UNIVERSIDADE NOVA DE LISBOA

MScCBBi

MASTER IN
COMPUTATIONAL BIOLOGY
& BIOINFORMATICS

Ana Carolina Borges Araújo
BSc in Biochemistry

Poking holes in membranes with coarse-grained simulations

Supervisor: Manuel Nuno Melo, PhD, ITQB-NOVA

Sep, 2023

ITQB NOVA
NOVA SCHOOL OF SCIENCE & TECHNOLOGY
NOVA MEDICAL SCHOOL
NOVA IMS Information Management School
DESDE 1902 INSTITUTO DE HIGIENE E MEDICINA TROPICAL
UNIVERSIDADE NOVA DE LISBOA

UNIVERSIDADE NOVA DE LISBOA

MScCBBi

MASTER IN
COMPUTATIONAL BIOLOGY
& BIOINFORMATICS

Ana Carolina Borges Araújo
BSc in Biochemistry

Poking holes in membranes with coarse-grained simulations

Supervisor: Manuel Nuno Melo, PhD, ITQB-NOVA

Sep, 2023

ITQB NOVA
NOVA SCHOOL OF SCIENCE & TECHNOLOGY
NOVA MEDICAL SCHOOL
NOVA IMS Information Management School
DESDE 1902 INSTITUTO DE HIGIENE E MEDICINA TROPICAL
UNIVERSIDADE NOVA DE LISBOA

Poking holes in membranes with coarse-grained simulations

This document was created with the (pdf/Xe/Lua)LaTeX processor and the NOVAthesis template (v6.10.17) [1].

Acknowledgements

If there is one thing I am certain of is that my achievements are a result of the support from the incredible people that have stood by me throughout my life. This work would not have been possible without the help of many.

I would like to start by acknowledging Dr Manuel N. Melo, who was an incredible supervisor and mentor with whom I learnt so much. Thank you for constantly offering me opportunities to challenge myself and to meet and work with extraordinary people. I hope I have fulfilled your expectations.

I have to thank Kasper Pedersen for his continuous availability, expertise, and fruitful discussions regarding my project.

To my colleagues at the Multiscale Modelling, Protein Modelling, and Molecular Simulation Labs, I am so grateful for how you welcomed me into this big family with open arms and for making every day at the lab that much more fun.

A special thank you to Fernando, who, from day one, offered me his unwavering support and always brought a little bit of unpredictability and chaos.

I also couldn't have overcome this year without my family and friends - people that play a big part in my life and to whom I owe a few words.

A big thank you to my close friends for forgiving me too many times for being absent on birthdays and celebrations because I was working on this thesis. It has been a joy to grow up with you.

To Inês and André, who have become like siblings to me, thank you for the earfuls and cheering-ups whenever you knew I needed them. I will forever look up to both of you. To Beatriz, I am so grateful for your ability to transform every negative into a hilarious positive and for always keeping me somewhat sane.

My parents, who have worked tirelessly for me to be where I am. Whoever I am today is only a reflection of your unconditional love and support. Thank you for never leaving my side, even when I am at fault.

Big cheers to my cat, who undoubtedly cares the most about this achievement.

To my older brother and best friend, I am forever indebted for how you constantly push me to be and do better. For myself and the world around me. I can only hope to become half the genius you are.

A huge thank you to Mariana, my big sister, for always offering me her best advice, friendship, and helping hand.

And lastly, my grandmother, who I know would have been incredibly proud to see me get here. I think of you every day.

Abstract

The lipid membrane is a crucial structure as it plays many roles in the cell, such as regulation of metabolite transport. Ion flux is typically controlled by membrane transporters but they can also leak through transient lipid pores. This regulation is essential for the maintenance of cell life. Antimicrobial peptides (AMPs), like alamethicin (Alm), form pores as an action mechanism against pathogens. Thus, pores are imperative structures to gain insight into these mechanisms. Coarse-grained (CG) molecular dynamics (MD) simulations allow us to study systems at time and sizescales otherwise inaccessible, at the expense of some atomic detail. In the Martini 3 CG force field, transient lipid pores are not well modeled — with higher energies of formation than those obtained by atomistic (AA) methods. Here, we studied Martini 3 simulations of transient pore formation in lipid bilayers and compared them to AA behavior, to try to improve lipid poration in the CG force field. We found a target — the lipid tail splay — that may affect formation behavior, and tested lipid prototypes with new tail topologies to analyze whether it improved. Whilst we did not obtain better free-energy profiles, we believe this target can still be causing incorrect modeling as we were not able to mimic the AA tail splay distribution with accuracy. In parallel, Alm's mechanism of action was studied using Martini 3. Alm poration was not observed with Martini 2, given its dependency on protein-protein interactions. With their improvement and Alm pores being less dependent on the distortion of the membrane topology, we hoped that these pores could be represented with Martini 3. Indeed, we were able to observe Alm pores with ion flux in Martini 3 lipid bilayers. We also saw Alm aggregate into clusters until ultimately forming a raft. Although the discrete levels of conductance were not recovered, our work is still an improvement over past CG studies.

Keywords: Pores, Lipid membranes, Antimicrobial peptides, Alamethicin, Coarse-grain, Martini

Resumo

A membrana lipídica é uma estrutura crucial pelos vários papéis que tem na célula, como a regulação do transporte de metabolitos. O fluxo de iões é tipicamente controlado por transportadores de membrana mas também ocorre fuga através de poros lipídicos transitórios. Esta regulação é essencial para a vida celular. Péptidos antimicrobianos (AMPs), como a alameticina (Alm), formam poros como mecanismo de ação contra patogénios. Logo, poros são estruturas imperativas para compreendermos estes mecanismos. Simulações de dinâmica molecular (MD) *coarse-grained* (CG) permitem-nos estudar sistemas a escalas de tempo e tamanho de outro modo inacessíveis, à custa de detalhe atómico. No campo de forças CG Martini 3, poros lipídicos transitórios não são bem modelados — com energias de formação mais elevadas do que as obtidas por métodos atomísticos (AA). Aqui, estudámos simulações Martini 3 da formação de poros transitórios e comparámo-las com o comportamento AA, para melhorar o modelo do campo de forças CG. Encontrámos um alvo - a abertura da cauda dos lípidos - que pode afetar o comportamento e estudámos protótipos de lípidos com novas topologias de cauda para analisar se este melhorava. Não obtivemos um melhor perfil de energia livre de formação destes poros, porém, acreditamos que este alvo pode ainda ser responsável, pois a distribuição AA não foi replicada em CG com exatidão. Paralelamente, estudámos o mecanismo de ação da Alm em Martini 3. Poros de Alm não eram replicados em Martini 2 dada a dependência de interações proteína-proteína. Com a melhoria destas e sendo os poros de Alm menos dependentes da distorção da topologia da membrana, esperávamos que estes poros pudessem ser representados no Martini 3. De facto, obtivemos poros de Alm com fluxo iónico e também observámos a Alm agregar em *clusters* até formar uma jangada. Apesar de não recuperarmos os níveis discretos de condutância, o nosso trabalho é uma melhoria em relação a estudos anteriores de CG.

Palavras-chave: Poros, Membrana lipídica, Péptidos antimicrobianos, Alameticina, Coarse-grain, Martini

Contents

List of Figures xiii

List of Tables xvii

Acronyms xix

Symbols xxi

Chemical Symbols xxiii

1 Introduction 1

1.1 Biological membranes . . . 1

1.1.1 The lipid bilayer . . . 2

1.1.2 Membrane proteins . . . 3

1.2 Pores in the biological membrane . . . 3

1.3 Lipid–lined pores . . . 5

1.4 Antimicrobial peptides . . . 7

1.4.1 Alamethicin . . . 8

1.5 Molecular Dynamics simulations . . . 11

1.5.1 Coarse-Grained vs All-Atom simulations . . . 11

1.5.2 The Martini 3 force field . . . 11

1.6 Objectives . . . 12

1.7 Other work - Cholesterol parameterization for the Martini 3 force field . . . 13

2 Materials and Methods 15

2.1 Molecular Mechanics . . . 15

2.1.1 Force fields . . . 15

2.1.1.1 Bonded interactions . . . 16

2.1.1.2 Non-bonded interactions . . . 17

2.2 Molecular Dynamics Simulations . . . 18

2.2.1 Energy minimization . . . 19

2.2.2 Temperature and Pressure coupling . . . 20

2.2.3 Periodic boundary conditions . . . 21

2.2.4 Coarse-grained simulations . . . 21
2.2.4.1 Martini 3 CG force field . . . 23
2.2.5 Umbrella sampling . . . 23
2.2.6 The lipid pore model for MD simulations . . . 24
2.3 Setup of lipid pore study MD simulations . . . 25
2.3.1 All-atom simulations . . . 26
2.3.2 Coarse-grained simulations . . . 26
2.4 Analysis of lipid pore MD simulations . . . 28
2.4.1 Weighted Histogram Analysis Method (WHAM) . . . 28
2.4.2 Lipid angle and distance distributions . . . 28
2.4.3 Lipid bead distances matrix . . . 29
2.4.4 Root mean square error (RMSE) . . . 29
2.4.5 Principal Component Analysis (PCA) . . . 29
2.4.6 Area per lipid distribution . . . 29
2.4.7 Membrane thickness analysis . . . 29
2.5 Setup of alamethicin MD simulations . . . 30
2.6 Analysis of alamethicin CG MD simulations . . . 31
2.6.1 Stamp plot . . . 31
2.6.1.1 Total number of pores per frame . . . 31
2.6.1.2 Total area of pore distribution . . . 31
2.6.2 Water contacts distribution . . . 31

3 Results and Discussion **33**
3.1 Martini 3 CG lipid–lined pore simulations . . . 33
3.1.1 Lipid–lined pore formation free-energy in CG simulations . . . 34
3.1.2 Pore lipid behavior . . . 35
3.1.3 Testing DLPC lipid topologies . . . 37
3.1.3.1 Prototype 1 - Added tail-tail bonds . . . 38
3.1.3.2 Prototype 2 - Added tail-tail and head-tail bonds . . . 40
3.1.3.3 Mixed Membrane . . . 42
3.2 Alamethicin pores . . . 46
3.2.1 Peptide aggregation analysis . . . 47
3.2.2 Reduced water-contact simulations . . . 49

4 Conclusion **53**

Bibliography **57**

Appendices

A Appendix **73**
A.1 Supplementary Tables . . . 73

A.2 Supplementary Videos 74
A.3 Supplementary Figures 74

Annexes

I Annex 1 **101**
I.1 Cholesterol parameterization 101

List of Figures

1.1 Representation of the fluid mosaic model for plasma membranes 2
1.2 Examples of integral, peripheral, and amphitropic membrane proteins . . . 4
1.3 Examples of protein pore formation and structures in biomembranes 5
1.4 Snapshots of the spontaneous closure of a lipid pore in a POPC membrane . 6
1.5 Examples of antimicrobial peptides' proposed mechanisms of action 8
1.6 Alamethicin's chemical structure. 9
1.7 Alamethicin's multiple conductance levels. 9
1.8 Alamethicin aggregation in atomistic and Martini 2 simulations 10

2.1 Examples of biomolecular force field terms 16
2.2 Energy minimization of a molecule's conformation 19
2.3 Periodic boundary conditions schematic in two dimensions 22
2.4 Examples of the Martini mapping 24
2.5 Flat-bottomed position restraint potentials 27

3.1 Standard Martini mapping of a phospholipid molecule 33
3.2 Potential Mean Force (PMF) distribution of lipid pore formation in atomistic and coarse-grained membrane systems 35
3.3 Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids 36
3.4 Distance matrices of the forward–mapped atomistic DLPC bead distances subtracted from the version 18 Martini 3 DLPC bead distances 37
3.5 Area per lipid and membrane thickness distributions of membranes built with the first lipid prototype 38
3.6 Free energy of lipid pore formation in the first lipid prototype membrane system 39
3.7 Distance matrices of the forward–mapped atomistic lipid's bead distances subtracted from the first lipid prototype lipid's bead distances 40
3.8 Free energy of lipid pore formation in the second lipid prototype membrane system. 41
3.9 Distance matrices of the forward–mapped atomistic lipid's bead distances subtracted from the second lipid prototype lipid's bead distances 42
3.10 Free energy of lipid pore formation in the mixed membrane system. 43

3.11 Distance matrices of the forward–mapped atomistic lipid's bead distances subtracted from the mixed membrane system lipid's bead distances 44
3.12 Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids 45
3.13 Distribution of the first principal component projection value 45
3.14 Martini 3 coarse-grained representation of alamethicin peptide 46
3.15 Selected snapshots of two distinct alamethicin pore structures in the high concentration, antiparallel peptide conformation system 47
3.16 Selected snapshot of the formation of the alamethicin raft in the high concentration, antiparallel peptide conformation system. 48
3.17 Connected component analysis results of the four alamethicin systems . . . 49
3.18 Selected snapshot of two distinct alamethicin pore structures in the high concentration, antiparallel peptide conformation system with water-peptide scaled interactions. 50
3.19 Number of water-peptide contacts throughout the simulated time 50
3.20 Connected component analysis results of the scaled water-peptide interactions systems 51

A.1 Distribution of the angle between beads C3A, PO4, and C3B of the lipids. . 74
A.2 Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids 75
A.3 Distribution of the dihedral angle of the lipids 75
A.4 Distribution of the angle formed between the glycerols and the PO4 beads of the lipids 76
A.5 Distribution of the radius of gyration of the lipids' head 76
A.6 Distribution of the angle formed between tail A, GL1 and PO4 77
A.7 Distribution of the angle between beads C1A, GL1, and GL2 of the lipids . . 77
A.8 Distribution of the angle between beads GL1, GL2, and C1B of the lipids . . 78
A.9 Distribution of the radius of gyration of the lipid's tails. 78
A.10 Distance matrices of the DLPC membrane and pore lipid's beads 79
A.11 Distance matrices of the DPPC membrane and pore lipid's beads 80
A.12 Standard deviation matrices of the forward-mapped atomistic DLPC bead distances subtracted from the version 18 Martini 3 DLPC bead distances . . 81
A.13 Distance matrices of the forward-mapped atomistic DPPC bead distances subtracted from the version 18 Martini 3 DPPC bead distances 81
A.14 Standard deviation matrices of the forward-mapped atomistic DPPC bead distances subtracted from the version 18 Martini 3 DPPC bead distances . . 82
A.15 Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids 82
A.16 Distribution of the angle between beads C3A, PO4, and C3B of the lipids. . 83
A.17 Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids 83
A.18 Distribution of the dihedral angle of the lipids 84

A.19 Distribution of the angle formed between the glycerols and the PO4 beads of the lipids 84
A.20 Distribution of the radius of gyration of the lipids' head 85
A.21 Distribution of the angle formed between tail A, GL1 and PO4 85
A.22 Distribution of the angle between beads C1A, GL1, and GL2 of the lipids . . 86
A.23 Distribution of the angle between beads GL1, GL2, and C1B of the lipids . . 86
A.24 Distribution of the radius of gyration of the lipid's tails. 87
A.25 Standard deviation matrices of the forward-mapped atomistic lipid's bead distances subtracted from the first lipid prototype lipid's bead distances . . 87
A.26 Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids 88
A.27 Distribution of the angle between beads C3A, PO4, and C3B of the lipids. . 88
A.28 Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids 89
A.29 Distribution of the dihedral angle of the lipids 89
A.30 Distribution of the angle formed between the glycerols and the PO4 beads of the lipids 90
A.31 Distribution of the radius of gyration of the lipids' head 90
A.32 Distribution of the angle formed between tail A, GL1 and PO4 91
A.33 Distribution of the angle between beads C1A, GL1, and GL2 of the lipids . . 91
A.34 Distribution of the angle between beads GL1, GL2, and C1B of the lipids . . 92
A.35 Distribution of the radius of gyration of the lipid's tails. 92
A.36 Standard deviation matrices of the forward-mapped atomistic lipid's bead distances subtracted from the second lipid prototype lipid's bead distances 93
A.37 Distribution of the angle between beads C3A, PO4, and C3B of the lipids. . 93
A.38 Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids 94
A.39 Distribution of the dihedral angle of the lipids 94
A.40 Distribution of the angle formed between the glycerols and the PO4 beads of the lipids 95
A.41 Distribution of the radius of gyration of the lipids' head 95
A.42 Distribution of the angle formed between tail A, GL1 and PO4 96
A.43 Distribution of the angle between beads C1A, GL1, and GL2 of the lipids . . 96
A.44 Distribution of the angle between beads GL1, GL2, and C1B of the lipids . . 97
A.45 Distribution of the radius of gyration of the lipid's tails. 97
A.46 Standard deviation matrices of the forward-mapped atomistic lipid's bead distances subtracted from the mixed membrane system lipid's bead distances 98
A.47 Principal component 1 distribution for the atomistic DLPC membrane system 98
A.48 Principal component 1 distribution for the Martini 3 in-development version 18 of the DLPC membrane system 99
A.49 Principal component 1 distribution for the mixed membrane system 99

List of Tables

3.1 Root mean square error values of the in–development Martini 3 DLPC bead distances. 37
3.2 Root mean square error values of the in–development and the Martini 3 DLPC prototype 1 bead distances. 40
3.3 Root mean square error values of the in–development and both Martini 3 DLPC prototypes bead distances. 42
3.4 Root mean square error values of the in–development, both Martini 3 DLPC prototypes, and the mixed membrane lipids' bead distances. 43

A.1 Root mean square error values of the in-development Martini 3 DPPC bead distances 73
A.2 Pore lipids' variance values explained by the first four principal components 73
A.3 Membrane lipids' variance values explained by the first four principal components 73

Acronyms

AA	All-atom
Alm	Alamethicin
AMP	Antimicrobial peptide
APL	Area per lipid
CG	Coarse-grained
CHARMM	Chemistry at HARvard Macromolecular Mechanics
DLPC	1,2-dilauroyl-*sn*-glycero-3-phosphocholine
DOFs	Degrees of freedom
DPPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine
ER	Endoplasmic reticulum
LJ	Lennard-Jones
MD	Molecular dynamics
MM	Molecular mechanics
PA	Phosphatidic acid
PBC	Periodic boundary conditions
PC	Phosphatidylcholine
PCA	Principal component analysis
PE	Phosphatidylethanolamine
PES	Potential energy surface
PI	Phosphatidylinositol
PMF	Potential mean force
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine
POPG	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1′-rac-glycerol)
PS	phosphatidylserine

QM Quantum mechanics

RC Reaction coordinate
RMSE Root mean square error

SM Sphingomyelin

VdW Van der Waals
VMD Visual molecular dynamics

WHAM Weighted histogram analysis method

Symbols

Å	Angstrom
$\delta_{\mathbf{s}}$	continuous indicator function
$\boldsymbol{d}_{\mathbf{s}}$	thickness of slices
γ	boundary line tension of pore edge
K	Kelvin
$\boldsymbol{N}_{\mathbf{s}}$	number of slices
π	the numerical value of pi
r	radius of pore
σ	lateral tension

Chemical Symbols

Aib Amino acid residue α-aminoisobutyric acid

Ca^{2+} Calcium positive ion
Cl^{-} Chlorine negative ion

Gln Amino acid residue glutamine

Na^{+} Sodium positive ion

Phl Amino acid residue phenylalaninol

1 Introduction

1.1 Biological membranes

Biological membranes are responsible for delimiting and defining the boundaries of the cell. They also allow for the compartmentalization of processes in eukaryotic cells, which contributes to their increased efficiency. Membranes have different permeability depending on the nature of the compounds — they are permeable to nonpolar solutes and impermeable to almost all polar or charged compounds. For this reason, biomembranes are deemed as "selective barriers". This selective permeability is what allows for the generation of the membrane potential which controls the movement of ions through the membrane and, thus, maintains the equilibrium of concentration [2]. Between these roles and multiple other chemical and physical processes in which the biological membrane is involved, it becomes clear that this structure is a crucial element in cell life.

Understanding the structure and composition of the biological membrane is imperative if we want to study its properties and mechanisms. In 1925, Gorter and Grendel first postulated that phospholipids aggregate to form a bilayer, with their hydrophilic groups facing the outside and their hydrophobic groups facing the inside [3]. This structure would become the key feature of every suggested model of biomembranes from then on out. Davson and Danielli built upon this proposal and, in 1935, suggested that biomembranes had proteins adsorbed to either side of the bilayer, in order to explain the selective transport of ions and apolar molecules through the membrane [4].

These models were the basis for Singer and Nicolson's proposal in 1972. Also taking into consideration the thermodynamic principles and the experimental data available at the time, they presented a new model that consisted of a fluid lipid bilayer with proteins integral to the membrane [5]. This model was named the fluid mosaic model (represented in Figure 1.1), since the phospholipid bilayer alternates with these globular proteins dynamically, with various lipids, like cholesterol, assuming the role of membrane fluidity regulators. Interestingly, at the time there were suggestions of potential aggregation of proteins into clusters, which has since been shown to be a common and relevant phenomenon — and not only with proteins [6, 7]. Singer and Nicolson also believed that their model suggested the existence of particular integral proteins that could span the entire thickness of the membrane. We know now that these proteins are quite common [8, 9], and many of them play major roles in the cell, such as metabolism regulation,

mediation of hormonal responses, etc.

To this day, the fluid mosaic model remains the most relevant model for biological membranes. However, it was an oversimplified model, and, naturally, it has since been refined to take into account the newest discoveries. An example of a "modernized" version, suggested by Nicolson in 2013 [10], now also accounts for the interactions between the extracellular matrix, membrane–associated cytoskeletal components, and cell membranes, amongst other features. Other authors have also suggested adding nano-raft domains and metastable protein oligomerization characteristics into the model [11].

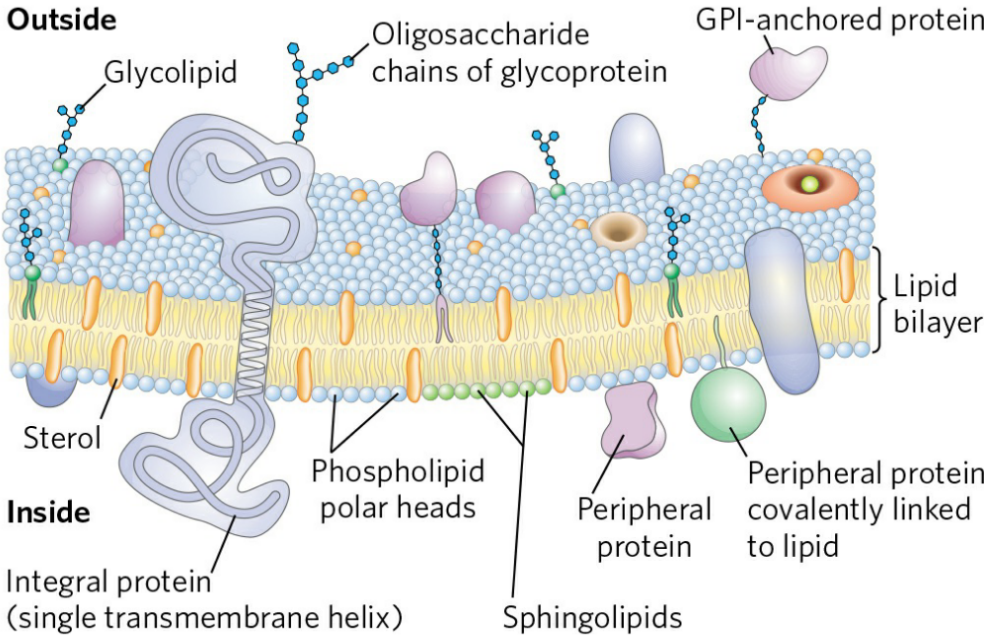

Figure 1.1: **Representation of the fluid mosaic model for plasma membranes.** Adapted from Lehninger Principles of Biochemistry [2].

1.1.1 The lipid bilayer

The basis of the structure of biological membranes is a lipid bilayer 3-4 nm (30-40 Å) thick. Membrane lipids — as well as membrane proteins — are mostly synthesized in the endoplasmic reticulum (ER) and are afterward transported to their target destination. It is during this trafficking that lipids and proteins suffer various modifications that determine their function and location within the cell [2].

Depending on the type of membrane, its lipid composition varies — both in type as well as relative proportions. In eukaryotic cells, their mitochondrial membranes contain little cholesterol and sphingolipids but are rich in cardiolipin, whilst the composition of their plasma membranes is the opposite [2]. As for prokaryotic plasma membranes, these are mostly comprised of anionic phospholipids like palmitoyloleoylphosphatidylethanolamine (POPE) and palmitoyloleoylphosphatidylglycerol (POPG) [12]. The

asymmetry in the lipid composition of the outer and inner leaflets is also a common feature of most biomembranes, and any changes to this distribution can lead to major biological consequences or signal the cell for programmed cell death [13, 14]. Typically, anionic lipids (phosphatidylserine (PS), phosphatidylethanolamine (PE), amongst others) are exclusively found in the inner leaflet. Other minor lipids like phosphatidylinositol (PI) and phosphatidic acid (PA) are also present in the eukaryotic inner leaflet. On the other hand, choline–containing lipids (like phosphatidylcholine (PC) and sphingomyelin (SM)), glycolipids, and sphingolipids are exclusively found in the outer leaflet [15–17].

Lateral heterogeneity is another prevalent feature of plasma membranes in both eukaryotic and prokaryotic cells [12, 18, 19]. It was in 1997 that the lipid raft hypothesis was first suggested [20] — which postulated that sphingolipids can be laterally segregated into membrane domains with different compositions and properties than their surrounding membrane. Since then, this first definition has changed and lipid rafts are now described as areas of the membrane that are stabilized by cholesterol molecules [21]. Lipid rafts have been associated with neurotrophic factor signaling [22, 23] as well as synaptic transmission [24], and they also play roles in signal transduction pathways [25]. To this day, the lipid raft theory is still a heavily discussed topic [26].

1.1.2 Membrane proteins

There are three types of membrane proteins — categorized according to how they are associated to the membrane (Figure 1.2). Integral membrane proteins are proteins that are inserted in the lipid bilayer. They can be monotopic, meaning that they interact with only one of the leaflets; or polytopic, with a polypeptide chain that crosses the bilayer multiple times or just once. Peripheral membrane proteins are proteins reversibly bound to the surface of the membrane through electrostatic interactions or hydrogen bonds. Lastly, amphitropic proteins are described as proteins reversibly associated with membranes [2]. We can mechanically distinguish membrane proteins by the different methods that make it possible to remove them from the membrane. Whilst integral proteins can be released using detergents or phospholipase C. treatments, peripheral proteins are extracted through changes in pH or ionic strength, removal of Ca^{2+} using chelating agents, or by adding urea or carbonate. In the case of amphitropic proteins, their removal process will depend on whether they are associated to the membrane or not [2].

1.2 Pores in the biological membrane

It is the amphiphilic character of lipids that grants the biomembrane its low permeability for most molecules, which is most important for the barrier role it plays in the cell. The transport of metabolites and solutes through the membrane can be done through several different mechanisms. Nonpolar compounds can cross the membrane through

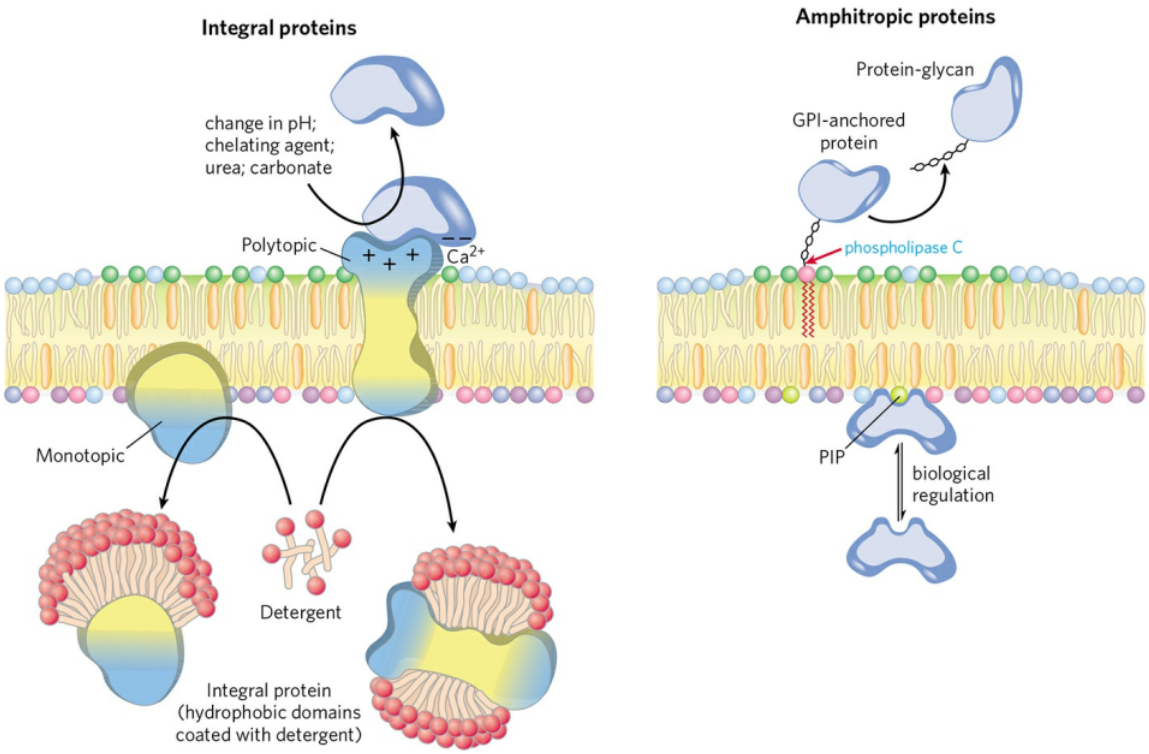

Figure 1.2: **Examples of integral, peripheral, and amphitropic membrane proteins.** Adapted from Lehninger Principles of Biochemistry [2].

simple diffusion, from the region of higher concentration to the region of lower concentration. Ion flux is generally controlled by transporters and channels that exist in the membrane, through passive transport or active transport (when against the gradient of concentration, electrical potential, or both) [2]. They can also leak through the membrane via transient lipid pores, which are responsible for the passive transport of water across the membrane [27]. The control over the metabolite concentration in the cell, as well as over the electrochemical gradient that is generated through their transport, is essential for the maintenance of cell life, as it plays a major role in the mechanisms of certain cell processes, like signal transduction and ATP synthesis. Thus, pores are crucial structures in any cell — since they allow for these exchanges between the intra and extracellular environments.

Membrane pores can be formed in multiple ways but they can be divided into essentially three categories: protein–lined pores (if they are made-up by proteins only [28, 29]), protein-lipid pores (if both lipids and protein partake in the structure [30, 31]) and pores formed solely by lipids [32, 33].

Pores formed by proteins — whether with the cooperation of lipids or not — can have a multitude of structures. Aquaporins are an example of protein channels that allow for water molecules (and other small molecules like glycerol and urea) to cross the membrane [28]. There are also many cases of peptides that aggregate to form ion–conducting pores [34–36], such as antimicrobial peptides, which use pore formation as a defense mechanism

[34, 36, 37]. Examples of the different protein pore structures can be seen in Figure 1.3.

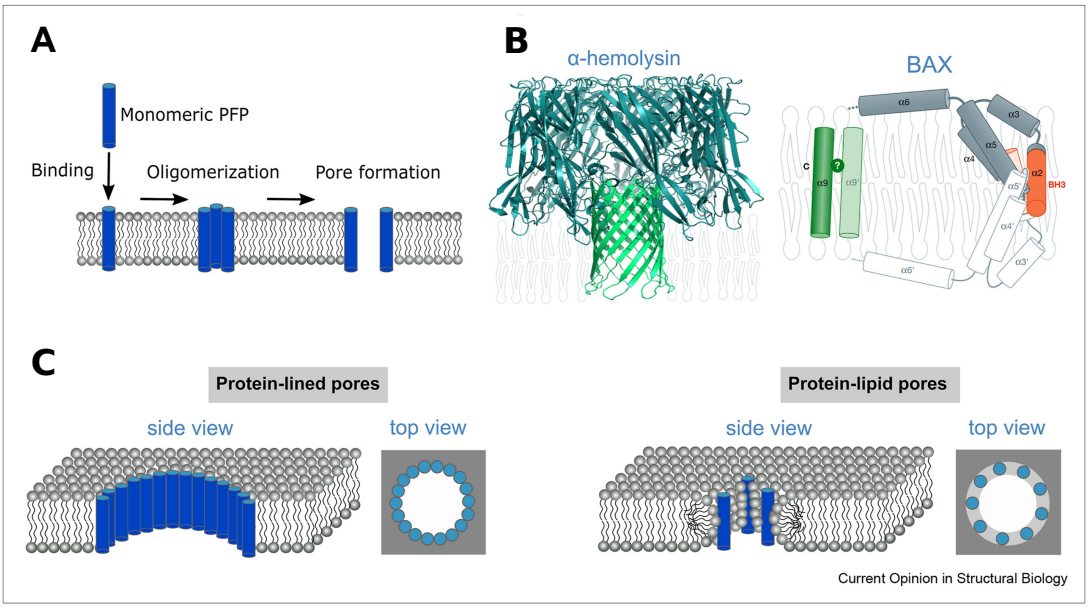

Figure 1.3: **Examples of protein pore formation and structures in biomembranes.** Adapted from Ros *et al.*[38]. (A) Binding of a monomeric pore–forming protein to the membrane and consequent oligomerization to form a protein–lined pore, (B) α-hemolysin protein–lined pore, (C) On the left, a schematic representation of protein–lined pores; on the right, a schematic representation of protein–lined pores.

In the case of pores formed only by lipids, although they are not very likely to occur, they can form in the membrane in response to stress [39]. These pores are usually transient, i.e. they do not stay permanently open. How the lipids rearrange themselves in order to form the pores is still not fully understood and has been a topic of discussion for several years. Snapshots of the closure of a lipid pore in a POPC membrane are shown in Figure 1.4.

1.3 Lipid–lined pores

Formation of lipid pores in the lipid bilayer can happen due to several physiological or pathological factors. The biomembrane is able to resist stress due to its flexibility, however, it is possible to reach a point where the force applied to the bilayer is too strong (high lateral tension) and, thus, lipid pores are formed [39]. The probability of this formation (which is generally low, given the high free energy cost) is inversely proportional to the necessary energy for the pore edge's lipids to form the defect. When this energy is related to the pore's perimeter, we obtain the line tension — i.e the necessary energy to expand the pore per unit of perimeter [40, 41].

In the classical lipid pore formation theory, the energy of a cylindrical and symmetrical pore is described by Equation 1.1

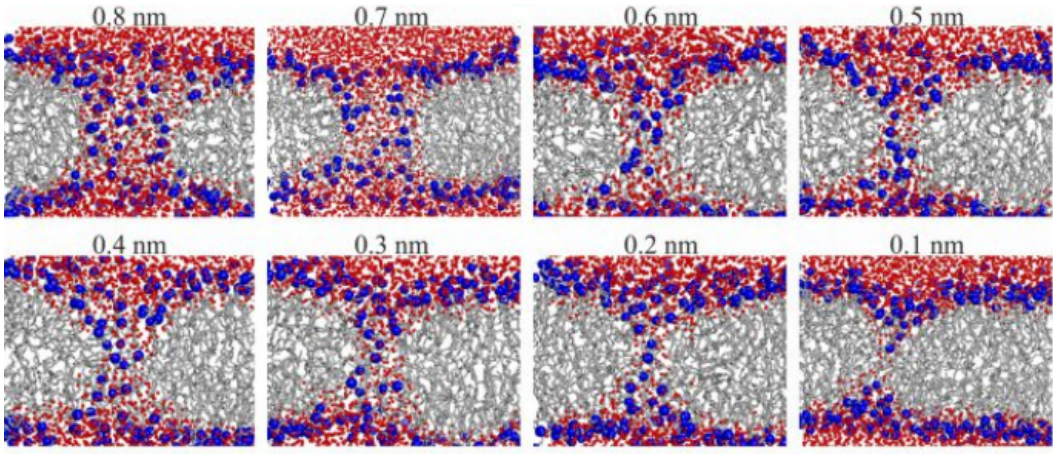

Figure 1.4: **Snapshots of the spontaneous closure of a lipid pore in a POPC membrane.** Adapted from Akimov *et al.*[40]. Polar atoms are represented by blue spheres, water molecules are represented by red spheres, and lipid tails are in grey. Radius of the pore in each snapshot is given in nanometers. Closure of the pore ran for 30 ns.

$$E(r) = 2\pi r\gamma - \pi r^2\sigma_0 \tag{1.1}$$

where r is the radius of the pore, γ is the boundary line tension of the pore edge, and σ is the lateral tension. In this theory, the energy of the lipid pore has its maximum at the critical radius of $r* = \frac{\gamma}{\sigma_0}$, where pores with a radius lower than this value can be closed and ones with a radius higher than the critical value can expand without limit — even leading to membrane rupture, given that as the size of the pore increases, the system's free energy decreases. Line tension is a characteristic that is dependent on the lipid composition of the biomembrane whilst lateral tension can actually vary throughout the length of the membrane [42].

Considering that in the classical theory the biomembrane is seen as an infinitely thin film with no internal structure, this theory fails to represent lipid pore formation in an intact biomembrane. Additionally, it also assumes line tension (which is defined by the elastic properties of the membrane) as a constant, not taking into account any possible dependencies on surface tension or pore radius [40].

The mechanism behind how lipid pores are formed as well as the free energy landscape of their formation are not clear. Lipid pores have been experimentally obtained through the application of mechanical stress [43] or an electrical field [44] in the membrane, or by inducing lipid flip-flop [45]. It is postulated that lipid pore formation begins with the penetration of a chain of water molecules in the bilayer, which then enlarges and is stabilized by the lipid headgroups that line the pore [45]. Nevertheless, there is still a lack of atomistic detail regarding this process, which, if obtained, could help us better understand it.

1.4 Antimicrobial peptides

Antimicrobial peptides (AMPs) are a class of peptides that play a major role in the first line of defense against not only bacteria, but also yeasts, fungi, and viruses [46, 47]. It has also been found that they exert some control over the inflammatory response in cases of infection [46]. AMPs greatly vary in their number of amino acids (6-100 amino acids) and can be found in various organisms, including humans and bacteria [46, 47]. These peptides have shown a lower susceptibility to bacterial resistance than their traditional therapeutic counterparts, thus, with the growing number of multidrug–resistant pathogens, the interest in AMPs as a possible solution has increased over the years [48].

AMPs are very diverse in their structure, therefore, classifying them is no easy task. In the past, AMPs have been classified in several different ways. Epand *et al.* divided them into six different classes: linear peptides that form amphipathic and hydrophobic helices; cyclic peptides and small proteins that form β-sheet structures; peptides with unique amino acid compositions; cyclic peptides with thioether groups in its ring; lipopeptides with a terminal amino alcohol group; and macrocyclic knotted peptides [49]. This contrasts with the much more simplistic classification done by Zhang *et al.*, which consists of three classes: α-helical; β-sheet; and extended AMPs [47]. Bahar had a similar classification system to Zhang, with the addition of a fifth class — the loop [46]. It is important to note that some antimicrobial peptides have different structural components and, thus, can belong to more than one class simultaneously [50]. Additionally, some AMPs even change their conformation when they interact with the lipid bilayer [51]. As such, it becomes apparent why it is so difficult to divide AMPs into structural classes, although, there is one common feature in most of their secondary structures: the existence of both a hydrophobic surface as well as a hydrophilic surface — i.e. an amphipathic character. This amphiphatic character, however, is a characteristic that is usually only seen when the peptide interacts with the membrane surface as part of its action mechanism [52].

For AMPs to perform their role in pathogenic defense, they need to interact with the membrane or cell wall. It is their amphipathic feature that allows for this to happen. Given that AMPs generally have a cationic surface (their hydrophilic surface), they can bind to lipid bilayers with anionic phospholipids in their inner leaflet. Then, it is this hydrophobic surface that allows for the peptide to enter the membrane and either form conducting pores or disrupt the bilayer. Evidently, for disruption to happen, it is necessary that the AMP concentration in the membrane overcomes a given threshold [53].

The mechanism behind how AMPs disrupt the membrane is not yet fully understood. Most commonly, AMPs form micelles or pores which will allow for the leakage of metabolites and other cell contents, thus causing cell death. In Figure 1.5 we can see the main suggested models of pores formed by AMPs. Whilst in the barrel-stave and toroidal models AMPs adopt a transmembrane topology and aggregate to form the pore, in the sinking-raft model they increase the curvature of the membrane by lying on its surface [54]. On the other hand, it has been shown that there is not an absolute correlation between the

peptides' ability to permeabilize membranes and their antimicrobial activity, meaning that the AMPs' mechanism of action does not always require the disruption of the membrane [52]. There are cases of AMPs that can simply cross the lipid bilayer and kill bacteria by inhibiting intracellular functions [55, 56]. Some AMPs even have the ability to inhibit the production of biofilm and perturb the already existent biofilms [47, 57, 58].

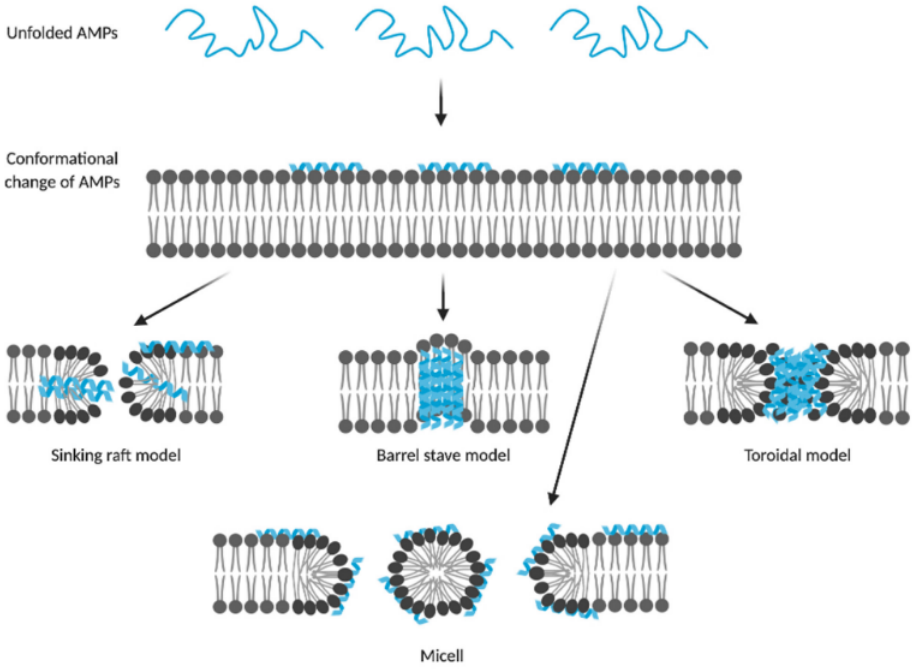

Figure 1.5: **Examples of antimicrobial peptides' proposed mechanisms of action.** Adapted from Dijksteel *et al.*[54].

1.4.1 Alamethicin

Alamethicin (Alm) is a 20-residue α-helical antimicrobial peptide produced by the fungus *Trichoderma viride* [59]. Alm contains several α-aminoisobutyric acid (Aib) residues (which induce its alpha-helical structure), making it part of the peptaibol family of peptides. Additionally, its C-terminal residue is a phenylalaninol (Phl) — a phenylalanine with a hydroxyl group in place of the carboxy terminus 1.6. Known for its activity against Gram-positive bacteria and fungi, Alm forms multi–conductance voltage–dependent pores in the lipid bilayer as its mechanism of action [60].

The mechanism behind the formation of Alm pores is yet to be fully understood. Previous studies suggest that, depending on the experimental conditions, Alm can stay parallel to the bilayer surface — in what is called the surface or S state — or it can insert itself vertically in the membrane — the inserted or I state. This insertion of Alm peptides in the membrane only happens at high concentrations of peptide, past a critical value that depends on temperature and nature of the lipids that make up the bilayer [62]. After insertion in the membrane, Alm peptides aggregate with their hydrophilic groups

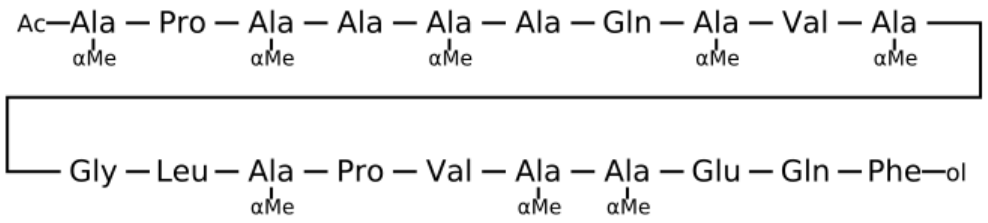

Figure 1.6: **Alamethicin's chemical structure.** Adapted from PubChem [61]. Ala–αMe residues correspond to Aib residues.

oriented towards the center and the hydrophobic groups facing the bilayer. This creates a water–filled and ion–conducting protein–lined barrel-stave pore structure (exemplified in Figure 1.5)[63–65]. For a matter of simplicity, from here on out, we will refer to these pores as peptide–induced pores but it is important to keep in mind that not all peptide–induced pores are protein–lined pores — they can also be protein–lipid pores (as seen in the toroidal model and micell examples of Figure 1.5). The number of alamethicin peptides that form these pores seems to vary between 6 to 12 peptides, depending on the degree of saturation as well as the length of the tails of the membrane lipids [62, 66]. The diversity in this number may justify the existence of various conductance levels that have been observed for Alm pores in lipid bilayers [65, 67]. On the other hand, in experimental studies done by Harriss *et al.* [67], it was possible to observe multiple conductance states originating from an individual Alm channel (represented in Figure 1.7). These results not only further proved that Alm conduction happens through discrete pores in the membrane but they also showed that a single Alm pore can be able to quickly switch between different conductance levels.

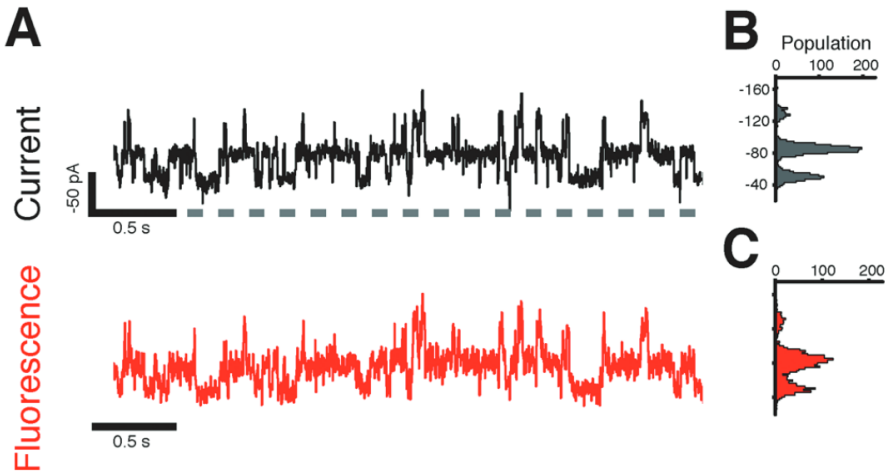

Figure 1.7: **Alamethicin's multiple conductance levels.** Adapted from Harriss *et al.* [67]. (A) Ca^{+2} and single-channel electrical recording (SCR) flux; (B) Histogram of the current data; (C) Histogram of the fluorescence data.

In past works, it was possible to see that the insertion of Alm peptides happens by firstly inserting the N-terminus (the most hydrophobic terminal of the two) of the peptide into the membrane [68–70]. However, this is only the case when the Alm pores consist of all parallel peptides, i.e when channel opening only happens at *cis* positive voltage — which helps the insertion of the N-terminus. When we have a case where channel opening is not dependent on voltage and the Alm peptides have traveled to the *trans* leaflet, it is possible to have antiparallel alamethicin pores. Rahaman *et al.* were able to study antiparallel alamethicin pores and concluded that, although these pores had lower effective energy, it was still important to consider the possibility of their formation in certain experimental conditions [71].

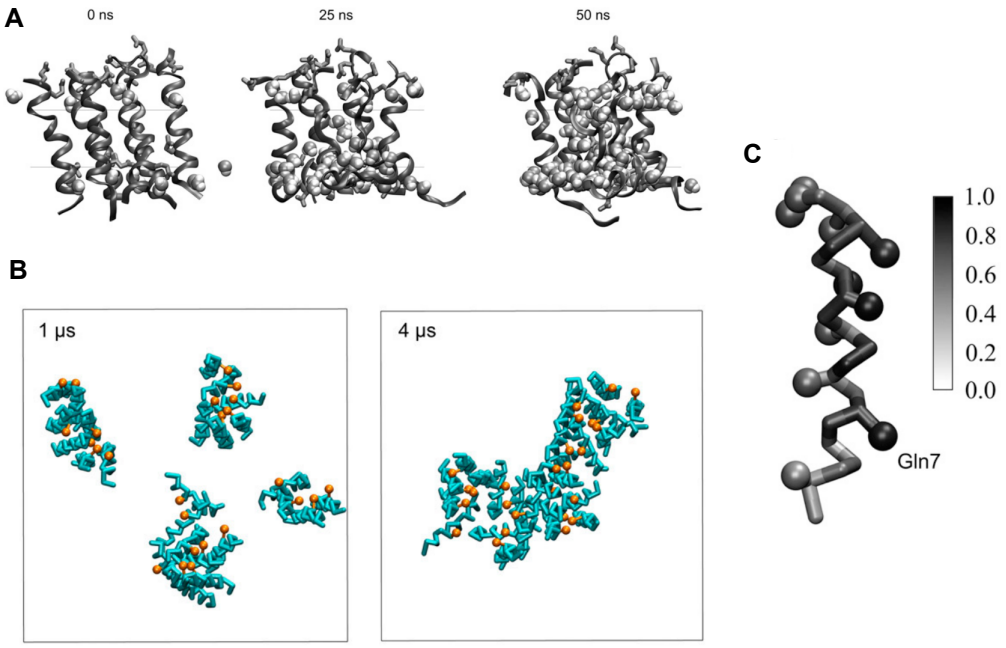

Figure 1.8: **Alamethicin aggregation in atomistic and Martini 2 simulations.** Adapted from Thøgersen *et al.* [69]. Alamethicin aggregation using (A) NAMD 2.6b1 and the CHARMM27 parameter set (alamethicin peptides represented in ribbon with their hydrophilic side-chains in licorice) and (B) the Martini 2 force field (peptide backbone represented in blue and Gln^7 side chains in orange). (C) Alamethicin peptide colored according to the number of interpeptide contacts established throughout the period of $t = 0.8 - 1$ ps using the Martini 2 force field. A probability equal to 1 means contact at all times and probability equal to 0 means no contact at any time.

Several molecular dynamics studies have been done with alamethicin peptides in lipid bilayers [68, 69]. These made it possible to not only observe the transitioning of the peptides from the S state to the I state, but also peptide aggregation into clusters. Thøgersen *et al.* saw that the Alm aggregates grew in size over the simulated time rather than consisting of a specific number of peptides [69]. The existence of this cluster organization is still highly debated and there are various results in favor of both sides [65, 72, 73]. Furthermore,

Thøgersen *et al.* were also able to study Alm pore behavior in MD (Figure 1.8). However, the pore behavior was only possible to be simulated in all-atom (AA) MD simulations, since with coarse-grained (CG) MD simulations, specifically with the Martini 2 force field, the water molecules could not pass through the transmembrane peptide assemblies, which was speculated to be due to the water beads' large size. Additionally, many of the Alm residues established interpeptide contacts throughout most of the simulation time that collapsed any would–be pore lumen (Figure 1.8) [69].

1.5 Molecular Dynamics simulations

Molecular dynamics (MD) simulations provide the ability to study a variety of biomolecular processes, like conformational changes and protein folding, and even predict the outcome of molecular protonation and phosphorylation. MD allows the study of these dynamics at a level of detail that experimentally would not be accessible [74].

This technique has become more prominent over the years, given its versatility and broad application in biological fields as well as the rise in computational power [75–77].

1.5.1 Coarse-Grained vs All-Atom simulations

In order to predict the movement of each atom in a system, MD Simulations rely on solving Newton's laws of motion to obtain the atom's coordinates. Thus, it is clear that, when dealing with large systems, the amount of calculations that have to be done may be limited by the current computational power.

As such, in cases where we can disregard some atomic detail, all-atom simulations that fully represent every atom in a system may not be the most efficient tool. Coarse-grained MD models decrease the computational cost of calculations by representing groups of atoms by effective beads, consequently decreasing the degrees of freedom (DOFs) and allowing for simulations of larger systems over longer periods of time. Thus, CG simulations can be the better alternative for modeling phenomena that happen on a large scale of time and space. An accurate and precise CG model for these systems is greatly relevant and can give us the possibility to study their energetics as well as their mechanisms.

CG simulations have been used to model and study a variety of biomolecules and processes, even allowing for the assembly of models with millions of beads. Most notably, in early 2023, a CG model of the minimal cell JCVI-syn3A was achieved [78].

1.5.2 The Martini 3 force field

Martini is the most widely used CG force field for molecular simulations [79]. It is a general–purpose field that uses a four-to-one mapping scheme — i.e. on average, four heavy atoms and their associated hydrogens are represented by one single CG bead. These

CG beads are then assigned a chemical class according to their nature, in terms of charge and polarity. From here, Martini molecules are assembled in a building–block manner.

As Martini became a staple tool, some faults in the force field became apparent, specifically in cases where the mappings were finer than 4-to-1. In previous Martini 2 simulations, Bennett and Tieleman tried to induce lipid pore formation by pulling a polar group (in this case, the head of a phospholipid) to the center of a bilayer. Although more stable defects were obtained with polarizable water and increased Lennard-Jones (LJ) interaction between the water beads and the choline and phosphate beads, it was never possible to observe water pores as seen with AA methods. This issue was postulated to be the fault of the water model of the Martini force field, calling attention to existent problems of high line tension of the pore and high bending modulus [80]. In other studies, it was also shown that particular molecules, such as membrane proteins and carbohydrates in solution, self-interacted too strongly and that the existent number of bead types and sizes did not allow for accurate implementation of Martini when simulating molecules like DNA [81–84].

In 2021, a newly improved version of Martini — Martini 3 [85] — was released that addressed some of these shortcomings, namely the excessive self-interactions. The Martini 3 force field had its nonbonded interaction terms reevaluated and introduced new bead types and sizes. The packing of molecules was improved and, as a consequence, it is now possible to obtain pores that remain open throughout the trajectory with aggregates of transmembrane proteins (much larger systems than AMP pores) [85]. For transmembrane proteins, it was also observed that, in Martini 3, their free energies and affinities were in better agreement with the previously obtained experimental results [85].

Although Martini 3 is a clear improvement from its previous version, there are still some limitations to the force field [86, 87]. Namely, previous studies have shown that transmembrane peptides do not stay inside the membrane which points to Martini 3 protein-membrane affinities potentially being inaccurate [86]. This behavior can, consequently, affect the study of protein-membrane systems and processes, such as protein–induced pore formation. Therefore, it is important to continue to work on the improvement of the force field and its interactions, as well as to find strategies that allow for better modeling of these molecules and systems. Since the release of this new iteration of Martini, a tireless effort has been and is continuously put towards better and more accurate models [88–90].

1.6 Objectives

Given the crucial role that pores play in cell life, it is important to be able to understand and describe their various dynamics and behaviors *in vivo*. With the rise of antibiotic–resistant bacteria, it has also become vital to find new forms of therapeutics that can be as effective and cost-efficient. Thus, MD simulations can offer great insight into the different types of pores that can be formed in the biomembrane — from lipid pores that assist with ion transport to peptide–induced pores that act as defense mechanisms.

In this thesis, I worked with two different types of pores: pores formed solely by lipids and peptide–induced pores.

Regarding lipid pores, our main objective was to identify the limitations of Martini 3 lipid pore simulations, using a specifically developed reaction coordinate (RC) [91] to describe the energetics of pore formation, and consequently try to improve them. To achieve this, several all-atom and coarse-grained simulations were assembled with the different Martini 3 lipid models that were developed by third-party collaborators throughout the time of this thesis, in order to compare energy measures and behavior of pore lipids and determine a probable cause of inaccurate lipid pore formation in Martini 3 simulations. Prototypes of lipid topologies were also created to test potential improvements.

In terms of peptide–induced pores, our objective was to understand whether alamethicin's pore formation could now be accurately replicated with the new Martini 3 force field as well as to study its formation behavior. For that, I ran Martini 3 simulations of Alm peptides inserted into lipid bilayers in various conditions and analyzed the characteristics of the alamethicin pores formed and their dynamics.

1.7 Other work - Cholesterol parameterization for the Martini 3 force field

During the time of this thesis, I also participated in other projects that went beyond the scope of this work.

With the release of the new Martini 3, there is an urgent demand to revisit and reparametrize formerly modeled biomolecules. One of those is cholesterol [92], an essential regulator of lipid membrane fluidity that also interacts with several membrane proteins. Cholesterol is also involved in the formation of lipid rafts in the mammalian cell membrane [93].

The reparameterization of the new cholesterol model for the Martini 3 force field was a collective effort. Our work focused on testing the recognition and binding of the cholesterol prototypes to proteins. To do this, we assembled several Martini 3 simulations of lipid membranes containing cholesterol alongside different proteins (both integral and peripheral membrane proteins) and compared our results to previously obtained atomistic, experimental, and Martini 2 results. This gave us an understanding of the quality of the cholesterol prototype and of what we should change and tweak, in order to obtain a more accurate and overall better model.

The final Martini 3 cholesterol model is now available on the Martini Force Field Initiative GitHub page and the preprint of the article detailing the work that was done has also been published [88].

2 Materials and Methods

2.1 Molecular Mechanics

Molecular Mechanics (MM) methods allow us to calculate and predict various molecular properties, from the structural conformation of biological systems to dipole moment and vibrational frequencies [94, 95]. In MM, biomolecules are treated as groups of atoms that are bonded by harmonic or non-harmonic forces [96].

In contrast with quantum mechanical methods (QM), MM disregards electronic motions, only considering nuclear positions to calculate the system's energy. This is done by employing several approximations that allow for fewer and faster calculations [97]. One of those approximations is the Born-Oppenheimer approximation. Since the relative mass of a nucleus is much larger than that of an electron, the velocity of its movement will be significantly lower — to the point where it is possible to consider electron motion as instantaneous. The Born-Oppenheimer approximation is coupled with a second approximation that treats electrons as single point charges collapsed in the center of the nucleus. Therefore, in MM, only the dynamic of the nuclei is taken into consideration in energy calculations [97–99]. Additionally, MM treats nuclei as classical particles, making it possible to calculate the forces applied to the system's atoms based on Newton's laws of motion [97, 99].

These approximations make it possible to perform the necessary energy calculations to predict the dynamic motion of biological systems over time, which is not feasible with QM methods.

2.1.1 Force fields

Force fields consist of mathematical expressions used to define the chemical space in a system. In MM (as well as in many other computational methods), the system's energetic potentials are calculated by the chosen force field's parameters [97, 100].

MM methods using conventional physics–based force fields are made up of potentials that are parameterized to describe the intra and intermolecular forces in systems (exemplified in Equations 2.1, 2.2, 2.3). All of these parameters are specific for each atom type. In many cases, like in AA molecular dynamics simulations, these potentials are an approximation to distance and vibrational frequency values that are derived from

experimental data or *ab initio* calculations. Most commonly, force fields include potentials that describe bond stretching, torsion angles, bond angle bending (Equation 2.2), and Van der Waals (VdW) and electrostatic interactions (Equation 2.3) [97, 100, 101] — exemplified in Figure 2.1.

$$U^{system} = U^{bonded} + U^{non-bonded} \tag{2.1}$$

$$U^{bonded} = U_{bonds} + U_{angles} + U_{dihedrals} \tag{2.2}$$

$$U^{non-bonded} = U_{VdW} + U_{electrostatic} \tag{2.3}$$

Depending on the goal and target of our study, we can use different force fields that account for different interactions, at varying degrees of detail. Some of the most used force fields include the Assisted Model Building with Energy Refinement (AMBER) [102], the Chemistry at HARvard Macromolecular Mechanics (CHARMM) [103], the Southamerican Initiative for a Rapid and Accurate Hamiltonian (SIRAH) [104, 105], and the Martini [79, 85] force fields.

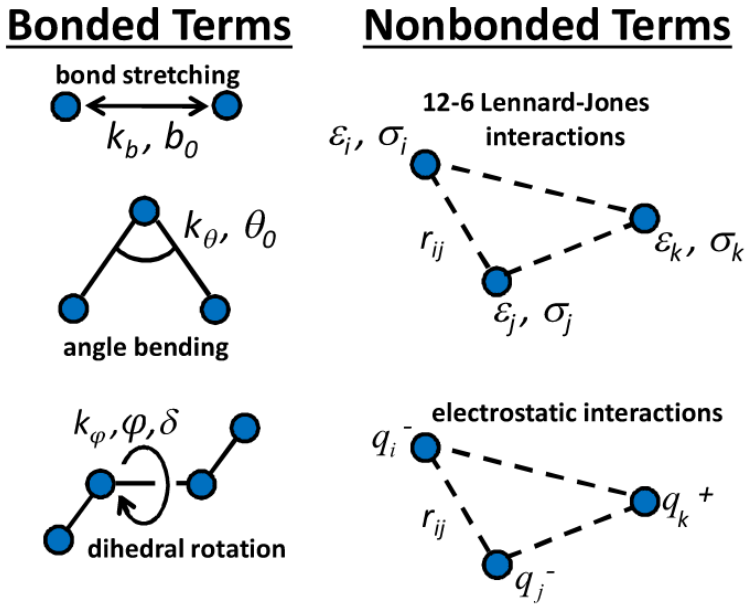

Figure 2.1: **Examples of biomolecular force field terms.** Adapted from R. A. Latour [106].

2.1.1.1 Bonded interactions

To represent the interactions between covalently bonded atoms, force fields include potentials that describe bonds, torsion angles, and angle bending.

Bond stretching (U_{bonds}) is frequently represented by harmonic potentials, which greatly reduce the time needed for computing calculations. However, these potentials don't allow us to study chemical processes as they implicitly describe bonds as unbreakable.

As seen in Equation 2.4, the potential energy increases with the deviation from b_{0_n}, which is the bond value when all other force field terms are zero [97, 100].

$$U_{bonds} = \sum_{bonds} \frac{1}{2} K_{b_n} (b_n - b_{0_n})^2 \tag{2.4}$$

Bond angles (U_{angles}) are also frequently described by harmonic potentials, as seen in Equation 2.5:

$$U_{angles} = \sum_{angles} \frac{1}{2} K_{\theta_n} (\theta_n - \theta_{0_n})^2 \tag{2.5}$$

or, as is the case of the Martini force field, by a simplified cosine based potential [107, 108]:

$$V_a(\theta_{ijk}) = \frac{1}{2} k^{\theta}_{ijk} (\cos(\theta_{ijk}) - \cos({\theta_{ijk}}^0))^2 \tag{2.6}$$

For molecules larger than four linearly displayed atoms, a potential that describes torsion angles ($U_{dihedrals}$) is often added. An accurate representation of dihedral angles is particularly important since these torsions are crucial for the rigidity of the molecule and the modeling of conformational changes and their stability [97, 100]. Dihedrals are most commonly expressed by cosine functions, as in Equation 2.7, where φ_n is the torsional angle, δ is the phase factor, and m_n is the multiplicity.

$$U_{dihedrals} = \sum_{n,dihed} \frac{1}{2} K_{\varphi_n} [1 + \cos(m_n \varphi_n - \delta_n)] \tag{2.7}$$

$$U_{improper} = \sum_{impropers} \frac{1}{2} K_{\xi_n} (\xi_n - \xi_{0_n})^2 \tag{2.8}$$

Often, a fourth term is added which represents improper dihedral angles by a harmonic potential ($U_{improper}$, exemplified in Equation 2.8, where ξ is the improper angle that represents the deviation from planarity). This potential is important to help maintain the planarity in biomolecules, since the proper dihedral potential is sometimes not sufficient to achieve this [100].

2.1.1.2 Non-bonded interactions

To represent non-bonded interactions between atoms that are more than three bonds apart, terms representing Van der Waals and electrostatic interactions are also added. Although it would be possible to calculate non-bonded interactions for neighbor atoms as well, it is assumed that the bonded parameters already account for these conformational representations. As such, typically, we exclude the overlap with non-bonded for atoms less than three bonds apart. [97, 100, 101].

To describe the Van der Waals interactions between atoms, 12-6 Lennard-Jones potentials are often used (U_{VdW}, seen in Equation 2.9). These potentials account for repulsive as well as attractive interactions between atoms i and j, represented by components $\left(\frac{\sigma}{r}\right)^{12}$ and $\left(\frac{\sigma}{r}\right)^{6}$, respectively. Here, σ is the collision diameter, i.e the value where the potential is equal to zero, ε is the depth of the well of the interaction, and r is the distance between the atoms [97, 100].

$$U_{VdW} = \sum_{i,j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{2.9}$$

As the approximations mentioned in Section 2.1 lead to the electrons being represented as point-charges, there are no explicit electronic clouds in MM. This means that the polarizing effect felt by the molecules resultant from their electronic clouds has to be represented by other means. Consequently, the dispersive term of the Van der Waals potential ($\left(\frac{\sigma}{r}\right)^{6}$) is also used to mimic this effect [107, 108].

Finally, Coulomb potentials are used to represent electrostatic interactions ($U_{electrostatic}$) — as exemplified in Equation 2.10. For this, nuclei are assigned partial atomic charges, q_i and q_j [97, 100].

$$U_{electrostatic} = \sum_{i,j} \left[\frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}} \right] \tag{2.10}$$

2.2 Molecular Dynamics Simulations

Molecular Dynamics (MD) simulations are a MM technique that allows us to predict the movement of every atom in a system, throughout a period of time. For this, the forces applied to each atom by every other atom of the system have to be calculated [74]. It is easy to understand that these calculations can quickly become time–consuming and require major computational power. As such, MD employs the same approximations as MM methods (mentioned in Section 2.1) to decrease the number of necessary calculations. It is then possible to obtain the forces applied to each atom by solving Newton's Laws of motion (Equations 2.11 and 2.12) [97, 99].

$$F_i = -\frac{dV}{dr_i} \tag{2.11}$$

$$F_i = \sum_j F_{ij} \tag{2.12}$$

Finally, knowing the applied forces, the acceleration of every atom is calculated using Equation 2.13.

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \tag{2.13}$$

These calculations are done in short time steps (integration time step) in the fs scale. Coarse-grained simulations use higher intervals than atomistic simulations, about 10-fold that of AA systems, since the atom mass of CG beads is larger and the used potentials (which represent average movement) also have smaller frequencies. In each step the forces are calculated and, from these, the new positions and velocities of the atoms are determined. During each step interval, it is assumed that the applied forces are constant. Thus, we are able to predict the trajectory of every atom in the system and sample the system's configuration [97].

In MD simulations, there are three stages that comprise its workflow: minimization, relaxation or equilibration, and production. The minimization step allows us to stabilize our system, giving us a starting structure with no steric clashes or inappropriate geometries. Afterward, with the equilibration step, we can first bring the temperature to the desired value and then apply pressure to the system to reach the proper density. Finally, with a well–equilibrated system, we are then able to run our simulation and produce data [107, 108].

2.2.1 Energy minimization

In MD methods, minimizing the energy of our system is particularly important. The minimum energy is where a system is at its most stable state and the adopted molecular conformations are energetically favorable. A biological system has a potential energy surface (PES) with multiple minima and a lowest energy state - the global energy minimum - which is rarely achieved (Figure 2.2) [97, 99].

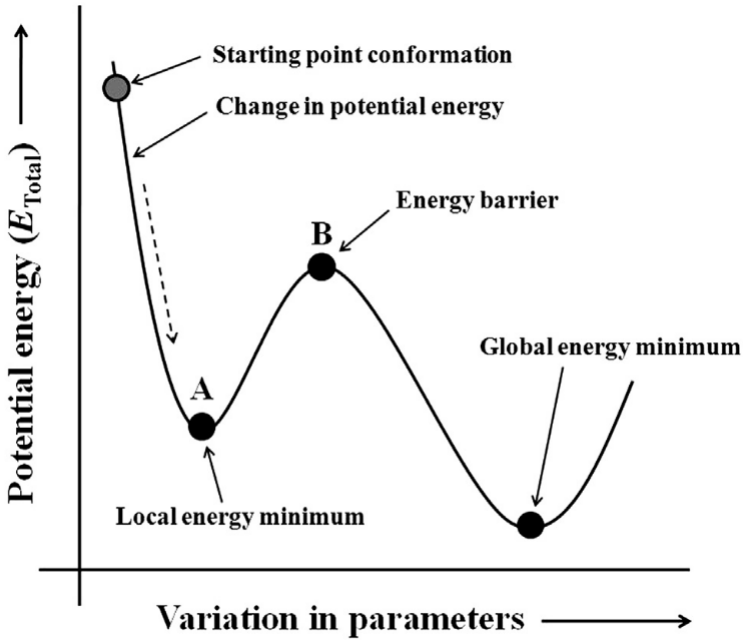

Figure 2.2: **Minimization of a molecule's conformation.** Adapted from Roy *et al* [99].

For the energy minimization step, a minimization algorithm is used, which can be

derivative or non-derivative, that "travels" through the PES until it finds a local energy minimum. The initial coordinates of the atoms (which are typically Cartesian coordinates) can be obtained through a variety of methods, such as NMR and X-ray crystallography, or even conformational search algorithms. From here, the minima are found by changing the atoms' coordinates to obtain configurations with progressively lower energies. Since most minimization algorithms cannot "travel" up on the energy surface, it is necessary to have several starting points from which the algorithm can explore the surface to try to find the global minimum [97, 99].

We want our minimization algorithm to be as efficient as possible, i.e for it to find a minimum quickly without an extensive amount of calculations. There is not a minimization algorithm that has proven to be best in every situation, thus, the choice depends on the problem at hand and the computational method we are using.

In this work, we used the steepest descents algorithm. This algorithm minimizes the structures in the opposite direction of the largest point of the surrounding potential energy surface. Every time a minimum is reached, the algorithm calculates new direction coordinates and searches for a new minimum [109, 110]. This is repeated until the maximum of the absolute values of the surrounding gradient is smaller than a value specified by the user or when the algorithm reaches a stipulated number of force evaluation steps [107, 108].

It is important to note that, in this minimization step, we only obtain the initial coordinates of the system for the simulation and not the initial velocities — these are usually randomly obtained from a Maxwell-Botzmann distribution of velocities coupled to the target temperature [111].

2.2.2 Temperature and Pressure coupling

In order to control and maintain the temperature and pressure of a system, MD methods use thermostats and barostats. Evidently, how and if we use these tools will depend on the conditions that we wish to replicate, thus, there are various ensembles that can be simulated. In this work, we used the isobaric-isothermal ensemble (NPT) — where the number of particles (N), pressure (P), and temperature (T) are fixed [111].

In cases where we wish to keep the temperature fixed around a desired value, the thermostat regulates the average temperature of the system by controlling the velocity of the particles. There are several algorithms that do this and the most appropriate choice depends on the type system that we are studying. In this work, we used two thermostats: the Nosé-Hoover [112] and the velocity-rescale [113] thermostats. The Nosé-Hoover thermostat considers the heat bath as part of the simulated system by representing it as a parameter in its equations that has its own momentum and equation of motion. The velocity-rescale thermostat is much simpler — it multiplies the velocities by a stochastic term that ensures a correct kinetic energy distribution and allows the system to stay within the desired temperature value [107, 108, 114, 115].

Regarding pressure coupling, barostats control the average system pressure by regulating the volume of the simulation box [111]. Once again, there are many algorithms available for this end, but the most frequently used in production steps is the Berendsen barostat [116]. This barostat rescales the coordinates as well as the box vectors at every time step by using a scaling matrix, which results in a relaxation effect that will lead the system towards the desired pressure value [107, 108]. Berendsen is often used in equilibration due to its robustness, however, given that it is also ill-defined, it is necessary to change to a more accurate barostat, like the Parrinello-Rahman [117], when we move onto production. In the Parrinello-Rahman barostat, the equation of motion is applied to the box vectors and the equations for the atoms are also changed, as is done for the Nosé-Hoover temperature coupling. Additionally, there are three different pressure coupling schemes that can be employed: the isotropic scheme, which controls the entire system pressure simultaneously; the anisotropic scheme, where the x,y, and z planes are controlled separately; and the semi-isotropic scheme, which controls the pressure in the xy and z planes independently, allowing for homogeneous pressure. In membrane studies, such as ours, the latter scheme is typically used [111].

2.2.3 Periodic boundary conditions

In MD simulations, periodic boundary conditions (PBC) are frequently used to solve artifacts caused by unwanted boundaries in our simulated system [107, 108].

PBC involves inserting our system into a box that is surrounded by copies of itself — as exemplified in Figure 2.3. When an atom leaves the simulation box through one side, the same atom will then appear on the opposite side, thus, the number of atoms in the box is constant [97, 107].

GROMACS [107, 108], which is the simulation engine mainly used in this work, uses PBC with the minimum image convention. This means that the cut-off radius of non-bonded interactions cannot be larger than half the shortest box vector, in order to avoid molecules 'sensing' the same atom from different sides.

There are many different simulation box types. Depending on what we are simulating, we should choose the box accordingly. In a case where our system consists of a molecule in solution, we may want to choose a box that will save us time spent on calculating solute positions and velocities [97, 107, 108]. GROMACS offers several box types, the most commonly used being the cubic, the rhombic dodecahedron, and the truncated octahedron boxes.

2.2.4 Coarse-grained simulations

As has been mentioned, when we perform MD simulations, it is necessary to calculate the position and velocity of every atom in the system. Processes like phase transitions and protein agglomeration, as well as systems like polymers and gels, occur on a large scale

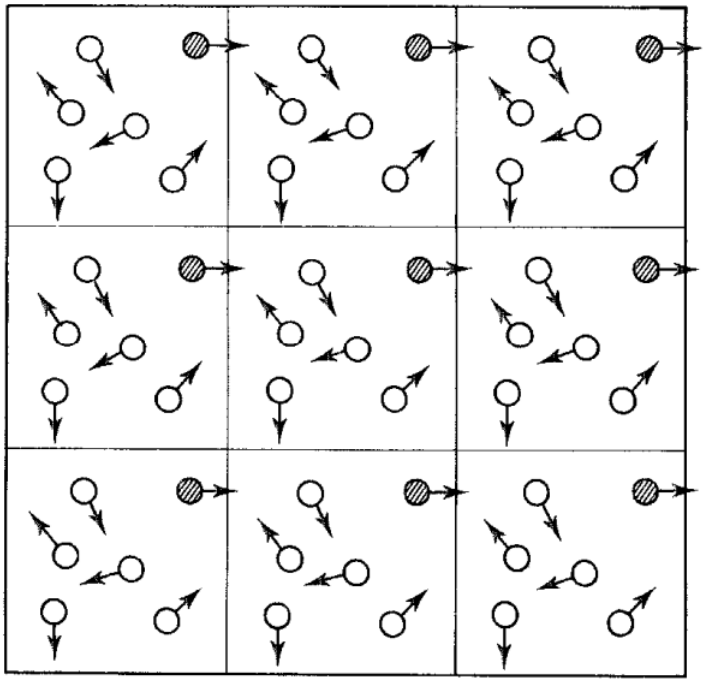

Figure 2.3: **Periodic boundary conditions schematic in two dimensions.** Adapted from L. R. Andrew [97].

and their dynamics can be very slow. As such, simulating them requires large time and size scales that, in all-atom simulations would be impractical and unfeasible [118].

Coarse-grained (CG) MD simulations allow us to model systems of up to micrometers in size, for milliseconds of time. In CG, groups of atoms are represented by beads (as exemplified in Figure 2.4) that interact with each other through the same potentials described in Section 2.1, but adapted to reflect this grouping. This representation can be done at various degrees, depending on what molecular properties and processes we wish to represent and study. By decreasing the number of particles in the system, the number of degrees of freedom also decreases, and, thus, some atomistic detail is lost (while typically retaining molecular identity). This decrease in DOF represents an advantage of CG in comparison with AA simulations, since it makes it possible to represent larger systems over longer periods of time without needing to perform as many calculations. Additionally, the reduced number of particles also leads to a smoothened potential energy surface, which increases the speed at which the minimization of our system is achieved and lets us use integration time steps of up to 20 fs (that, as mentioned in Section 2.2, are much longer than in AA) [118].

These characteristics of CG simulations are what allow us to simulate larger–sized systems for longer time scales since the amount of computing needed is significantly decreased.

To parameterize CG models — i.e to define the potentials that will lead a CG model to faithfully reproduce the structural configurations and interaction behaviour of a given

molecule — there are two different methods we can adopt: we can model molecules with the objective of mimicking and fitting features from AA or QM models (bottom-up approach); or we can calibrate our model to reproduce experimental data regarding molecular properties, such as thermodynamic data (top-down approach) [119–121]. The CG parameterization diverges from the AA method, as there are less parallels between CG beads and real atoms from which we can extract experimental measures and data, than between AA and real atoms.

There are many CG force fields that have been developed throughout the years. From general–purpose fields, like SIRAH [104, 105] and Martini [79, 85], to more specifically applied fields, such as UNRES [122–125] for modelling proteins, coarse-graining has allowed for the study of many biological processes, specifically protein and lipid related processes [126–128].

2.2.4.1 Martini 3 CG force field

The Martini force field is a general–purpose force field that relies on a four-to-one mapping scheme. This means that, on average, four heavy atoms and their associated hydrogens are represented by one single bead (as seen on Figure 2.4). The parameterization of its models is done through a combination of top-down and bottom-up approaches, which are typically implemented via Lennard-Jones and bonded terms, respectively [129].

There are four main classes of Martini CG bead types: C beads, which represent apolar groups; N beads that represent nonpolar groups; P beads for polar groups; and Q beads which represent charged chemical groups. In the new and improved version of the Martini force field, Martini 3, besides the labeling according to the degree of polarity or charge, these beads can now come in three different sizes — regular (R), small (S) and tiny (T) — so as to allow for finer mappings than the classical 4-1 scheme of Martini [79, 85].

Martini uses the same potentials mentioned in Section 2.1 to describe the chemical space. It treats solvent explicitly, which means that Martini systems are simulated in water and ions, although there is also a version of the force field — the Dry Martini — that is available for lipid membrane simulations with implicit solvent [131]. In regular Martini, for every 4 water molecules, 1 Martini bead is used to represent them. Whilst Martini is mostly used with GROMACS [107], it has also been implemented in NAMD [132], LAMMPS [133], openMM [134], and other engines.

2.2.5 Umbrella sampling

The umbrella sampling method is used to overcome an energy barrier through the introduction of a bias potential. This technique models conformational transitions by defining multiple windows along a chosen reaction coordinate — coordinates that describe the reaction dynamics and measure its progress. A harmonic potential is then added to the Hamiltonian potential of the system, in every window:

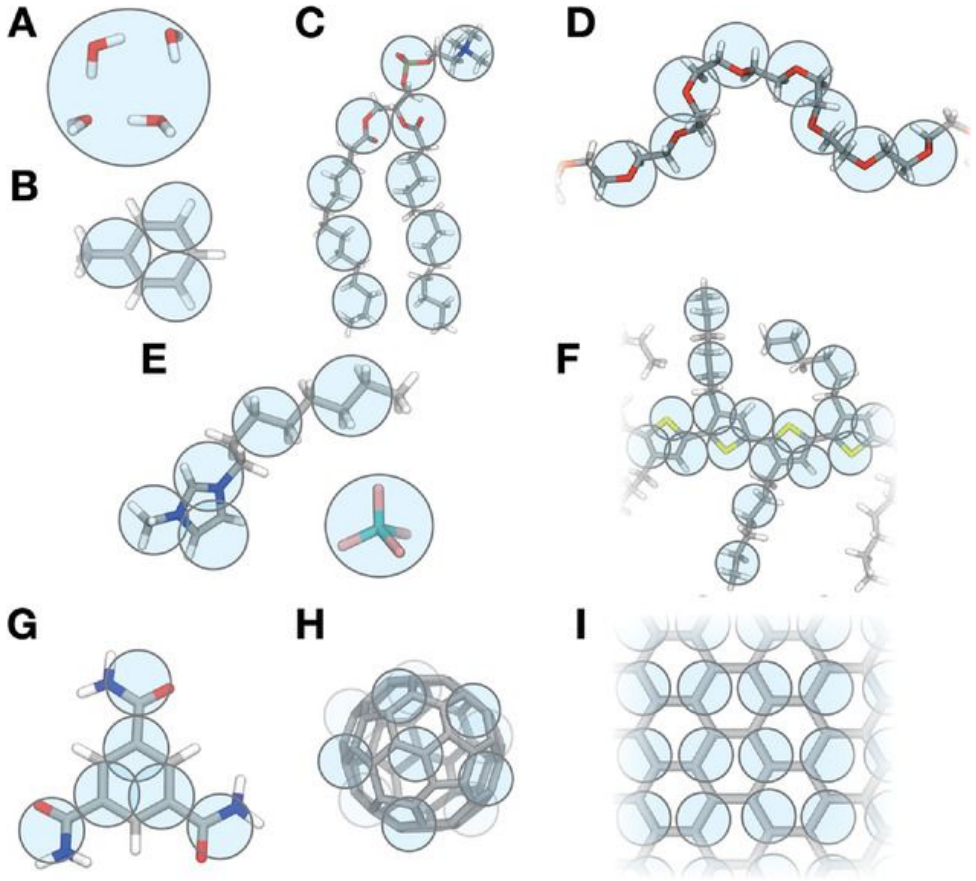

Figure 2.4: **Example of the Martini mapping.** Martini CG beads overlay the atomistic structure and are represented as blue transparent beads. (A) Standard water bead, (B) Organic solvent toluene, (C) Dimyristoylphosphatidylcholine (DMPC), (D) Polyethylene oxide (PEO), (E) 1-octyl-3-methylimidazolium tetrafluoroborate $[C8mim]^+[BF4]^-$ ionic liquid, (F) Poly(3-hexylthiophene) (P3HT), (G) 1,3,5-benzenetricarboxamide (BTA), (H) Fullerene, (I) Surface of graphene. Adapted from Riccardo Alessandri *et al.* [130].

$$\Delta V_i(q) = \frac{k}{2} \times (s(q) - s_i)^2 \tag{2.14}$$

where k is a force constant, that is chosen taking into consideration the potential energy at each reference point s_i along the RC $s(q)$. Finally, MD simulations are run for each window, until the explored space of neighboring windows has significant overlap. This allows for the complete free-energy profiling of the system [135–138]. However, it should be taken into consideration that these are biased simulations. We have thus to perform analyses that will allow us to obtain the unbiased distribution, such as a weighted histogram analysis (WHAM).

2.2.6 The lipid pore model for MD simulations

When we want to model the mechanism of a molecular reaction, we have to be able to describe the transition between populated intermediate states along reaction paths. To do

that, we need to not only identify these transition states of the molecular reaction but we also need accurate RCs [139]. The RC concept described in the previous section (Section 2.2.5) can also be applied to lipid poration, amongst other processes.

Taking into consideration the fact that there is still a lack of knowledge regarding important aspects of pore formation, it is not surprising that accurate RC — that are needed in order to calculate free energies — for pore formation are difficult to define. This shortcoming makes it challenging to study pore structures and accurately calculate energetics by means of MD simulations.

In 2017, Hub *et al.* [91] proposed a new model for pore formations in MD simulations. Here, the RCs are described by a membrane–spanning cylinder that is split into horizontal N_s slices, with d_s thickness — so as to allow the formation of stable hydrogen bonds between pairs of polar atoms in subsequent slices. The reaction coordinates are defined as follows:

$$\xi = N_s^{-1} \sum_{s=0}^{N_s-1} \delta_s(N_s^{(p)}) \tag{2.15}$$

where $N_s^{(p)}$ is the number of polar heavy items — which include water oxygen atoms and the four oxygen atoms that are part of the lipids' phosphate groups — in slice s; and δ_s is a continuous indicator function that equals to zero when there are no polar atoms in slice$s$ and is close to unity when there are one or more polar atoms in slice s. Thus, in this model, the RC indicates the state of the pore — where $\xi = 1$ is fully open — and allows us to distinguish cases where the polar defect is only partially penetrating the membrane.

When calculating the potential mean force of pore formation along this new RC in AA simulations, a barrier at $\xi \approx 0.8$ is visible, which was not evident in previous works [80, 140]. This may explain why, with this RC, the pores did not spontaneously close, as had been the case with other models. Furthermore, it showcased no hysteresis between pore–opening and pore–closing simulations, a frequent problem with other RCs, and it was possible to sample the transition state. For these reasons, the Hub *et al.* RC is currently the most well-suited model in MD simulations to induce and study pore formation.

2.3 Setup of lipid pore study MD simulations

For both AA and CG simulations, we used the GROMACS simulation package [107, 108] version 2018 to run the lipid pore umbrella sampling simulations and version 2021 for the lipid in water and lipid in alkane melt systems. To model lipid pore formation, we followed the method described by Hub *et al.* and performed umbrella samplings using the open-source, community–developed PLUMED library [141], version 2.5.0 [142]. Every lipid pore formation was centered in the membrane.

2.3.1 All-atom simulations

The CHARMM36m force field for biomolecular simulations [143] was used for all AA simulations. The topologies used are all available in CHARMM-GUI [144–146]. We made use of the CHARMM-GUI membrane builder tool [144, 147–149] to assemble the initial lipid membrane structure. The DLPC membrane was built with 144 lipids per leaflet and solvated with ~13300 water molecules. For the DPPC membrane study, we used previously in–house built systems of the same size. For the lipid pore simulation, a 9.5×9.5×7.5 cubic box was used, whilst for the lipid in solution and in melt systems, a 6.5x6.5x6.5 dodecahedron box was used. All AA systems were solvated using the TIP3P water model and neutralized with Na^+ cations and Cl^- anions.

For every simulation, we used the fast smooth Particle-Mesh Ewald (SPME) summation [150] to compute electrostatics and Van der Waals interactions were switched off smoothly from 1.0 to 1.2 nm. The Verlet list scheme was used to update the particle neighbor list. For the DLPC pore simulations, we made use of a Nosé-Hoover thermostat [112] with a coupling time of 1.0 ps to maintain the temperature at 300 K, whilst for the lipid in solution and in alkane melt, we used a v-rescale thermostat [113] with the same coupling time to maintain the same temperature. For the membrane simulations, we used a semi-isotropic constant pressure, whilst for the rest of the systems we used an isotropic constant pressure. In both cases, they were coupled to 1.0 bar using the Berendsen barostat [116] with a relaxation time of 5.0 ps in the equilibration step, whilst for the production step we used the Parrinello-Rahman barostat [117] with the same pressure and relaxation time settings. The simulations were run at a 10 ps time-step, succeeding the initial energy minimization along with the temperature and pressure equilibration runs. We ran the in–water and in–alkane melt systems simulations for at least 2 µs and the membrane systems were run for at least 130 ns per umbrella window.

2.3.2 Coarse-grained simulations

The Martini 3 CG force field for biomolecular simulations [79, 85] was used for all CG simulations. The topologies used are all available with the force field, with the exception of lipid topologies, which were those of models currently being developed for better representation of bilayer properties, in collaboration with several other Martini developers. In every membrane simulation, we made use of the insane.py CG building tool [151] to assemble the lipid membranes with 144 lipids per leaflet. CG systems were solvated using regular Martini water beads and neutralized with Na^+ cations and Cl^- anions. For the lipid pore simulations, a 9×9×13 cubic box was used, whilst for the lipid in solution and in melt systems, a 6.5x6.5x6.5 dodecahedron box was used.

We developed various prototype topologies, that involved modifying the base topology of Martini 3 DLPC, to try to better describe the conformational space visited by AA lipids when forming pores. To do that, bonds were created between the tails' beads and between the last bead of each tail and the PO4, GL1, and GL2 beads. These bonds were described

by a harmonic potential. In order to create intermediate states between a no-bias and a biased topology, we used the GROMACS mechanism that allows us to interpolate between two values. Here, the lambda parameter (λ) represents the weight the first value has on the topology behavior versus the second value. The distance values between the CG lipid beads and between the forward–mapped AA lipid beads, obtained from bead distance matrices, were set as the minimum and maximum bond parameters, respectively. For the first prototype, we built three different membrane systems with different lambda values (0, 0.5 and 1.0), whilst the second prototype was tested at a lambda of 1.0 only. A force of 700 kJ was used in every case. Finally, a membrane consisting of the in-development lipids and of our second lipid prototype (20 per leaflet) was built and we used flat-bottomed position restraints with a cylindrical geometry (parallel to z) of 500 kJ of force and 3.0 nm radius (Figure 2.5 A) to corral the lipid prototype molecules in the center of the membrane, and, thus, in the vicinity of the pore.

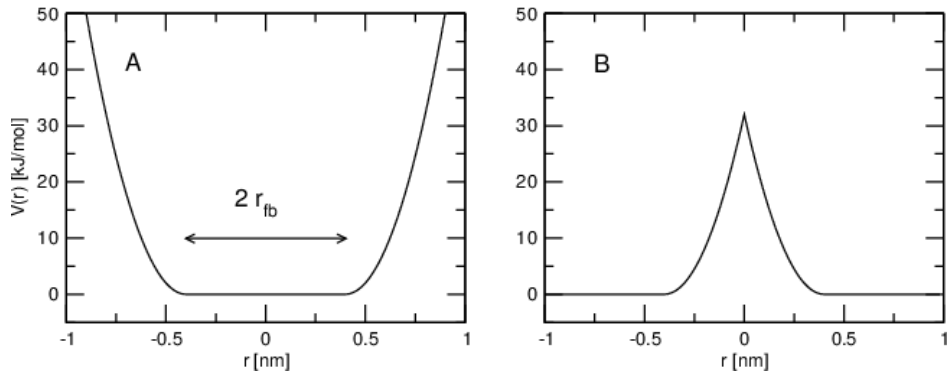

Figure 2.5: **Flat-bottomed position restraint potentials.** (A) Not inverted potential. (B) Inverted potential. Adapted from GROMACS user manual [107].

For every simulation, the nonbonded interactions had a cut-off distance of 1.1 nm and we used reaction-field electrostatics [152] with a dielectric constant of 15 and an infinite reaction-field dielectric constant to treat the Coulombic interactions. The Verlet list scheme was used to update the particle neighbor list. We made use of a v-rescale thermostat [113] with a coupling time of 1.0 ps to maintain the temperature at 323 K for DPPC simulations and at 300 K for DLPC simulations. For the membrane simulations we used a semi-isotropic constant pressure, whilst for the rest we used an isotropic constant pressure. In both cases, they were coupled to 1.0 bar using the Berendsen barostat [116] with a relaxation time of 3.0 ps in the equilibration step, whilst for the production step we used the Parrinello-Rahman barostat [117] with the same pressure but relaxation time of 12.0 ps. All simulations were run at a 10 ps time-step, with the exception of the in–development version 18 DPPC and the prototype simulations which were run at a 500 ps time-step, succeeding the initial energy minimization along with the temperature and pressure equilibration runs. We ran the in–water and in–alkane melt systems simulations for at least 19 µs and the membrane systems were ran for at least 2 µs per umbrella window.

2.4 Analysis of lipid pore MD simulations

All systems were analyzed with in-house and third–party developed Python 3 programs using the MDAnalysis package [153]. The NumPy [154], matplotlib [155], seaborn [156] packages were also used for scientific computing purposes and VMD [157] was used for visualization and rendering of simulations. The WHAM [158] and the Principal Component Analysis (PCA) PCALipids [159] software were also used for analysis purposes. PBC artifacts were fixed, for both CG and AA trajectories, using the GROMACS tool gmx trjconv, and the AA trajectories were forward-mapped to CG using an in-house developed script and third-party index.

Unless stated otherwise, for analysis purposes, we considered involved in pore formation (pore lipids) the lipids within a 8 Å cut-off distance of the center of the pore, at $\xi = 0.7$; the behavior of non-pore lipids (membrane lipids) was taken over the entire membrane, over the last segment of the trajectories at $\xi = 0.12$, where no pore is present (the last 5 µs were used for most systems, but the last 250 ns were used for trajectories with the in-development version 18 DPPC, prototypes, and mixed membrane systems, and the complete trajectories were used for the in–water and in–alkane melt systems).

2.4.1 Weighted Histogram Analysis Method (WHAM)

The WHAM package [158, 160] was used to obtain the free-energy surface of the systems. This method uses the raw trajectories of the multiple umbrella sampling windows to obtain the unnormalized probability histogram and, subsequently, calculate the unbiased potential mean force (PMF) profile. WHAM is considered to be converged when, in each umbrella sampling window, the free energy value does not vary more than a set tolerance value on consecutive iterations. For every simulation, we used a tolerance value of 0.0001. For the DLPC systems, the temperature was set to 300 K, whilst, for the DPPC systems, the temperature used was 323 K. To perform Monte Carlo bootstrap error analysis, 1000 fake data sets were generated for every system.

2.4.2 Lipid angle and distance distributions

In order to analyze lipid behavior in lipid pore formation, several Python 3 programs were developed using the MDAnalysis [153] and numpy [154] packages to calculate lipid angle distributions.

A schematic representation of the Martini 3 DLPC structure and corresponding bead names can be seen in Figure 3.1.

The *lipid tails splay angle* distribution was calculated from the angle between the positions of the C3A, PO4, and C3B beads. The *curling of lipid tails* was determined from the radius of gyration of each tail's beads. The *lipid head behavior in relation to glycerol* was obtained from the radius of gyration of the PO4, GL1, GL2, and NC3 beads. The *lipid dihedral angle* was calculated using the numpy calc_dihedrals function, with the positions

of the C3A, GL1, GL2, and C3B beads. The *C1A-GL1-GL2 and GL1-GL2-C1B splay angles* were determined from the angle between the positions of the C1A, GL1, and GL2 beads, and between the positions of the GL1, GL2, C1B beads, respectively. The *Tail A-GLs-Tail B splay angle* was obtained from the angle between the position of the C3A bead, the center of geometry of the GL1 and GL2 beads, and the position of the C3B bead. The *Tails-GLs-PO4 splay angle* distribution was calculated from the angle between the center of mass of the selection of every tail bead, the center of geometry of the GL1 and GL2 beads, and the position of the PO4 bead. The *GL1-G2-PO4 splay angle* was determined by calculating the angle between the positions of the GL1, GL2, and PO4 beads. The *Tail A-GL1-PO4 splay angle* distribution was obtained from the calculation of the angle between the positions of the C3A, GL1, and PO4 beads.

2.4.3 Lipid bead distances matrix

To compare CG and AA pore lipid structure behavior, a third–party developed script was used to obtain the bead distances matrices. The MDAnalysis package [153] was used to calculate distance arrays between the beads of the CG lipids' and of the forward–mapped AA lipids' structures. The seaborn package [156] was then used to plot the final heatmaps with the distances and standard deviations.

2.4.4 Root mean square error (RMSE)

The root mean square error (RMSE) was calculated for the different lipid topologies' bead distances using the numpy package [154]. The forward–mapped atomistic bead distances were used as the target value.

2.4.5 Principal Component Analysis (PCA)

The Principal Component Analysis PCALipids [159] software was also used to obtain and compare the lipid's conformational changes. Eigenvectors and eigenvalues were calculated using a trajectory of selected lipids, where 25% corresponded to CG pore lipids, 25% to CG membrane lipids, 25% to forward–mapped AA pore lipids, and 25% to forward–mapped AA membrane lipids.

2.4.6 Area per lipid distribution

The area per lipid distribution throughout the entirety of the simulated time was calculated using the MDAnalysis package [153]. For each frame, we used the box dimensions to calculate the area of the system, then divided it by the number of lipids in one leaflet.

2.4.7 Membrane thickness analysis

The membrane thickness throughout the entirety of the simulated time was calculated using the MDAnalysis [153] package. Membrane thickness was deemed as the distance

from the center of geometry of the upper lipids' PO4 bead to the center of geometry of the PO4 bead of the lower leaflet lipids.

2.5 Setup of alamethicin MD simulations

We used the GROMACS simulation package [107] version 2021 for all simulations.

The Martini 3 CG force field for biomolecular simulations [79] was used for all CG simulations. All of the topologies that were used are available with the force field, with the exception of the alamethicin topology.

The Alm peptide structure was built on PyMol [161], where the Aib and Phl residues were replaced by alanine and phenylalanine residues, respectively, and no C-terminal acetyl group was represented. We used Martinize2 to convert the Alm model obtained from PyMol into a CG structure, with an elastic network of a 700 kJ/mol bond force and a cutoff distance of 0.8 nm, as well as applied side-chain corrections [162]. The Martini beads corresponding to the side chains of the Aib residues were changed from TC3 to SC3, in order to represent the second β-carbon that is present in Aib residues but not in alanine. Finally, to represent the Phl residue and acetyl group at the terminals of Alm, we changed the backbone beads to P2 beads.

In all simulations, we made use of the insane.py CG building tool [151] to assemble the lipid membranes and insert the peptides into the membrane. The CG systems were solvated using regular Martini water beads and neutralized with Na^+ cations and Cl^- anions. A membrane of 75% POPE and 25% POPG lipid composition was used, along with a 49×49×12 box, for all simulations. Flat-bottomed position restraints, with a layer parallel to z geometry of 2.0 nm of width and 700 kJ of force, were applied to both the lipids and the peptides in order for the membrane to not undulate and the peptides to remain inside the membrane [86], respectively. To the lipids, a not–inverted potential was applied (Figure 2.5 A), selecting the two glycerol beads, whilst to the peptides we applied an inverted potential (Figure 2.5 B), selecting the first and last backbone beads. For the antiparallel systems, a script was developed to turn peptides 180 °, in a checkerboard pattern. In the reduced water-contact simulations, the interactions between the water and the peptide beads were weakened by 10%, as done in previous works [163].

For every simulation, the nonbonded interactions had a cut-off distance of 1.1 nm and we used reaction-field electrostatics [152] with a dielectric constant of 15 and an infinite reaction-field dielectric constant to treat the Coulombic interactions. The Verlet list scheme was used to update the particle neighbor list. We made use of a v-rescale thermostat [113] with a coupling time of 3.0 ps to maintain the temperature at 300 K. A semi-isotropic constant pressure was used for all simulations. They were coupled to 1.0 bar using the Parrinello-Rahman barostat [117] with a relaxation time of 12.0 ps. The simulations were run at a 20 fs time-step, succeeding the initial energy minimization along with the temperature and pressure equilibration runs. We ran the CG systems simulations for at least 3.7 µs.

2.6 Analysis of alamethicin CG MD simulations

The systems were analyzed using in-house developed Python 3 programs with the MDAnalysis package [153]. VMD [157] was used for visualization and rendering of simulations. The NumPy [154], matplotlib [155], and scikit-image [164] packages were also used for scientific computing purposes. The full trajectories were used for all analyses.

2.6.1 Stamp plot

To analyze peptide aggregation behavior, we developed what we deemed as the 'stamp plot'. The stamp plot consists of analyzing the xy position of each peptide and lipid throughout the simulation run and marking it on a plot. To do that, an array of zeros with the dimensions of the universe is built and a different-sized 'stamp' (a circle of ones with a corresponding radius) is assigned to each peptide's/lipid's bead size (regular, small, or tiny — radius assigned according to the same rules as for determining Martini 3 solvent accessible surfaces [165]). Stamps were then projected on an area corresponding to a top-view of the system, at a resolution of 0.1 Å per pixel. PBC artifacts were taken into consideration when marking the 'stamps' — i.e. the part of the 'stamp' that went beyond the PBC boundary was marked on the opposite side. We then essentially performed a connected component analysis of white (unmarked) pixels, using the skimage.measure.label module, which takes a binary image as input and distinguishes between objects according to a criterion that says that pixels with the same value belong to the same object. For this, the connectivity level, i.e. the maximum distance from our reference at which we still consider pixels as neighbors, was set to 2. The np.bincount function was then used to obtain the area of every pore, in each frame. The developed code can be found in the Supplementary Material available with this thesis.

2.6.1.1 Total number of pores per frame

The total number of pores per frame by pore area (in pixels) is a histogram distribution of the number of pores with varying areas. This distribution was calculated from the data obtained with the stamp plot.

2.6.1.2 Total area of pore distribution

The distribution of the total area of pore throughout the simulation run was also plotted. To calculate this, we used the data derived from the stamp plot and created an array containing the sum of the values in the data, for each frame. A threshold value of 10 pixels was used, in order to avoid considering spurious low-density regions as pores.

2.6.2 Water contacts distribution

The number of peptide-water contacts per frame was calculated using the MDAnalysis [153] package. A distance cut-off of 4 Å was used to represent a contact.

3 Results and Discussion

3.1 Martini 3 CG lipid–lined pore simulations

With the Martini 2 force field, lipid–lined pore formation with regular water could only be observed under specific situations and, although it was possible to obtain water defects in membranes, these quickly became unstable [80, 166]. In the new Martini 3 force field, the free-energy distribution of pore formation is much higher than that of the atomistic models. Since the release of this new iteration, several efforts have been put into the reparametrization and improvement of Martini 3 biomolecule models. Phospholipids (Martini mapping exemplified in Figure 3.1) are an example of that, and with this comes the opportunity to address any shortcomings of its previous model — such as incorrect lipid pore behavior.

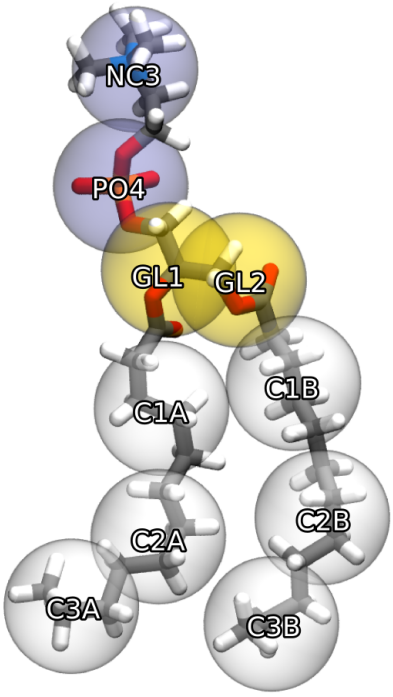

Figure 3.1: **Standard Martini mapping of a phospholipid molecule.** DLPC atomistic and Martini CG models are superimposed. The atomistic model is represented by sticks whilst the Martini model is represented by beads.

To test lipid pore formation in Martini 3 CG simulations [91], using the Hub *et al.* model, we assembled three different systems: a lipid membrane of 144 lipids per leaflet; one lipid in water; and one lipid in alkane melt. These systems were replicated with DPPC and DLPC lipids.

Throughout the course of developing this thesis, several models of Martini 3 lipids were developed by third–party collaborators . In this work, we tested lipid pore formation behavior with both the original Martini 3 lipids as well as versions 15 and 18 of the in–development new Martini 3 lipids, which were shared with us in an effort to test them and, consequently, possibly lead to their improvement.

3.1.1 Lipid–lined pore formation free-energy in CG simulations

We started by analyzing the free energy of lipid–lined pore formation in Martini 3 lipid membranes and comparing them with the values of atomistic simulations of the same membrane systems. It is evident in Figure 3.2 that, for both DLPC and DPPC membranes, there is a large difference in PMF profiles.

In the DPPC profile (Figure 3.2, left) we can see that both the minimum free energy and the the nucleation barrier (which is characteristic of the Hub *et al.* RC and is where the pores are in their transition state between a closed and a continuous transport state) are achieved earlier in every Martini simulation than with AA methods. This may be due to the fact that the RCs used for the CG and AA simulations had slightly different parameters that could be contributing to this disparity. Additionally, in CG models, the penetration of water beads into the membrane is smaller than in AA, which could also contribute to the membrane reaching the flat intact state at a smaller ξ value. Notwithstanding the varying RC parameters, what stands out the most is the fact that the nucleation barrier in Martini, whether with its original or the in–development lipids, is significantly higher than in AA simulations (a difference of around 60-70 kJ/mol) — although there was an improvement from the original lipids to the latest in-development version. This means that it is much harder to reach a metastable pore in Martini than it is with AA methods. Despite the fact that the AA value we obtained is underestimated in comparison with past PMF calculations (where it was $\sim$ 78 kJ/mol) [45], we are dealing with a double-digit difference which is still noteworthy. Finally, the barrier is not as well defined in the CG force field, i.e the free energy continues to quickly increase instead of stabilizing and slowly decreasing, as is seen in AA simulations. This suggests that Martini pores may not be metastable and could spontaneously close, hence why it is not possible to obtain lipid–pore formation in the force field without using restraints and additional forces.

In the DLPC PMF profile (Figure 3.2, right), similar behavior to the DPPC pores can be seen. Although the minimum energy of the profile is now reached at around the same reaction coordinate, the nucleation barrier is still higher in Martini than in AA (this time, a difference of around 25-50 kJ/mol) and the same undefined profile is also clear. Interestingly, our AA DLPC profile does not have a well-defined barrier either. We suspect

this may be due to not having enough sampling of the umbrella windows which did not allow for convergence of the PMF. Nonetheless, in previous calculations [45], it was seen that in DLPC membranes the free energy of the open pore was of ~ 17 kJ/mol, which only increases the disparity between these values.

PMF values can be affected by a number of factors: initial conformations, reaction coordinates, sampling of the conformational space in the umbrella sampling windows, etc [167]. Since we are using the same RC for both AA and CG simulations, it is clear that the cause of this difference most likely lies in the lipid pore conformation in the Martini force field. Overall, the inaccurate replication of pore formation in the Martini force field may result from the parametrization of the lipids themselves or from how lipid-water interactions are modeled. As such, our first objective was to find the most probable cause of this issue.

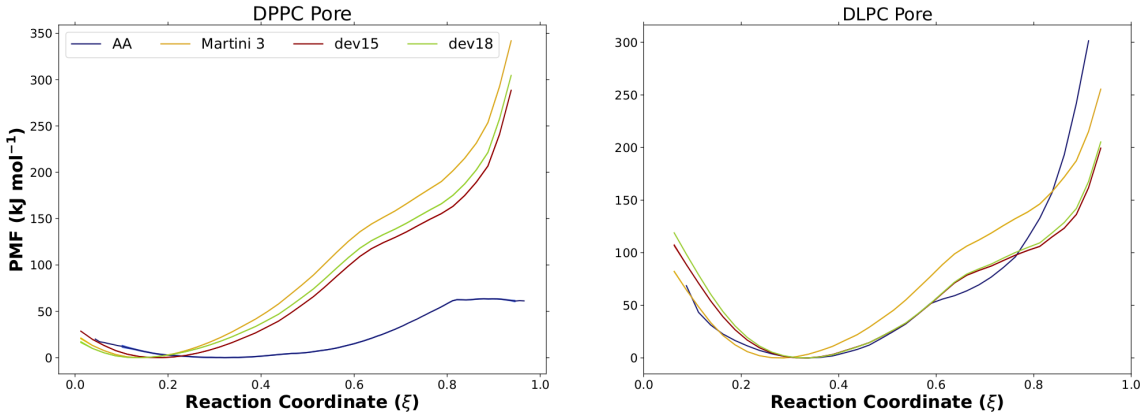

Figure 3.2: **Potential Mean Force (PMF) distribution of pore formation in atomistic and coarse-grained membrane systems.** On the left is the PMF distribution of DPPC pores; on the right is the PMF distribution of DLPC pores. The atomistic distributions are in dark blue; the distribution of membranes built with the original Martini 3 lipids are in yellow; and distributions of versions 15 and 18 of the in–development Martini 3 lipids are in red and green, respectively. Error bars are represented by shading in the same colors along the corresponding curve, for the systems from which WHAM was able to calculate the values. The shading is on the same scale as the plot curves' thickness.

3.1.2 Pore lipid behavior

In order to pinpoint a possible target that could be hindering lipid pore formation, we analyzed several intramolecular properties of Martini 3 the lipids involved in pore formation (referred to as pore lipids, moving forward) and compared them to their forward–mapped atomistic counterparts.

From the various angle and distance distributions that were calculated (available in appendix), the most promising target seemed to be the splay angle formed by the lipids' tails and the glycerol beads' center of mass. As we can see in Figure 3.3, in DLPC membrane lipids (i.e non pore–forming lipids) this angle seems to be very similar in AA and both

Martini 3 simulations. On the other hand, in DPPC membranes there was a much bigger difference between the AA lipid's splay distribution and the version 15 lipid's distribution. This issue was subsequently addressed in the version 18 lipids and now we no longer have a significant difference between the AA and Martini 3 lipids' tail splay angle. Whilst the latest Martini version and the atomistic lipids' behaviors in the membrane now seem to be almost identical, Martini 3 pore lipids still showcase significantly different behavior than their AA counterpart. Both DLPC and DPPC Martini models show a big increase in their tail splay when forming pores. In the remaining water and alkane melt systems, it seems that the opposite happens: the angle distributions for the CG lipids are narrower than their AA counterparts. This further indicates that the bigger tail splay may be an exclusive characteristic of Martini lipids when forming pores and, consequently, a potential cause of the inaccurate lipid pore behavior in Martini 3 simulations.

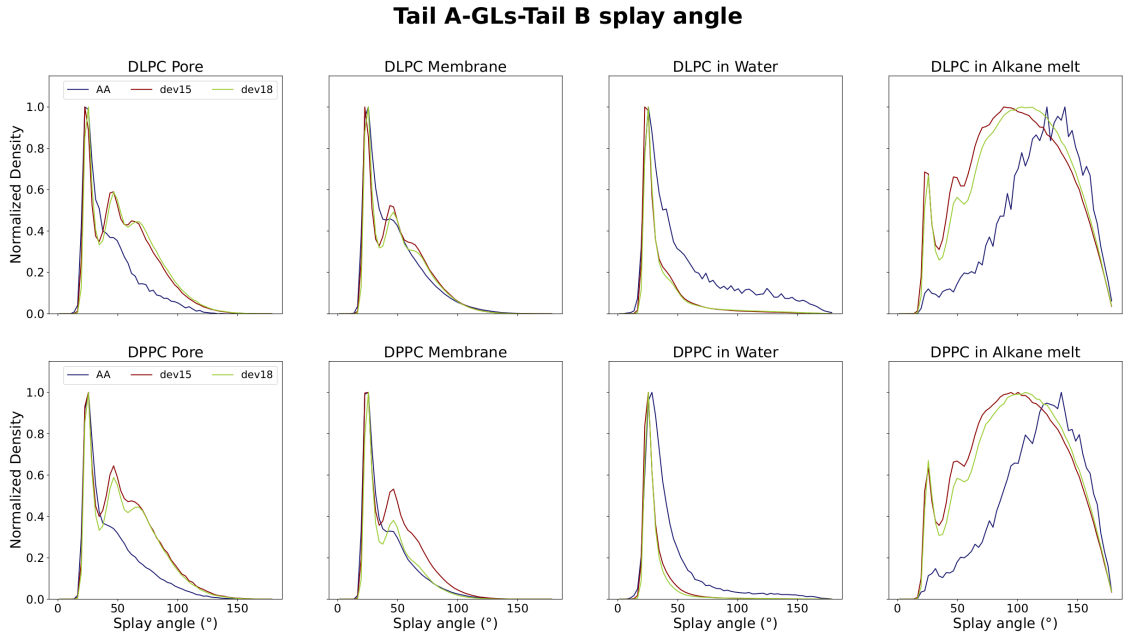

Figure 3.3: **Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in–development Martini 3 lipids are in red and green, respectively.

This clear difference in the tail splay angle is also evident in the lipid bead distance matrices analysis that we performed for version 18 of Martini 3 DLPC and DPPC lipids (Supplementary Figures A.10 and A.11, respectively). In these matrices, we can once again see that the DLPC tail splay is much bigger in CG pore lipids than in AA lipids, since the distance between the tail A and tail B beads is considerably larger. We are also able to note that, in the version 18 DLPC pore lipids, there is a bigger distance between the C2 and C3 beads but a smaller distance between the C1 beads, when compared to the AA DLPC lipids bead distances.

Additionally, there is also an asymmetry in the Martini DLPC tails that is not present in AA lipids. This is made evident by Supplementary Figure A.10 as well as our second

distance matrix analysis — this time of the difference between Martini 3 and AA lipids distances, i.e. the atomistic distances subtracted from the Martini distances, seen in Figure 3.4. Martini DLPC lipids have their tail B slightly curled into tail A, with the C3B bead at a shorter distance from the C1A bead in comparison with C3A and C1B. All of these discrepancies in value can be quantitatively seen when calculating the RMSE values for the three versions of Martini 3 lipids, with the AA results as the target value (Table 3.1).

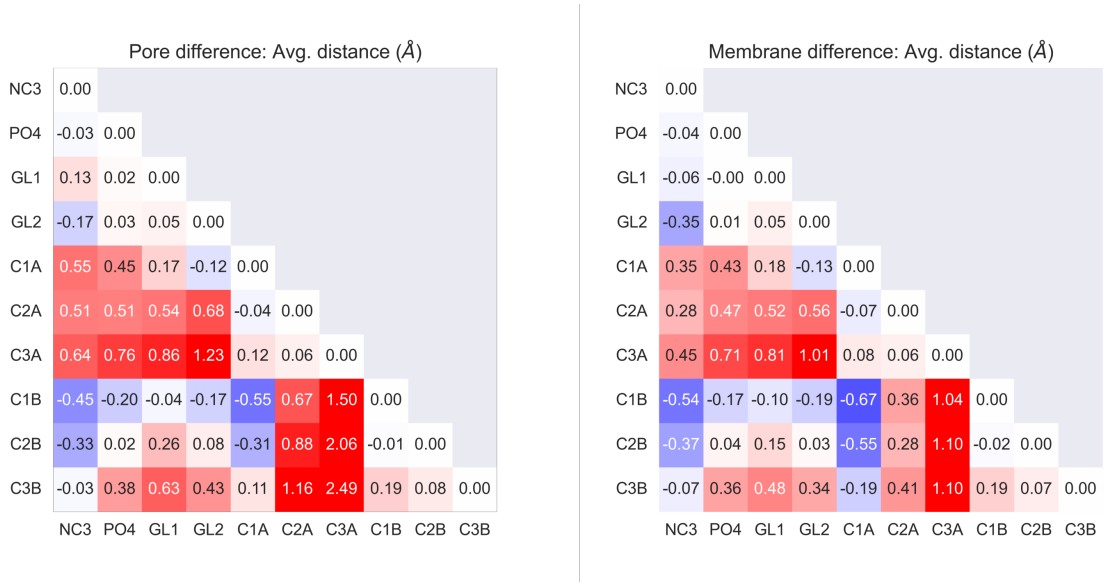

Figure 3.4: **Distance matrices of the forward–mapped atomistic DLPC bead distances subtracted from the version 18 Martini 3 DLPC bead distances.** On the left is the distance matrix corresponding to pore lipids; on the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

Table 3.1: **Root mean square error values of the in–development Martini 3 DLPC bead distances in Å.** The forward–mapped atomistic distances were used as the target value.

	Pore lipids (DLPC)	Membrane Lipids (DLPC)
Martini dev15	0.834	0.655
Martini dev18	0.665	0.436

Taking these results into consideration, we deemed this property as the most promising target to achieve a better and more accurate lipid behavior when forming pores. As such, we went on to tweak and test various DLPC topologies — since these lipids tend to form metastable pores naturally [168].

3.1.3 Testing DLPC lipid topologies

To understand if the difference in tail splay was the main cause of the incorrect lipid pore formation behavior in Martini 3, we created multiple prototypes of version 18 of

Martini's DLPC topology with added bonds that would address the differences between the AA and CG models that we saw in our analysis.

3.1.3.1 Prototype 1 - Added tail-tail bonds

We started by testing a Martini 3 DLPC topology that had new intramolecular bonds between both tails' beads, i.e. bonds between C1A and C1B, C2A and C2B, and so on. We built three lipid bilayers consisting entirely of this DLPC topology, where each had the lipid's new bonds with different lambda values — 0, 0.5, and 1.0. As such, we had a control group where the topology behaves as purely CG (system with lambda at 0); a middle stage of interpolation between a purely CG topology and a topology with the distance between the tails with AA values (system with lambda at 0.5); and, finally, a topology where the new bonds have purely AA distance values. These distance values were obtained from bead distance matrices and the maximum bond parameter was set as the forward–mapped AA bead distance.

To ensure that the new bonds did not harm any of the crucial lipid properties, we analyzed the membrane thickness as well as the area per lipid (APL). Figure 3.5 shows that, although there were a few alterations in values, the lamellar structure of the membrane was not compromised with the added bonds, for either lambda value. DLPC membranes are typically 3.14 nm thick [169], therefore, a thickness of ≈ 3.5 nm can be considered reasonable. Furthermore, since we are decreasing the distance between the tails, it is expected that our membrane lipids become more tightly compacted and, thus, the APL decreases slightly.

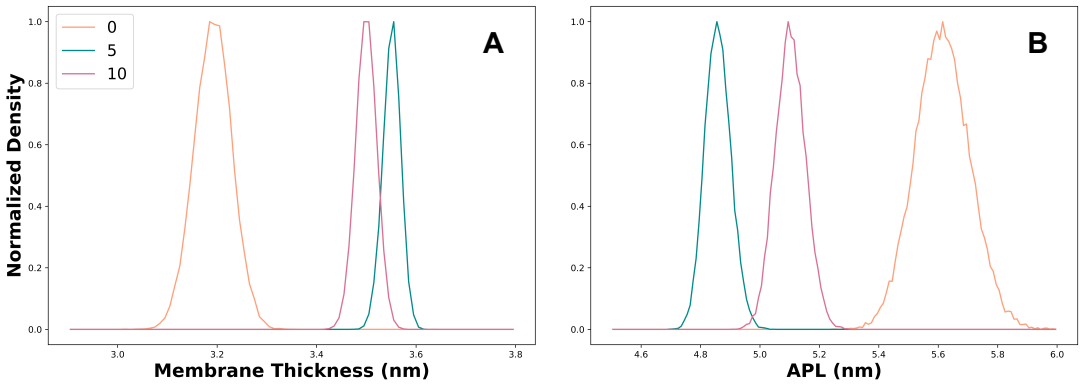

Figure 3.5: **Area per lipid and membrane thickness distributions of membranes built with the first lipid prototype.** (A) is the membrane thickness distribution throughout the simulated time; (B) is the APL distribution throughout the simulated time. Regular lipid distribution is in orange, whilst the distributions of our lipid prototype with the new bonds at 50% and 100% of maximum force are in blue and pink, respectively.

As we can see in Figure 3.6, the added bonds did not lead to a better pore formation PMF profile, with either lambda value. In fact, the maximum free energy even increased

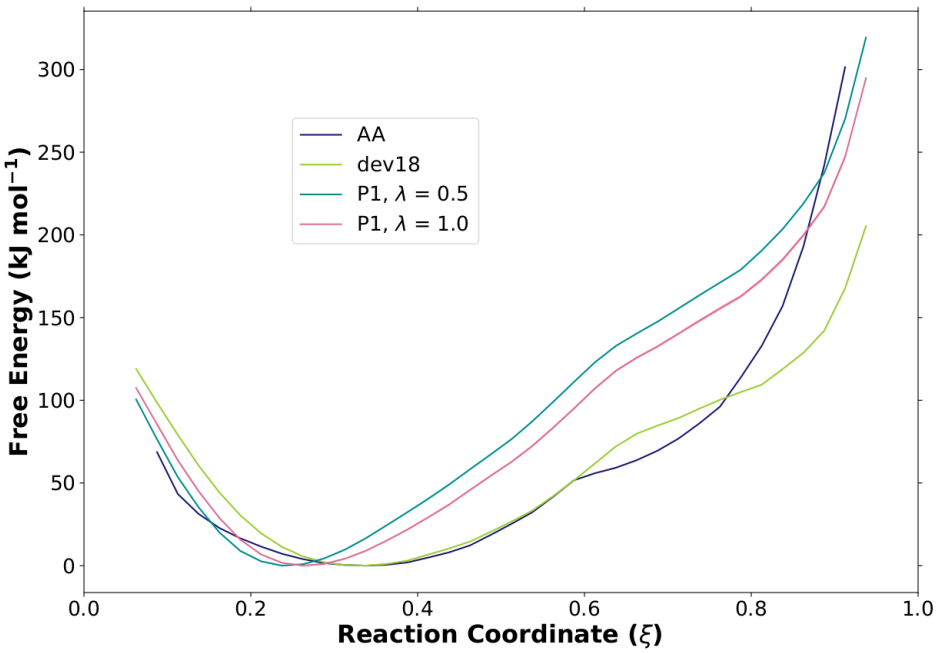

Figure 3.6: **Free energy of lipid pore formation in the first lipid prototype membrane system.** The atomistic distribution is in dark blue, the distribution of version 18 of the in–development Martini 3 lipids is in green, and the distributions of our first lipid prototype at lambda values of 0.5 and 1.0 are in light blue and light pink, respectively. Error bars are represented by shading in the same colors along the corresponding curve, for the systems from which WHAM was able to calculate the values. The shading is on the same scale as the plot curves' thickness.

in both cases, when compared with the Martini 3 original topology's value, which further distanced us from the target AA PMF behavior. This led us to believe that either the tail splay was not the only lipid property that was potentially causing the inaccurate lipid pore behavior or it did not have any significant impact on this behavior. From this point onward, all analyses were done with the prototypes at lambda equal to 1.0.

In Figure 3.7, it is clear that the asymmetry of the CG lipids tails (i.e the difference in curling) is slightly better. The tails of the membrane lipids now also showcase a much shorter splay, as seen in Figure A.15, which did not happen with the in–development version 18. This suggests that this shorter splay is most likely the result of the added bonds restraining the movement of the tails, influencing the behavior of the lipids in both the membrane and the pore. Although the distance matrices do not look so promising, when we look at the RMSE results in Table 3.2, we can see that there actually was a big decrease in value for pore lipids (although this same value increased for membrane lipids).

However, with the distance matrices in 3.7 we were also able to identify a new potential target. Our prototype's distances between the beads of the head and the last bead of both tails now have a big disparity from its AA counterpart. Whilst previously we had another asymmetry in this distance (tail A appeared to curl less than tail B), in our new prototype

it seems that this no longer happens and, instead, both tails do not curl as much as the AA model. This can be further evidenced by Figure A.24, where the AA lipids have a broader angle distribution in comparison with the CG model. Nonetheless, this distance was still present in one of the tails of version 18 lipids, therefore, combined with the tail splay issue, it could also be affecting the pore lipid behavior. As such, we considered this another pore lipid property that seemed to be a promising target.

Table 3.2: **Root mean square error values of the in–development and the Martini 3 DLPC prototype bead distances.** The forward–mapped atomistic distances were used as the target value.

	Pore lipids (DLPC)	Membrane Lipids (DLPC)
Martini dev18	0.665	0.436
Prototype 1	0.508	0.490

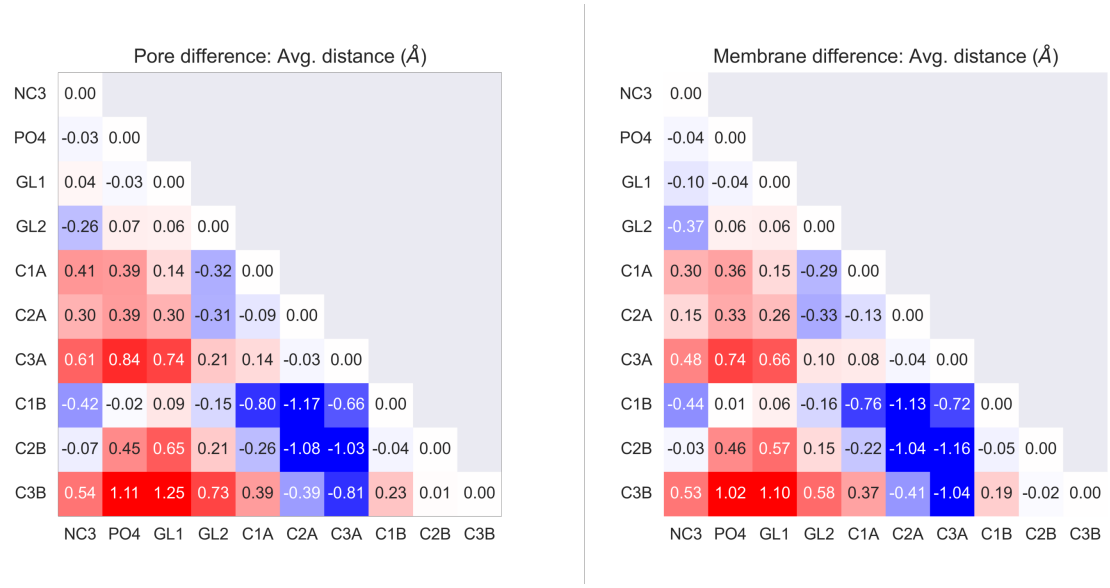

Figure 3.7: **Distance matrices of the forward–mapped atomistic lipid's bead distances subtracted from the first lipid prototype lipid's bead distances.** On the left is the distance matrix corresponding to pore lipids and on the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

3.1.3.2 Prototype 2 - Added tail-tail and head-tail bonds

Taking this new target into consideration, we built a second DLPC topology that maintained the extra tail bonds we had in our previous prototype and had new added bonds between the head beads and the last bead of each tail, i.e GLs and C3A, GLs and C3B, PO4 and C3A, PO4 and C3B, etc. These lipids were then tested in a membrane system with a lambda equal to 10.

This time, the lipid pore formation free energy did lower slightly, however, it is still significantly higher than that of the AA simulations and there is no distinguishable nucleation barrier (Figure 3.8).

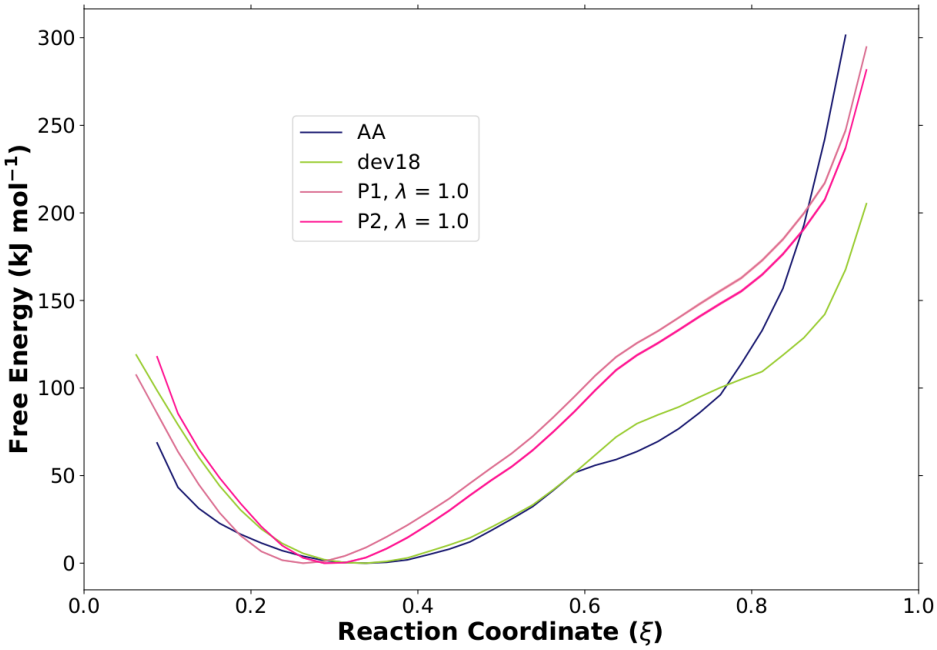

Figure 3.8: **Free energy of lipid pore formation in the second lipid prototype membrane system.** The atomistic distribution is in dark blue, the distribution of version 18 of the in–development Martini 3 lipids is in green, the distribution of our first lipid prototype at a lambda value of 1.0 is in light pink, and the distribution of our second lipid prototype at a lambda value of 1.0 is in dark pink, respectively. Error bars are represented by shading in the same colors along the corresponding curve, for the systems from which WHAM was able to calculate the values. The shading is on the same scale as the plot curves' thickness.

From the distance matrices of the difference between the AA system's and the second prototype's distances (Figure 3.9), we can see that our second prototype has very similar behavior to the AA model. We were even able to further decrease the RMSE value of the pore lipids as well as correct the previous increase in the membrane lipids' RMSE (Table 3.3). Nonetheless, the CG tail splay is still considerably smaller than that of the AA lipids for both pore and membrane lipids, just like in the case of the first prototype, showcasing a possible lack of freedom in the movement of the tails caused by the new bonds that is affecting both membrane and pore lipid behavior. This lack of freedom becomes quite apparent in Figure A.26, where we can see that both pore and membrane lipids have lost that specific distribution behavior and are, instead, restricted to a specific value.

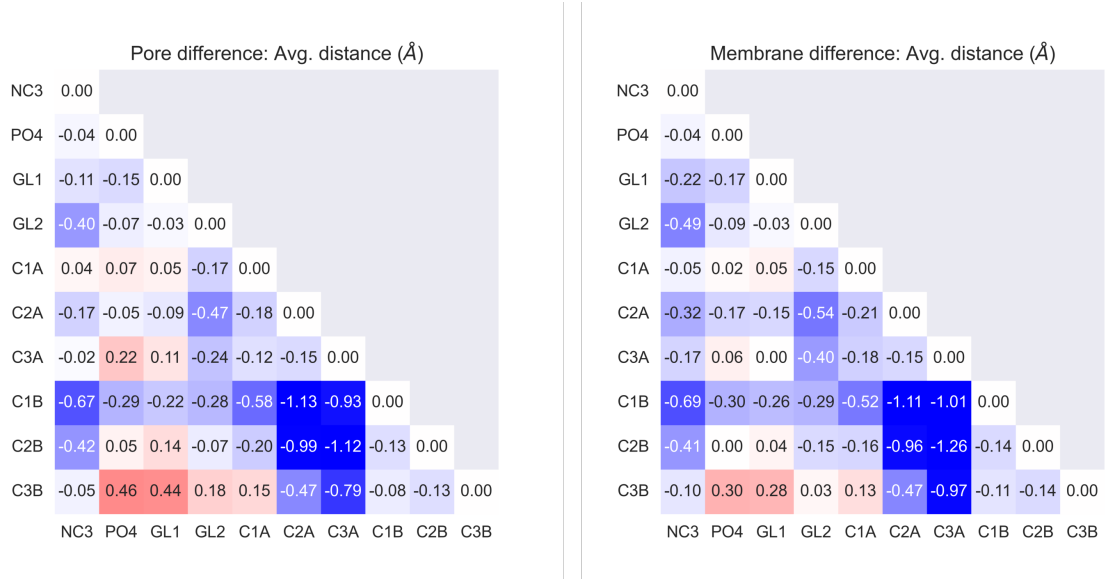

Figure 3.9: **Distance matrices of the forward–mapped atomistic lipid's bead distances subtracted from the second lipid prototype lipid's bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

Table 3.3: **Root mean square error values of the in–development and both Martini 3 DLPC prototypes bead distances.** The forward–mapped atomistic distances were used as the target value.

	Pore lipids (DLPC)	Membrane Lipids (DLPC)
Martini dev18	0.665	0.436
Prototype 1	0.508	0.490
Prototype 2	0.391	0.412

3.1.3.3 Mixed Membrane

Given that the Martini version 18 membrane lipids behaved very similarly to their atomistic counterpart and only the pore lipids showcased differences in tail splay, we decided to build a membrane where the pore was formed only by DLPC lipids that followed our second prototype topology whilst the rest of the membrane consisted of version 18 DLPC.

With this mixed membrane, the free energy value lowered once again, although it is more or less the same as version 18 of Martini lipids, as we can see in Figure 3.10. Additionally, we also still lack a nucleation barrier in our energy profile. This points to the possibility that neither the tail splay nor the curling of the tails is causing the inaccurate energy profile of pore formation.

Still, when we look at the distance matrices of the difference between the distances of the AA system and the mixed membrane system (Figure 3.11), we can see that the lipid behavior is now mostly identical, both in the case of the membrane lipids as well

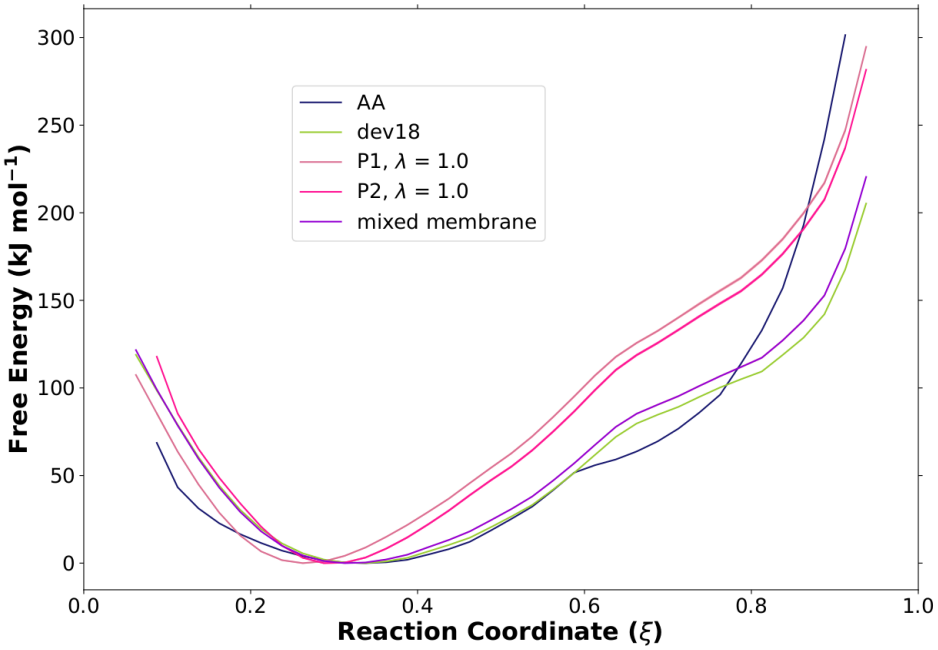

Figure 3.10: **Free energy of lipid pore formation in the mixed membrane system.** The atomistic distribution is in dark blue, the distribution of version 18 of the in–development Martini 3 lipids is in green, the distribution of our first lipid prototype at a lambda value of 1.0 is in light pink, the distribution of our second lipid prototype at a lambda value of 1.0 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively. Error bars are represented by shading in the same colors along the corresponding curve, for the systems from which WHAM was able to calculate the values. The shading is on the same scale as the plot curves' thickness.

as the pore lipids. This is also shown quantitatively by the RMSE values in Table 3.4. The RMSE has significantly decreased for both pore and membrane lipids, in comparison with prototypes 1 and 2. In fact, they even reached lower values than those of the Martini in–development version 18. This suggests that, by using both our prototype and Martini version 18's lipid topologies in our membrane system, we were able to almost perfectly mimic the AA model's behavior regarding bead distances.

Table 3.4: **Root mean square error values of the in–development, both Martini 3 DLPC prototypes, and the mixed membrane lipids' bead distances.** The forward–mapped atomistic distances were used as the target value.

	Pore lipids (DLPC)	Membrane Lipids (DLPC)
Martini dev18	0.665	0.436
Prototype 1	0.508	0.490
Prototype 2	0.391	0.412
Mixed Membrane	0.386	0.338

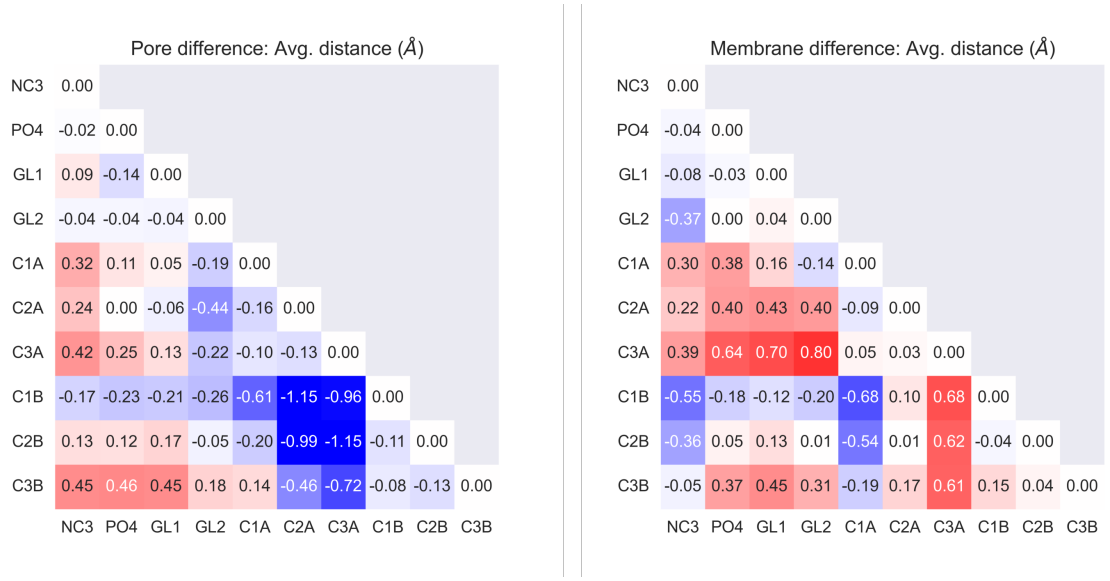

Figure 3.11: **Distance matrices of the forward–mapped atomistic lipid's bead distances subtracted from the mixed membrane system lipid's bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

In Figure 3.12, however, we can see that the tail splay angle of the pore lipids is restricted to the average angle of the AA lipids' distribution. This is most likely due to the force we are applying to the bonds being too strong, which leads to constrained movement of the tails and loss of the behavior that was previously seen in the AA model and the in–development version 18 of the Martini 3 model' distributions. This is further evidenced by the principal component analysis (PCA) seen in Figure 3.13. Here, the first principal component (PC) represents the tail splay movement. The pore lipids of the mixed membrane have a significantly smaller variance in PC1 than that of the atomistic system. In fact, most of their variance is actually described by PC2 and PC3 — as seen in Supplementary Table A.2. This difference is not as significant in Martini 3 version 18 lipids, which has a marginally higher variance for PC1. Additionally, in Figure 3.12, we can also see a peak in the membrane lipids' splay around the atomistic average value, which is most likely due to our selection not being restrictive and, thus, also accounting for the lipid prototype's splay.

Taking all of these results into consideration, it is possible that the root of the issue is not in the lipid parameterization but rather in how water-lipid interactions are modeled in the force field. Since four water molecules are represented by a single Martini bead, with only one Lennard-Jones site, there is a loss of fine structural detail which may be affecting the molecule's interaction with the pore lipids and leading to an inaccurate representation of water transport through lipid–lined pores in the biomembrane (and, consequently, of the PMF profile of pore formation) [85].

Nonetheless, given that the added bonds in our prototypes seem to be too restrictive and we were not able to accurately reproduce the distribution profile of the AA model,

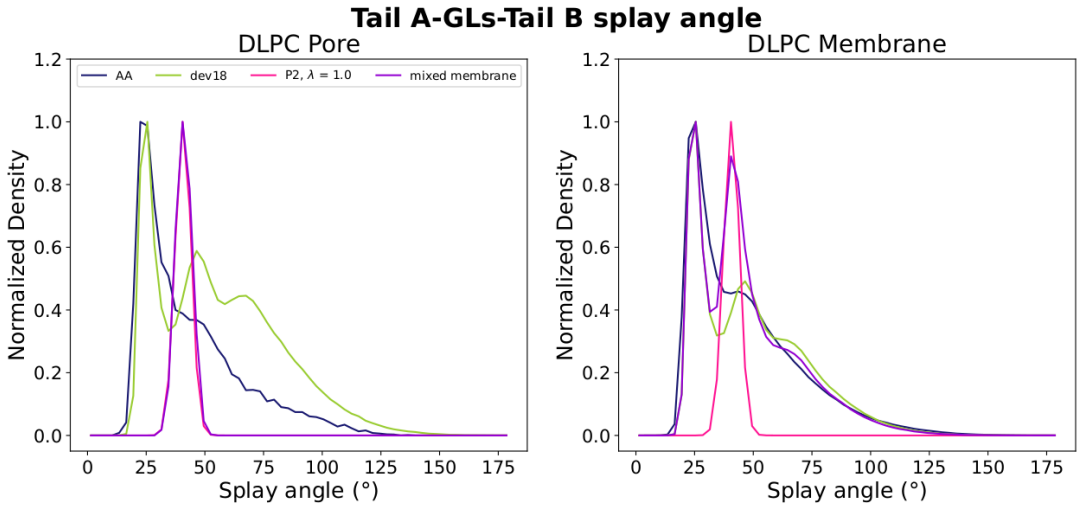

Figure 3.12: **Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in–development Martini 3 lipids is in green, the distribution of our second lipid prototype at a lambda value of 1.0 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

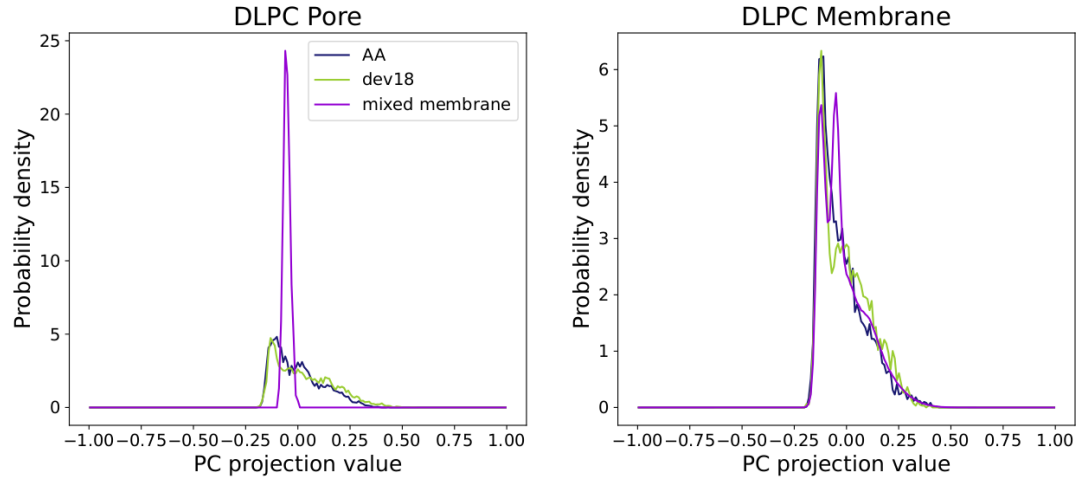

Figure 3.13: **Distribution of the first principal component projection value.** The atomistic distribution is in dark blue, the distribution of version 18 of the in–development Martini 3 lipids is in green, and the distribution of the mixed membrane system is in purple, respectively.

we cannot completely disregard the tail splay as a possible cause for the incorrect pore lipid behavior.

Although we were not able to come to a definite conclusion and pinpoint exactly the cause of our problem, we now have potential targets that may be affecting lipid pore behavior. In future work, we will test these prototypes with weaker forces on the added bonds so as to understand if, with more freedom of movement of the tails, we can obtain a better outcome.

3.2 Alamethicin pores

In previous Martini 2 simulations [69], it was not possible to obtain alamethicin pore formation in biomembranes, since the Alm peptides (sequence represented in Figure 3.14) had several residues that established continuous interpeptide contacts throughout the simulated time and the water molecules could not travel between the peptides.

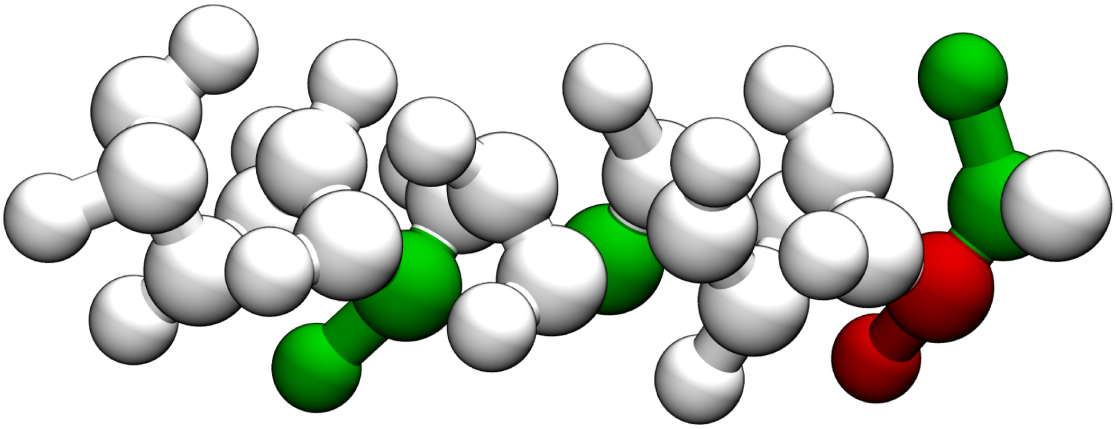

Figure 3.14: **Martini 3 coarse-grained representation of alamethicin peptide.** Residue sequence is described at the bottom. Apolar residues are represented in white; polar residues are in green; and acidic residues are in red.

However, with the new version of the Martini force field and its improved protein-protein interactions [85], we are hopeful that these issues have been addressed and, consequently, we can now be able to observe alamethicin–induced pores in coarse-grained simulations.

To study Alm pore behavior in Martini 3 CG simulations, we prepared four different systems. A higher peptide concentration system, with around 1 peptide per 16 lipids, and a smaller peptide concentration system of 1 peptide to 48 lipids ratio. Both of these conditions were replicated with the peptides in parallel and in antiparallel conformations.

In all four systems, we were able to obtain stable Alm pores (seen in Figure 3.15), with water and ion transport. In past Martini 2 simulations [69], several of alamethicin's residues established peptide-peptide contacts whilst forming the pore, throughout a significant portion of the simulation period. These contacts could have been a contributing factor as to why the Alm pore behavior could not be properly simulated. But now, with the improved protein-protein interactions of Martini 3, it seems that this issue has been overcome and we're able to simulate alamethicin's pore behavior more accurately. Additionally, it was postulated that the size of the water beads could be another reason for the incorrect modeling. Given that the size of the water beads was not changed between Martini versions, it becomes clear that this was not, in fact, affecting pore behavior.

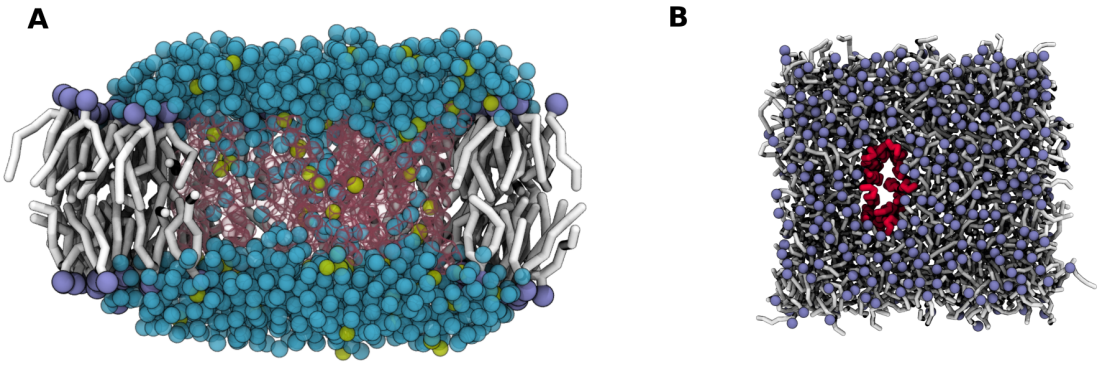

Figure 3.15: **Selected snapshots of two distinct alamethicin pore structures in the high concentration, antiparallel peptide conformation system.** Lipid heads are represented in lilac, lipid tails in white, the alamethicin backbone in red, water molecules in blue and ions in yellow. (A) Side view of an alamethicin cluster with multiple temporary pores (with peptide backbone represented in see-through). (B) Top view of a single alamethicin pore.

In Supplementary Video A.1, we can see the Alm peptides forming increasingly bigger clusters — like the one in Figure 3.16. These clusters aggregated until they ultimately formed a single large raft with several temporary peptide–induced pores. Pieta *et al.* [65] as well as Thøgersen *et al.* [69] obtained similar clusters in their experimental and computational works to the ones we saw in the beginning of our simulations. This leads us to believe that these clusters may be an important property of Alm channels that contributes to the peptide's antimicrobial activity. Nonetheless, cluster formation by Alm peptides is still up for debate and has yet to be proven. To our knowledge, the formation of a large raft structure like the one we obtained — and that we nicknamed "the island" — has not been seen either experimentally or in other AA or CG MD simulations. This is an interesting phenomenon and it may be a relevant structure of alamethicin's mechanism of action that has yet to be discussed.

In future work, it would be worth developing specific parametrization of the Aib and Phl residues in order to represent alamethicin's structure and its interactions as accurately as possible. Although this is a limitation of our model, we believe Alm behavior was still faithfully reproduced.

3.2.1 Peptide aggregation analysis

To quantitatively analyze our results, we started by doing a connected component analysis. From this analysis, we were able to plot the total area of alamethicin pores throughout the simulated time — Figure 3.17 A. We expected to obtain a behavior similar to what Harriss *et al.* achieved in their work [67], where multiple different pore conductance states were evident, however, we were not able to replicate this in CG. Instead, our results show no distinctive levels in any of our systems. Generally, these different conductance

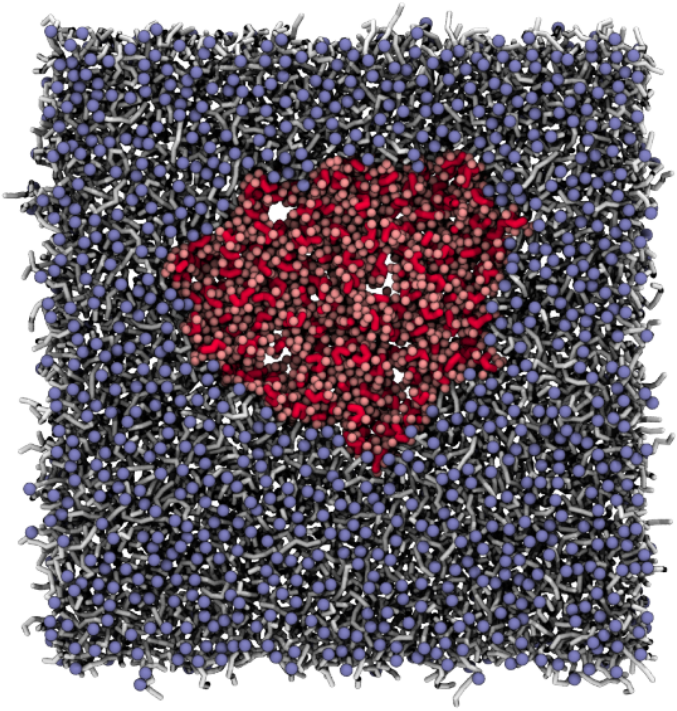

Figure 3.16: **Selected snapshot of the formation of the alamethicin raft in the high concentration, antiparallel peptide conformation system.** Alamethicin peptides aggregated into large clusters, such as this one, with multiple temporary peptide–induced pores, until they finally formed a raft (nicknamed the island). Lipid heads are represented in lilac, the alamethicin backbone in red, and the side chains in pink.

states are attributed to pores formed with different numbers of peptides. As such, the lack of discrete levels may suggest that we did not obtain pore populations of well-defined and different sizes. Conversely, if this behavior can be attributed to single alamethicin pores (as suggested by Harris *et al.*), it indicates that this characteristic is possibly not well replicated in Martini 3.

Although we were not able to replicate those different states along the simulated time, it seems like the most favorable system, regarding Alm pore formation, was the high–concentration system with the peptides organized in a parallel conformation. This indicates the possibility of Alm peptides having a preference for exclusively forming pores with peptides that have the same orientation, although antiparallel pores still formed, as was postulated by Rahaman *et al.* [71].

The connected component analysis also allowed us to plot the distribution of Alm pores per frame, binned by pore area (Figure 3.17 B). It is clear in Figure 3.17 that, independently of the concentration and orientation of the peptides, as the size of the open pores increases, the number of existent pores decreases. This further corroborates the existence of alamethicin pore structures of different sizes until the peptides ultimately aggregate into the island pore structure.

It would have been interesting to analyze the number of peptides that make up the

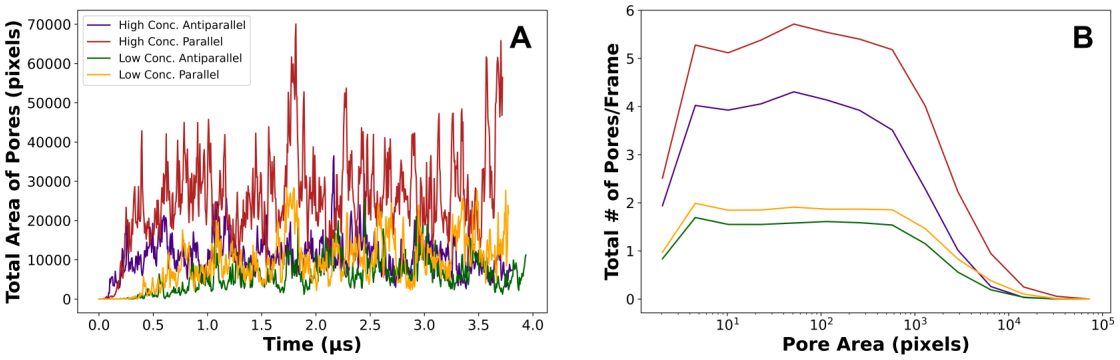

Figure 3.17: **Connected component analysis results of the four alamethicin systems.** (A) is the distribution of the total area of alamethicin pores in pixels over the simulated time. (B) is the distribution of the total number of alamethicin pores per frame by pore area in pixels.

pores' structures throughout the simulated time since it would have made it possible for us to know with certainty that we obtained Alm pore structures formed by different numbers of peptides. We could then relate this to the size of the pore area as well as the conductance states, in order to come to more accurate conclusions. These are some examples that can constitute future analysis work as we study Alm poration further.

3.2.2 Reduced water-contact simulations

Given that in Martini 3 the protein-water interactions had to be scaled up in order to faithfully represent the experimental dimensions of soluble and intrinsically disordered proteins [85], we scaled these interactions to 90% of their original value (as done in previous works to try to combat the peptides' excessive tendency to leave the membrane [163]) and built the same four systems. This would help us understand whether the obtained Alm pore behavior was a product of overestimated water-peptide interactions (and potentially fix the problem with the peptides escaping the membrane [86], discussed in Section 2.5).

With scaled water interactions, the aforementioned problem of peptides leaving the transmembrane state persisted, thus, we had to maintain the same position restraints. Nonetheless, with these restraints, we were able to obtain similar behavior to our previous systems — where we could see the Alm peptides aggregating to form various and different–sized transmembranar structures (Figure 3.18 B).

Regarding quantitative analysis, we started by calculating the number of water-peptide contacts for both the scaled and the regular systems. As we can see in Figure 3.19, the number of contacts for each scaled system was cut to around half, when compared to their regular counterpart. Although a slight decrease is expected, since we are scaling down the interaction value, this huge difference suggests that we no longer have significant water transport in our systems.

As seen in Figure 3.20 A, there is also a significant decrease in the total area of

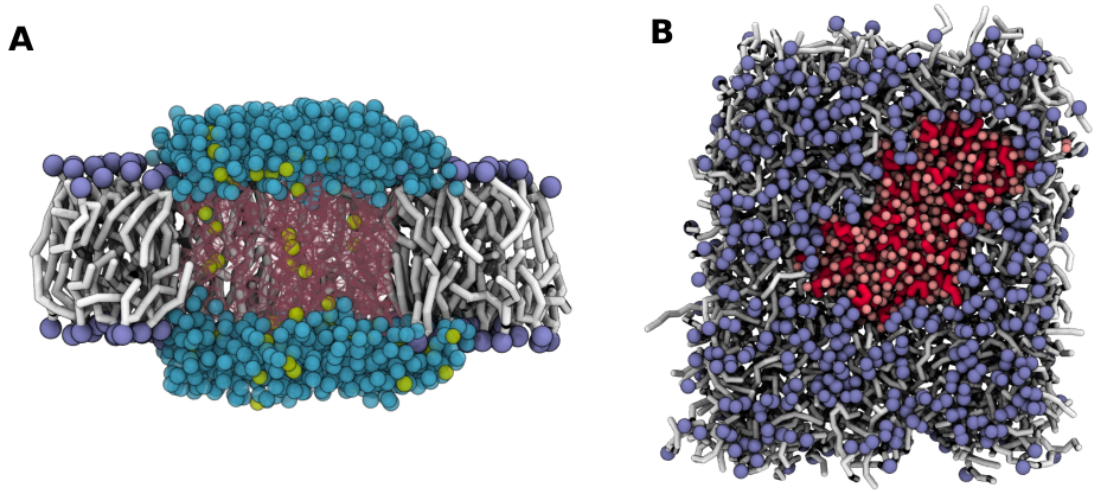

Figure 3.18: **Selected snapshot of two distinct alamethicin pore structures in the high concentration, antiparallel peptide conformation system with water-peptide scaled interactions.** Lipid heads are represented in lilac, lipid tails in white, the alamethicin backbone in red, water molecules in blue, and ions in yellow. (A) Side view of an alamethicin cluster (with peptide backbone represented in see-through). (B) Top view of an alamethicin cluster.

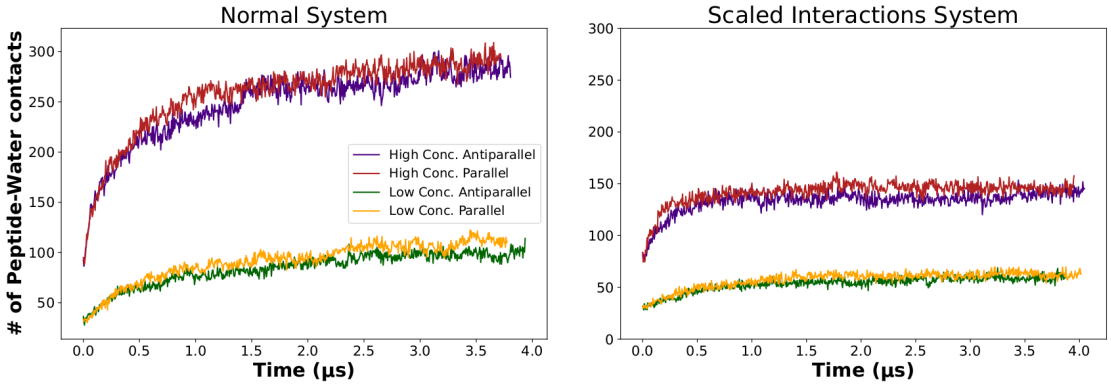

Figure 3.19: **Number of water-peptide contacts throughout the simulated time.** A distance cut-off of 4 Åwas used to represent a contact. On the left is the distribution for the normal systems, and on the right is the distribution for the water-peptide scaled interactions systems.

Alm pores that further shows that the transportation of water through the alamethicin channels, in all systems, was compromised. Additionally, there seems to no longer be a significant difference in conductance between systems, which was noticeable in our regular systems. The size of the Alm pores, as well as the number of aggregates, also appear to have decreased (Figure 3.20 B). Although the peptides still aggregated into large clusters, which can be seen in 3.18 B as well as in Supplementary Video A.2, they did not form the island structure that we previously saw. Given that we are reducing the

interactions between water and peptides, it is to be expected that alamethicin behaves more hydrophobic and, thus, does not aggregate into such a large water-conducting raft structure as seen in the regular systems. Nonetheless, there is also a possibility that, with enough simulation time, the peptides would eventually form the island structure.

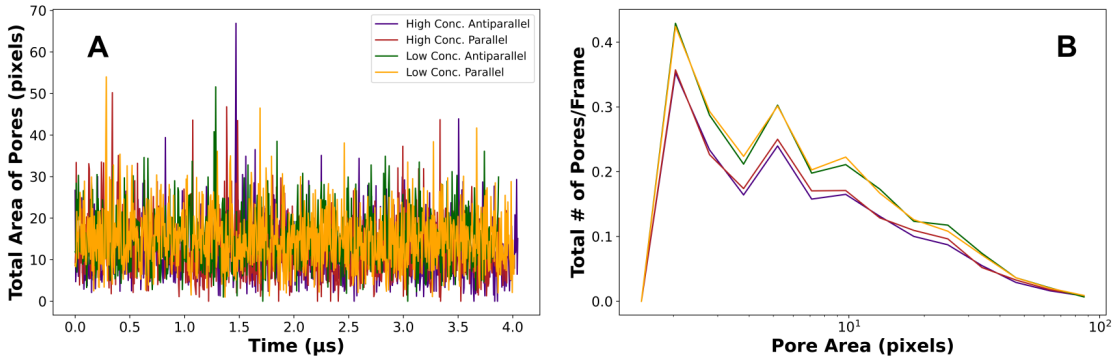

Figure 3.20: **Connected component analysis results of the scaled water-peptide interactions systems.** (A) is the distribution of total area of alamethicin pores in pixels over the simulated time. (B) is the distribution of the total number of alamethicin pores per frame by pore area in pixels.

With these results, it becomes evident that scaling down the water-peptide interactions does not majorly impact the cluster formation property that we saw in the regular systems. This further implies that this is, in fact, a characteristic of Alm pores that may be crucial for its antimicrobial activity. However, the scaling factor led to a loss of pore conductance, which indicates that the adjustment in the Martini 3 force field helped with AMP pore behavior simulation. Thus, bluntly scaling down protein-water interactions is not a viable option for simulating peptide–induced pore formation in Martini 3.

4 Conclusion

Pores play major roles in the cell which makes them crucial for its life and maintenance. An example of this is the part they play in the control of metabolite concentration and, thus, in the generation and dissipation of electrochemical gradients which are essential for many cell processes [2]. In the case of pores formed by antimicrobial peptides, these can even serve as defense mechanisms against fungi and bacteria [46]. For these reasons, the study of the mechanisms as well as the factors that lead to poration are of utmost importance as these can allow us to develop strategies to induce or inhibit pore formation.

In this work, we studied two different types of pore formation in lipid membranes.

Firstly, we studied lipid–lined pore formation in Martini 3 coarse-grained simulations. Here, the objective was to identify the limitations of lipid pore simulations with the Hub *et al*. reaction coordinates in the Martini 3 force field, and subsequently try to improve them. For that, we analyzed several pore lipid properties in order to determine a possible cause for the inaccurate pore behavior in Martini 3 simulations. We were able to identify that pore lipids had a much bigger tail splay when forming pores than their atomistic model counterpart. This led us to build several prototypes of DLPC lipids as a way to test whether pore behavior improved when we restricted the tail movement.

Although we were not able to obtain a free energy landscape similar to the ones of previous AA and CG simulations [91], we saw that our final pore lipid prototype behaved very similarly to the AA model in regard to intramolecular configurations. This could be an indication that this property is not a probable cause for inaccurate pore formation in Martini simulations. However, since our prototype's tail splay was limited to the AA's average value and, consequently, lost its original distribution behavior, we cannot say that with certainty. It is still possible that the tail splay influences pore behavior and the lack of freedom in the tails of our prototypes is not allowing us to obtain a better free energy profile.

As such, as future perspectives, we plan on testing our prototypes with a smaller force on the added distance restraints. Furthermore, we could also further experiment with stricter selections when running the analysis. When we made our pore lipid selection, it is possible that other lipids that were close to the pore but not actually forming it were also selected, which would influence the distributions — consequently leading to biased results and possibly masking existing faults in the lipid models' behavior. These strategies may help us come to a conclusion concerning the tail splay and possibly even pinpoint

other target properties that might contribute to the pore lipid's incorrect behavior.

Nonetheless, it is also possible that the problem does not lie in the parametrization of the Martini 3 lipids but in the way water-lipid interactions are modeled in the force field. The loss of fine detail caused by the 4-1 mapping rule of Martini [85] can be affecting the ability of the water beads to interact with the lipids that form the channels, which would, in turn, influence the transport through the lipid pores and result in inaccurate modeling.

Lastly, we also studied peptide–induced pores — in particular, pores formed by the antimicrobial peptide alamethicin. Regarding Alm pore formation, our main objective was to understand if its pore formation mechanism could now be studied with the improved Martini 3 force field as well as to study the peptide's poration behavior and characteristics. To do this, we assembled various membrane systems with different peptide concentrations and dispositions.

Contrary to what happened in Martini 2 simulations, with the new iteration of the force field, we were able to, for the first time in Martini history, successfully obtain Alm pore formation with ion and water transport. In every system, the Alm peptides formed several clusters until they all aggregated into one large raft — a structure we nicknamed "island" — with several temporary pores.

When we rebuilt the same systems with weakened water-peptide interactions, however, we lost pore conductance. This led us to believe that the new Martini 3 protein-water interactions are crucial to accurately simulate peptide–induced pore behavior. Although in these systems we did not obtain the island structure that we saw previously, the peptides still formed large clusters, as had been seen in previous experimental studies [65]. This suggests that these clusters may be a crucial characteristic of alamethicin's mechanism of action.

AMPs are promising potential alternatives for antibiotics, in an era where drug–resistant bacteria are becoming a more prominent issue. Therefore, further studying alamethicin's antimicrobial activity is of significant relevance. In future work, it would be interesting to quantitatively analyze the number of peptides that formed Alm pores and even their orientation in the antiparallel pores, in order for us to relate this to the pores' conductance states as well as dive into the particular characteristics of the alamethicin pore structure.

In summary, the present work allowed us to disregard as well as identify possible causes of the pitfalls in the Martini 3 lipid pore simulation model. Although we were not able to improve the model entirely, our results show that the lipid tail splay may be a promising target that can lead to a better representation of poration in Martini 3 lipid membranes. Furthermore, our study of alamethicin–induced pore formation not only behaved in agreement with past atomistic and experimental studies — which indicates that Alm pore formation can now be faithfully represented with the new version of the Martini force field — but it also revealed interesting characteristics of the peptide's mechanism that may be of relevance for the development of future therapeutic strategies. Overall, the results of this work revealed new and optimistic avenues to be further explored that can lead to important

improvements of the Martini force field, as well as clear enhancements in version 3 that allowed for the discovery of possible alamethicin pore formation mechanisms.

Bibliography

[1] J. M. Lourenço. *The NOVAthesis LaTeX Template User's Manual*. NOVA University Lisbon. 2021. URL: https://github.com/joaomlourenco/novathesis/raw/master/template.pdf.

[2] D. L. Nelson and M. M. Cox. *Lehninger Principles of Biochemistry, Fourth Edition*. Ed. by Freeman. Fourth Edition. 2004.

[3] E. Gorter and F. Grendel. "On Bimolecular Layers of Lipoids on the Chromocytes of the Blood". In: *Journal of Experimental Medicine* 41.4 (1925), pp. 439–443. DOI: 10.1084/jem.41.4.439.

[4] J. F. Danielli and H. Davson. "A contribution to the theory of permeability of thin films". In: *Journal of Cellular and Comparative Physiology* 5.4 (1935), pp. 495–508. DOI: 10.1002/jcp.1030050409.

[5] S. J. Singer and G. L. Nicolson. "The Fluid Mosaic Model of the Structure of Cell Membranes". en. In: *Science* 175.4023 (1972-02), pp. 720–731. DOI: 10.1126/science.175.4023.720.

[6] L. Johannes et al. "Clustering on Membranes: Fluctuations and More". en. In: *Trends Cell Biol.* 28.5 (2018-05), pp. 405–415. DOI: 10.1016/j.tcb.2018.01.009.

[7] K. Simons and E. Ikonen. "Functional rafts in cell membranes". en. In: *Nature* 387.6633 (1997-06), pp. 569–572. DOI: 10.1038/42408.

[8] V. Shoshan-Barmatz et al. "VDAC1, mitochondrial dysfunction, and Alzheimer's disease". In: *Pharmacological Research* 131 (2018), pp. 87–101. ISSN: 1043-6618. DOI: 10.1016/j.phrs.2018.03.010.

[9] D. M. Rosenbaum, S. G. F. Rasmussen, and B. K. Kobilka. "The structure and function of G-protein-coupled receptors". In: *Nature* 459.7245 (2009-05), pp. 356–363. DOI: https://doi.org/10.1038/nature08144.

[10] G. L. Nicolson. "Update of the 1972 Singer-Nicolson Fluid-Mosaic Model of Membrane Structure". In: *Discoveries* (2013-12). DOI: 10.15190/d.2013.3.

[11] K. G. Suzuki and A. Kusumi. "Refinement of Singer-Nicolson fluid-mosaic model by microscopy imaging: Lipid rafts and actin-induced membrane compartmentalization". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1865.2 (2023), p. 184093. ISSN: 0005-2736. DOI: 10.1016/j.bbamem.2022.184093.

[12] H. Strahl and J. Errington. "Bacterial Membranes: Structure, Domains, and Function". In: *Annual Review of Microbiology* 71.1 (2017), pp. 519–538. DOI: 10.1146/annurev-micro-102215-095630.

[13] E. Agmon and B. R. Stockwell. "Lipid homeostasis and regulated cell death". In: *Current Opinion in Chemical Biology* 39 (2017), pp. 83–89. DOI: 10.1016/j.cbpa.2017.06.002.

[14] P. F. Devaux. "Static and dynamic lipid asymmetry in cell membranes". In: *Biochemistry* 30.5 (1991), pp. 1163–1173. DOI: 10.1021/bi00219a001.

[15] M. Ikeda, A. Kihara, and Y. Igarashi. "Lipid Asymmetry of the Eukaryotic Plasma Membrane: Functions and Related Enzymes". In: *Biological and Pharmaceutical Bulletin* 29.8 (2006-08), pp. 1542–1546. DOI: 10.1248/bpb.29.1542.

[16] D. L. Daleke. "Regulation of transbilayer plasma membrane phospholipid asymmetry". In: *Journal of Lipid Research* 44.2 (2003), pp. 233–242. DOI: 10.1194/jlr.r200019-jlr200.

[17] C. Sohlenkamp and O. Geiger. "Bacterial membrane lipids: diversity in structures and pathways". In: *FEMS Microbiology Reviews* 40.1 (2015), pp. 133–159. DOI: 10.1093/femsre/fuv008.

[18] E. J. Shimshick et al. "Lateral phase separations in membranes". In: *Journal of Supramolecular Structure* 1.4–5 (1973), pp. 285–294. DOI: 10.1002/jss.400010406.

[19] J. J. Kinnun et al. "Lateral heterogeneity and domain formation in cellular membranes". In: *Chemistry and Physics of Lipids* 232 (2020), p. 104976. DOI: 10.1016/j.chemphyslip.2020.104976.

[20] K. Simons and E. Ikonen. "Functional rafts in cell membranes". In: *Nature* 387.6633 (1997), pp. 569–572. DOI: 10.1038/42408.

[21] S. Sonnino and A. Prinetti. "Membrane Domains and the Lipid Raft Concept". In: *Current Medicinal Chemistry* 20.1 (2012), pp. 4–21. DOI: 10.2174/092986731132001 0003.

[22] G. Paratcha and C. F. Ibáñez. "Lipid rafts and the control of neurotrophic factor signaling in the nervous system: variations on a theme". In: *Current Opinion in Neurobiology* 12.5 (2002), pp. 542–549. DOI: 10.1016/s0959-4388(02)00363-x.

[23] G. Nagappan and B. Lu. "Activity-dependent modulation of the BDNF receptor TrKb: Mechanisms and implications". In: *Trends in Neurosciences* 28.9 (2005), pp. 464–471. DOI: 10.1016/j.tins.2005.07.003.

[24] B. A. Tsui-Pierchala et al. "Lipid rafts in neuronal signaling and function". In: *Trends in Neurosciences* 25.8 (2002), pp. 412–417. DOI: 10.1016/s0166-2236(02)02215-4.

[25] A. Prinetti et al. "Sphingolipid-enriched Membrane Domains from Rat Cerebellar Granule Cells Differentiated in Culture". In: *Journal of Biological Chemistry* 275.16 (2000), pp. 11658–11665. DOI: 10.1074/jbc.275.16.11658.

[26] I. Levental, K. R. Levental, and F. A. Heberle. "Lipid Rafts: Controversies Resolved, Mysteries Remain". In: *Trends in Cell Biology* 30.5 (2020), pp. 341–353. DOI: 10.1016/j.tcb.2020.01.009.

[27] D. W. Deamer and J. Bramhall. "Permeability of lipid bilayers to water and ionic solutes". In: *Chemistry and Physics of Lipids* 40.2 (1986). Special Issue: Liposomes, pp. 167–188. ISSN: 0009-3084. DOI: 10.1016/0009-3084(86)90069-1.

[28] K. TAKATA. "Aquaporins: water channel proteins of the cell membrane". In: *Progress in Histochemistry and Cytochemistry* 39.1 (2004), pp. 1–83. DOI: 10.1016/j.proghi.2004.03.001.

[29] G. Menestrina, M. Dalla Serra, and G. Prévost. "Mode of action of β-barrel pore-forming toxins of the staphylococcal α-hemolysin family". In: *Toxicon* 39.11 (2001), pp. 1661–1672. ISSN: 0041-0101. DOI: 10.1016/S0041-0101(01)00153-2.

[30] G. Fuertes et al. "Role of Membrane Lipids for the Activity of Pore Forming Peptides and Proteins". In: *Advances in Experimental Medicine and Biology* (2010), pp. 31–55. DOI: 10.1007/978-1-4419-6327-7_4.

[31] U. Ros and A. J. García-Sáez. "More Than a Pore: The Interplay of Pore-Forming Proteins and Lipid Membranes". In: *The Journal of Membrane Biology* 248.3 (2015), pp. 545–561. DOI: 10.1007/s00232-015-9820-y.

[32] E. Karatekin et al. "Cascades of Transient Pores in Giant Vesicles: Line Tension and Transport". In: *Biophysical Journal* 84.3 (2003), pp. 1734–1749. DOI: 10.1016/s0006-3495(03)74981-9.

[33] S. D. Pritzl et al. "Photolipid Bilayer Permeability is Controlled by Transient Pore Formation". In: *Langmuir* 36.45 (2020), pp. 13509–13515. DOI: 10.1021/acs.langmuir.0c02229.

[34] R. Smith et al. "Structure and Orientation of the Pore-forming Peptide Melittin, in Lipid Bilayers". In: *Journal of Molecular Biology* 241.3 (1994), pp. 456–466. ISSN: 0022-2836. DOI: 10.1006/jmbi.1994.1520.

[35] J. M. Rausch, J. R. Marks, and W. C. Wimley. "Rational combinatorial design of pore-forming β-sheet peptides". In: *Proceedings of the National Academy of Sciences* 102.30 (2005), pp. 10511–10515. DOI: 10.1073/pnas.0502013102.

[36] K. E. Ridyard and J. Overhage. "The Potential of Human Peptide LL-37 as an Antimicrobial and Anti-Biofilm Agent". In: *Antibiotics* 10.6 (2021), p. 650. DOI: 10.3390/antibiotics10060650.

[37] B. A. Johnson, H. Anker, and F. L. Meleney. "Bacitracin: A New Antibiotic Produced by a Member of the *B. subtilis* Group". In: *Science* 102.2650 (1945), pp. 376–377. DOI: 10.1126/science.102.2650.376.

[38] U. Ros, L. Pedrera, and A. J. Garcia-Saez. "Techniques for studying membrane pores". In: *Current Opinion in Structural Biology* 69 (2021), pp. 108–116. ISSN: 0959-440X. DOI: 10.1016/j.sbi.2021.03.012.

[39] D. V. Zhelev and D. Needham. "Tension-stabilized pores in giant vesicles: determination of pore size and pore line tension". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1147.1 (1993), pp. 89–104. DOI: 10.1016/0005-2736(93)90319-u.

[40] S. A. Akimov et al. "Pore formation in lipid membrane I: Continuous reversible trajectory from intact bilayer through hydrophobic defect to transversal pore". In: *Scientific Reports* 7.1 (2017). DOI: 10.1038/s41598-017-12127-7.

[41] S. A. Akimov et al. "Line tension and structure of through pore edge in lipid bilayer". In: *Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology* 8.4 (2014), pp. 297–303. DOI: 10.1134/s1990747814050018.

[42] R. J. Molotkovsky and S. A. Akimov. "Calculation of line tension in various models of lipid bilayer pore edge". In: *Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology* 3.2 (2009), pp. 223–230. DOI: 10.1134/s1990747809020160.

[43] D. P. Tieleman et al. "Simulation of Pore Formation in Lipid Bilayers by Mechanical Stress and Electric Fields". In: *Journal of the American Chemical Society* 125.21 (2003), pp. 6382–6383. DOI: 10.1021/ja029504i.

[44] R. A. Böckmann et al. "Kinetics, Statistics, and Energetics of Lipid Membrane Electroporation Studied by Molecular Dynamics Simulations". In: *Biophysical Journal* 95.4 (2008), pp. 1837–1850. DOI: 10.1529/biophysj.108.129437.

[45] W. F. D. Bennett, N. Sapay, and D. P. Tieleman. "Atomistic Simulations of Pore Formation and Closure in Lipid Bilayers". In: *Biophysical Journal* 106.1 (2014), pp. 210–219. DOI: 10.1016/j.bpj.2013.11.4486.

[46] A. Bahar and D. Ren. "Antimicrobial Peptides". In: *Pharmaceuticals* 6.12 (2013), pp. 1543–1575. DOI: 10.3390/ph6121543.

[47] L.-j. Zhang and R. L. Gallo. "Antimicrobial peptides". In: *Current Biology* 26.1 (2016). DOI: 10.1016/j.cub.2015.11.017.

[48] S. Majd et al. "Applications of biological pores in nanomedicine, sensing, and nanoelectronics". In: *Current Opinion in Biotechnology* 21.4 (2010), pp. 439–476. DOI: 10.1016/j.copbio.2010.05.002.

[49] R. M. Epand and H. J. Vogel. "Diversity of antimicrobial peptides and their mechanisms of action". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1462.1-2 (1999), pp. 11–28. DOI: 10.1016/s0005-2736(99)00198-4.

[50] M. Uteng et al. "Three-Dimensional Structure in Lipid Micelles of the Pediocin-like Antimicrobial Peptide Sakacin P and a Sakacin P Variant That Is Structurally Stabilized by an Inserted C-Terminal Disulfide Bridge". In: *Biochemistry* 42.39 (2003), pp. 11417–11426. DOI: 10.1021/bi034572i.

[51] A. Rozek, C. L. Friedrich, and R. E. Hancock. "Structure of the Bovine Antimicrobial Peptide Indolicidin Bound to Dodecylphosphocholine and Sodium Dodecyl Sulfate Micelles". en. In: *Biochemistry* 39.51 (2000-12), pp. 15765–15774. DOI: 10.1021/bi000714m.

[52] M. Wu et al. "Mechanism of Interaction of Different Classes of Cationic Antimicrobial Peptides with Planar Bilayers and with the Cytoplasmic Membrane of *Escherichia coli*". en. In: *Biochemistry* 38.22 (1999-06), pp. 7235–7242. DOI: 10.1021/bi9826299.

[53] M. N. Melo, R. Ferre, and M. A. R. B. Castanho. "Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations". In: *Nature Reviews Microbiology* 7.3 (2009-03), pp. 245–250. DOI: 10.1038/nrmicro2095.

[54] G. S. Dijksteel et al. "Review: Lessons Learned From Clinical Trials Using Antimicrobial Peptides (AMPs)". In: *Frontiers in Microbiology* 12 (2021). DOI: 10.3389/fmicb.2021.616979.

[55] M. Taniguchi et al. "Pyrrhocoricin, a proline-rich antimicrobial peptide derived from insect, inhibits the translation process in the cell-free *Escherichia coli* protein synthesis system". In: *Journal of Bioscience and Bioengineering* 121.5 (2016), pp. 591–598. DOI: 10.1016/j.jbiosc.2015.09.002.

[56] M. Graf et al. "Proline-rich antimicrobial peptides targeting protein synthesis". In: *Natural Product Reports* 34.7 (2017), pp. 702–711. DOI: 10.1039/c7np00020k.

[57] L. Segev-Zarko et al. "Mechanisms of biofilm inhibition and degradation by antimicrobial peptides". In: *Biochemical Journal* 468.2 (2015), pp. 259–270. DOI: 10.1042/bj20141251.

[58] L. S. Amer, B. M. Bishop, and M. L. van Hoek. "Antimicrobial and antibiofilm activity of cathelicidins and short, synthetic peptides against *Francisella*". In: *Biochemical and Biophysical Research Communications* 396.2 (2010), pp. 246–251. DOI: 10.1016/j.bbrc.2010.04.073.

[59] C. E. Meyer and F. Reusser. "A polypeptide antibacterial agent isolated from *Trichoderma viride*". In: *Experientia* 23.2 (1967), pp. 85–86. DOI: 10.1007/bf02135929.

[60] D. S. Cafiso. "Alamethicin: A Peptide Model for Voltage Gating and Protein-Membrane Interactions". In: *Annual Review of Biophysics and Biomolecular Structure* 23.1 (1994), pp. 141–165. DOI: 10.1146/annurev.bb.23.060194.001041.

[61] *Alamethicin*. URL: https://pubchem.ncbi.nlm.nih.gov/compound/Alamethicin.

[62] K. He et al. "Neutron Scattering in the Plane of Membranes: Structure of Alamethicin Pores". In: *Biophysical Journal* 70.6 (1996), pp. 2659–2666. DOI: 10.1016/s0006-3495(96)79835-1.

[63] S. Qian et al. "Structure of the Alamethicin Pore Reconstructed by X-Ray Diffraction Analysis". In: *Biophysical Journal* 94.9 (2008), pp. 3512–3522. DOI: 10.1529/biophysj.107.126474.

[64] M. Mathew and P. Balaram. "A helix dipole model for alamethicin and related transmembrane channels". In: *FEBS Letters* 157.1 (1983), pp. 1–5. DOI: 10.1016/0014-5793(83)81105-3.

[65] P. Pieta, J. Mirza, and J. Lipkowski. "Direct visualization of the alamethicin pore formed in a planar phospholipid matrix". In: *Proceedings of the National Academy of Sciences* 109.52 (2012), pp. 21223–21227. DOI: 10.1073/pnas.1201559110.

[66] T. Salditt, C. Li, and A. Spaar. "Structure of antimicrobial peptides and lipid membranes probed by interface-sensitive X-ray scattering". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1758.9 (2006), pp. 1483–1498. DOI: 10.1016/j.bbamem.2006.08.002.

[67] L. M. Harriss et al. "Imaging Multiple Conductance States in an Alamethicin Pore". In: *Journal of the American Chemical Society* 133.37 (2011), pp. 14507–14509. DOI: 10.1021/ja204275t.

[68] D. Tieleman, H. Berendsen, and M. Sansom. "Voltage-Dependent Insertion of Alamethicin at Phospholipid/Water and Octane/Water Interfaces". In: *Biophysical Journal* 80.1 (2001), pp. 331–346. DOI: 10.1016/s0006-3495(01)76018-3.

[69] L. Thøgersen et al. "Peptide Aggregation and Pore Formation in a Lipid Bilayer: A Combined Coarse-Grained and All Atom Molecular Dynamics Study". In: *Biophysical Journal* 95.9 (2008), pp. 4337–4347. DOI: 10.1529/biophysj.108.133330.

[70] T. Okazaki et al. "Ion Channels of Alamethicin Dimer N-Terminally Linked by Disulfide Bond". In: *Biophysical Journal* 85.1 (2003), pp. 267–273. DOI: 10.1016/s0006-3495(03)74472-5.

[71] A. Rahaman and T. Lazaridis. "A thermodynamic approach to alamethicin pore formation". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1838.1 (2014), pp. 98–105. DOI: 10.1016/j.bbamem.2013.09.012.

[72] D. Marsh et al. "TOAC Spin Labels in the Backbone of Alamethicin: EPR Studies in Lipid Membranes". In: *Biophysical Journal* 92.2 (2007), pp. 473–481. DOI: 10.1529/biophysj.106.092775.

[73] E. S. Salnikov et al. "Alamethicin Topology in Phospholipid Membranes by Oriented Solid-state NMR and EPR Spectroscopies: a Comparison". In: *The Journal of Physical Chemistry B* 113.10 (2009), pp. 3034–3042. DOI: 10.1021/jp8101805.

[74] S. A. Hollingsworth and R. O. Dror. "Molecular Dynamics Simulation for All". In: *Neuron* 99.6 (2018), pp. 1129–1143. DOI: 10.1016/j.neuron.2018.08.011.

[75] L. Wang et al. "Molecular Dynamics Insight into the CO2 Flooding Mechanism in Wedge-Shaped Pores". In: *Molecules* 28.1 (2022), p. 188. DOI: 10.3390/molecules28010188.

[76] L. S. Adamson et al. "Pore structure controls stability and molecular flux in engineered protein cages". In: *Science Advances* 8.5 (2022). DOI: 10.1126/sciadv.abl7346.

[77] A. Bunker and T. Róg. "Mechanistic Understanding From Molecular Dynamics Simulation in Pharmaceutical Research 1: Drug Delivery". In: *Frontiers in Molecular Biosciences* 7 (2020). DOI: 10.3389/fmolb.2020.604770.

[78] J. A. Stevens et al. "Molecular dynamics simulation of an entire cell". In: *Frontiers in Chemistry* 11 (2023-01). DOI: 10.3389/fchem.2023.1106495.

[79] S. J. Marrink et al. "The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations". In: *The Journal of Physical Chemistry B* 111.27 (2007), pp. 7812–7824. DOI: 10.1021/jp071097f.

[80] W. D. Bennett and D. P. Tieleman. "Water Defect and Pore Formation in Atomistic and Coarse-Grained Lipid Membranes: Pushing the Limits of Coarse Graining". In: *Journal of Chemical Theory and Computation* 7.9 (2011), pp. 2981–2988. DOI: 10.1021/ct200291v.

[81] R. Alessandri et al. "Pitfalls of the Martini Model". In: *Journal of Chemical Theory and Computation* 15.10 (2019), pp. 5448–5460. DOI: 10.1021/acs.jctc.9b00473.

[82] M. Javanainen, H. Martinez-Seara, and I. Vattulainen. "Excessive aggregation of membrane proteins in the Martini model". In: *PLOS ONE* 12.11 (2017). DOI: 10.1371/journal.pone.0187936.

[83] P. S. Schmalhorst et al. "Overcoming the Limitations of the MARTINI Force Field in Simulations of Polysaccharides". In: *Journal of Chemical Theory and Computation* 13.10 (2017), pp. 5039–5053. DOI: 10.1021/acs.jctc.7b00374.

[84] J. J. Uusitalo et al. "Martini Coarse-Grained Force Field: Extension to DNA". In: *Journal of Chemical Theory and Computation* 11.8 (2015), pp. 3932–3945. DOI: 10.1021/acs.jctc.5b00286.

[85] P. C. Souza et al. "Martini 3: a general purpose force field for coarse-grained molecular dynamics". In: *Nature Methods* 18.4 (2021), pp. 382–388. DOI: 10.1038/s41592-021-01098-3.

[86] J. K. Spinti, F. Neiva Nunes, and M. N. Melo. "Room for improvement in the initial martini 3 parameterization of peptide interactions". In: *Chemical Physics Letters* 819 (2023), p. 140436. ISSN: 0009-2614. DOI: 10.1016/j.cplett.2023.140436.

[87] F. E. Thomasen et al. "Improving Martini 3 for Disordered and Multidomain Proteins". In: *Journal of Chemical Theory and Computation* 18.4 (2022), pp. 2033–2041. DOI: 10.1021/acs.jctc.1c01042.

[88] L. Borges-Araújo et al. "Martini 3 Coarse-Grained Force Field for cholesterol". In: *Journal of Chemical Theory and Computation* (2023-in press).

[89] F. Grünewald et al. "Martini 3 Coarse-Grained Force Field for Carbohydrates". In: *Journal of Chemical Theory and Computation* 18.12 (2022), pp. 7555–7569. DOI: 10.1021/acs.jctc.2c00757.

[90] L. Borges-Araújo et al. "Improved Parameterization of Phosphatidylinositide Lipid Headgroups for the Martini 3 Coarse-Grain Force Field". In: *Journal of Chemical Theory and Computation* 18.1 (2021), pp. 357–373. DOI: 10.1021/acs.jctc.1c00615.

[91] J. S. Hub and N. Awasthi. "Probing a Continuous Polar Defect: A Reaction Coordinate for Pore Formation in Lipid Membranes". In: *Journal of Chemical Theory and Computation* 13.5 (2017), pp. 2352–2366. DOI: 10.1021/acs.jctc.7b00106.

[92] M. N. Melo, H. I. Ingólfsson, and S. J. Marrink. "Parameters for Martini sterols and hopanoids based on a virtual-site description". In: *The Journal of Chemical Physics* 143.24 (2015), p. 243152. DOI: 10.1063/1.4937783.

[93] J. M. Crane and L. K. Tamm. "Role of Cholesterol in the Formation and Nature of Lipid Rafts in Planar and Spherical Model Membranes". In: *Biophysical Journal* 86.5 (2004), pp. 2965–2979. DOI: 10.1016/s0006-3495(04)74347-7.

[94] Y. K. Ng et al. "Deep eutectic solvents as green and efficient media for biocatalytic processes". In: *Current Developments in Biotechnology and Bioengineering* (2022), pp. 161–180. DOI: 10.1016/b978-0-323-99905-2.00001-7.

[95] T. M. Parker and C. D. Sherrill. "Assessment of Empirical Models versus High-Accuracy Ab Initio Methods for Nucleobase Stacking: Evaluating the Importance of Charge Penetration". In: *Journal of Chemical Theory and Computation* 11.9 (2015), pp. 4197–4204. DOI: 10.1021/acs.jctc.5b00588.

[96] E. M. Engler, J. D. Andose, and P. V. Schleyer. "Critical evaluation of molecular mechanics". In: *Journal of the American Chemical Society* 95.24 (1973), pp. 8005–8025. DOI: 10.1021/ja00805a012.

[97] L. R. Andrew. *Molecular modelling principles and applications*. 2nd ed. Pearson Education Limited, 2001.

[98] D. M. Hanson et al. *10.1: The Born-Oppenheimer approximation*. 2022-04. URL: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Book%3A_Quantum_States_of_Atoms_and_Molecules_(Zielinksi_et_al)/10%3A_Theories_of_Electronic_Molecular_Structure/10.01%3A_The_Born-Oppenheimer_Approximation.

[99] K. Roy, S. Kar, and R. N. Das. "Computational Chemistry". In: *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment* (2015), pp. 151–189. DOI: 10.1016/b978-0-12-801505-6.00005-3.

[100] M. González. "Force fields and molecular dynamics simulations". In: *École thématique de la Société Française de la Neutronique* 12 (2011), pp. 169–200. DOI: 10.1051/sfn/201112009.

[101] S. Lifson and A. Warshel. "Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules". In: *The Journal of Chemical Physics* 49.11 (1968), pp. 5116–5129. DOI: 10.1063/1.1670007.

[102] D. A. Pearlman et al. "AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules". In: *Computer Physics Communications* 91.1-3 (1995), pp. 1–41. DOI: 10.1016/0010-4655(95)00041-d.

[103] B. R. Brooks et al. "CHARMM: The Biomolecular Simulation Program". In: *Journal of Computational Chemistry* 30.10 (2009), pp. 1545–1614. DOI: 10.1002/jcc.21287.

[104] L. Darré et al. "SIRAH: A Structurally Unbiased Coarse-Grained Force Field for Proteins with Aqueous Solvation and Long-Range Electrostatics". In: *Journal of Chemical Theory and Computation* 11.2 (2015), pp. 723–739. DOI: 10.1021/ct5007746.

[105] M. R. Machado et al. "The SIRAH 2.0 Force Field: Altius, Fortius, Citius". In: *Journal of Chemical Theory and Computation* 15.4 (2019), pp. 2719–2733. DOI: 10.1021/acs.jctc.9b00006.

[106] R. A. Latour. "Perspectives on the simulation of protein–surface interactions using empirical force field methods". In: *Colloids and Surfaces B: Biointerfaces* 124 (2014), pp. 25–37. DOI: 10.1016/j.colsurfb.2014.06.050.

[107] M. J. Abraham et al. "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". In: *SoftwareX* 1-2 (2015), pp. 19–25. ISSN: 2352-7110. DOI: 10.1016/j.softx.2015.06.001.

[108] W. F. Van Gunsteren and H. J. Berendsen. "A Leap-frog Algorithm for Stochastic Dynamics". In: *Molecular Simulation* 1.3 (1988), pp. 173–185. DOI: 10.1080/08927028808080941.

[109] K. B. Wiberg. "A Scheme for Strain Energy Minimization. Application to the Cycloalkanes". In: *Journal of the American Chemical Society* 87.5 (1965), pp. 1070–1078. DOI: 10.1021/ja01083a024.

[110] H. B. Curry. "The method of steepest descent for non-linear minimization problems". In: *Quarterly of Applied Mathematics* 2.3 (1944), pp. 258–261. DOI: 10.1090/qam/10667.

[111] N. Liguori et al. "Molecular dynamics simulations in photosynthesis". In: *Photosynthesis Research* 144.2 (2020), pp. 273–295. DOI: 10.1007/s11120-020-00741-y.

[112] D. J. Evans and B. L. Holian. "The Nosé–Hoover Thermostat". In: *The Journal of Chemical Physics* 83.8 (1985), pp. 4069–4074. DOI: 10.1063/1.449071.

[113] G. Bussi, D. Donadio, and M. Parrinello. "Canonical sampling through velocity rescaling". In: *The Journal of Chemical Physics* 126.1 (2007), p. 014101. DOI: 10.1063/1.2408420.

[114] J. E. Basconi and M. R. Shirts. "Effects of Temperature Control Algorithms on Transport Properties and Kinetics in Molecular Dynamics Simulations". In: *Journal of Chemical Theory and Computation* 9.7 (2013), pp. 2887–2899. DOI: 10.1021/ct400109a.

[115] P. H. Hünenberger. "Thermostat Algorithms for Molecular Dynamics Simulations". In: *Advanced Computer Simulation: Approaches for Soft Matter Sciences I*. Ed. by C. Dr. Holm and K. Prof. Dr. Kremer. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 105–149. ISBN: 978-3-540-31558-2. DOI: 10.1007/b99427.

[116] H. J. Berendsen et al. "Molecular dynamics with coupling to an external bath". In: *The Journal of Chemical Physics* 81.8 (1984), pp. 3684–3690. DOI: 10.1063/1.448118.

[117] M. Parrinello and A. Rahman. "Polymorphic transitions in single crystals: A new molecular dynamics method". In: *Journal of Applied Physics* 52.12 (1981), pp. 7182–7190. DOI: 10.1063/1.328693.

[118] S. Y. Joshi and S. A. Deshmukh. "A review of advancements in coarse-grained molecular dynamics simulations". In: *Molecular Simulation* 47.10-11 (2020), pp. 786–803. DOI: 10.1080/08927022.2020.1828583.

[119] S. Riniker, J. R. Allison, and W. F. van Gunsteren. "On developing coarse-grained models for biomolecular simulation: A Review". In: *Physical Chemistry Chemical Physics* 14.36 (2012), p. 12423. DOI: 10.1039/c2cp40934h.

[120] W. G. Noid. "Perspective: Coarse-grained models for biomolecular systems". In: *The Journal of Chemical Physics* 139.9 (2013). DOI: 10.1063/1.4818908.

[121] J. Jin et al. "Bottom-up Coarse-Graining: Principles and Perspectives". In: *Journal of Chemical Theory and Computation* 18.10 (2022). PMID: 36070494, pp. 5759–5791. DOI: 10.1021/acs.jctc.2c00643.

[122] A. Liwo et al. "A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data". In: *J. Comput. Chem.* 18 (1997), pp. 849–873.

[123] A. Liwo et al. "A united-residue force field for off-lattice protein-structure simulations. II. Parameterization of short-range interactions and determination of weights of energy terms by Z-score optimization". In: *Journal of Computational Chemistry* 18.7 (1997), pp. 874–887. DOI: 10.1002/(SICI)1096-987X(199705)18:7<849::AID-JCC1>3.0.CO;2-R.

[124] A. Liwo et al. "UNRES: a united-residue force field for energy-based prediction of protein structure—orgin and significance of multibody terms". In: *Proceedings of the fourth annual international conference on Computational molecular biology* (2000). DOI: 10.1145/332306.332544.

[125] A. Liwo et al. "A unified coarse-grained model of biological macromolecules based on mean-field multipole–multipole interactions". In: *Journal of Molecular Modeling* 20.8 (2014). DOI: 10.1007/s00894-014-2306-5.

[126] B. West, F. L. Brown, and F. Schmid. "Membrane-Protein Interactions in a Generic Coarse-Grained Model for Lipid Bilayers". In: *Biophysical Journal* 96.1 (2009), pp. 101–115. ISSN: 0006-3495. DOI: 10.1529/biophysj.108.138677.

[127] M. Baaden and S. J. Marrink. "Coarse-grain modelling of protein–protein interactions". In: *Current Opinion in Structural Biology* 23.6 (2013), pp. 878–886. DOI: 10.1016/j.sbi.2013.09.004.

[128] J. Yoo and Q. Cui. "Membrane-Mediated Protein-Protein Interactions and Connection to Elastic Models: A Coarse-Grained Simulation Analysis of Gramicidin A Association". In: *Biophysical Journal* 104.1 (2013), pp. 128–138. DOI: 10.1016/j.bpj.2012.11.3813.

[129] H. I. Ingólfsson et al. "The power of coarse graining in biomolecular simulations". In: *WIREs Computational Molecular Science* 4.3 (2013), pp. 225–248. DOI: 10.1002/wcms.1169.

[130] R. Alessandri, F. Grünewald, and S. J. Marrink. "The Martini Model in Materials Science". In: *Advanced Materials* 33.24 (2021), p. 2008635. DOI: 10.1002/adma.202008635.

[131] C. Arnarez et al. "Dry Martini, a Coarse-Grained Force Field for Lipid Membrane Simulations with Implicit Solvent". In: *Journal of Chemical Theory and Computation* 11.1 (2014), pp. 260–275. DOI: 10.1021/ct500477k.

[132] J. C. Phillips et al. "Scalable molecular dynamics on CPU and GPU architectures with NAMD". In: *The Journal of Chemical Physics* 153.4 (2020), p. 044130. DOI: 10.1063/5.0014475.

[133] A. P. Thompson et al. "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales". In: *Comp. Phys. Comm.* 271 (2022), p. 108171. DOI: 10.1016/j.cpc.2021.108171.

[134] P. Eastman et al. "OpenMM 7: Rapid development of high performance algorithms for Molecular Dynamics". In: *PLOS Computational Biology* 13.7 (2017). DOI: 10.1371/journal.pcbi.1005659.

[135] Q. Liao. "Enhanced sampling and free energy calculations for protein simulations". In: *Computational Approaches for Understanding Dynamical Systems: Protein Folding and Assembly* (2020), pp. 177–213. DOI: 10.1016/bs.pmbts.2020.01.006.

[136] A. Mukherjee and W. D. Sasikala. "Drug–DNA Intercalation: From Discovery to the Molecular Mechanism". In: *Dynamics of Proteins and Nucleic Acids* (2013), pp. 1–62. DOI: 10.1016/b978-0-12-411636-8.00001-8.

[137] G. Saladino, C. Estarellas, and F. Gervasio. "Recent Progress in Free Energy Methods". In: *Comprehensive Medicinal Chemistry III* (2017), pp. 34–50. DOI: 10.1016/b978-0-12-409547-2.12356-x.

[138] F. Musiani and A. Giorgetti. "Protein Aggregation and Molecular Crowding: Perspectives From Multiscale Simulations". In: *International Review of Cell and Molecular Biology* (2017), pp. 49–77. DOI: 10.1016/bs.ircmb.2016.08.009.

[139] R. B. Best and G. Hummer. "Reaction coordinates and rates from transition paths". In: *Proceedings of the National Academy of Sciences* 102.19 (2005), pp. 6732–6737. DOI: 10.1073/pnas.0408098102.

[140] N. Awasthi and J. S. Hub. "Simulations of Pore Formation in Lipid Membranes: Reaction Coordinates, Convergence, Hysteresis, and Finite-Size Effects". In: *Journal of Chemical Theory and Computation* 12.7 (2016), pp. 3261–3269. DOI: 10.1021/acs.jctc.6b00369.

[141] T. P. consortium. "Promoting transparency and reproducibility in enhanced molecular simulations". In: *Nature Methods* 16.8 (2019), pp. 670–673. DOI: 10.1038/s41592-019-0506-8.

[142] M. Bonomi et al. "PLUMED: A portable plugin for free-energy calculations with molecular dynamics". In: *Computer Physics Communications* 180.10 (2009), pp. 1961–1972. DOI: 10.1016/j.cpc.2009.05.011.

[143] J. Huang et al. "CHARMM36M: An improved force field for folded and intrinsically disordered proteins". In: *Nature Methods* 14.1 (2016), pp. 71–73. DOI: 10.1038/nmeth.4067.

[144] S. Jo et al. "CHARMM-GUI: A web-based graphical user interface for CHARMM". In: *Journal of Computational Chemistry* 29.11 (2008), pp. 1859–1865. DOI: 10.1002/jcc.20945.

[145] B. R. Brooks et al. "CHARMM: The siomolecular simulation program". In: *Journal of Computational Chemistry* 30.10 (2009), pp. 1545–1614. DOI: 10.1002/jcc.21287.

[146] J. Lee et al. "CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field". In: *Journal of Chemical Theory and Computation* 12.1 (2015), pp. 405–413. DOI: 10.1021/acs.jctc.5b00935.

[147] E. L. Wu et al. "CHARMM-GUI *Membrane Builder* toward realistic biological membrane simulations". In: *Journal of Computational Chemistry* 35.27 (2014), pp. 1997–2004. DOI: 10.1002/jcc.23702.

[148] S. Jo et al. "CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes". In: *Biophysical Journal* 97.1 (2009), pp. 50–58. DOI: 10.1016/j.bpj.2009.04.013.

[149] J. Lee et al. "CHARMM-GUI *Membrane Builder* for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans". In: *Journal of Chemical Theory and Computation* 15.1 (2018), pp. 775–786. DOI: 10.1021/acs.jctc.8b01066.

[150] T. Darden, D. York, and L. Pedersen. "Particle mesh Ewald: An $N \cdot \log(n)$ method for Ewald sums in large systems". In: *The Journal of Chemical Physics* 98.12 (1993), pp. 10089–10092. DOI: 10.1063/1.464397.

[151] T. A. Wassenaar et al. "Computational Lipidomics with *insane*: A Versatile Tool for Generating Custom Membranes for Molecular Simulations". In: *Journal of Chemical Theory and Computation* 11.5 (2015), pp. 2144–2155. DOI: 10.1021/acs.jctc.5b00209.

[152] D. H. de Jong et al. "Martini straight: Boosting performance using a shorter cutoff and GPUs". In: *Computer Physics Communications* 199 (2016), pp. 1–7. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.09.014.

[153] N. Michaud-Agrawal et al. "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations". In: *Journal of Computational Chemistry* 32.10 (2011), pp. 2319–2327. DOI: 10.1002/jcc.21787.

[154] C. R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

[155] J. D. Hunter. "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/mcse.2007.55.

[156] M. L. Waskom. "seaborn: statistical data visualization". In: *Journal of Open Source Software* 6.60 (2021), p. 3021. DOI: 10.21105/joss.03021.

[157] W. Humphrey, A. Dalke, and K. Schulten. "VMD: Visual molecular dynamics". In: *Journal of Molecular Graphics* 14.1 (1996), pp. 33–38. DOI: 10.1016/0263-7855(96)00018-5.

[158] A. Grossfield. *WHAM: the weighted histogram analysis method*. URL: http://membrane.urmc.rochester.edu/wordpress/?page_id=126.

[159] K. Mustafin, P. Buslaev, and I. Gushchin. *PCALipids: Scripts for PCA and related analyses of lipid motions*. URL: https://github.com/membrane-systems/PCAlipids.

[160] S. Kumar et al. "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method". In: *Journal of Computational Chemistry* 13.8 (1992), pp. 1011–1021. DOI: 10.1002/jcc.540130812.

[161] Schrödinger, LLC. "The PyMOL Molecular Graphics System, Version 1.8". 2015-11.

[162] Marrink-Lab. *Marrink-Lab/Vermouth-Martinize: Describe and apply transformation on molecular structures and Topologies*. URL: https://github.com/marrink-lab/vermouth-martinize.

[163] A. C. Cabezudo et al. "Scaling protein-water interactions in the Martini 3 coarse-grained force field to simulate transmembrane helix dimers in different lipid environments". In: *bioRxiv* (2022). DOI: 10.1101/2022.09.09.506752.

[164] S. van der Walt et al. "scikit-image: image processing in Python". In: *PeerJ* 2 (2014). DOI: 10.7717/peerj.453.

[165] R. Alessandri et al. "Martini 3 Coarse-Grained Force Field: Small Molecules". In: *Advanced Theory and Simulations* 5.1 (2021). DOI: 10.1002/adts.202100391.

[166] S. J. Marrink and D. P. Tieleman. "Perspective on the Martini model". In: *Chemical Society Reviews* 42.16 (2013), p. 6801. DOI: 10.1039/c3cs60093a.

[167] W. You, Z. Tang, and C.-e. A. Chang. "Potential Mean Force from Umbrella Sampling Simulations: What Can We Learn and What Is Missed?" In: *Journal of Chemical Theory and Computation* 15.4 (2019), pp. 2433–2443. DOI: 10.1021/acs.jctc.8b01142.

[168] C. L. Ting et al. "Metastable Prepores in Tension-Free Lipid Bilayers". In: *Physical Review Letters* 120.12 (2018). DOI: 10.1103/physrevlett.120.128103.

[169] D. Marsh. "Section II: PHOSPHOLIPIDS". In: *Handbook of Lipid Bilayers*. CRC Press, 1990, p. 380.

[170] B. P. Weiser et al. "Computational Investigation of Cholesterol Binding Sites on Mitochondrial VDAC". In: *The Journal of Physical Chemistry B* 118.33 (2014), pp. 9852–9860. DOI: 10.1021/jp504516a.

[171] S. Dadsena et al. "Ceramides bind VDAC2 to trigger mitochondrial apoptosis". In: *Nature Communications* 10.1 (2019). DOI: 10.1038/s41467-019-09654-4.

[172] G. Hedger et al. "Cholesterol Interaction Sites on the Transmembrane Domain of the Hedgehog Signal Transducer and Class F G Protein-Coupled Receptor Smoothened". In: *Structure* 27.3 (2019). DOI: 10.1016/j.str.2018.11.003.

[173] T. Naito and Y. Saheki. "GRAMD1-mediated accessible cholesterol sensing and transport". In: *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 1866.8 (2021), p. 158957. DOI: 10.1016/j.bbalip.2021.158957.

[174] F. Fornasier et al. "Lipophilicity of Coarse-Grained Cholesterol Models". In: *Journal of Chemical Information and Modeling* 60.2 (2020), pp. 569–577. DOI: 10.1021/acs.jcim.9b00830.

[175] S. Thallmair et al. "Nonconverged Constraints Cause Artificial Temperature Gradients in Lipid Bilayer Simulations". In: *The Journal of Physical Chemistry B* 125.33 (2021), pp. 9537–9546. DOI: 10.1021/acs.jpcb.1c03665.

A Appendix

A.1 Supplementary Tables

Table A.1: **Root mean square error values of the in-development Martini 3 DPPC bead distances.** The forward-mapped atomistic distances were used as the target value.

	Pore lipids (DPPC)	Membrane Lipids (DPPC)
Martini 3 dev15	1.001	0.757
Martini 3 dev18	1.086	0.508

Table A.2: **Pore lipids' variance values explained by the first four principal components.**

	Atomistic DLPC	Martini 3 dev18 DLPC	Mixed Membrane DLPC
PC1	0.01365	0.01941	0.00027
PC2	0.00813	0.00789	0.00671
PC3	0.00654	0.00694	0.00448
PC4	0.00423	0.00385	0.00540

Table A.3: **Membrane lipids' variance values explained by the first four principal components.**

	Atomistic DLPC	Martini 3 dev18 DLPC	Mixed Membrane DLPC
PC1	0.01219	0.01327	0.01219
PC2	0.00753	0.00733	0.00764
PC3	0.00643	0.00721	0.00699
PC4	0.00421	0.00393	0.00387

A.2 Supplementary Videos

Supplementary Videos A.1 and A.2 can be found in the Supplementary Material for this work.

The heads of the lipids are represented in blue, the peptide backbones in red, and the background is in black. In both videos, we are able to clearly see the opening and closure of the pores as well as the aggregation of alamethicin peptides into clusters. In the case of the system with regular water-peptide interactions, we can also see the formation of the island pore structure.

A.3 Supplementary Figures

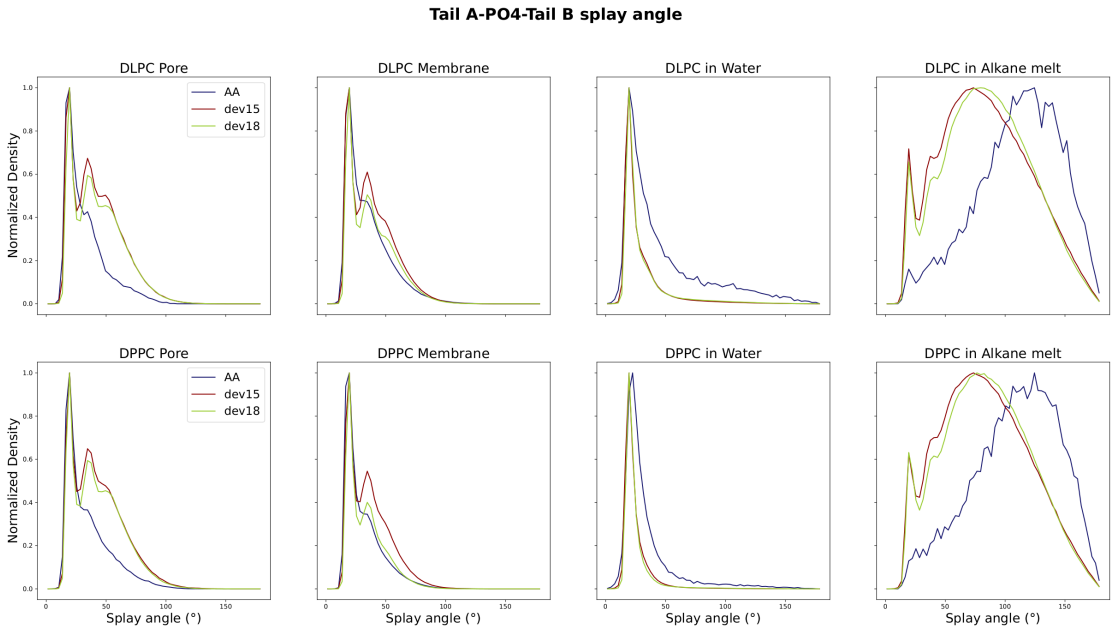

Figure A.1: **Distribution of the angle between beads C3A, PO4, and C3B of the lipids.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

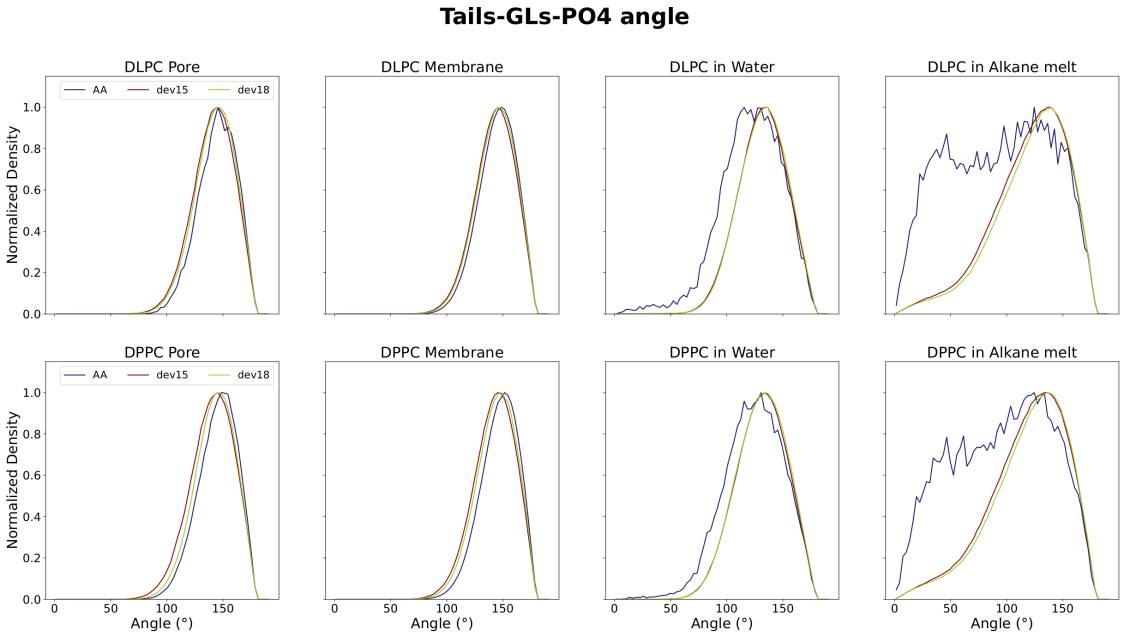

Figure A.2: **Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

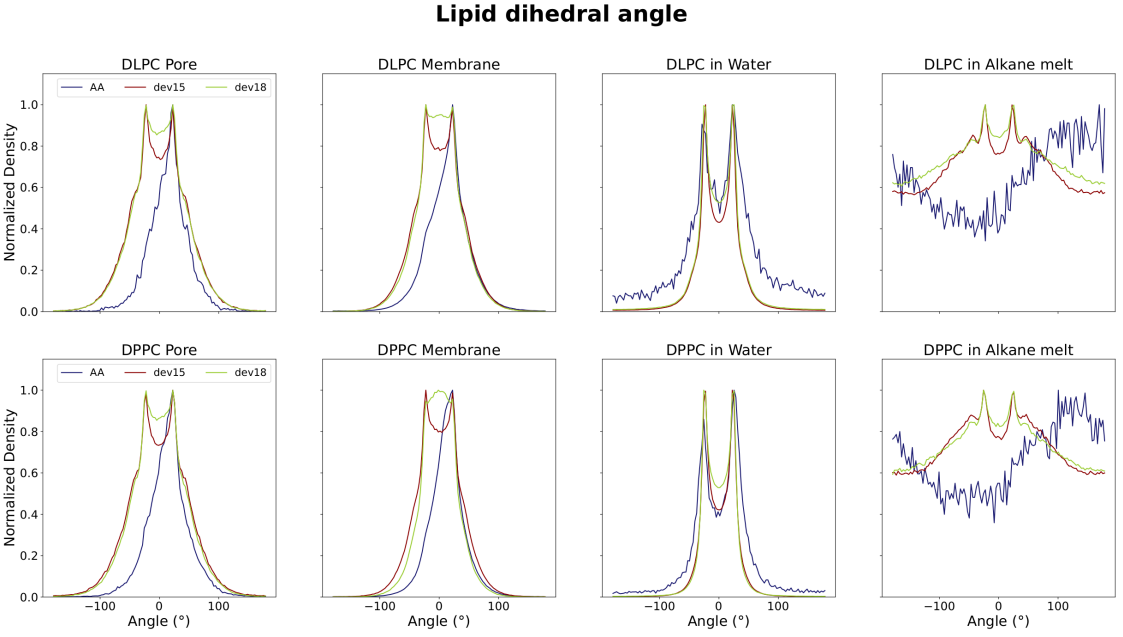

Figure A.3: **Distribution of the dihedral angle of the lipids.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

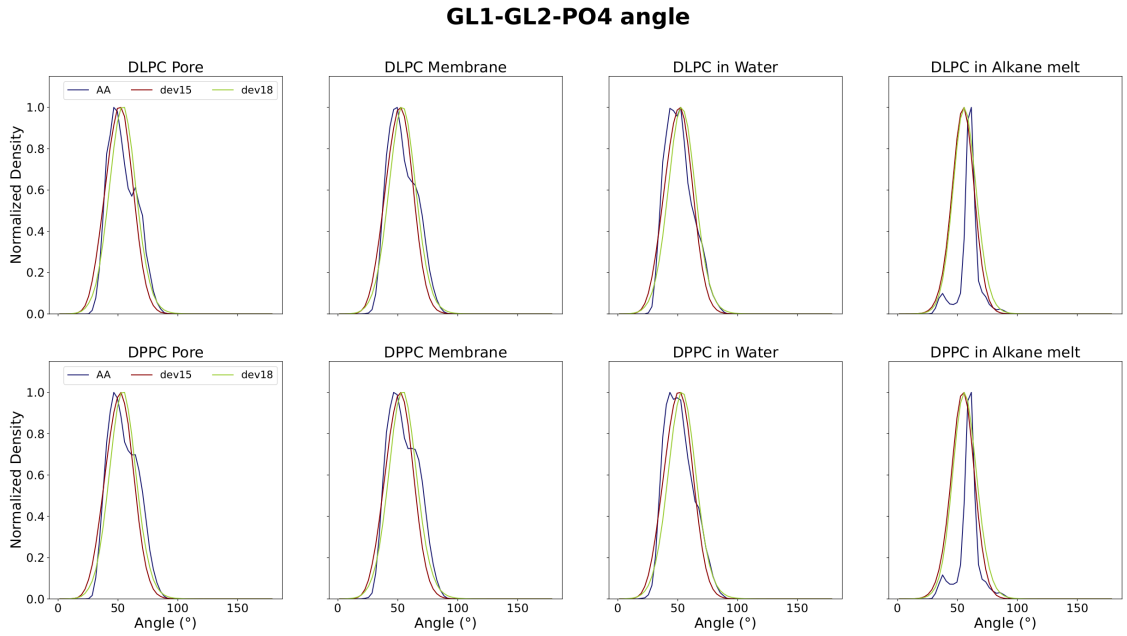

Figure A.4: **Distribution of the angle formed between the glycerols and the PO4 beads of the lipids.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

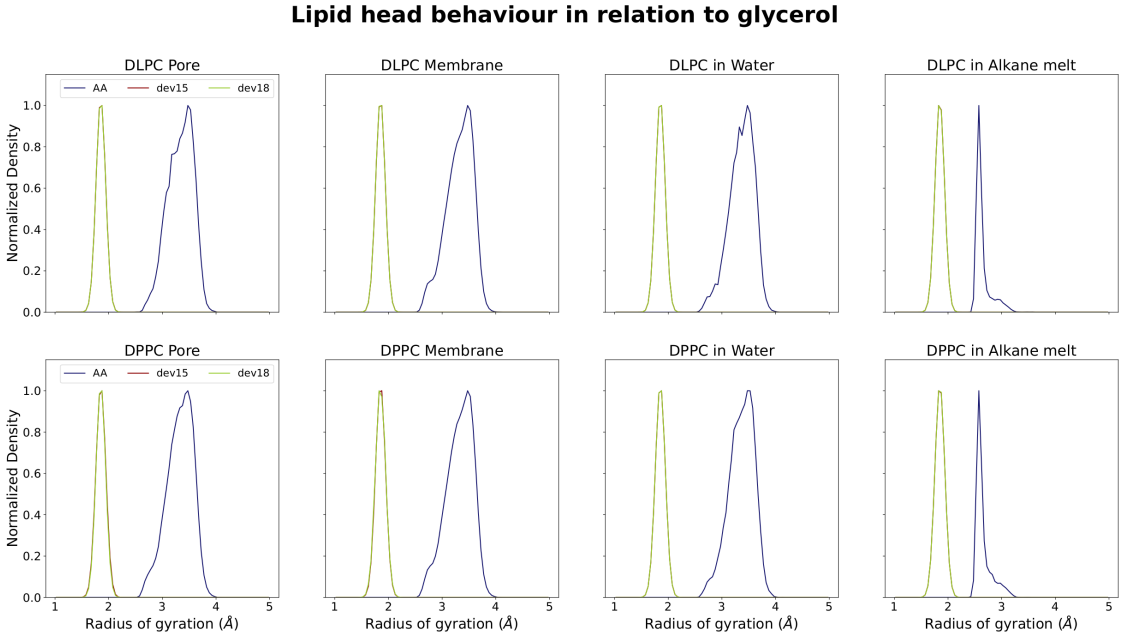

Figure A.5: **Distribution of the radius of gyration of the lipids' head.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

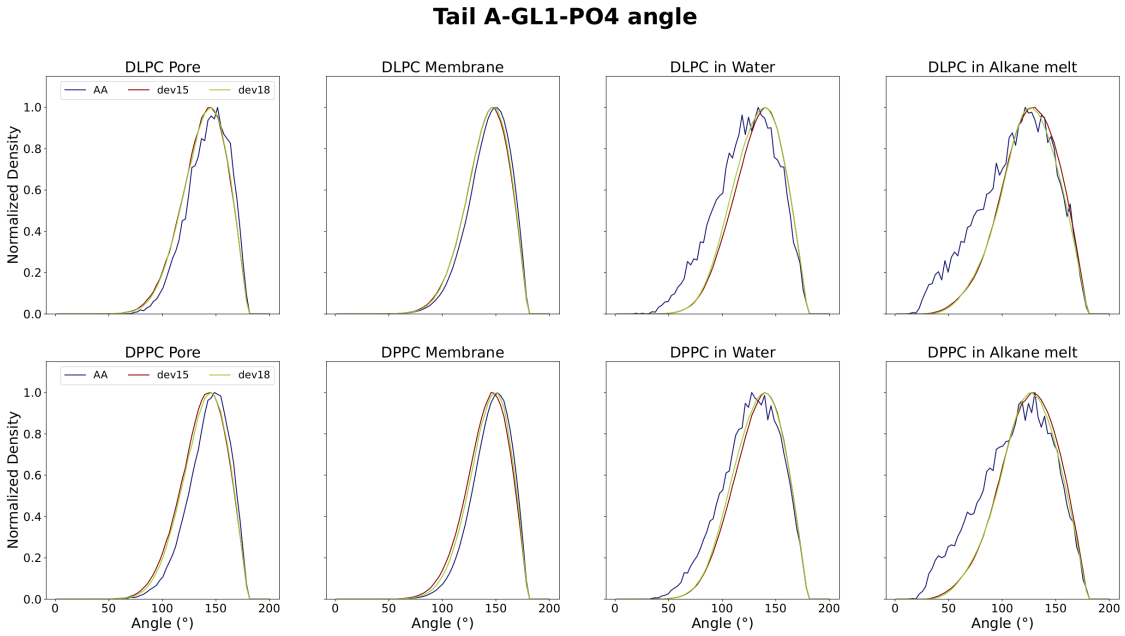

Figure A.6: **Distribution of the angle formed between tail A, GL1 and PO4.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

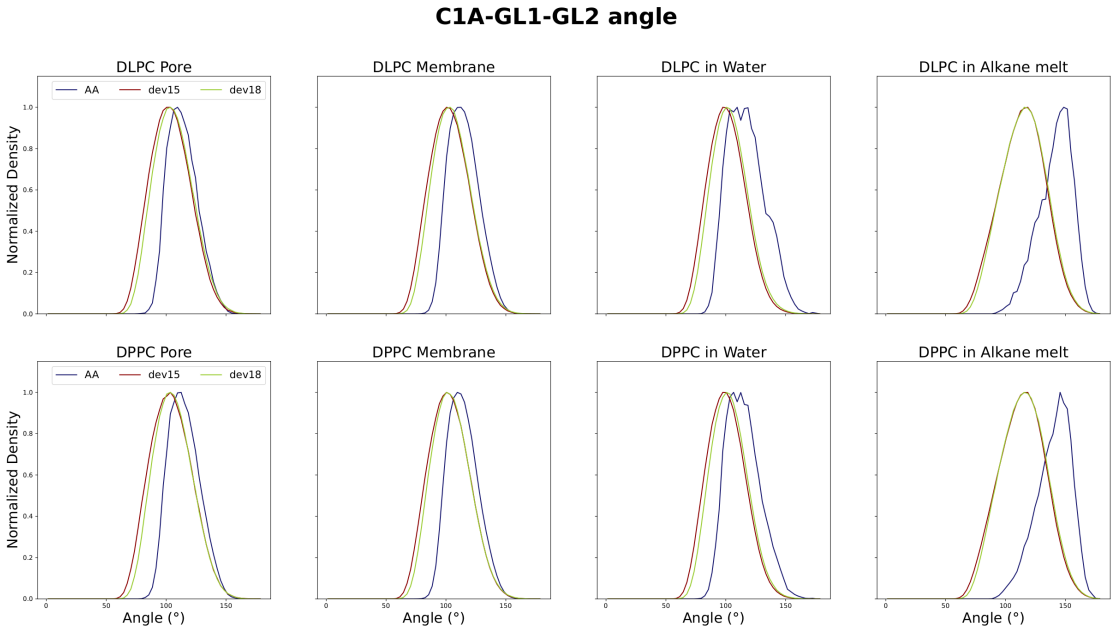

Figure A.7: **Distribution of theangle between beads C1A, GL1, and GL2 of the lipids.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

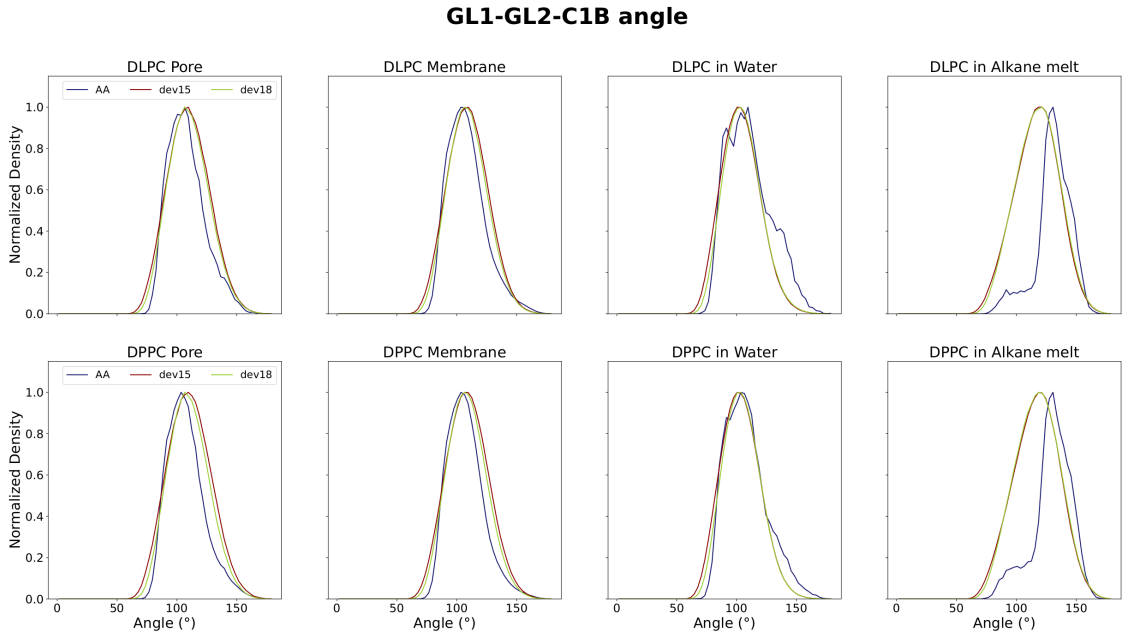

Figure A.8: **Distribution of the angle between beads GL1, GL2, and C1B of the lipids** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

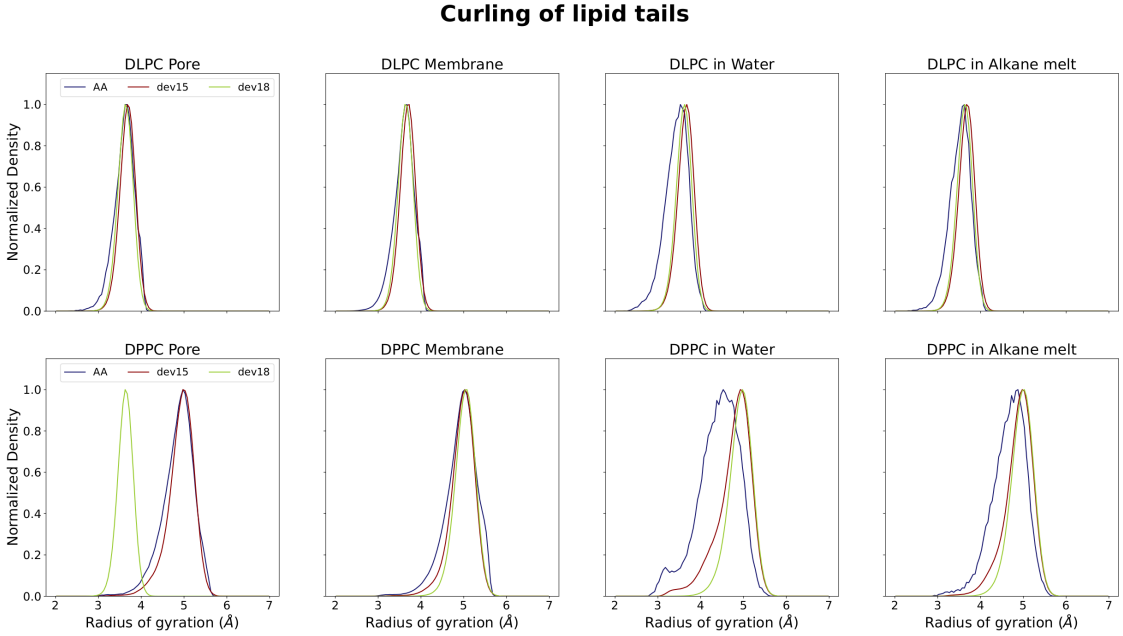

Figure A.9: **Distribution of the radius of gyration of the lipid's tails.** The atomistic distributions are in dark blue and distributions of versions 15 and 18 of the in-development Martini 3 lipids are in red and green, respectively.

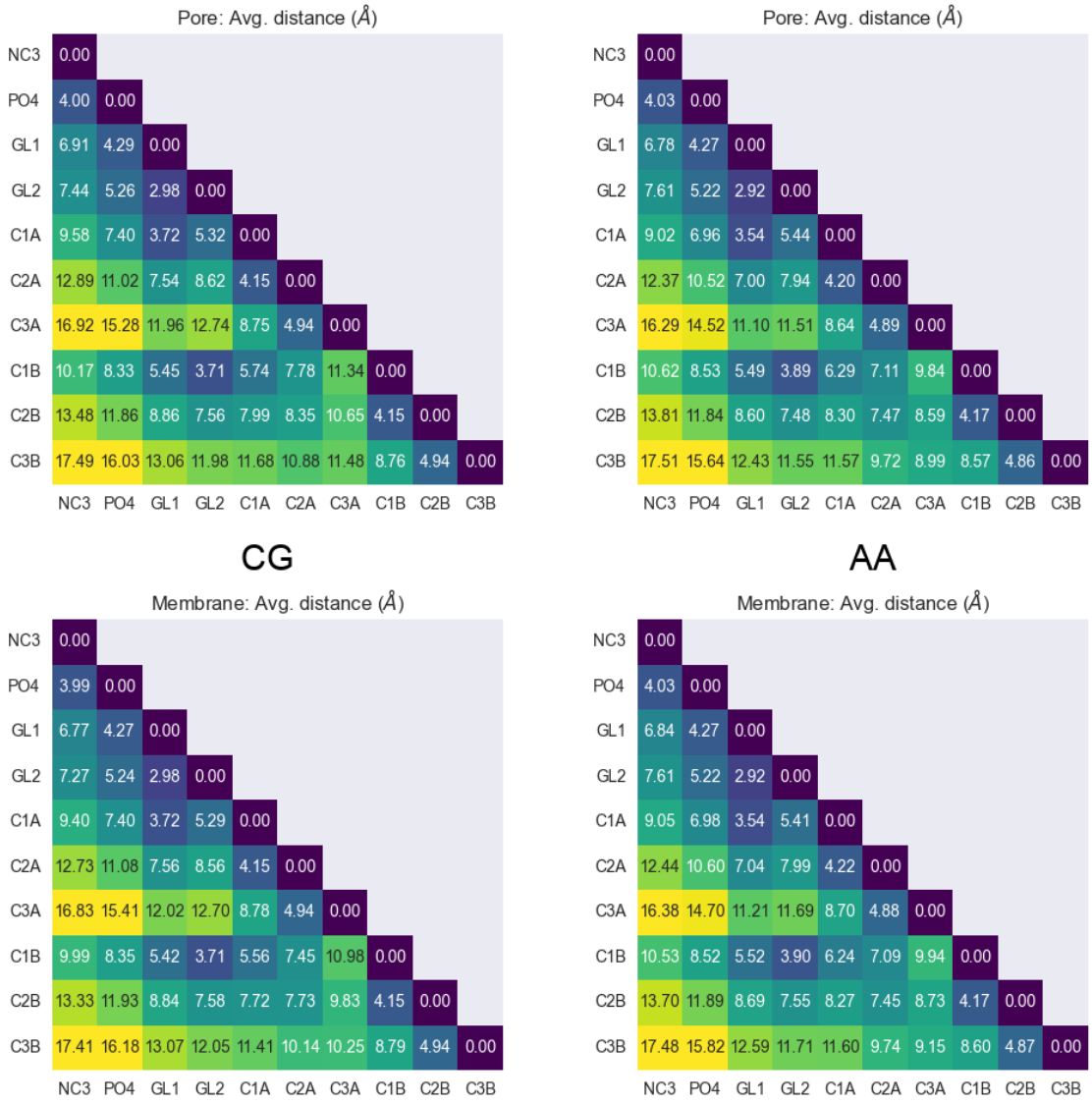

Figure A.10: **Distance matrices of the DLPC membrane and pore lipid's beads.** The matrices on the left side correspond to version 18 of the in-development Martini 3 lipids, and the matrices on the right side correspond to the forward-mapped atomistic lipids. The average distances are in Angstrom (Å).

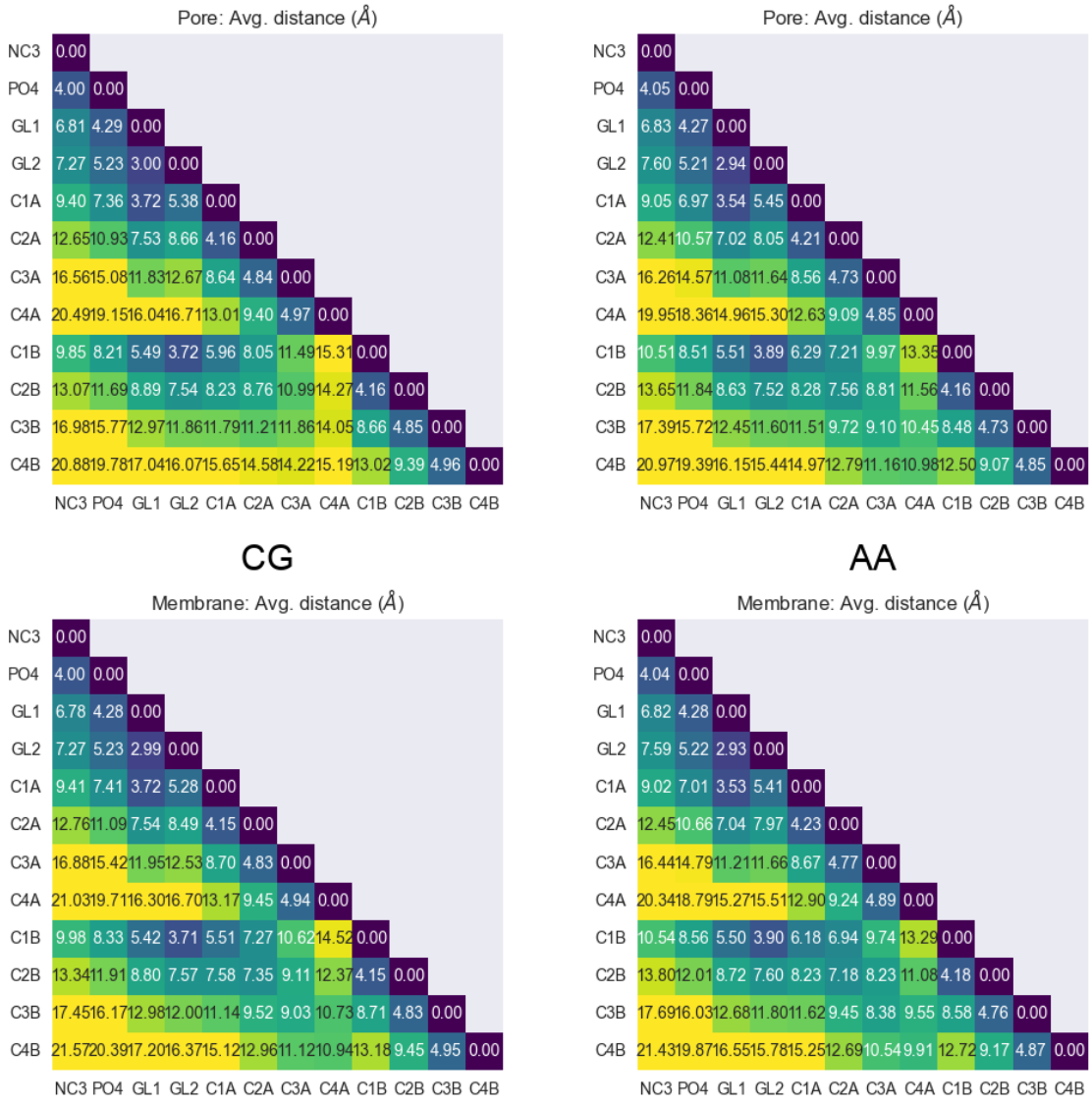

Figure A.11: **Distance matrices of the DPPC membrane and pore lipid's beads.** The matrices on the left side correspond to version 18 of the in-development Martini 3 lipids, and the matrices on the right side correspond to the forward-mapped atomistic lipids. The average distances are in Angstrom (Å).

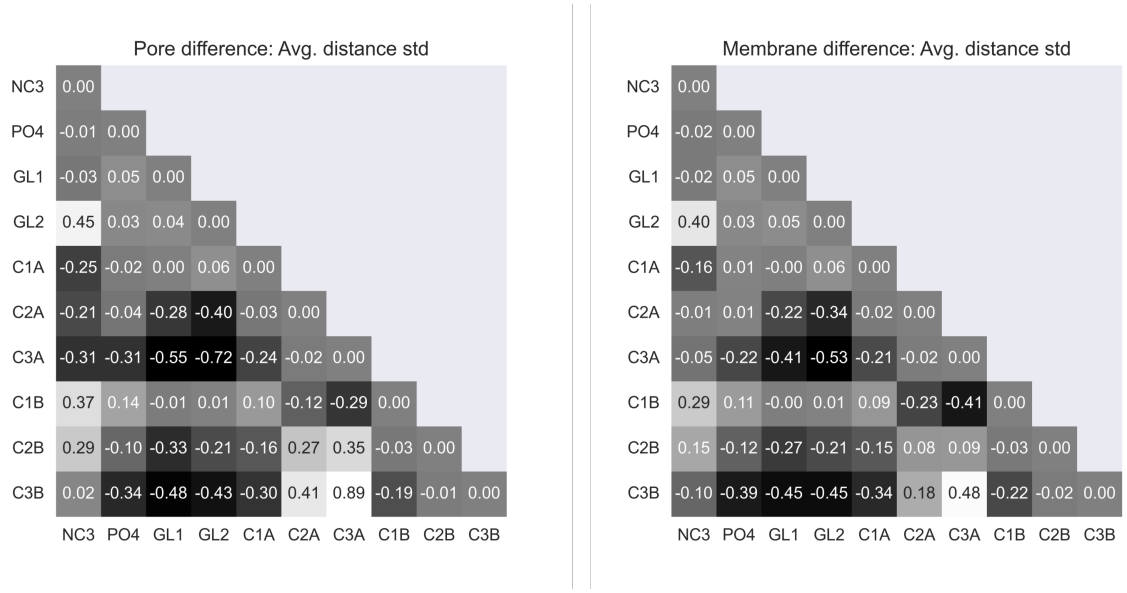

Figure A.12: **Standard deviation matrices of the forward-mapped atomistic DLPC bead distances subtracted from the version 18 Martini 3 DLPC bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

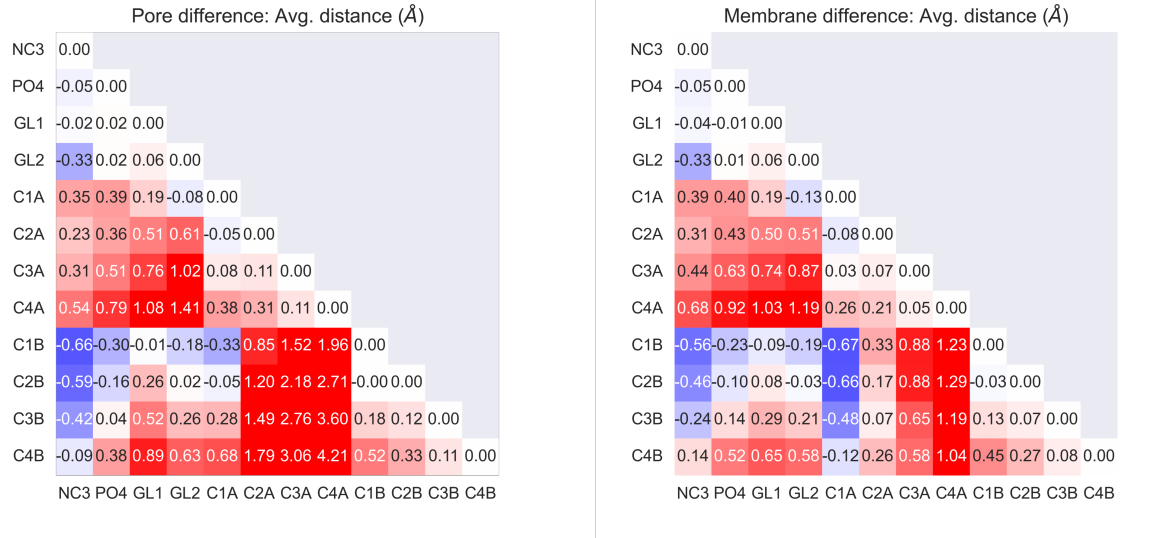

Figure A.13: **Distance matrices of the forward-mapped atomistic DPPC bead distances subtracted from the version 18 Martini 3 DPPC bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

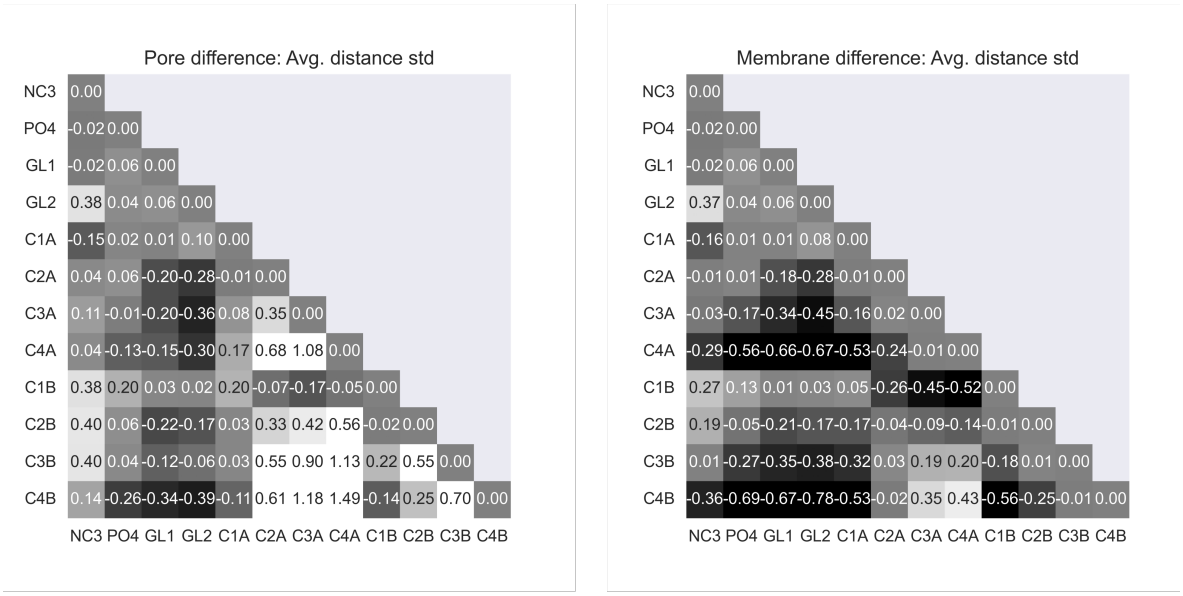

Figure A.14: **Standard deviation matrices of the forward-mapped atomistic DPPC bead distances subtracted from the version 18 Martini 3 DPPC bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

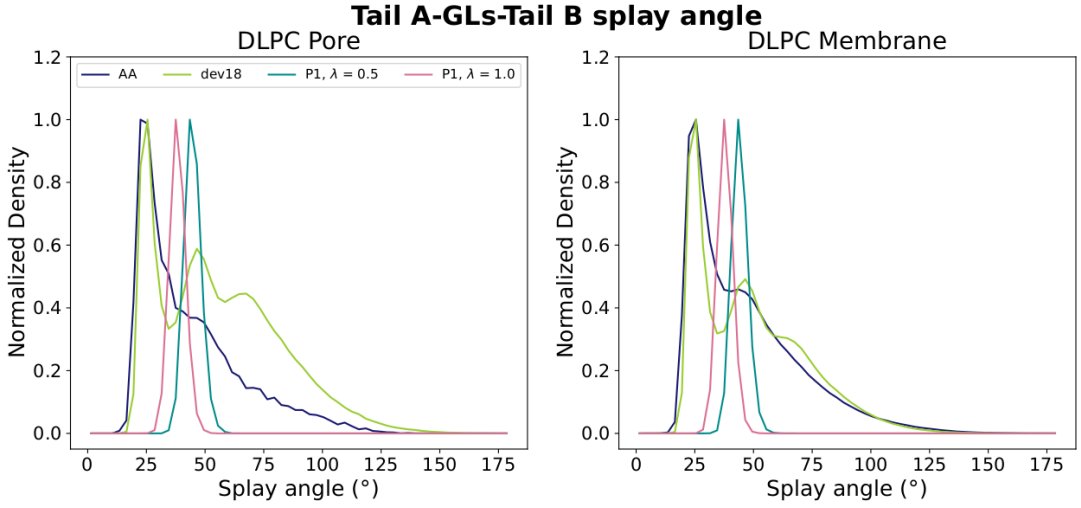

Figure A.15: **Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

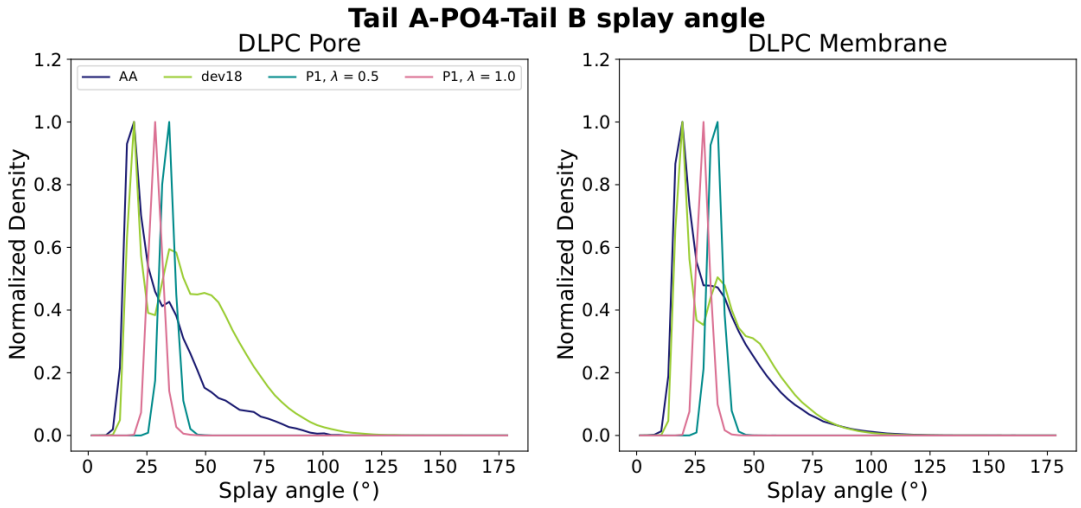

Figure A.16: **Distribution of the angle between beads C3A, PO4, and C3B of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

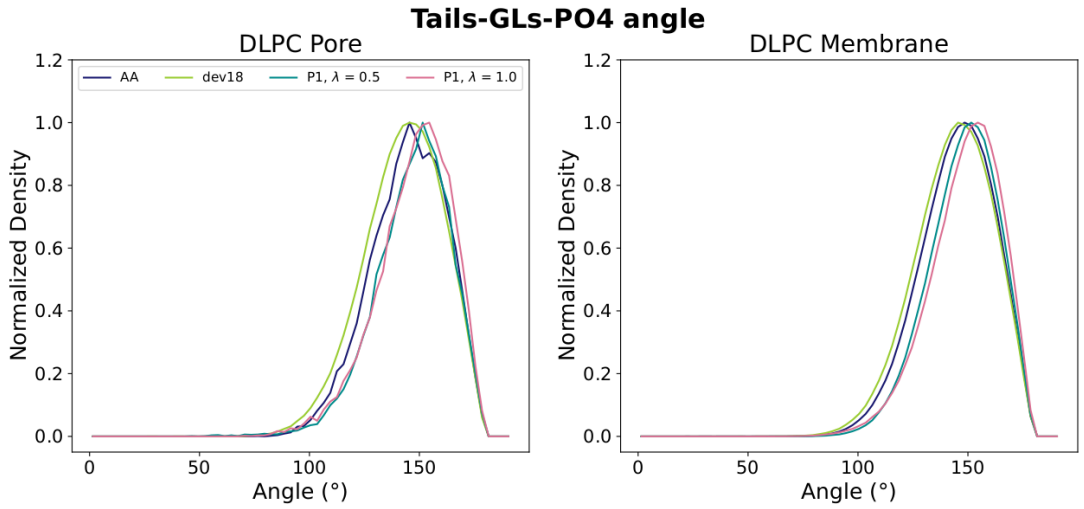

Figure A.17: **Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

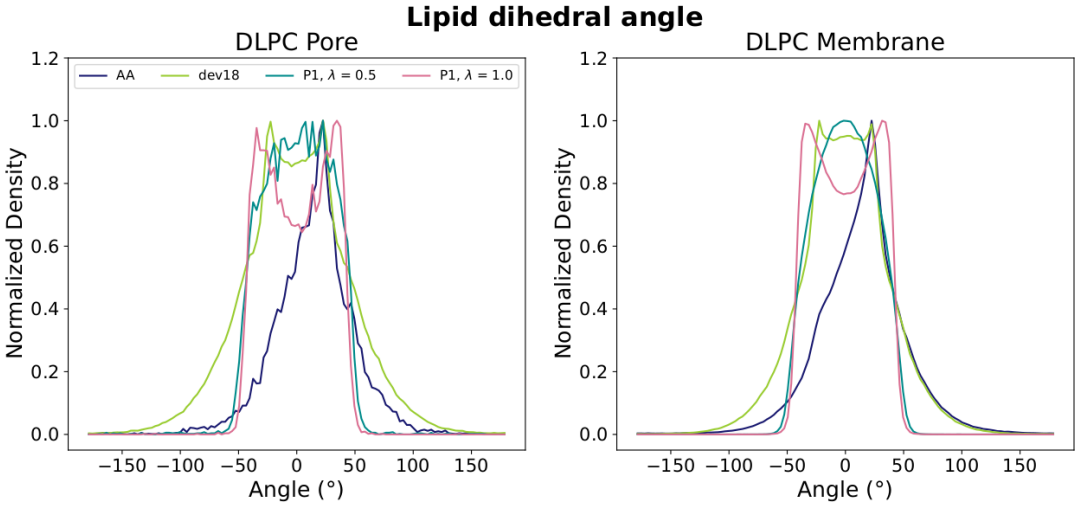

Figure A.18: **Distribution of the dihedral angle of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

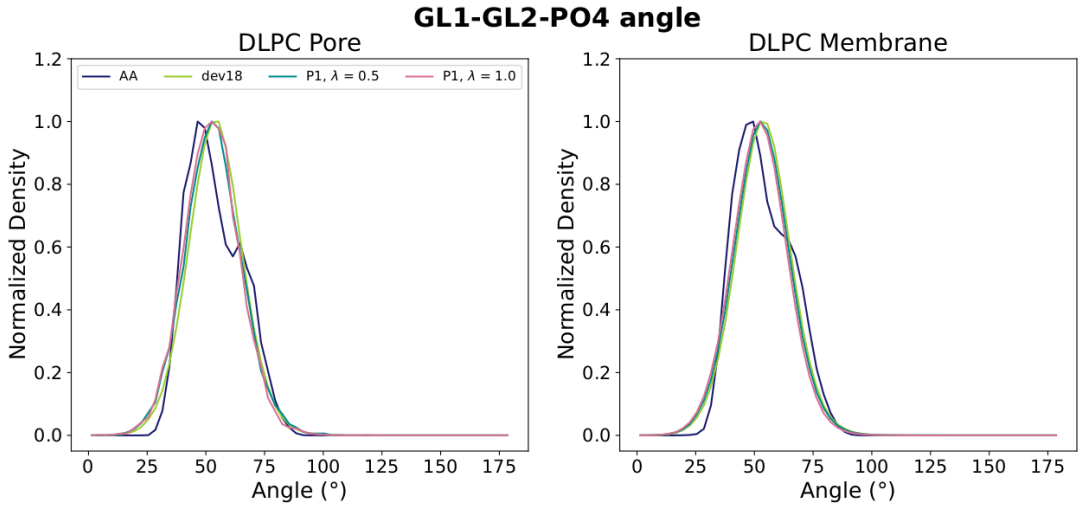

Figure A.19: **Distribution of the angle formed between the glycerols and the PO4 beads of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

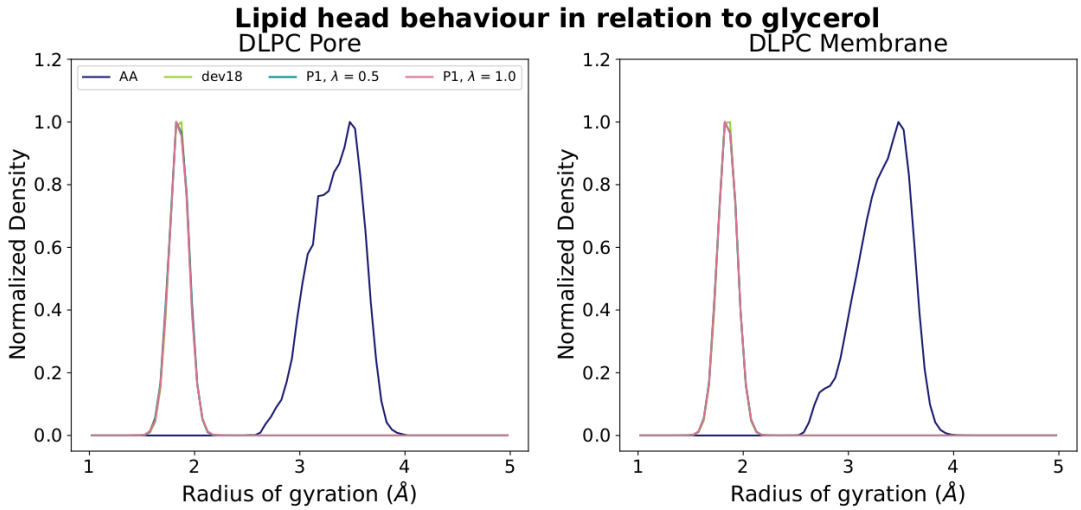

Figure A.20: **Distribution of the radius of gyration of the lipids' head.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

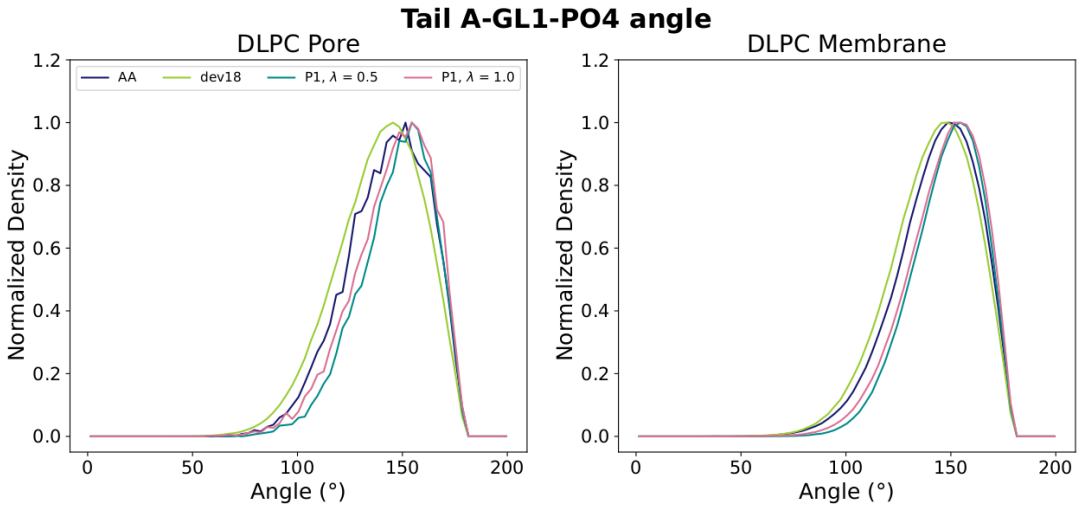

Figure A.21: **Distribution of the angle formed between tail A, GL1 and PO4.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

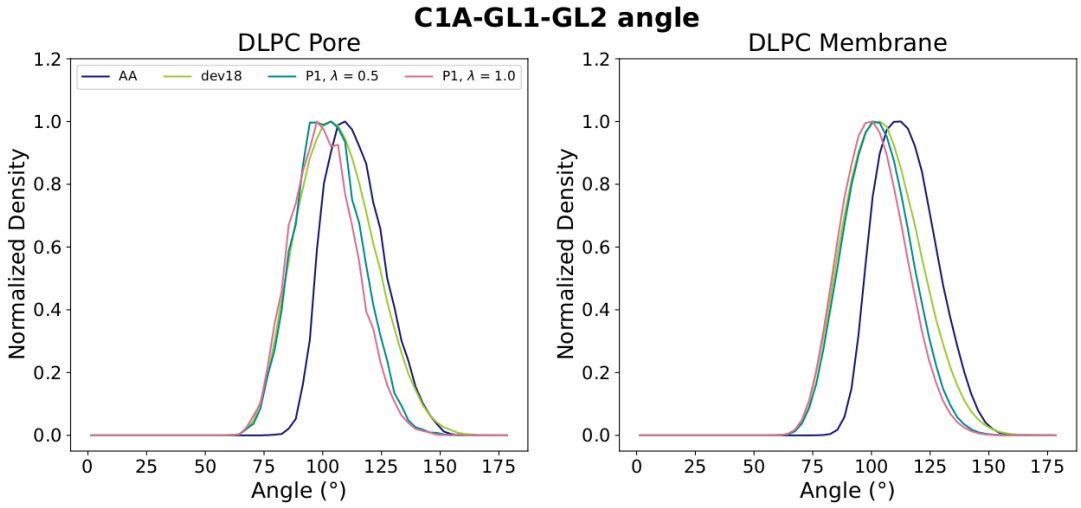

Figure A.22: **Distribution of theangle between beads C1A, GL1, and GL2 of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

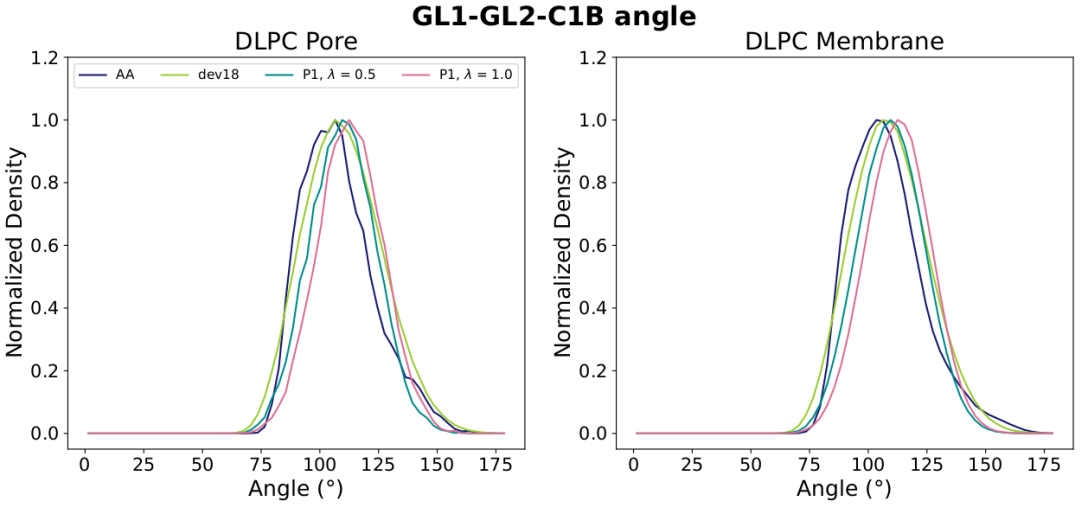

Figure A.23: **Distribution of the angle between beads GL1, GL2, and C1B of the lipids.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

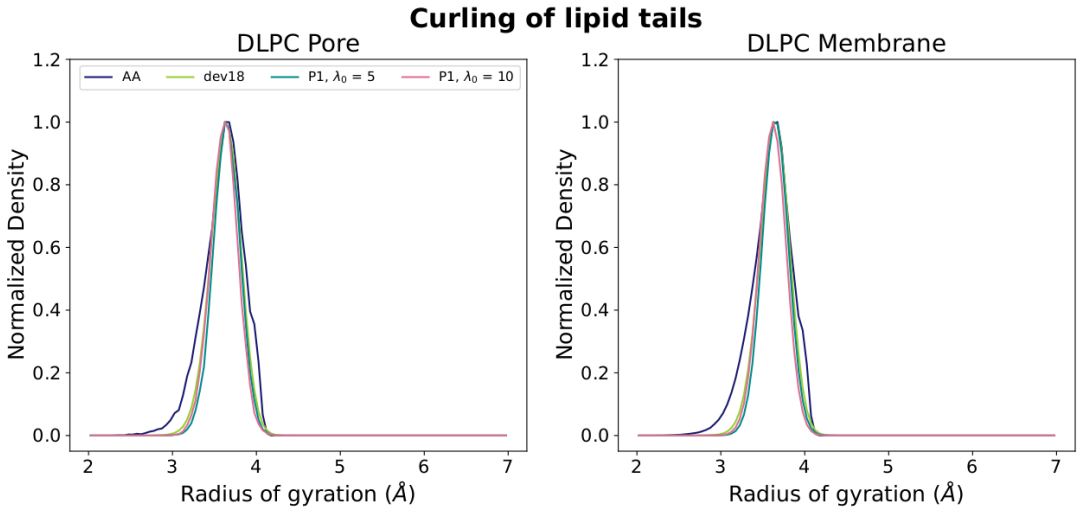

Figure A.24: **Distribution of the radius of gyration of the lipid's tails.** The atomistic distributions are in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, and the distributions of our lipid prototype at lambda equal to 0.5 and 1.0 are in light blue and light pink, respectively.

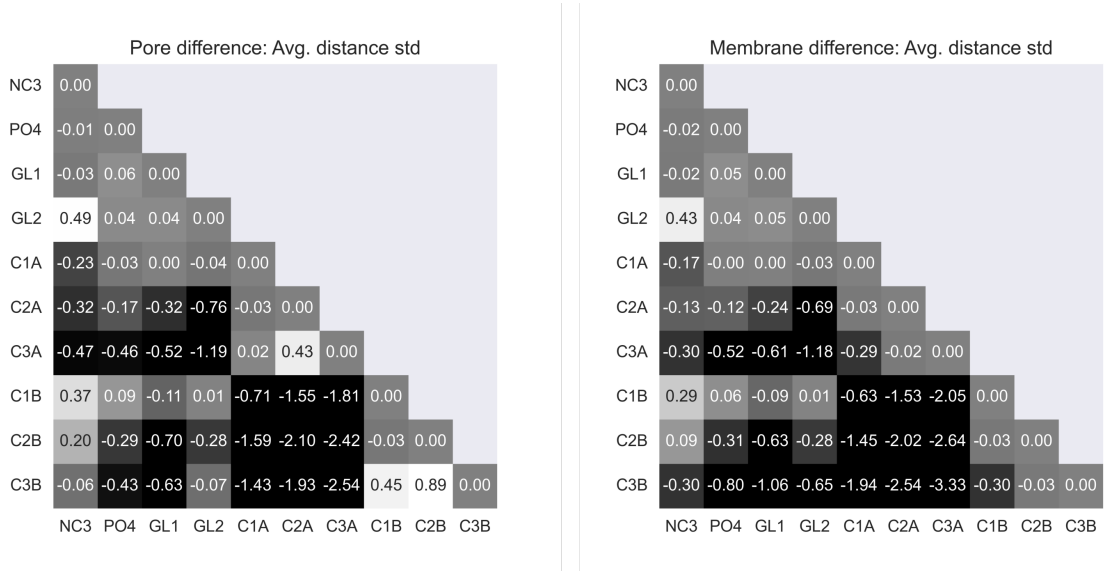

Figure A.25: **Standard deviation of the forward-mapped atomistic lipid's bead distances subtracted from the first lipid prototype lipid's bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

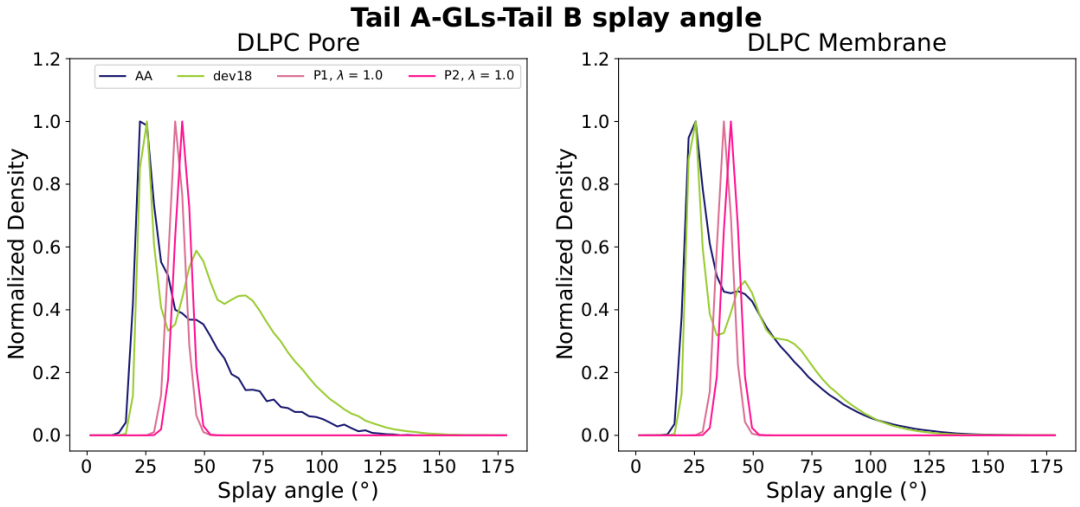

Figure A.26: **Distribution of the angle between Tail A, the glycerols, and Tail B of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

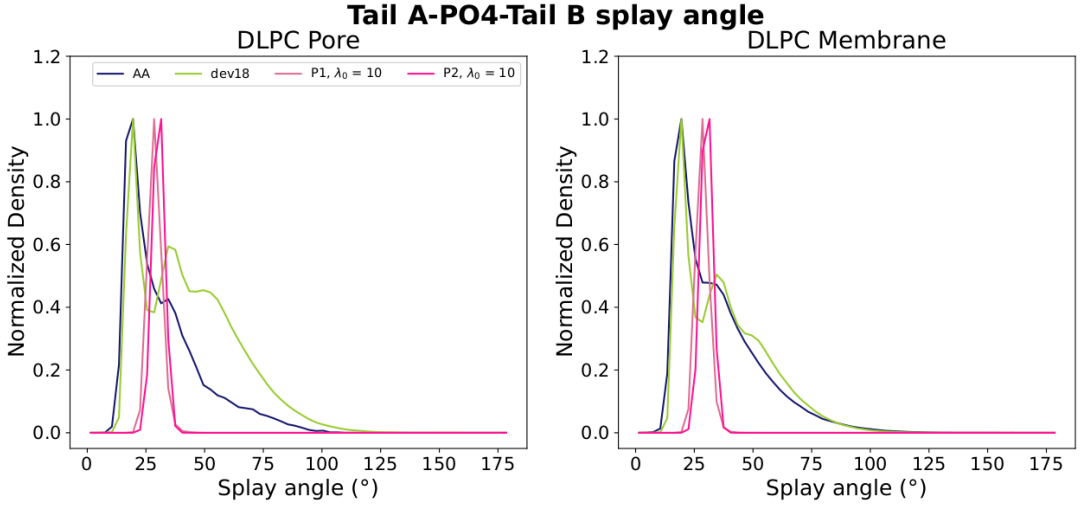

Figure A.27: **Distribution of the angle between beads C3A, PO4, and C3B of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

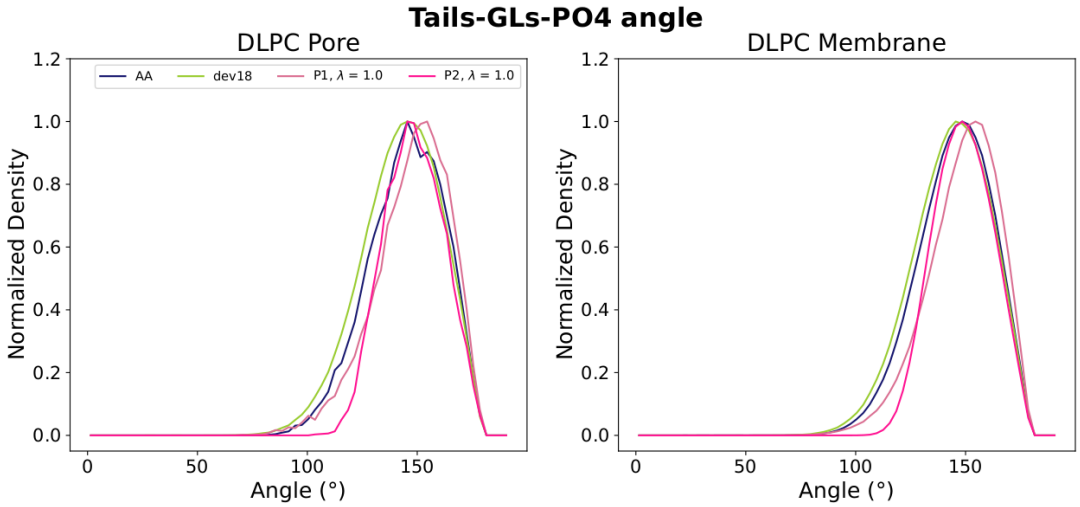

Figure A.28: **Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

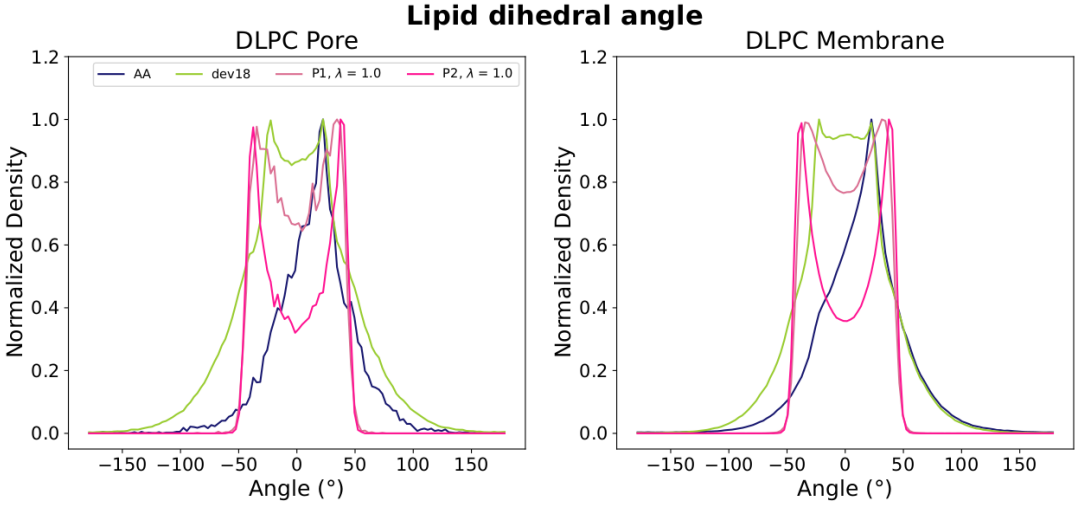

Figure A.29: **Distribution of the dihedral angle of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

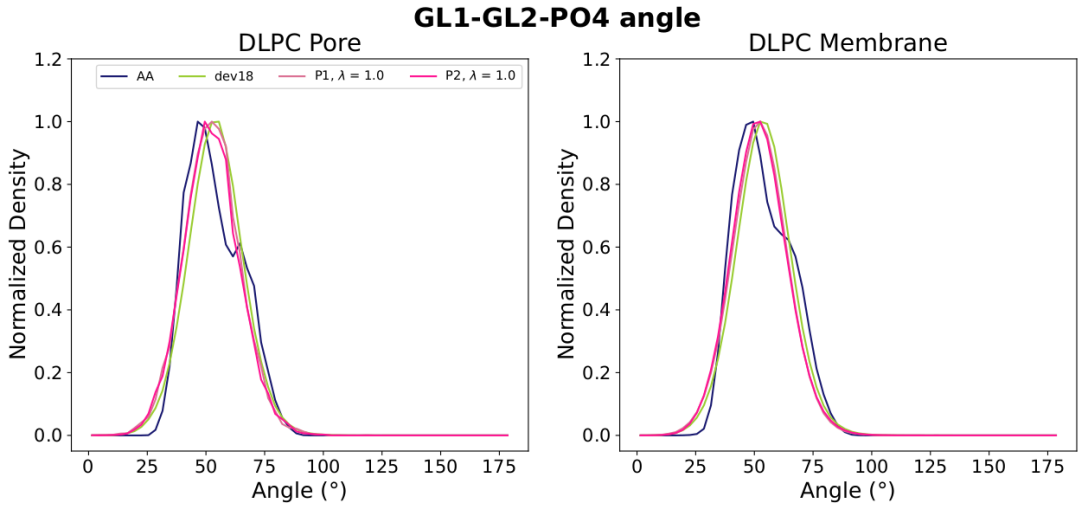

Figure A.30: **Distribution of the angle formed between the glycerols and the PO4 beads of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

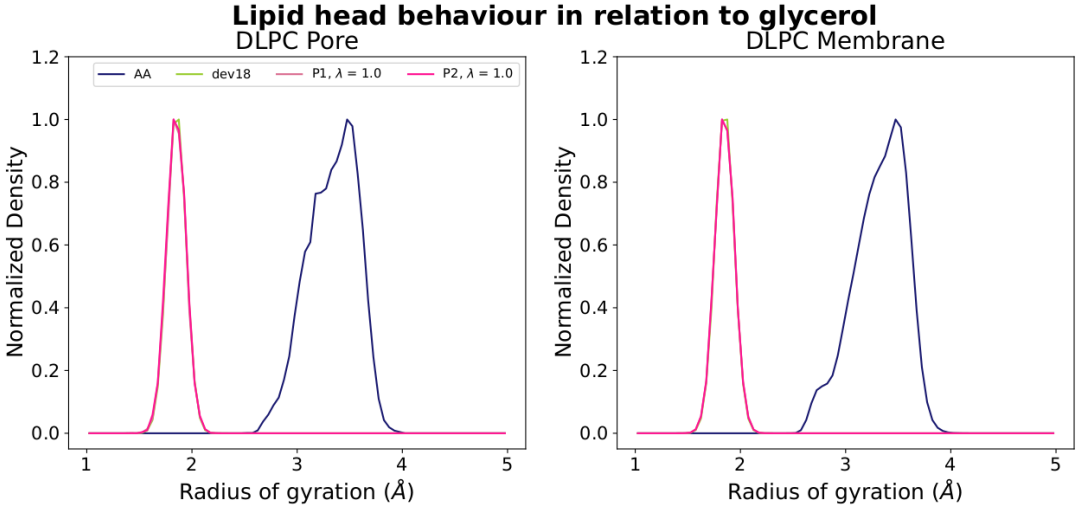

Figure A.31: **Distribution of the radius of gyration of the lipids' head.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

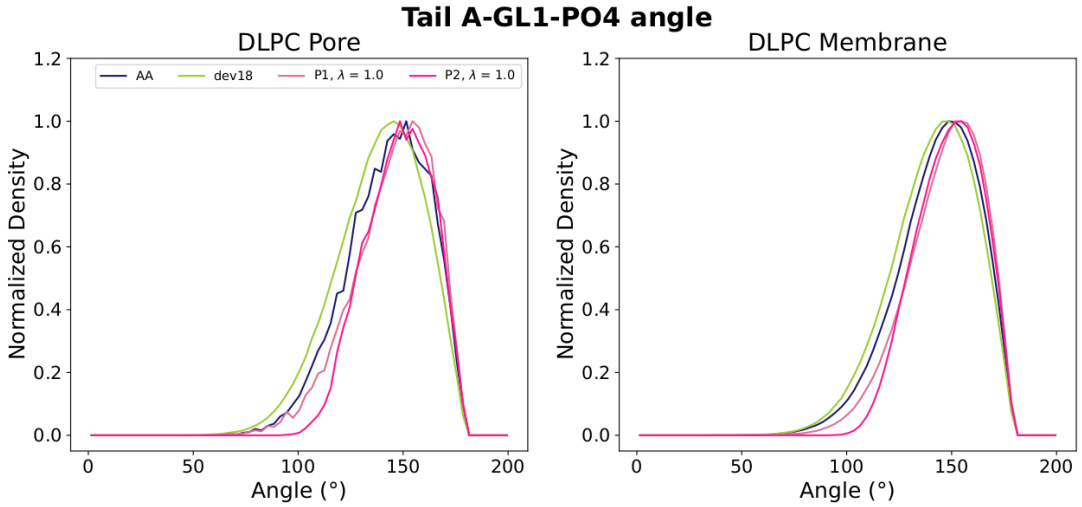

Figure A.32: **Distribution of the angle formed between tail A, GL1 and PO4.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

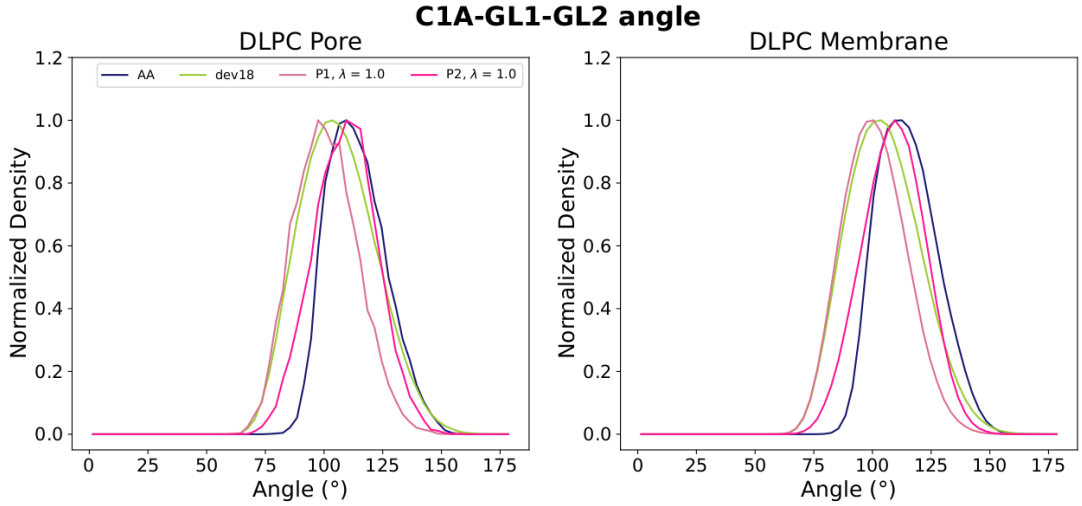

Figure A.33: **Distribution of theangle between beads C1A, GL1, and GL2 of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

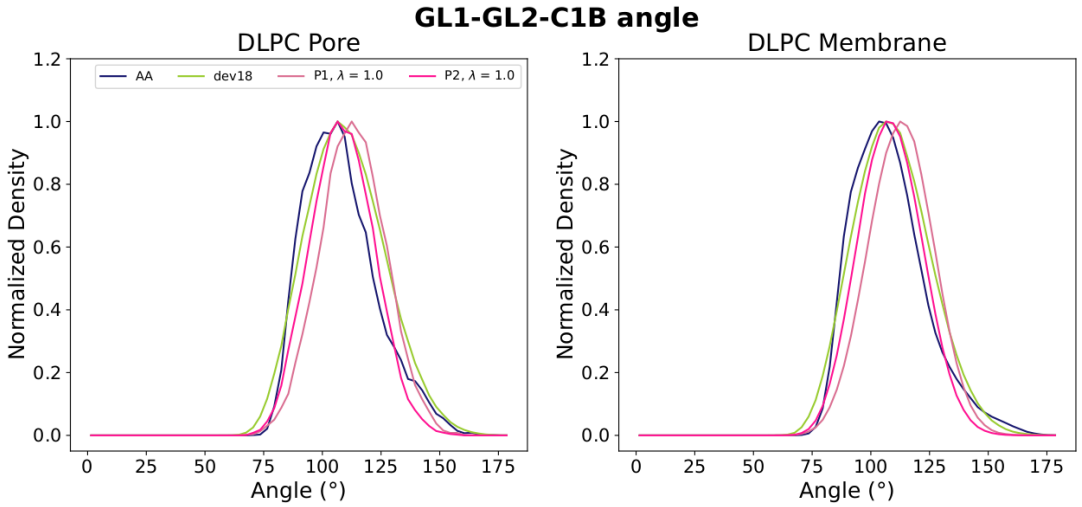

Figure A.34: **Distribution of the angle between beads GL1, GL2, and C1B of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

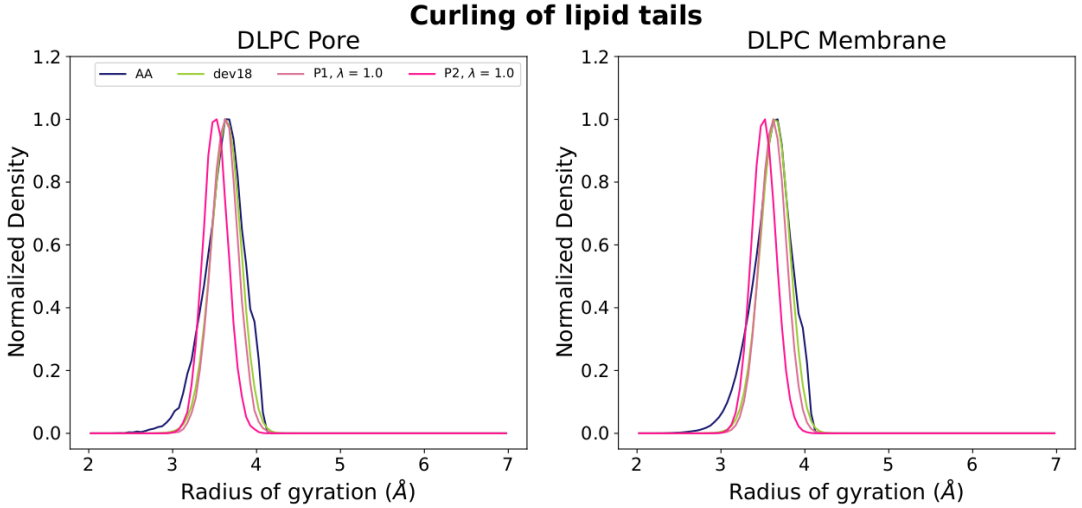

Figure A.35: **Distribution of the radius of gyration of the lipid's tails.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our first lipid prototype at an initial lambda state value of 10 is in light pink, and the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, respectively.

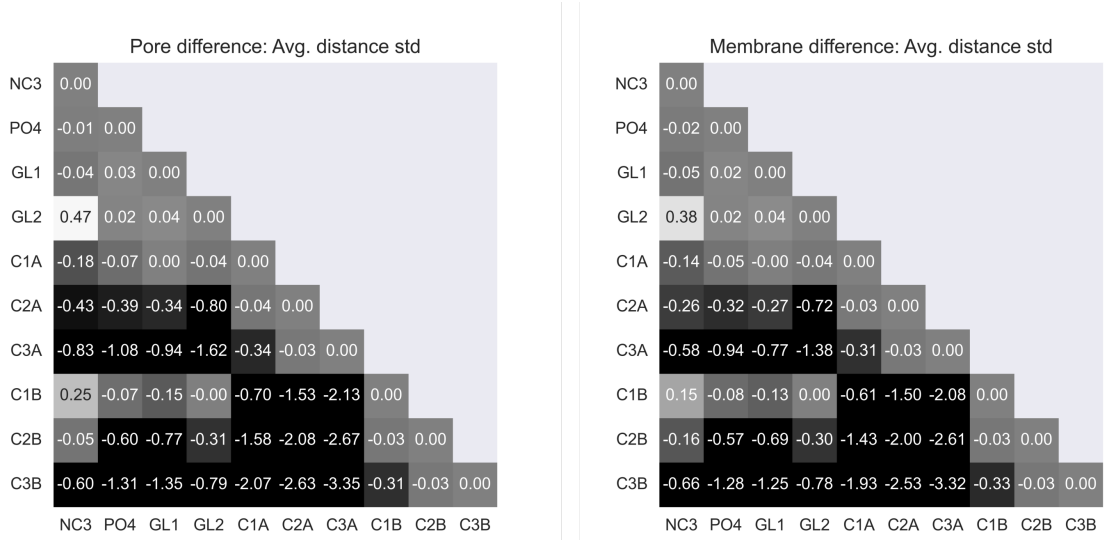

Figure A.36: **Standard deviation of the forward-mapped atomistic lipid's bead distances subtracted from the second lipid prototype lipid's bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

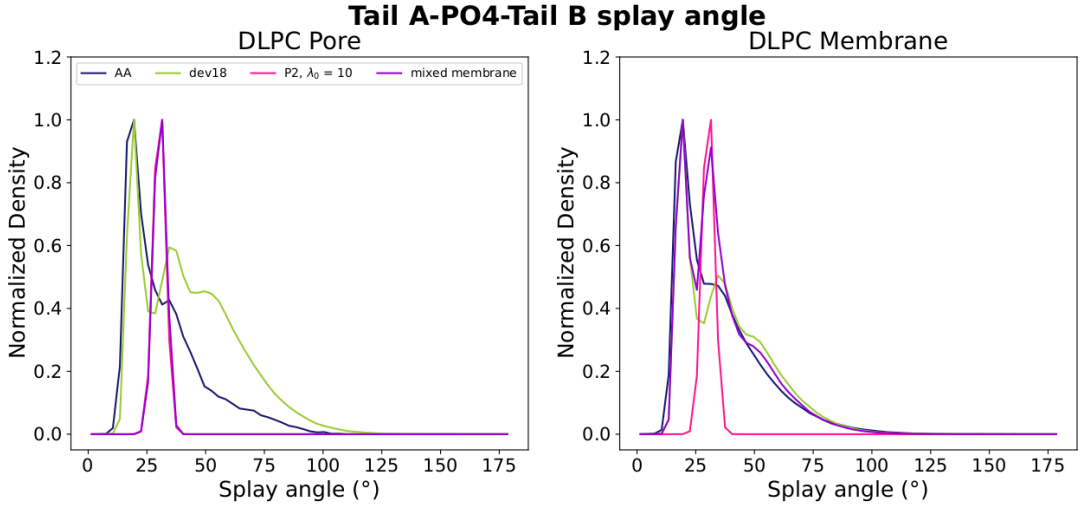

Figure A.37: **Distribution of the angle between beads C3A, PO4, and C3B of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

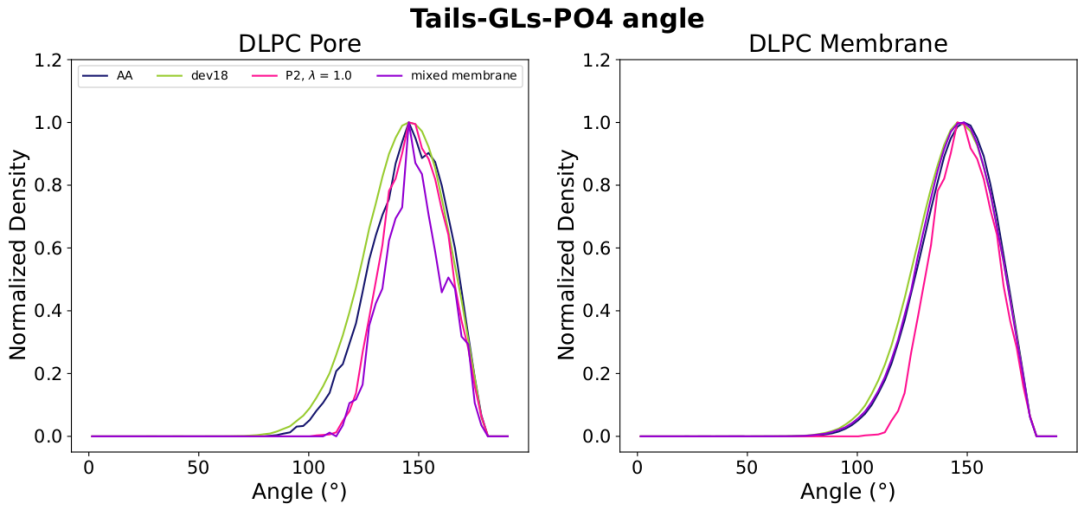

Figure A.38: **Distribution of the angle formed between the tails, the glycerols, and the PO4 beads of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

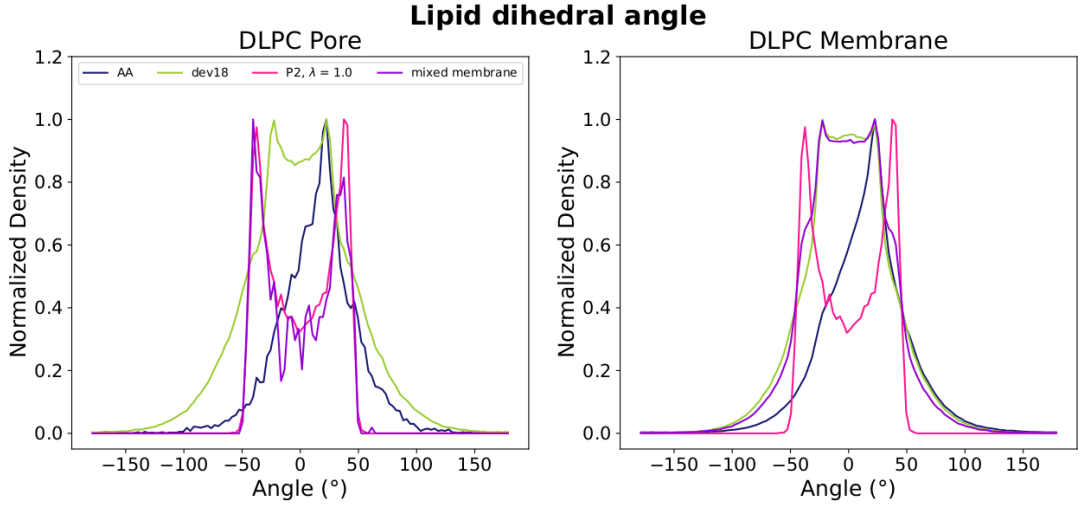

Figure A.39: **Distribution of the dihedral angle of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

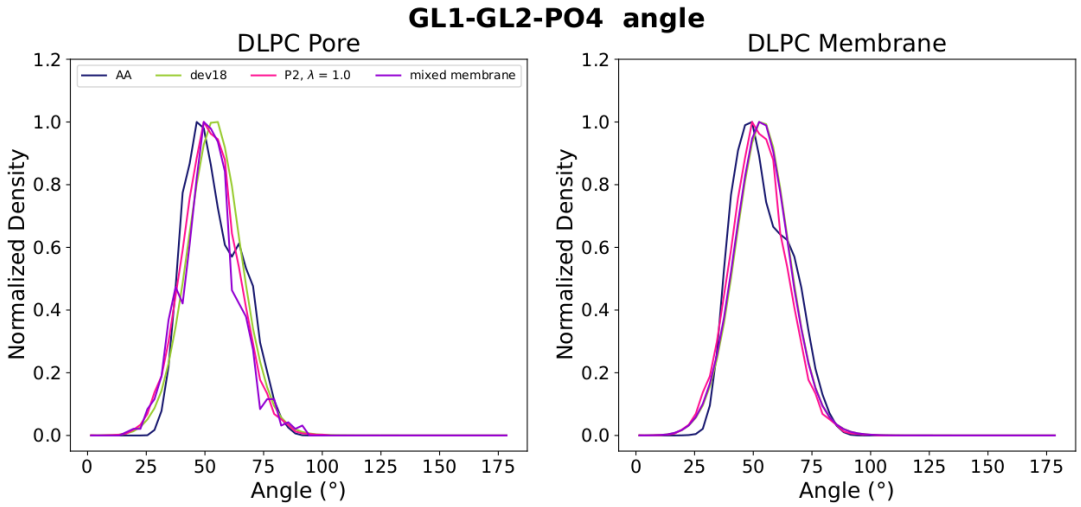

Figure A.40: **Distribution of the angle formed between the glycerols and the PO4 beads of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

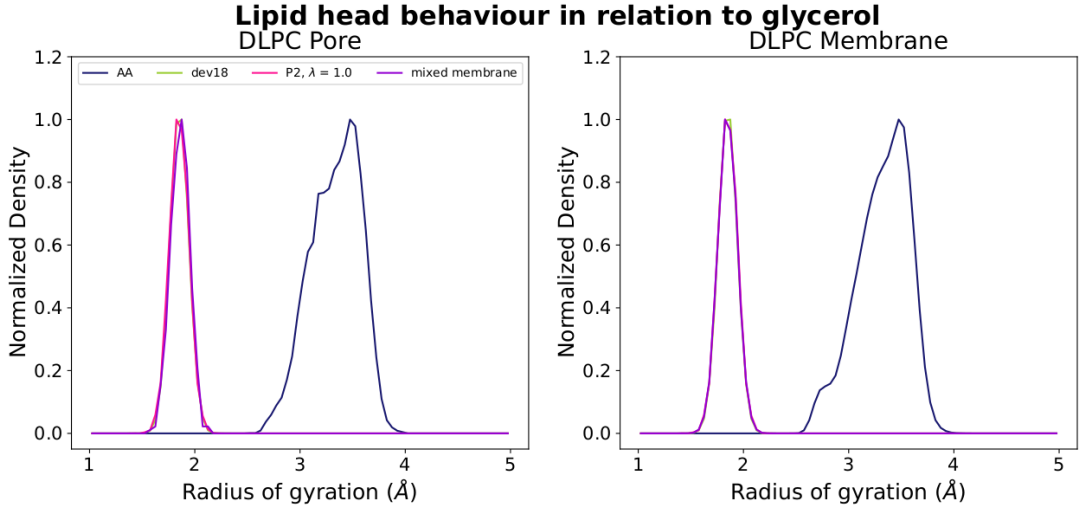

Figure A.41: **Distribution of the radius of gyration of the lipids' head.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

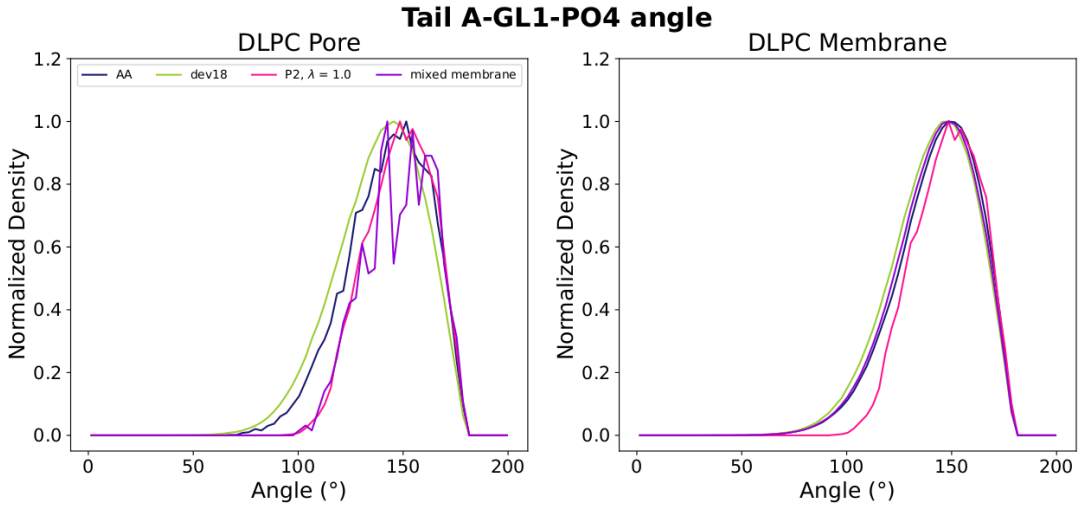

Figure A.42: **Distribution of the angle formed between tail A, GL1 and PO4.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

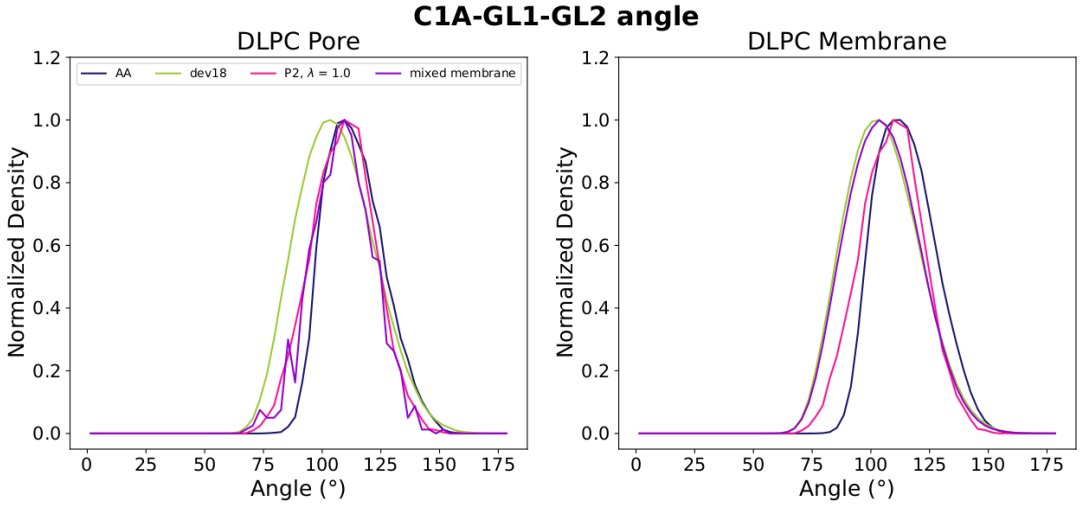

Figure A.43: **Distribution of theangle between beads C1A, GL1, and GL2 of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

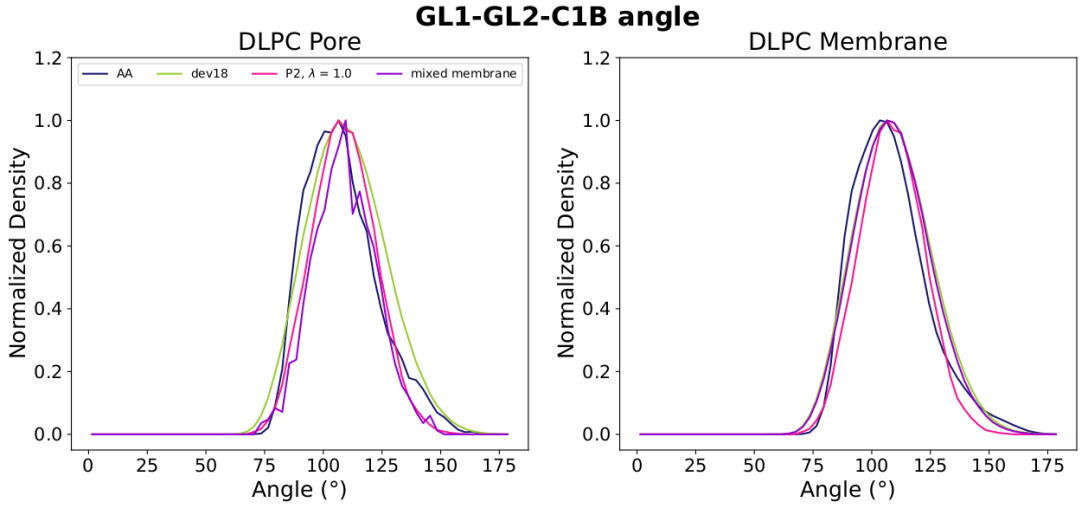

Figure A.44: **Distribution of the angle between beads GL1, GL2, and C1B of the lipids.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

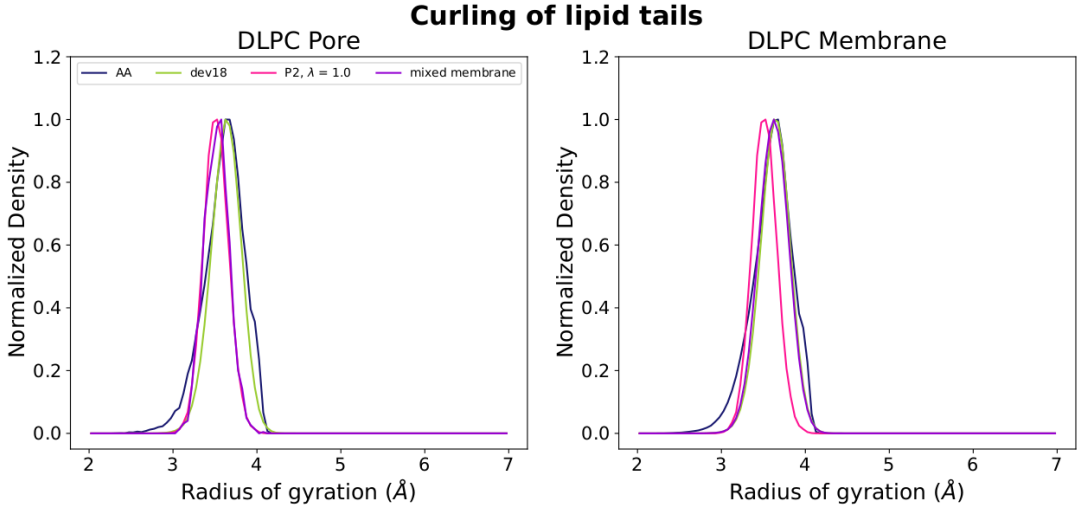

Figure A.45: **Distribution of the radius of gyration of the lipid's tails.** The atomistic distribution is in dark blue, the distribution of version 18 of the in-development Martini 3 lipids is in green, the distribution of our second lipid prototype at an initial lambda state value of 10 is in dark pink, and the distribution of the mixed membrane system is in purple, respectively.

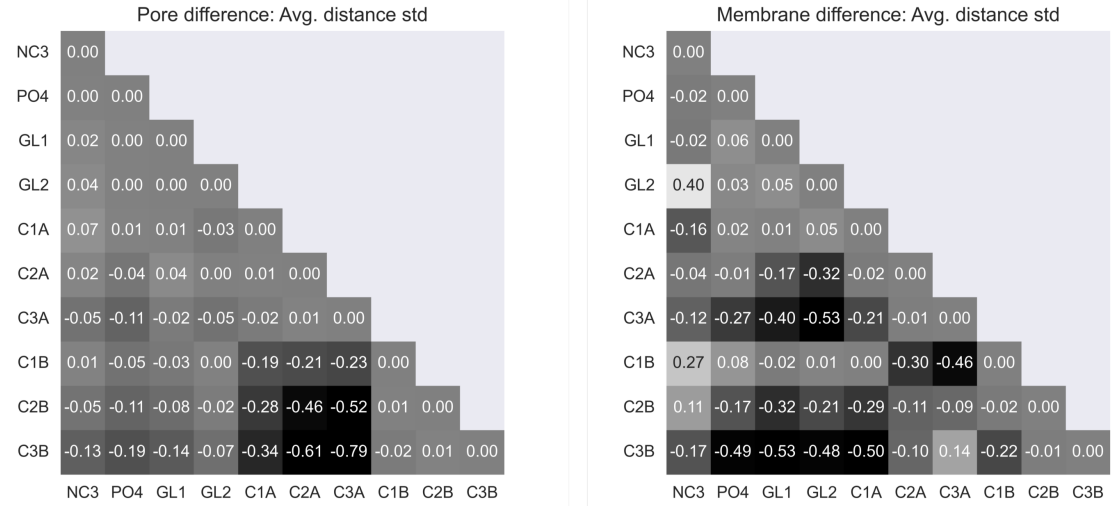

Figure A.46: **Standard deviation matrices of the forward-mapped atomistic lipid's bead distances subtracted from the mixed membrane system lipid's bead distances.** On the left is the distance matrix corresponding to pore lipids; On the right is the matrix corresponding to membrane lipids. The average distances are in Angstrom (Å).

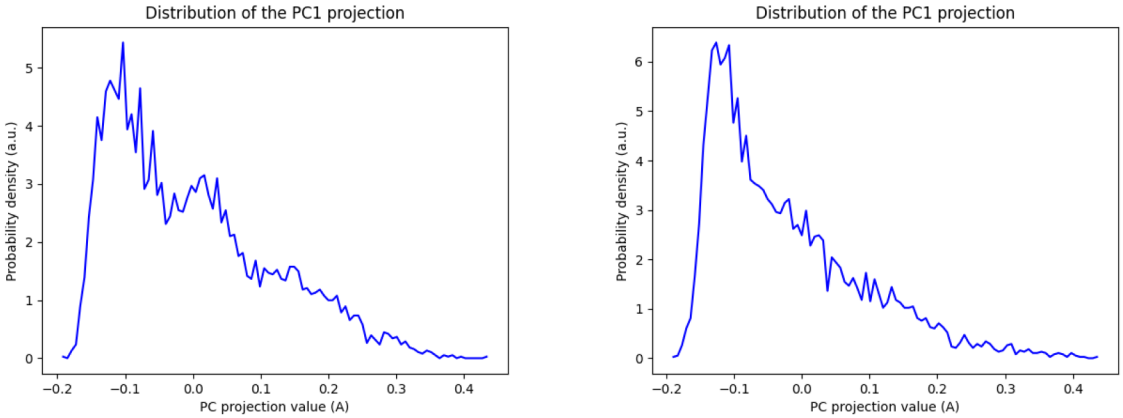

Figure A.47: **Principal component 1 distribution for the atomistic DLPC membrane system.** On the left is the PC1 distribution of the pore lipids, and on the right is the PC1 distribution of the membrane lipids.

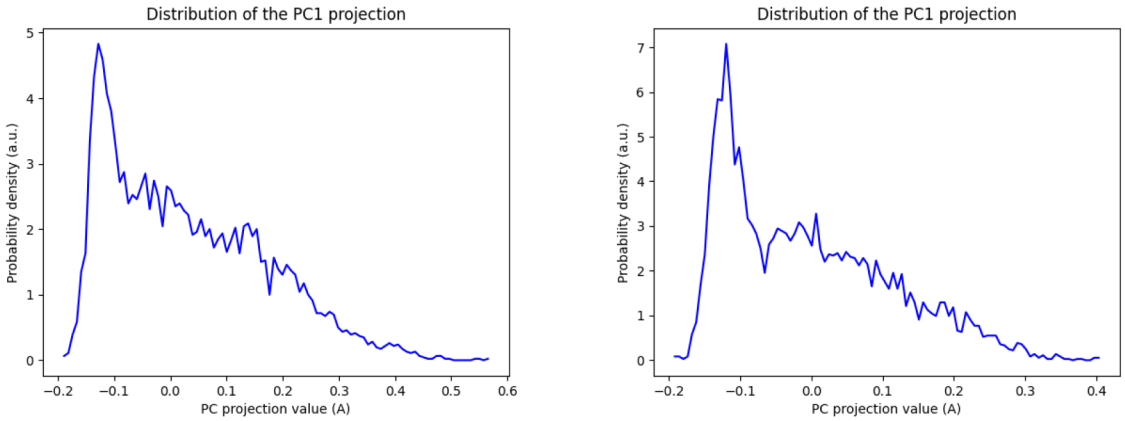

Figure A.48: **Principal component 1 distribution for the Martini 3 in-development version 18 of the DLPC membrane system.** On the left is the PC1 distribution of the pore lipids, and on the right is the PC1 distribution of the membrane lipids.

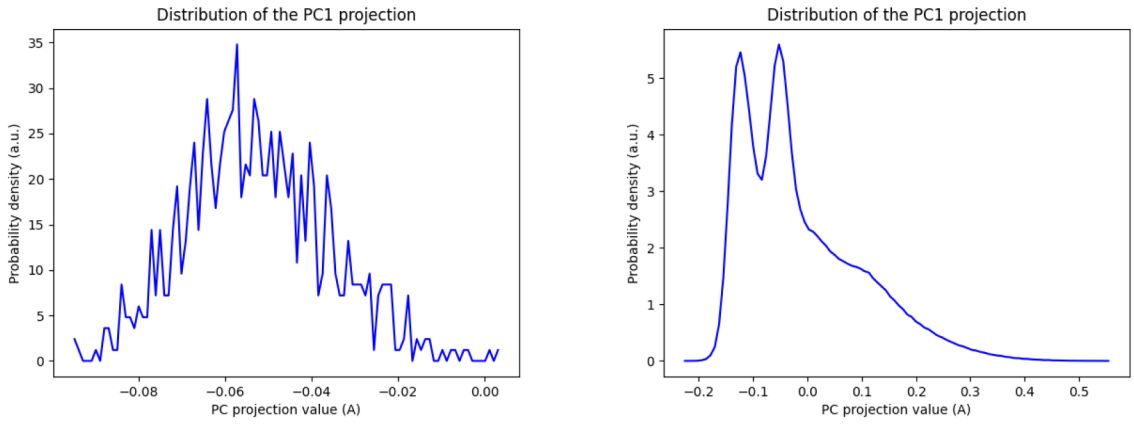

Figure A.49: **Principal component 1 distribution for the mixed membrane system.** On the left is the PC1 distribution of the pore lipids, and on the right is the PC1 distribution of the membrane lipids.

I Annex 1

I.1 Cholesterol parameterization for the Martini 3 force field

Cholesterol is an important biomolecule in cell life, specifically for animal cells where it plays crucial roles in the regulation of membrane fluidity, rigidity, and permeability [2]. Additionally, cholesterol also interacts with numerous membrane proteins [170–173].

With the release of the Martini 3 force field came a need to revisit and reparametrize the previously modeled biomolecules - one of those molecules is cholesterol. The previous Martini 2 cholesterol model [92] was found to be too lipophilic [174] and there were issues with artificial temperature gradients [175]. All of these shortcomings of the model contributed to incorrect cholesterol behavior in the membrane and affected its interactions with other molecules. Therefore, with the new iteration of the force field and its consequent need of reparametrization, there was also the opportunity to address these problems in the previous Martini 2 cholesterol in order to obtain a much improved model.

For the reparameterization of the new cholesterol model for the Martini 3 force field, our efforts focused mostly on testing the recognition and binding of the various prototypes to proteins. We built various CG MD simulations of lipid membranes containing cholesterol as well as an integral or a peripheral protein. For this, we used two proteins: VDAC1, and the transmembrane domain of the Class F G-Protein coupled receptor (GPCR) smoothened (SMO). VDAC1 is a voltage-dependent anion channel with a multitude of binding sites for cholesterol and it is responsible for the regulation of ion and metabolite flux between the cytosol and the mitochondrial intermembrane space [171]. The GPCR smoothened is a vital component of the Hedgehog signaling cascade, which controls various processes, such as the regulation of adult stem cells. It is suggested that this protein is regulated via interactions of its transmembrane domain with cholesterol [172]. Three replicates of each system were used for analysis purposes. Our simulations run with the new Martini 3 cholesterol model recovered binding pockets and modes in agreement with experimental, all-atom, and Martini 2 studies, thus providing positive validation for the new cholesterol model.

The article detailing how the model was built as well as all of the cholesterol-protein interactions analyses and other validations has since been accepted for publication [88]. However, the pre-print is already available on ChemRxiv: https://doi.org/10.26434/chemrxiv-2023-lh7bq-v2.

UNIVERSIDADE NOVA
DE LISBOA