.

- Shi, Y. and Massagué, J. (2003). Mechanisms of TGF- β Signaling from Cell Membrane to the Nucleus. *Cell*, 113(6):685–700.
- Shore, A. C., Deasy, E. C., Slickers, P., et al. (2011). Detection of Staphylococcal Cassette Chromosome *mec* Type XI Carrying Highly Divergent *mecA*, *mecI*, *mecR1*, *blaZ*, and *ccr* Genes in Human Clinical Isolates of Clonal Complex 130 Methicillin-Resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 55(8):3765–3773.
- Sieprawska-Lupa, M., Mydel, P., Krawczyk, K., et al. (2004). Degradation of human antimicrobial peptide LL-37 by *Staphylococcus aureus*-derived proteinases. *Antimicrobial agents and chemotherapy*, 48(12):4673–9.
- Silhavy, T. J., Kahne, D., and Walker, S. (2010). The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology*, 2(5):a000414–a000414.
- Silva, R. F., Araújo, D. R., Silva, E. R., et al. (2013). L-Diphenylalanine Microtubes As a Potential Drug-Delivery System: Characterization, Release Kinetics, and Cytotoxicity. *Langmuir*, 29(32):10205–10212.
- Singh, P., Pandit, S., Mokkapati, V. R. S. S., et al. (2018). Gold Nanoparticles in Diagnostics and Therapeutics for Human Cancer. *International journal of molecular sciences*, 19(7).
- Skjervik, Å. A., Madej, B. D., Dickson, C. J., et al. (2015). All-atom lipid bilayer self-assembly with the AMBER and CHARMM lipid force fields. *Chemical Communications*, 51(21):4402–4405.
- Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiology Reviews*, 40(1):133–159.
- Song, C., de Groot, B. L., and Sansom, M. S. (2019). Lipid Bilayer Composition Influences the Activity of the Antimicrobial Peptide Dermcidin Channel. *Biophysical Journal*, 116(9):1658–1666.
- Sørensen, S. J., Bailey, M., Hansen, L. H., et al. (2005). Studying plasmid horizontal transfer in situ: a critical review. *Nature Reviews Microbiology*, 3(9):700–710.

- Spaar, A., Münster, C., and Salditt, T. (2004). Conformation of peptides in lipid membranes studied by x-ray grazing incidence scattering. *Biophysical journal*, 87(1):396–407.
- Spänig, S. and Heider, D. (2019). Encodings and models for antimicrobial peptide classification for multi-resistant pathogens. *BioData Mining*, 12(1):7.
- Spector, A. A. and Yorek, M. A. (1985). Membrane lipid composition and cellular function. *Journal of lipid research*, 26(9):1015–35.
- Starke-Peterkovic, T. and Clarke, R. J. (2009). Effect of headgroup on the dipole potential of phospholipid vesicles. *European Biophysics Journal*, 39(1):103–110.
- Strebhardt, K. and Ullrich, A. (2008). Paul Ehrlich's magic bullet concept: 100 years of progress. *Nature Reviews Cancer*, 8(6):473–480.
- Stremper, N., Strehmel, J., and Overhage, J. (2014). Potential Application of Antimicrobial Peptides in the Treatment of Bacterial Biofilm Infections. *Current Pharmaceutical Design*, 21(1):67–84.
- Strøm, M. B., Haug, B. E., Rekdal, O., et al. (2002). Important structural features of 15-residue lactoferricin derivatives and methods for improvement of antimicrobial activity. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 80(1):65–74.
- Sun, D., Forsman, J., and Woodward, C. E. (2015). Atomistic Molecular Simulations Suggest a Kinetic Model for Membrane Translocation by Arginine-Rich Peptides. *The Journal of Physical Chemistry B*, 119(45):14413–14420.
- Sun, R., Dama, J. F., Tan, J. S., et al. (2016). Transition-Tempered Metadynamics Is a Promising Tool for Studying the Permeation of Drug-like Molecules through Membranes. *Journal of Chemical Theory and Computation*, 12(10):5157–5169.
- Sun, W.-J., Suter, R. M., Knewton, M. A., et al. (1994). Order and disorder in fully hydrated unoriented bilayers of gel-phase dipalmitoylphosphatidyl-choline. *Physical Review E*, 49(5):4665–4676.

- Sunde, M., Serpell, L. C., Bartlam, M., et al. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology*, 273(3):729–739.
- Swoboda, J. G., Campbell, J., Meredith, T. C., et al. (2009). Wall Teichoic Acid Function, Biosynthesis, and Inhibition. *ChemBioChem*, 11(1):35–45.
- Takeuchi, K., Takahashi, H., Sugai, M., et al. (2004). Channel-forming membrane permeabilization by an antibacterial protein, sapecin: determination of membrane-buried and oligomerization surfaces by NMR. *The Journal of biological chemistry*, 279(6):4981–7.
- Tang, M. and Hong, M. (2009). Structure and mechanism of β -hairpin antimicrobial peptides in lipid bilayers from solid-state NMR spectroscopy. *Molecular BioSystems*, 5(4):317.
- The Johns Hopkins University (2017). Hopkins ABX Guide. <http://www.hopkins-abxguide.org/>.
- Tieleman, D. and Marrink, S.-J. (2006). Lipids Out of Equilibrium: Energetics of Desorption and Pore Mediated Flip-Flop. *Journal of the American Chemical Society*, 128(38):12462–12467.
- Tieleman, D. P. (2004). The molecular basis of electroporation. *BMC Biochemistry*, 5(1):10.
- Tieleman, D. P. and Berendsen, H. J. C. (1998). Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *The Journal of Chemical Physics*, 105(11):4871.
- Tironi, I. G., Sperb, R., Smith, P. E., et al. (1995). A generalized reaction field method for molecular dynamics simulations. *The Journal of Chemical Physics*, 102(13):5451–5459.
- Tomita, M., Takase, M., Bellamy, W., et al. (1994). A review: the active peptide of lactoferrin. *Acta paediatrica Japonica : Overseas edition*, 36(5):585–91.
- Torres, M. D., Sothiselvam, S., Lu, T. K., et al. (2019). Peptide Design Principles for Antimicrobial Applications. *Journal of Molecular Biology*.

- Torrie, G. and Valleau, J. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199.
- Tsai, C.-W., Hsu, N.-Y., Wang, C.-H., et al. (2009). Coupling Molecular Dynamics Simulations with Experiments for the Rational Design of Indolicidin-Analogous Antimicrobial Peptides. *Journal of Molecular Biology*, 392(3):837–854.
- Tsutsumi, A., Javkhlanlugs, N., Kira, A., et al. (2012). Structure and orientation of bovine lactoferrampin in the mimetic bacterial membrane as revealed by solid-state NMR and molecular dynamics simulation. *Biophysical journal*, 103(8):1735–43.
- ULC, C. C. G. (2018). Molecular operating environment (moe), 2013.08. 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7.
- Ulmschneider, J. P. (2017). Charged Antimicrobial Peptides Can Translocate across Membranes without Forming Channel-like Pores. *Biophysical Journal*, 113(1):73–81.
- Ulrich, A. and Watts, A. (1994). Molecular response of the lipid head-group to bilayer hydration monitored by ^2H -NMR. *Biophysical Journal*, 66(5):1441–1449.
- Ulrich, A. S., Volke, F., and Watts, A. (1990). The dependence of phospholipid head-group mobility on hydration as studied by deuterium-NMR spin-lattice relaxation time measurements. *Chemistry and Physics of Lipids*, 55(1):61–66.
- Uppulury, K., Coppock, P. S., and Kindt, J. T. (2015). Molecular Simulation of the DPPE Lipid Bilayer Gel Phase: Coupling between Molecular Packing Order and Tail Tilt Angle. *J. Phys. Chem. B*, 119:6.
- van Gunsteren, W. F., Bakowies, D., Baron, R., et al. (2006). Biomolecular Modeling: Goals, Problems, Perspectives. *Angewandte Chemie International Edition*, 45(25):4064–4092.

- van Gunsteren, W. F., Dolenc, J., and Mark, A. E. (2008). Molecular simulation as an aid to experimentalists. *Current Opinion in Structural Biology*, 18(2):149–153.
- Van Lehn, R. C. and Alexander-Katz, A. (2014). Membrane-Embedded Nanoparticles Induce Lipid Rearrangements Similar to Those Exhibited by Biological Membrane Proteins. *The Journal of Physical Chemistry B*, 118(44):12586–12598.
- Van Lehn, R. C., Ricci, M., Silva, P. H., et al. (2014). Lipid tail protrusions mediate the insertion of nanoparticles into model cell membranes. *Nature Communications*, 5(1):4482.
- van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008). Membrane lipids: where they are and how they behave. *Nature reviews. Molecular cell biology*, 9(2):112–24.
- Vargiu, A. V. and Nikaido, H. (2012). Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50):20637–42.
- Veltri, D., Kamath, U., and Shehu, A. (2018). Deep learning improves antimicrobial peptide recognition. *Bioinformatics*, 34(16):2740–2747.
- Venable, R. M., Ingólfsson, H. I., Lerner, M. G., et al. (2017). Lipid and Peptide Diffusion in Bilayers: The Saffman-Delbrück Model and Periodic Boundary Conditions. *The Journal of Physical Chemistry B*, 121(15):3443–3457.
- Vermeer, L. S., de Groot, B. L., Réat, V., et al. (2007). Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with ${}^2\text{H}$ NMR experiments. *European Biophysics Journal*, 36(8):919–931.
- Vernier, P. T., Ziegler, M. J., Sun, Y., et al. (2006). Nanopore Formation and Phosphatidylserine Externalization in a Phospholipid Bilayer at High Transmembrane Potential. *Journal of the American Chemical Society*, 128(19):6288–6289.

- Vivcharuk, V., Tomberli, B., Tolokh, I. S., et al. (2008). Prediction of binding free energy for adsorption of antimicrobial peptide lactoferricin B on a POPC membrane. *Physical Review E*, 77(3):031913.
- Vogel, A. and Huster, D. (2017). Combining NMR Spectroscopy and Molecular Dynamics Simulation to Investigate the Structure and Dynamics of Membrane-Associated Proteins. pages 311–350. Springer, Cham.
- Vögele, M., Köfinger, J., and Hummer, G. (2018). Hydrodynamics of Diffusion in Lipid Membrane Simulations. *Physical Review Letters*, 120(26):268104.
- Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2):149–167.
- Vuong, C., Kocanova, S., Voyich, J. M., et al. (2004a). A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion, and virulence. *The Journal of biological chemistry*, 279(52):54881–6.
- Vuong, C., Voyich, J. M., Fischer, E. R., et al. (2004b). Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. *Cellular microbiology*, 6(3):269–75.
- Wakabayashi, H., Abe, S., Okutomi, T., et al. (1996). Cooperative anti-*Candida* effects of lactoferrin or its peptides in combination with azole antifungal agents. *Microbiology and immunology*, 40(11):821–5.
- Wang, G., Li, X., Wang, Z., et al. (2016a). APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Research*, 44(D1):D1087–D1093.
- Wang, H., Zhang, H., Liu, C., et al. (2012a). Coarse-grained molecular dynamics simulation of self-assembly of polyacrylamide and sodium dodecylsulfate in aqueous solution. *Journal of Colloid and Interface Science*, 386(1):205–211.
- Wang, J., Chou, S., Xu, L., et al. (2015). High specific selectivity and Membrane-Active Mechanism of the synthetic centrosymmetric α -helical peptides with Gly-Gly pairs. *Scientific reports*, 5(1):15963.

- Wang, J., Wolf, R. M., Caldwell, J. W., et al. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174.
- Wang, L., Bose, P. S., and Sigworth, F. J. (2006a). Using cryo-EM to measure the dipole potential of a lipid membrane. *Proceedings of the National Academy of Sciences*, 103(49):18528–18533.
- Wang, X., McGrath, S. C., Cotter, R. J., et al. (2006b). Expression cloning and periplasmic orientation of the Francisella novicida lipid A 4'-phosphatase LpxF. *The Journal of biological chemistry*, 281(14):9321–30.
- Wang, Y., Chen, C. H., Hu, D., et al. (2016b). Spontaneous formation of structurally diverse membrane channel architectures from a single antimicrobial peptide. *Nature Communications*, 7:13535.
- Wang, Y., Schlamadinger, D. E., Kim, J. E., et al. (2012b). Comparative molecular dynamics simulations of the antimicrobial peptide CM15 in model lipid bilayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1818(5):1402–1409.
- Wang, Y., Zhao, T., Wei, D., et al. (2014). How reliable are molecular dynamics simulations of membrane active antimicrobial peptides? *BBA - Biomembranes*, 1838(9):2280–2288.
- Wang, Z. Q., Lin, H. N., Li, S., et al. (1994). Calorimetric studies and molecular mechanics simulations of monounsaturated phosphatidylethanolamine bilayers. *The Journal of biological chemistry*, 269(38):23491–9.
- Warshaviak, D. T., Muellner, M. J., and Chachisvilis, M. (2011). Effect of membrane tension on the electric field and dipole potential of lipid bilayer membrane. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(10):2608–2617.
- Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., et al. (2015). Computational Lipidomics with *insane* : A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *Journal of Chemical Theory and Computation*, 11(5):2144–2155.
- Wikipedia (2015). Liposome.

- Wilson, J. R., Clark, R. B., Banderali, U., et al. (2011). Measurement of the membrane potential in small cells using patch clamp methods. *Channels*, 5(6):530–537.
- Wipf, P., Xiao, J., and Stephenson, C. R. J. (2009). Peptide-Like Molecules (PLMs): A Journey from Peptide Bond Isosteres to Gramicidin S Mimetics and Mitochondrial Targeting Agents. *CHIMIA International Journal for Chemistry*, 63(11):764–775.
- Wishart, D. S., Feunang, Y. D., Guo, A. C., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research*, 46(D1):D1074–D1082.
- Woodford, N. and Johnson, A. P. (2013). Global spread of antibiotic resistance: the example of New Delhi metallo- β -lactamase (NDM)-mediated carbapenem resistance. *Journal of Medical Microbiology*, 62(4):499–513.
- Wright, G. (2005). Bacterial resistance to antibiotics: Enzymatic degradation and modification. *Advanced Drug Delivery Reviews*, 57(10):1451–1470.
- Wu, D. and Yang, X. (2012). Coarse-Grained Molecular Simulation of Self-Assembly for Nonionic Surfactants on Graphene Nanostructures. *The Journal of Physical Chemistry B*, 116(39):12048–12056.
- Wu, W., Hsiao, S. C., Carrico, Z. M., et al. (2009). Genome-free viral capsids as multivalent carriers for taxol delivery. *Angewandte Chemie (International ed. in English)*, 48(50):9493–7.
- Yamaguchi, S., Huster, D., Waring, A., et al. (2001). Orientation and Dynamics of an Antimicrobial Peptide in the Lipid Bilayer by Solid-State NMR Spectroscopy. *Biophysical Journal*, 81(4):2203–2214.
- Yang, L., Harroun, T. A., Weiss, T. M., et al. (2001). Barrel-stave model or toroidal model? A case study on melittin pores. *Biophysical journal*, 81(3):1475–85.
- Yeaman, M. R. and Yount, N. Y. (2003). Mechanisms of Antimicrobial Peptide Action and Resistance. *Pharmacological Reviews*, 55(1):27–55.
- Yeates, T. O. (2019). Protein assembles into Archimedean geometry. *Nature*, 569(7756):340–342.

- Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D., et al. (2010). Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Computational Biology*, 6(6):e1000810.
- Yingchoncharoen, P., Kalinowski, D. S., and Richardson, D. R. (2016). Lipid-Based Drug Delivery Systems in Cancer Therapy: What Is Available and What Is Yet to Come. *Pharmacological reviews*, 68(3):701–87.
- Yonezawa, A., Kuwahara, J., Fujii, N., et al. (1992). Binding of tachyplesin I to DNA revealed by footprinting analysis: significant contribution of secondary structure to DNA binding and implication for biological action. *Biochemistry*, 31(11):2998–3004.
- Yoo, J.-W., Irvine, D. J., Discher, D. E., et al. (2011). Bio-inspired, bio-engineered and biomimetic drug delivery carriers. *Nature Reviews Drug Discovery*, 10(7):521–535.
- Zasloff, M. (2002). Antimicrobial peptides of multicellular organisms. *Nature*, 415(6870):389–395.
- Zhang, F., Wen, Y., and Guo, X. (2014a). CRISPR/Cas9 for genome editing: progress, implications and challenges. *Human Molecular Genetics*, 23(R1):R40–R46.
- Zhang, L., Rozek, A., and Hancock, R. E. (2001). Interaction of cationic antimicrobial peptides with model membranes. *The Journal of biological chemistry*, 276(38):35714–22.
- Zhang, Q., Jiang, Q., Li, N., et al. (2014b). DNA Origami as an *In Vivo* Drug Delivery Vehicle for Cancer Therapy. *ACS Nano*, 8(7):6633–6643.
- Zhao, L., Cao, Z., Bian, Y., et al. (2018). Molecular Dynamics Simulations of Human Antimicrobial Peptide LL-37 in Model POPC and POPG Lipid Bilayers. *International Journal of Molecular Sciences*, 19(4):1186.
- Zhao, R., Babani, S., Gao, F., et al. (2000). The mechanism of transport of the multitargeted antifolate (MTA) and its cross-resistance pattern in cells with markedly impaired transport of methotrexate. *Clinical cancer research: an official journal of the American Association for Cancer Research*, 6(9):3687–95.

- Zhu, J., Alexov, E., and Honig, B. (2005). Comparative Study of Generalized Born Models: Born Radii and Peptide Folding. *The Journal of Physical Chemistry B*, 109(7):3008–3022.