

Electronic Thesis and Dissertation Repository

12-15-2011 12:00 AM

Molecular Dynamics Studies of Interactions of Phospholipid Membranes with Dehydroergosterol and Penetrating Peptides

Amir Mohsen Pourmousa Abkenar
The University of Western Ontario

Supervisor
Mikko Karttunen
The University of Western Ontario

Graduate Program in Applied Mathematics
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Amir Mohsen Pourmousa Abkenar 2011

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Biological and Chemical Physics Commons](#)

Recommended Citation

Pourmousa Abkenar, Amir Mohsen, "Molecular Dynamics Studies of Interactions of Phospholipid Membranes with Dehydroergosterol and Penetrating Peptides" (2011). *Electronic Thesis and Dissertation Repository*. 350.

<https://ir.lib.uwo.ca/etd/350>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

MOLECULAR DYNAMICS STUDIES OF INTERACTIONS OF
PHOSPHOLIPID MEMBRANES WITH DEHYDROERGOSTEROL AND
PENETRATING PEPTIDES

(Spine title: Plib)

(Thesis format: Monograph)

by

Amir Mohsen Pourmousa Abkenar

Graduate Program in Applied Mathematics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Amir Mohsen Pourmousa Abkenar 2011

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor:

.....
Dr. Mikko Karttunen

Supervisory Committee:

.....
Dr. Colin Denniston

.....
Dr. James Choy

Examiners:

.....
Dr. Geoff Wild

.....
Dr. Colin Denniston

.....
Dr. Lars Konermann

.....
Dr. Apichart Linhananta

The thesis by

Amir Mohsen Pourmoussa Abkenar

entitled:

**Molecular Dynamics Studies of Interactions of Phospholipid Membranes with
Dehydroergosterol and Penetrating Peptides**

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

.....
Date

.....
Chair of the Thesis Examination Board

Abstract

We have performed molecular modeling of membrane and peptide systems by employing the classical molecular dynamics method and force field parameterizations. In this thesis, our main interest is the interaction of sterols as well as peptides with membranes. The thesis consists of three related projects.

The first project focuses on cholesterol (CHOL) and dehydroergosterol (DHE) interacting with palmitoyl-oleoyl phosphatidylcholine (POPC) lipid bilayers. We study the effects of both sterols on the bilayer and compare them with each other. We first study the condensing and ordering effect of these sterols. Then, we study their orientations within the bilayer and relate them to their interactions with the bilayer.

In the second project, we study the interaction of a cell-penetrating-peptide, namely transportan, with dipalmitoyl phosphatidylcholine (DPPC) lipid bilayers with the aim of understanding the physical mechanism of the penetration process. By analyzing 10 simulations, we shed light on the behavior of the peptide and membrane when they are associated with each other. We also analyze four simulations in which the peptide is directly inserted in the bilayer core so as to study the behavior of the peptide when it is deep in the bilayer. By using umbrella sampling method and performing 41 biased simulations, we also try to find the free energy profile of the system as a function of the distance between the peptide and the center of mass of the bilayer.

In the third project, we study the interaction of another cell-penetrating peptide, namely penetratin, with DPPC lipid bilayers. In this research, the spontaneous binding of a penetratin peptide as well as the role of different residues in binding are investigated. The behavior of penetratin is then compared with that of transportan and important differences are highlighted.

Keywords: MD simulation, gromacs, cell-membrane, DPPC, POPC, cholesterol, dehydroergosterol, transportan, penetratin

Contents

Certificate of Examination	ii
Abstract	iii
List of Figures	viii
List of Tables	xvi
List of Abbreviations	xvii
1 Biological Background	1
1.1 Cell and Membrane	1
1.1.1 Historical background	1
1.1.2 Lipid bilayer	3
1.2 Proteins and peptides	7
1.2.1 Amino acids	7
1.2.2 Peptide bond	9
1.3 The structure of this thesis	11
2 Physical Background	13
2.1 Brief review of thermodynamics and statistical mechanics	13
2.1.1 Thermodynamic equilibrium	13
2.1.2 Classical ensemble theory	14
2.1.3 Ergodicity	17
2.2 Phase behavior of model membranes	18
2.3 Protein conformation	21
2.3.1 α -helix	23
2.3.2 3_{10} -helix	23
2.3.3 π -helix	24
2.3.4 Helical wheel diagram representation	24

2.4	Electrostatics of peptide-membrane systems	25
2.4.1	Hydrogen bonding	25
2.4.2	Charge-pair interaction	26
2.5	Methods	28
2.5.1	The molecular dynamics methods	28
2.5.2	MD integrators	30
2.5.3	Force field	32
	Choosing the force field	34
	Simple Charge Model (SPC) for water molecules	36
2.5.4	Temperature and pressure coupling	37
	Berendsen thermostat	38
	Berendsen barostat	40
2.5.5	Constraint dynamics	41
	LINear Constraint Solver (LINCS)	42
2.5.6	Long-Range interactions	44
	The Ewald summation of Coulomb energy	45
	Particle-Mesh-Ewald (PME)	49
2.5.7	Free energy calculation	50
	Umbrella sampling method	50
3	Interaction of POPC bilayers with Cholesterol and Dehydroergosterol	54
3.1	Introduction	54
3.2	Methods	56
3.3	Results	59
3.3.1	Area per molecule and membrane thickness	59
3.3.2	Location and orientation of sterols in the bilayer	60
3.3.3	Order and conformation of acyl chains	62
3.3.4	Conformation of sterols	64
3.3.5	Packing of chains	67
3.3.6	Packing of atoms relative to sterol ring atoms	70
3.3.7	Membrane-water interface	71
3.3.8	Lateral pressure	73
3.3.9	Discussion	75
3.3.10	Conclusion	78
4	DPPC Bilayer Interacting With Transportan	79
4.1	Introduction	79

4.2	Simulation Details	82
4.3	Analysis	84
4.4	Results	85
4.4.1	Unbiased Simulations	85
	Distance between several system components and the middle of the bilayer	85
	Minimal distance between lysine residues and bilayer's phosphate groups	86
	Hydrogen bonding between lysine residues and bilayer's phosphate groups	87
	Area per Lipid and Thickness	89
	Order parameter	91
	P-N angle	92
	Peptide structure	93
4.4.2	Insertion of transportan into bilayer	93
4.4.3	Biased Simulations	94
4.5	Discussion	98
4.6	Conclusion	104
5	DPPC Bilayer Interacting With Penetratin	105
5.1	Introduction	105
5.2	Simulation Details	106
5.3	Analysis	107
5.3.1	Results	108
	Distance between several system components and the middle of the bilayer	108
	Hydrogen bonds	111
	Membrane thickness	112
	Order Parameter	113
	Peptide conformation	114
5.4	Discussion and Conclusion	116
6	Conclusion and Proposal for Future Work	117
6.1	Conclusion	117
6.2	Proposal for Future Work	118
	Bibliography	120

List of Figures

1.1	(a) Cell structure of a bacterium which is prokaryotic. (b) Structure of a typical animal cell which is eukaryotic. The pictures have been released into the public domain. Sources: wikipedia:prokaryote (a) and wikipedia:eukaryote (b).	2
1.2	Fluid mosaic model for membrane structure. The membrane lipids form a fluid 2D solvent where the integral proteins float laterally in the plane of the bilayer. The picture has been released into the public domain. Source: wikipedia:cell membrane.	3
1.3	A schematic picture of a lipid and a bilayer formed by assembling of lipids. The picture has been released into the public domain. Source: wikipedia:phospholipid.	4
1.4	Three classes of membrane lipids: phosphoglycerides (a), sphingolipids (b), and cholesterol (c). Examples of phosphoglycerides include phosphotidylcholine (PC), phosphotidylethanolamine (PE), phosphotidylserine (PS), and phosphatidylglycerol (PG). Sphingomyelin (SM) is an example of a sphingolipids.	5
1.5	The generic structure of an amino acid. Different amino acids differ in which side chain, or R group, is attached to them. The nitrogen atom together with two hydrogen atoms are called amino group. The carbon atom together with the oxygen atoms are called carboxyl group. The carbon atom attached to the R group is called an α -carbon. The picture has been released into the public domain. Source: wikipedia:amino acid.	7
1.6	Peptide bond occurring between two amino acids. Each amino acid unit is called a residue. The repeating part of the polypeptide, i.e. everything except the R groups, is called the backbone or the main chain. At neutral pH, the carboxyl group (C-terminus) loses one proton, making it negatively charged, while the amino group (N-terminus) accepts one proton, making it positively charged. Thus the whole polypeptide is polar. The picture has been released into the public domain. Source: wikipedia:amino acid.	9
1.7	Illustration of dihedral angles ϕ and ψ . The angle of rotation about the bond between the nitrogen and the α -carbon atoms is called phi (ϕ). The angle of rotation about the bond between the α -carbon and the carbonyl carbon atoms is called psi (ψ). The picture has been released into the public domain. Source: wikipedia:protein structure.	10

1.8	Ramachandran diagram shows that in a polypeptide, because two atoms cannot be in the same place at the same time, not all ϕ and ψ values are possible. The most favorable regions are shown; The picture is licensed under the Creative Commons license and can be found at http://en.wikipedia.org/wiki/File:Ramachandran_plot_original_outlines.jpg .	11
2.1	(<i>left</i>) A non-ergodic system in which the trajectory does not pass through all possible states in the phase space. The <i>red</i> line encases the sub-phase space of the non-ergodic trajectory. (<i>right</i>) An ergodic system in which the trajectory passes through all possible states in phase space.	17
2.2	A stack of lipid bilayers forming a multi-bilayer. The picture is taken from Ref. [77] © Elsevier Limited.	18
2.3	A typical DSC graph from an experiment on Dimyristoyl phosphatidylcholine (DMPC) lipids. Each curve corresponds to a different sample preparation and hydration. Temperatures at which the peaks occur correspond to different phase transitions. This research was originally published in Journal Name. Author(s). Title. <i>The Journal of Biological Chemistry</i> . 1979; 254:6068-6078. © the American Society for Biochemistry and Molecular Biology.	19
2.4	Temperature-composition phase diagram of hydrated DMPC multi-bilayers. The structures of different phases L_{α} , $L_{\beta'}$, and $P_{\beta'}$ are shown in the middle. The packing lattice (viewed from the top) is hexagonal for the $P_{\beta'}$ and distortedly hexagonal for the $L_{\beta'}$ phase. The picture is taken from Ref. [77] © Elsevier Limited.	20
2.5	Primary, secondary, tertiary and quaternary structures of proteins. The picture has been released into the public domain. Source: wikipedia:protein structure.	22
2.6	A representation of the 3D structure of the myoglobin protein. Alpha helices are shown in colour. The picture has been released into the public domain. Source: wikipedia:myoglobin.	22
2.7	Side view of an α -helix (a), a 3_{10} -helix (b), and a π -helix of alanine. The hydrogen bonds are shown in <i>magenta</i> . Oxygens are in <i>red</i> , hydrogens in <i>white</i> , carbons in <i>gray</i> and nitrogens in <i>blue</i> . The pictures are licensed under the Creative Commons license.	23
2.8	The helical wheel diagram of transportan. Each residue is categorized according to the following color scheme: <i>white</i> (hydrophobic in neutral pH), <i>cyan</i> (hydrophilic and neutral in neutral pH), <i>red</i> (hydrophilic and charged in neutral pH). There are 27 residues in total.	24

2.9	Hydrogen bond criterion. The donor/acceptor molecules are shown in white/black spheres. The cutoff values of α and r used in this simulation are 30° and 0.35 nm respectively. $r = 0.35$ nm corresponds to the first minimum of the RDF of SPC water molecules [11, 104, 50, 90].	25
2.10	Charge pair interaction formed between adjacent lipids: charge-pair between methyl group and carbonyl group (left), and charge-pair between methyl group and phosphate group (right).	27
2.11	Illustration of bond stretching (left), bond bending (middle) and torsional terms appearing in eq. 2.27.	32
2.12	The Lennard-Jones interaction: ϵ is the depth of the potential, r is the distance between two particles, and σ is the distance at which the inter-particle potential is zero. The regions where the slope of the potential is negative (positive) corresponds to a repulsion (an attraction) interaction.	34
2.13	Schematic illustration of the united atom model.	35
2.14	The SPC water model.	37
2.15	The three position-updates needed for one time step. The dashed line is the old bond of length d , the solid lines are the new bonds. In the left picture, the updated positions are not constrained yet. In the middle one, the projection of the new bond on the old bond is set to zero, and in the right one, a correction is applied to set the bond's length equal to d . $l = d / \cos \theta$ and $p = (2d^2 - l^2)^{\frac{1}{2}}$	45
3.1	Chemical structure of POPC.	56
3.2	Chemical structure of CHOL and DHE together while differences are indicated in <i>red</i> (a), representation of CHOL (b) and DHE (c).	57
3.3	A snapshot of the simulation of the POPC-DHE system at 0 ns. Water is shown in <i>red</i> , DHE molecules are represented in <i>yellow</i> while POPC molecules are shown as transparent beads in <i>light blue</i> . Nitrogen and phosphorus atoms are shown as beads in <i>blue</i> and <i>light grey</i> respectively.	57
3.4	temperature vs time for pure POPC (a), POPC+CHOL (b) and POPC+DHE (c).	59
3.5	Total energy (top two graphs) and potential energy (bottom two graphs) for POPC+CHOL (black and red) and POPC+DHE (blue and green) systems.	59
3.6	Area per lipid vs time for pure POPC (a), POPC+CHOL (b) and POPC+DHE(c).	60
3.7	Partial density profiles along the bilayer normal for POPC (<i>black</i>), water (<i>blue</i>) and sterol molecules (<i>red</i>) for pure POPC (<i>solid line</i>), POPC+CHOL (<i>dashed line</i>) and POPC+DHE (<i>dash-dotted</i>)	61

3.8	The density profiles of the lower leaflet of pure POPC (<i>solid line</i>), POPC-CHOL (<i>dashed line</i>) and POPC-DHE (<i>dashed-dotted</i>). The density profile represents the nitrogen group (NH_4^+), the phosphate group (PO_4^-), the beginning of POPC chains (C_{12} , C_{13} , and C_{32}), carbonyl groups ($\text{C}=\text{O}$), the end of POPC chains (C_{50} , and C_{52}), sterols' head groups (OH), sterols' ring system, and sterols' tails.	62
3.9	Order parameter S_{mol} of the palmitoyl acyl chains in pure POPC (<i>solid line</i>), POPC+CHOL (<i>dashed line</i>) and POPC+DHE (<i>dash-dotted</i>).	63
3.10	Normalized distribution of the angle θ between the P-N vector of POPC lipids and the z -axis in pure POPC (<i>solid line</i>), POPC+CHOL (<i>dashed line</i>) and POPC+DHE (<i>dash-dotted</i>).	64
3.11	Normalized distribution of the tilt angle of palmitoyl (<i>black</i>) and oleoyl (<i>red</i>) acyl chains in pure POPC (<i>solid line</i>), POPC+CHOL (<i>dashed line</i>) and POPC+DHE (<i>dash-dotted</i>). Here, the angle is defined to be the angle between the z -axis and the vector connecting the two ends of POPC tails, namely C_{50} to C_{34} for palmitoyl, and C_{52} to C_{15} for oleoyl acyl chain (see Fig. 3.1 for atom numbering).	65
3.12	Normalized distribution of the tilt angle of CHOL ring structure (<i>solid line</i>) and DHE ring structure (<i>dashed line</i>). Here, the tilt angle is defined to be the angle between the z -axis and the vector connecting the two ends of the sterols' ring structure, i.e. C_3 and C_{17} (see Fig. 3.2 for atom numbering).	66
3.13	Normalized distribution of the tilt angle of CHOL tail (<i>solid line</i>) and DHE tail (<i>dashed line</i>). Here, the tilt angle is defined to be the angle between the z -axis and the vector connecting C_{17} to C_{25} (see Fig. 3.2 for atom numbering).	67
3.14	The angle between the vector connecting C_{20} to C_{21} and the plane formed by three methyl groups C_{18} , C_{13} and C_{20} in CHOL (<i>solid black line</i>) and DHE (<i>dashed black line</i>). The graph in <i>red</i> relates to the extra methyl group C_{29} and is the angle between the vector connecting C_{24} to C_{29} and the plane formed by three methyl groups C_{18} , C_{13} and C_{24} . For atom numbering see Fig. 3.2.	68
3.15	Schematic illustration of CHOL and DHE. The sterols are divided into two main parts, a ring system and a tail. The other methyl groups are budded out of the body of sterols in the way which is shown. The extra methyl group of DHE (C_{29}) is shown in <i>red</i>	68
3.16	Two-dimensional RDFs of COM of CHOL- (<i>black</i>) and DHE-ring (<i>red</i>) versus themselves (a), and of COM of acyl chains (palmitoyl or oleoyl) versus sterols' ring structure (C_1 to C_{17}) (b).	69
3.17	Three-dimensional RDFs of carbon atoms of POPC molecules versus sterol ring (C_1 to C_{17}).	69

3.18	Number of neighbors of sterol ring methyl groups located on α -side (<i>dashed line</i>), β -side (<i>dashed-dotted</i>) and both sides (total) (<i>solid line</i>) for CHOL (<i>black</i>) and DHE (<i>gray</i>).	70
3.19	Number of neighbors of sterol tails of CHOL (<i>black</i>) and DHE (<i>gray</i>). Here, the tail of CHOL is numbered from methyl groups C ₂₀ to C ₂₇ . DHE has one more methyl group, namely C ₂₉ , which is shown in a larger font size to emphasize that there is no C ₂₈ in the numbering of DHE.	71
3.20	Lateral pressure for pure POPC (<i>solid line</i>), POPC+CHOL (<i>dashed line</i>), and POPC+DHE (<i>dashed-dotted</i>). The horizontal axis corresponds to the bilayer normal. Important maxima and minima of the profiles are indicated by <i>red</i> arrows and are explained in the text.	73
3.21	Different sterols that are similar to CHOL (<i>black</i>) in structure. Differences to CHOL are shown in different colors. DCHOL does not have the two methyl groups C ₁₈ and C ₁₉ . DESMO has an extra double bond between C ₂₄ and C ₂₅ . 7-DHC has an extra double bond between C ₇ and C ₈ . ERG has a double bond between C ₇ and C ₈ , a double bond between C ₂₂ and C ₂₃ , and an extra methyl group attached to C ₂₄	76
4.1	Schematic representation of DPPC.	82
4.2	First configuration (a), and last configuration at 780 ns (b) of the longest simulation. The peptide is shown in <i>yellow</i> , water molecules in light <i>blue</i> , DPPC lipids in <i>gray</i> and phosphorus atoms are shown as balls in <i>tan</i>	83
4.3	The helical wheel diagram of transportan. Each residue is categorized according to the following color scheme: <i>white</i> (hydrophobic in neutral pH), <i>cyan</i> (hydrophilic and neutral in neutral pH), <i>red</i> (hydrophilic and charged in neutral pH).	85
4.4	Distance of several components of the system from the middle of the bilayer. The COM of the whole bilayer is set to zero. Color coding: <i>yellow</i> (LEU27), <i>blue</i> (THR3), <i>green</i> (TRP2), <i>red</i> (GLY1), <i>black</i> (lipids' carbonyl group), <i>brown</i> (lipids' phosphorus atoms), <i>violet</i> (lipids' nitrogen atoms).	87
4.5	Distance of GLY1 (<i>red</i>), LEU27 (<i>yellow</i>) and phosphorus atoms (<i>brown</i>) from the middle of the bilayer. These graphs demonstrate which end of the peptide associates with the membrane first in each simulation. For more information regarding the insertion depth, see Table 4.1.	88
4.6	Minimal distance between NH ₃ ⁺ of lysine side chains and phosphate groups. Color coding: <i>black</i> (LYS13), <i>red</i> (LYS17), <i>green</i> (LYS24) and <i>blue</i> (LYS25).	89
4.7	Hydrogen bond criterion. The donor/acceptor molecules are shown in white/black spheres. The values of α and r used in this simulation are 30° and 0.35 nm respectively.	89

4.8	Hydrogen bonds between LYS13, LYS17, LYS24 and LYS25 with lipid phosphate groups. Color coding: <i>black</i> (LYS13), <i>red</i> (LYS14), <i>green</i> (LYS24) and <i>blue</i> (LYS25).	90
4.9	Area per lipid (a), and the thickness of neighboring (<i>black</i>) and distant (<i>red</i>) lipids (b). As defined in section 4.3, the neighboring lipids are no more than 0.6 nm away from the peptide.	91
4.10	z -coordinates of nitrogen atoms of neighboring (<i>black</i>) and distant (<i>red</i>) lipids. The peptide is associated with the upper monolayer. As defined in section 4.3, the neighboring lipids are no more than 0.6 nm away from the peptide.	91
4.11	Representation of θ_n as used in computing S_{mol} . z -axis is defined to be in the direction of bilayer normal.	92
4.12	Order parameter S_{mol} of the <i>sn</i> -1 (a) and <i>sn</i> -2 (b) acyl chains of pure DPPC (<i>solid line</i>), peptide's neighboring (<i>dashed line</i>) and distant lipids (<i>dotted line</i>) obtained from the time interval of 600 ns to 780 ns of simulation #2.	93
4.13	The distribution of the angle between P-N vector and the bilayer normal (z -axis) for pure DPPC (<i>solid line</i>), peptide's neighboring (<i>dashed line</i>) and distant (<i>dotted line</i>) lipids of the peptide-associated monolayer. The mean values are shown in <i>red</i> with the same line style.	94
4.14	Evolution of secondary structures of transportan as defined by DSSP [68]. The residues start from GLY1 (N-terminus) to LEU27 (C-terminus). Different colors correspond to different secondary structures as shown in legends.	95
4.15	Transportan inserted into the bilayer's center: initial configuration of the system with transportan inside the bilayer (a), last configuration of the simulation #1 after 172 ns (b), last configuration of the simulation #2 after 172 ns (c), last configuration of the simulation #3 after 164 ns (d), and last configuration of the simulation #4 after 134 ns (e). Phosphorus atoms are shown as balls in <i>tan</i> , lysine residues in <i>black</i> , hydrophilic residues in <i>blue</i> and hydrophobic residues in <i>green</i> . The overall conformation of the peptide is shown as a tube in <i>red</i> .	96
4.16	Order parameter, S_{mboxmol} of <i>sn</i> -1 acyl chains of DPPC in the inserted-transportan simulations obtained from the last 30 ns of the simulations.	97
4.17	Free energy as a function of the peptide's distance from the bilayer's center. The bilayer's center is set to zero. The water-membrane interface, defined as the average position of phosphorus atoms in the unbiased simulations, is shown as a <i>dashed line</i> in <i>red</i> . There are three important energy minima at $z = 0.4$ nm, $z = 0.8$ nm and $z = 1.2$ nm as indicated by the arrows.	98

4.18	Snapshots of the biased simulations when the peptide is restrained in the middle of the bilayer (a), 0.4 nm from the bilayer's center (b), 0.8 nm from the bilayer's center (c), and 1.2 nm from the bilayer's center (d). Phosphorus atoms are shown as balls in <i>tan</i> , lysine residues in <i>black</i> , hydrophilic residues in <i>blue</i> and hydrophobic residues in <i>green</i> . The overall conformation of the peptide is shown as a tube in <i>red</i>	99
4.19	Two-dimensional density map of water (a), and phosphate-group (PO_4^{-1}) (b) when the peptide is inserted into the bilayer's center. The figures correspond to simulation #2 of the inserted-transportan simulation (Fig. 4.15c). The unit is nm^{-3} , and the vertical axis is along the bilayer normal.	102
4.20	Two-dimensional density map of water ((a) and (c)) and phosphate-group (PO_4^{-1}) ((b) and (d)). In (a) and (b) the peptide is restrained at $z=0$ nm, and in (c) and (d) the peptide is restrained at $z=0.4$ nm. The unit is nm^{-3} , and the vertical axis is along the bilayer normal.	104
5.1	First configuration (a), and the configuration at 57 ns (b) of the the simulation #4. The peptide is shown in <i>yellow</i> , water molecules in light <i>blue</i> , DPPC lipids in <i>gray</i> and phosphorus atoms are shown as balls in <i>tan</i>	106
5.2	The helical wheel diagram of penetratin as obtained from Gromacs. Each residue is categorized according to the following color scheme: <i>white</i> (hydrophobic in neutral pH), <i>cyan</i> (hydrophilic and neutral in neutral pH), <i>red</i> (hydrophilic and charged in neutral pH).	107
5.3	Distance of several components of the system from the middle of the bilayer for eight simulations. From these figures one can see which terminal touches the membrane first and which residues go deeper into the bilayer. For more information, see Table 5.1. . .	109
5.4	Distance between several components of the system from the middle of the bilayer for eight simulations. From these figures one can see which residue touches the membrane first and which residues go deeper into the bilayer. For more information, see Table 5.1.	110
5.5	Number of H-bonds as a function of time for simulations #1 and #4. The donor groups are ARG, LYSH, GLN, ASN, and TRP residues. The acceptor groups are the C=O and PO_4^- groups of DPPC lipids.	112
5.6	z -coordinates of nitrogen atoms of neighboring (<i>black</i>) and distant (<i>red</i>) lipids in simulations #1 (a), and #4 (b). In (a) the peptide is associated with the lower monolayer, and in (b) the peptide is associated with the upper monolayer.	113
5.7	Membrane thickness for simulations #1 (a), and #4 (b), computed for the neighbor (<i>black</i>) and distant (<i>red</i>) lipids.	113

5.8 Molecular order parameter of *sn-1* (a) and *sn-2* (b) acyl chains. The associated monolayer is shown in *black* and the free monolayer is shown in *red*. The presence of penetratin induces disorder in *sn-1* acyl chains. 114

5.9 Evolution of secondary structures of transportan as defined by DSSP [68]. The residues start from ARG1 (N-terminus) to LYSH16 (C-terminus). Different colors correspond to different secondary structures as shown in legends. 115

List of Tables

1.1	Major lipid components of selected biomembranes. PC = phosphatidylcholine; PE = phosphatidylethanolamine; PS = phosphatidylserine; SM = sphingomyelin. The data are presented in mol% and are obtained from Ref. [87].	6
1.2	Amino acids with their 3-letter and 1-letter abbreviations.	8
3.1	Comparison of different properties of pure POPC, CHOL and DHE.	61
3.2	Hydrogen bonds and charge-pair interactions per sterol in the membrane/water interface. Atom numbering is shown in Fig. 3.1.	72
4.1	Details of different simulations. Insertion depth is defined to be the distance between the center of mass of the peptide and phosphorus atoms. N-terminus corresponds to GLY1 residue and C-terminus corresponds to LEU27 residue. Insertion depth is defined to be the average distance between the center of mass of the peptide and phosphorus atoms within the indicated time interval.	86
4.2	Area per lipid and thickness of the bilayer when transportan is directly inserted in the bilayer's core.	94
5.1	Details of different simulations. N-terminus corresponds to ARG1 residue and C-terminus corresponds to LYSH16 residue. Insertion depth is defined to be the average distance between the center of mass of the peptide and phosphorus atoms within the indicated time interval. A negative insertion depth means that the peptide has not penetrated. When it is zero, it means that the the penetration is about to occur. A positive insertion depth means that the peptide has penetrated.	111

List of Abbreviations

DPPC	dipalmitoyl phosphatidylcholine
POPC	palmitoyl-oleoyl phosphatidylcholine
DMPC	Dimyristoyl phosphatidylcholine
PC	phosphatidylcholine
PE	phosphatidylethanolamine
PG	phosphatidylglycerol
PS	phosphatidylserine
SM	sphingomyelin
CHOL	cholesterol
DHE	dehydroergosterol
ERG	ergosterol
DCHOL	demethylated cholesterol
DESMO	desmosterol
7-DHC	7-dehydrocholesterol
CPP	cell-penetrating peptide
DSC	Differential Scanning Calorimetry
RDF	Radial Distribution Function
GROMACS	GRONingen MACHine for Chemical Simulations
RESP	Restrained ElectroStatic Potential
SPC	SimPle Charge model (for water molecules)
PMF	Potential of Mean Force
OPLS	Optimized Potentials for Liquid Simulations
LINCS	LINear Constraint Solver (for molecular simulations)
FFT	Fast Fourier Transform
MD	Molecular Dynamics
MC	Monte Carlo
NMR	Nuclear Magnetic Resonance
PME	Particle Mesh Ewald
P ³ M	Particle-Particle Particle-Mesh
NPT	systems of constant particle Number, Pressure and Temperature
NVT	systems of constant particle Number, Volume and Temperature
H-Bond	hydrogen bond
COM	Center Of Mass

Chapter 1

Biological Background

1.1 Cell and Membrane

Cells are the structural and functional units of all living organisms. There are about 200 cell types in nature [16] which can be divided into two general categories: *eukaryotes* and *prokaryotes* [2]. Cells in plant, animal, and fungi are eukaryotic. Bacteria are prokaryotes. In eukaryotes, DNA is packed inside a space called the nucleus, a membrane-enclosed organelle. DNA in prokaryotes is not encapsulated in a specific compartment. Pictures of eukaryotes and prokaryotes cell are shown in Fig. 1.1.

Inside the cell is the cytoplasm where the cytosol provides an aqueous environment dissolving biomolecules such as enzymes, messenger RNA, amino acids, nucleotides and metabolites. The content of the cell is encapsulated with the plasma membrane (Fig. 1.1).

The plasma membrane serves as a barrier that separates the contents of the cell from its surroundings. The organelles of almost all other non-bacterial cells are also encapsulated with membranes. Membranes have remarkable physical properties of being flexible, self-sealing, and selectively permeable to specific polar solutes [107, 106]. Their flexibility permits shape changes essential to cellular processes such as cell growth. Also, due to their ability to rupture and reseal without inducing gross leaks, two membrane systems can fuse, as in exocytosis, and a membrane system can be formed through budding from another one, as in endocytosis and cell division.

1.1.1 Historical background

The name "cell" is attributed to Robert Hooke as he named the small compartments he observed in cork tissue in 1665 [58]. In 1855, Carl Nägeli realized the existence of a semipermeable layer around plant cells and then called it the plasma membrane [157]. Between 1895 and 1902, Ernest Overton managed to measure cell membrane permeability for many compounds

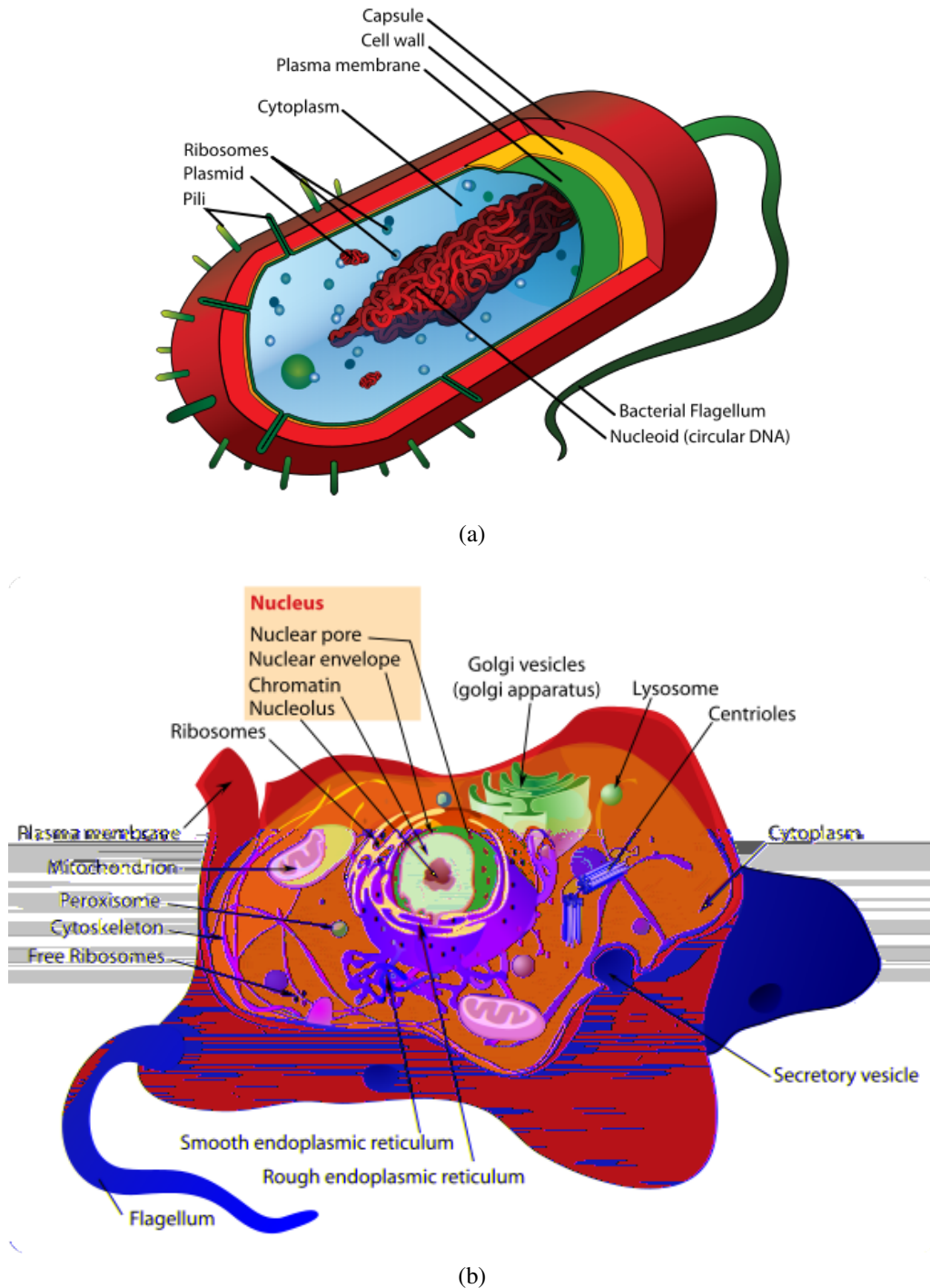


Figure 1.1: (a) Cell structure of a bacterium which is prokaryotic. (b) Structure of a typical animal cell which is eukaryotic. The pictures have been released into the public domain. Sources: wikipedia:prokaryote (a) and wikipedia:eukaryote (b).

and proposed that it is the lipids that form the thin film of membranes through which some substances can pass the cell. In 1925, Evert Gorter and F. Grendel extracted lipids from mem-

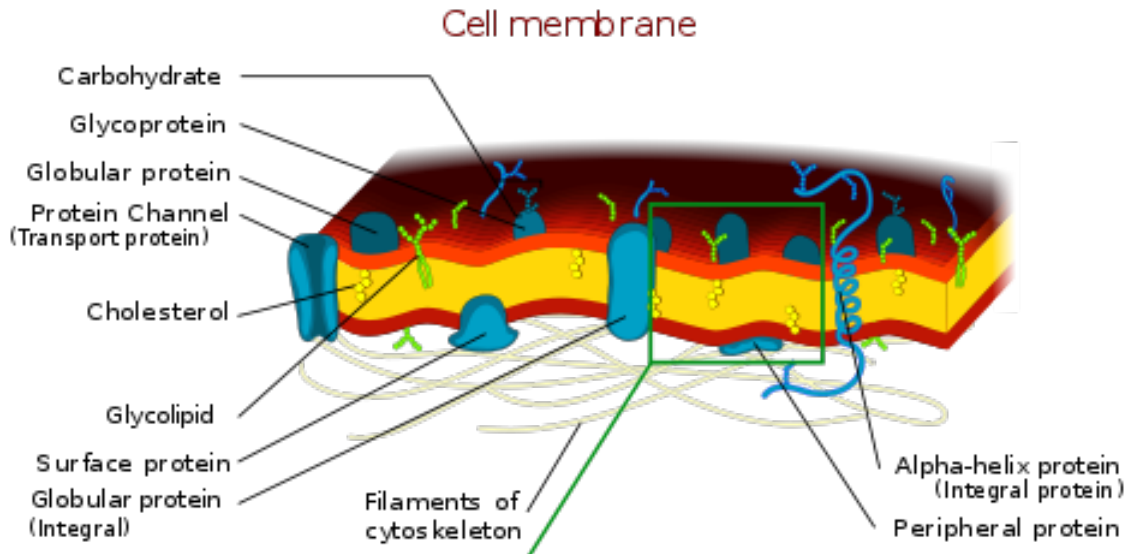


Figure 1.2: Fluid mosaic model for membrane structure. The membrane lipids form a fluid 2D solvent where the integral proteins float laterally in the plane of the bilayer. The picture has been released into the public domain. Source: wikipedia:cell membrane.

branes and proposed that lipids in membranes are arranged in a double layer. The progress in new experimental techniques such as electron microscopy led to a better view of membranes. In 1972, S. Jonathan Singer and Garth L. Nicolson proposed the famous fluid mosaic model [140] (Fig. 1.2), which describes the membrane as a fluid-like lipid bilayer in which embedded globular proteins diffuse freely. This model has proven to be very useful in explaining many phenomena taking place in biomembranes such as the rapid diffusion of proteins. The *dynamically structured mosaic model* [155] is a modified version of the fluid mosaic model proposed by Vereb *et al.* In this model, non-random co-distributed patterns of specific kind of proteins are supported to form small-scale clusters at the molecular level and large-scale clusters at the submicrometer level. Compared to fluid mosaic model, this model gives a more realistic description of the structure and functional properties of biomembranes.

1.1.2 Lipid bilayer

Lipids constitute about 50% of the mass of membranes in animal cells, while the rest of the mass is mainly proteins [107]. In the most general description, a lipid consists of a polar head group and two tails (Fig. 1.3). Most lipids in cell membranes are amphiphilic since they have a charged or polar hydrophilic headgroup and one or two hydrophobic tails.

In an aqueous environment, the hydrophilic head group tends to interact with water molecules to form hydrogen bonds. They are energetically favorable since the strength of hydrogen bonds

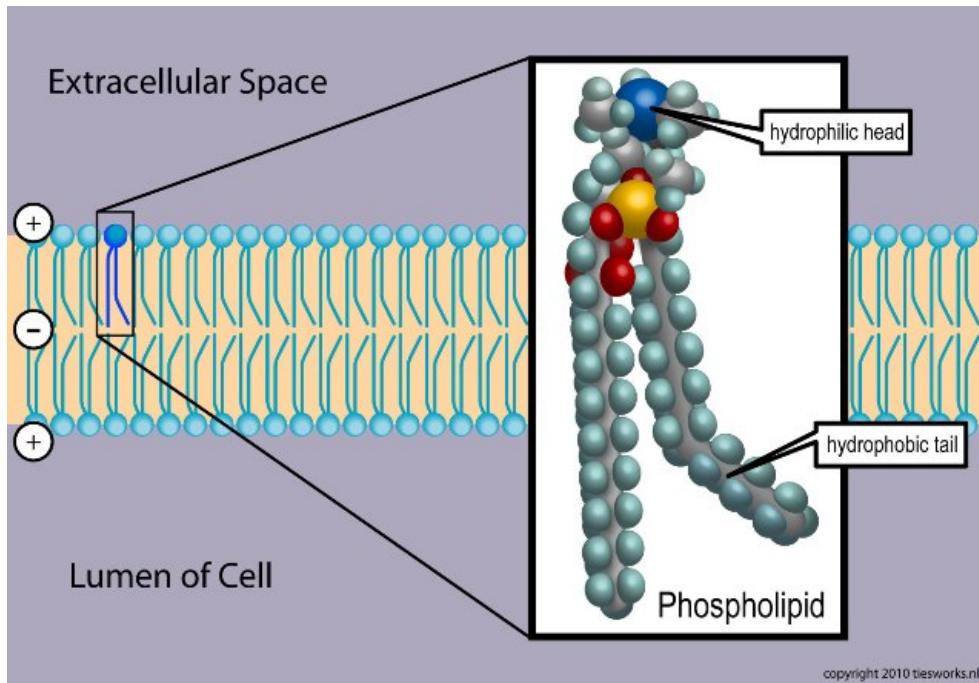


Figure 1.3: A schematic picture of a lipid and a bilayer formed by assembling of lipids. The picture has been released into the public domain. Source: wikipedia:phospholipid.

is typically 10-40 kJ/mol [63], which is larger than the thermal energy being ≈ 4 kJ/mol at room temperature. The hydrophobic tails prefer to avoid the water molecules and to form a hydrocarbon core in which the van der Waals interactions are maximized in the presence of entropic motion. Depending on their shape, lipids can either form spherical micelles, with tails pointing inward, or they can form a bilayer, which is the preferable structure for lipids with a cylindrical shape.

Membrane lipids can be generally divided into three categories [87]. These are *phosphoglycerides* or *phospholipids*, *sphingolipids*, and *cholesterol*, as shown in Fig. 1.4.

Phospholipids (Fig. 1.4a) are the abundant membrane lipids. Most phosphoglycerides are derivatives of glycerol 3-phosphate containing two esterified fatty acyl chains, constituting the hydrophobic tail or acyl chains and a polar head group esterified to the phosphate. The fatty acids can vary in length and be saturated (no double bonds) or unsaturated (one, two, or three double bonds). In phosphatidylcholine (PC), the head group is choline. In Fig. 1.4a three other common phosphoglycerides are shown: phosphatidylethanolamine (PE), phosphatidylserine (PS), and phosphatidylglycerol (PG). Among these lipids, PS and PG are negatively charged while others are neutral but having dipolar head group and are often called zwitterionic. Two PC lipids that are involved in this thesis are dipalmitoyl phosphatidylcholine (DPPC) and palmitoyl-oleoyl phosphatidylcholine (POPC). DPPC has two acyl chains each of which being

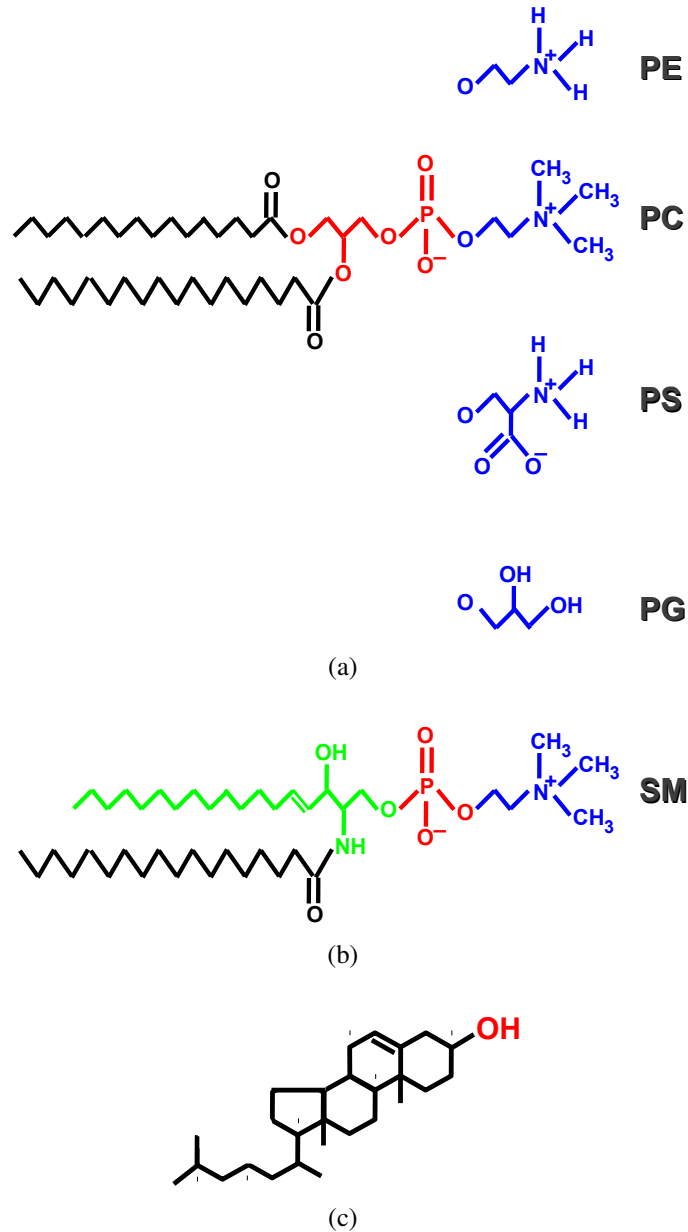


Figure 1.4: Three classes of membrane lipids: phosphoglycerides (a), sphingolipids (b), and cholesterol (c). Examples of phosphoglycerides include phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), and phosphatidylglycerol (PG). Sphingomyelin (SM) is an example of a sphingolipids.

fully saturated (without any double bond) and having 16 carbon atoms. POPC has two acyl chains one of which being unsaturated (with a double bond) having 18 carbon atoms, and the other being fully saturated having 16 carbon atoms. The structure of these PC lipids will be shown in Chapters 3 and 4.

Sphingolipids (Fig. 1.4b) are derivatives of sphingosine, an amino alcohol with a long hy-

Table 1.1: Major lipid components of selected biomembranes. PC = phosphatidylcholine; PE = phosphatidylethanolamine; PS = phosphatidylserine; SM = sphingomyelin. The data are presented in mol% and are obtained from Ref. [87].

Source/Location	PC	PE + PS	SM	Cholesterol
Plasma membrane (human erythrocytes)	21	29	21	26
Myelin membrane	16	37	13	34
Plasma membrane (<i>E. coli</i>)	0	85	0	0
Endoplasmic reticulum membrane (rat)	54	26	5	7
Golgi membrane (rat)	45	20	13	13
Inner mitochondrial membrane (rat)	45	45	2	7
Outer mitochondrial membrane (rat)	34	46	2	11
Primary leaflet location	Exoplasmic	Cytosolic	Exoplasmic	both

drocarbon chain. Various fatty acyl chains are connected to sphingosine by an amide bond. The sphingomyelins (SM), which contain a phosphocholine head group, are phospholipids. Other sphingolipids are glycolipids in which a single sugar residue or branched oligosaccharide is attached to the sphingosine backbone. For instance, the simple glycolipid glucosylcerebroside (GlcCer) has a glucose head group.

Eukaryotic plasma membranes usually contain a large amount of cholesterol. As shown in Fig. 1.4c, cholesterol has a polar hydroxyl group, a rigid steroid ring as its middle part and a non-polar hydrocarbon tail. In a lipid bilayer, the hydroxyl groups of cholesterol molecules stay close to the head groups of lipid molecules with hydrocarbon tails within the hydrocarbon core of the lipid bilayer. With such an orientation, their rigid steroid structures partly immobilize the upper parts of their neighboring lipids so that this region of lipid bilayer is less deformable and permeable to small water-soluble molecules. Although cholesterol tends to reduce the fluidity of membranes, at high concentrations it also prevents the hydrocarbon chains in lipid components from packing together and undergoing phase transition to the gel phase. In this way, it inhibits possible phase transitions and regulates the fluidity of membranes [107].

A typical cell contains myriad types of membranes, each with unique properties bestowed by its particular mix of lipids and proteins. The data in Table 1.1 illustrate the variation in lipid composition among different biomembranes. Several phenomena contribute to these differences. For instance, differences between membranes in the endoplasmic reticulum (ER) and the Golgi are largely explained by the fact that phospholipids are synthesized in the ER, whereas sphingolipids are synthesized in the Golgi. As a result, the proportion of sphingomyelin as a percentage of total membrane lipid phosphorus is about six times as high in Golgi membranes as it is in ER membranes. In other cases, the translocation of membranes from one cellular compartment to another can selectively enrich membranes in certain lipids.

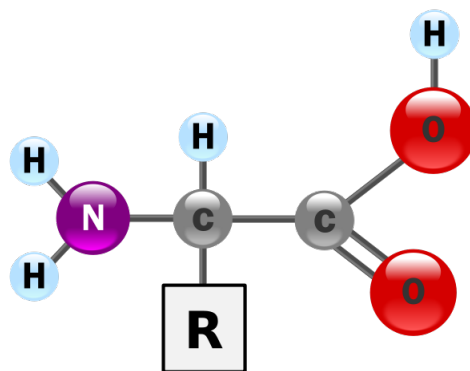


Figure 1.5: The generic structure of an amino acid. Different amino acids differ in which side chain, or R group, is attached to them. The nitrogen atom together with two hydrogen atoms are called amino group. The carbon atom together with the oxygen atoms are called carboxyl group. The carbon atom attached to the R group is called an α -carbon. The picture has been released into the public domain. Source: wikipedia:amino acid.

1.2 Proteins and peptides

Proteins are essential parts of organisms and participate in virtually every process within cells. The name “protein” was first proposed by Gerardus Johannes Mulder in 1839 [102]. In his paper he claimed that “The organic substances which are present in all constituents of the animal body, also as we shall soon see, in the plant kingdom, could be named *protein*”. Each protein polymer consists of a sequence formed from 20 amino acids (see below). For chains under 40 residues the term *peptide* is frequently used instead of protein. Protein structures range in size from ten to several thousand residues [19].

1.2.1 Amino acids

As mentioned above, amino acids are building blocks of proteins and peptides. The key elements of an amino acid are carbon, hydrogen, oxygen, and nitrogen. The generic structure of an amino acid is shown in Fig. 1.5. The *amino* group (the nitrogen together with two bonded hydrogen atoms) is attached to the *carboxyl* group (the carbon together with its two bonded oxygen atoms) through a carbon atom called α -carbon. The various amino acids differ in which *side chain* (R group in Fig. 1.5) is attached to their α -carbon, and can vary in size from just one hydrogen atom in glycine to a large heterocyclic group in tryptophan.

Amino acids in solution at neutral pH exist predominantly as dipolar ions (also called *zwitterions*). In the dipolar form, the amino group is protonated ($-\text{NH}_3^+$) and the carboxyl group is deprotonated ($-\text{COO}^-$). The ionization state of an amino acid varies with pH. In acidic

Table 1.2: Amino acids with their 3-letter and 1-letter abbreviations.

amino acids	3-letter name	1-letter name
alanine	ALA	A
arginine	ARG	R
asparagine	ASN	N
aspartic acid	ASP	D
cysteine	CYS	C
glutamic acid	GLU	E
glutamine	GLN	Q
glycine	GLY	G
histidine	HIS	H
isoleucine	ILE	I
leucine	LEU	L
lysine	LYS	K
methionine	MET	M
phenylalanine	PHE	F
proline	PRO	P
serine	SER	S
threonine	THR	T
tryptophan	TRP	W
tyrosine	TYR	Y
valine	VAL	V

solution (e.g., pH 1), the amino group is protonated ($-\text{NH}_3^+$) and the carboxyl group is not dissociated ($-\text{COOH}$). As the pH is raised, the carboxylic acid is the first group to give up a proton [14]. The dipolar form persists until the pH approaches 9, when the protonated amino group loses a proton. In conditions corresponding to neutral pH, i.e. pH 7, three amino acids have positively charged side chains, namely arginine and lysine and histidine. The latter is conditional and, depending on its environment, can be charged or neutral. Two amino acids have negatively charged side chains, namely aspartic acid and glutamic acid. The rest have a net charge of zero, but their side chains can be either polar or non-polar.

Twenty-two amino acids are naturally incorporated into polypeptides and are called proteinogenic or natural amino acids [28]. Of these, 20 are encoded by the universal genetic code. The remaining two, *selenocysteine* and *pyrrolysine*, are incorporated into proteins by unique synthetic mechanisms. Amino acids are often designated by either a three-letter abbreviation or a one-letter symbol. These abbreviations and symbols are an integral part of the vocabulary of biochemists. Table 1.2 shows a list of all 20 amino acids.

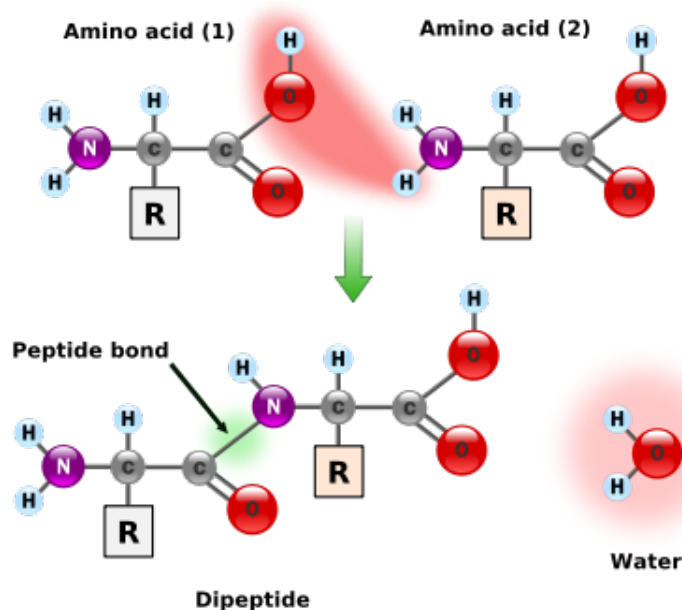


Figure 1.6: Peptide bond occurring between two amino acids. Each amino acid unit is called a residue. The repeating part of the polypeptide, i.e. everything except the R groups, is called the backbone or the main chain. At neutral pH, the carboxyl group (C-terminus) loses one proton, making it negatively charged, while the amino group (N-terminus) accepts one proton, making it positively charged. Thus the whole polypeptide is polar. The picture has been released into the public domain. Source: wikipedia:amino acid.

1.2.2 Peptide bond

Peptides are formed by linking of amino acids. Two amino acids can covalently bond to each other when the carboxyl group of one molecule reacts with the amino group of the other molecule, causing the release of a water molecule (H_2O). The resulting C-N bond is called a peptide bond. Figure 1.6 shows the chemical reaction leading to formation of a peptide consisting of two amino acids.

Each amino acid unit in a polypeptide is called a *residue*. A polypeptide chain consists of a regularly repeating part, called the *main chain* or *backbone*, and a variable part, comprising the distinctive side chains. A polypeptide chain has polarity because its ends are different, with an amino group at one end and a carboxyl group at the other, and these ends are positively and negatively charged respectively at neutral pH. By convention, the amino end is taken to be the beginning of a polypeptide chain, and so the sequence of amino acids in a polypeptide chain is written starting with the amino terminal residue. The amino terminal residue is called the *N – terminus* and the other end of the polypeptide, i.e. the carboxyl terminal is called the *C – terminus*.

Examination of the geometry of the protein backbone reveals several important features.

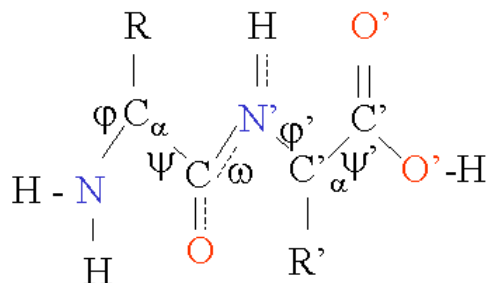


Figure 1.7: Illustration of dihedral angles ϕ and ψ . The angle of rotation about the bond between the nitrogen and the α -carbon atoms is called phi (ϕ). The angle of rotation about the bond between the α -carbon and the carbonyl carbon atoms is called psi (ψ). The picture has been released into the public domain. Source: wikipedia:protein structure.

First, the peptide bond is essentially planar. Thus, for a pair of amino acids linked by a peptide bond, six atoms lie in the same plane: the α -carbon atom and CO group from the first amino acid and the NH group and α -carbon atom from the second amino acid. The nature of the chemical bonding within a peptide explains this geometric preference. The peptide bond has considerable *double-bond character*, which prevents rotation about this bond. The inability of the bond to rotate constrains the conformation of the peptide backbone and accounts for the bond's planarity.

The double-bond character of the peptide-bond is also expressed in the length of the bond between the CO and NH groups. The C-N distance in a peptide bond is typically 1.32 Å [14], which is between the values expected for a C-N single bond (1.49 Å [14]) and a C=N double bond (1.27 Å [14]). Therefore, it is better to say that the peptide bond has *partial* double-bond character, and the π electrons of the carbonyl group are delocalized and are shared with the electrons of the amide nitrogen.

In contrast to the peptide bond, the bonds between the amino group and the α -carbon atom and between the α -carbon atom and the carbonyl group are pure single bonds. The two adjacent rigid peptide units may rotate about these bonds, taking on various orientations. This freedom of rotation about two bonds of each amino acid allows proteins to fold in many different ways. The rotations about these bonds can be specified by dihedral angles (Fig. 1.7). The angle of rotation about the bond between the nitrogen and the α -carbon atoms is called phi (ϕ). The angle of rotation about the bond between the α -carbon and the carbonyl carbon atoms is called psi (ψ). Therefore, a polypeptide chain may be considered as a series of planes with two angles of rotation between each plane.

Not all combinations of ϕ and ψ are possible. G. N. Ramachandran recognized that many combinations are forbidden because of steric collisions between atoms, that is, because two

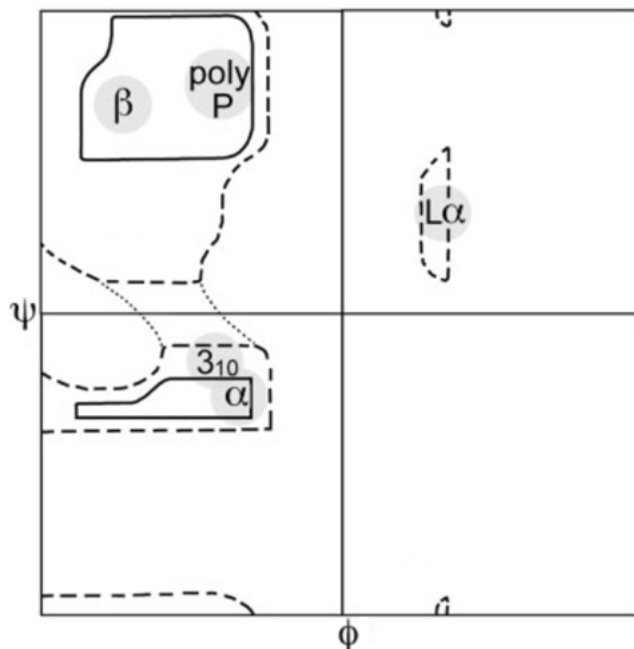


Figure 1.8: Ramachandran diagram shows that in a polypeptide, because two atoms cannot be in the same place at the same time, not all ϕ and ψ values are possible. The most favorable regions are shown; The picture is licensed under the Creative Commons license and can be found at http://en.wikipedia.org/wiki/File:Ramachandran_plot_original_outlines.jpg.

atoms cannot be in the same place at the same time. The allowed values can be visualized on a two-dimensional plot called a Ramachandran diagram (Fig. 1.8). Three quarters of the possible (ϕ , ψ) combinations are excluded simply by local steric clashes .

Finally, the polypeptide backbone is rich in hydrogen-bonding potential. Each residue contains a carbonyl group, which is a good hydrogen-bond acceptor and an NH group, which is a good hydrogen-bond donor. These hydrogen bonds are responsible for the secondary structures of the polypeptides such as α -helix and β -sheet which will be discussed in the next chapter.

1.3 The structure of this thesis

In this thesis, we employ atomistic molecular dynamics simulations to study the interaction of lipid membranes with two kinds of sterols, namely cholesterol (CHOL) and dehydroergosterol (DHE), as well as two cell-penetrating peptides, namely transportan and penetratin. We aimed to understand the underlying physics behind relevant biological systems at microscopic levels. In the first study, we study pure POPC, CHOL-POPC and DHE-POPC lipid bilayers to compare

the effect of CHOL with that of DHE on lipid bilayers. This is the first computational study of DHE interacting with a model bilayer. In the second project, we study the interaction of transportan with a DPPC bilayer to understand the possible mechanism of penetration of such a peptide through membranes. This is also the first study of transportan interacting with a model membrane. In the third project, we study the interaction of penetratin with a DPPC bilayer. We mostly focus on the binding of this peptide with the membrane and compare the results with those obtained for transportan.

This thesis is organized as follows. Chapter 1 gives introduction to biological background of membranes as well as peptides. Chapter 2 gives introduction to physical background of the thesis, ranging from statistical mechanics to peptide conformations and membrane transitions. Chapter 3 presents the first original research of this thesis on the interaction of sterols with POPC lipid bilayers. Chapter 4 presents the second research of this thesis on the interaction of transportan with DPPC lipid bilayers. Chapter 5 presents the third research of this thesis on the interaction of penetratin with DPPC lipid bilayers. In all these three chapters, we start by presenting the motivation of the study, then we proceed by describing the simulation parameters and forcefields, then we present the results and discussion, and finally we summarize the study with a conclusion. In Chapter 6, we give a conclusion to the whole thesis and propose the possible future works.

Chapter 2

Physical Background

2.1 Brief review of thermodynamics and statistical mechanics

Statistical mechanics connects the microscopic properties of a system to macroscopic properties, such as temperature, pressure, etc., which can be measured in experiments. Computer simulations provide us with detailed microscopic information about a system, and hence, there is a need to use statistical mechanics to infer macroscopic information from computer simulations. Here, we explain briefly some important aspects of statistical mechanics as well as thermodynamics which are relevant to our study. For more comprehensive discussions see Refs. [113, 122, 59].

2.1.1 Thermodynamic equilibrium

A system should fulfill the following conditions in thermodynamic equilibrium with its environment:

1- The net force \vec{F}_i and torque $\vec{\tau}_i$ on it is zero. This is called mechanical equilibrium and can be written as

$$\sum_{i=1} \vec{F}_i = 0, \quad (2.1)$$

and

$$\sum_{i=1} \vec{\tau}_i = 0. \quad (2.2)$$

2- Its temperature is the same as the temperature of the environment. This is called thermal

equilibrium and can be written as

$$T_{\text{system}} = T_{\text{environment}} \quad (2.3)$$

3- The chemical potentials of all substances μ_i in the system are constant in time and space. This is called chemical equilibrium and can be written as

$$\mu_i = \text{const.} \quad (2.4)$$

Put it in a more thermodynamical context, in an equilibrium state, relevant thermodynamic functions of the whole system should be an extremum. For example, a system with a constant number of particles and at a constant volume and temperature evolve in such a way that it minimizes the Helmholtz free energy A defined as

$$A(N, V, T) = U - TS, \quad (2.5)$$

where U is the internal energy of the system and S is the entropy. The minimum of A corresponds to the equilibrium state of the system. For a system with a constant number of particles and at constant pressure and temperature, the relevant thermodynamic potential is the Gibbs free energy G which is defined as

$$G(N, p, T) = U - TS + pV. \quad (2.6)$$

Similarly, this system evolves in such way that it minimizes the Gibbs free energy.

The process in which the system with a given initial state reaches equilibrium is called *thermalization* or *equilibration*. After thermalization, the observables of the system sample a stationary distribution. In MD simulations, to determine whether the system is in equilibrium or not, one should examine those properties of the system which are slow to converge to a steady-state value. After the convergence of these quantities, it is assured that the system has passed the thermalization. In the study of membrane-water systems, for example, one can measure the value of the area per lipid or the number of hydrogen bonds. These quantities converge slowly.

2.1.2 Classical ensemble theory

Consider a system consisting of N classical particles which is in thermodynamic equilibrium with its environment. The generalized coordinate and velocity of each particle are (q_1, q_2, q_3) and (p_1, p_2, p_3) , respectively. We denote the coordinates and momenta of all the particles by

q_1, q_2, \dots, q_{3N} and p_1, p_2, \dots, p_{3N} , respectively. The indexes run from 1 to $3N$ because the particles are described by three dimensional coordinates and velocities. Thus, these parameters ($\{q_i, p_i\}$) evolve with time within a $6N$ dimensional space called *phase space*. A point in the phase space represents a state of the entire N -particle system, and is referred to as the *representative point*. In other words, a representative point corresponds to a specific set of $\{q_i, p_i\}$ called a *microstate*. An *ensemble* consists of all possible microstates available to the system.

The phase space may conveniently be described by a density function $\rho(p, q, t)$, where (p, q) is an abbreviation for $(q_1, \dots, q_{3N}; p_1, \dots, p_{3N})$, so that

$$\rho(p, q, t) d^{3N} q d^{3N} p$$

is the number of representative points that at time t are contained in the infinitesimal volume element $d^{3N} q d^{3N} p$ of the phase space centered about the point (p, q) . An ensemble is completely specified by $\rho(p, q, t)$.

Given $\rho(p, q, t)$ at any time t , its subsequent values are determined by the dynamics of molecular motion. Let the Hamiltonian of a system in the ensemble be $\mathcal{H}(q_1, \dots, q_N; p_1, \dots, p_N)$. The equations of motion for a system are given by

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \quad (i = 1, \dots, 3N), \quad (2.7)$$

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad (i = 1, \dots, 3N). \quad (2.8)$$

These will tell us how a representative point moves in phase space as time evolves.

Liouville's Theorem Liouville's theorem states that if we follow the motion of a representative point in the phase space, we find that the density of representative points in its neighborhood is constant. This implies that the volume element in generalized coordinates, i.e. $d^{3N} q d^{3N} p$ is invariant under a Hamiltonian or canonical transformation. Mathematically, Liouville's theorem is

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{d\rho}{dp_i} \dot{p}_i + \frac{d\rho}{dq_i} \dot{q}_i \right) = 0. \quad (2.9)$$

The observed value of a dynamical quantity O of the system, which is generally a function of the coordinates and momenta, is supposed to be its averaged value taken over a suitably chosen ensemble:

$$\langle O \rangle = \frac{\int d^{3N} q d^{3N} p O(p, q) \rho(p, q, t)}{\int d^{3N} q d^{3N} p \rho(p, q, t)} \quad (2.10)$$

This is called the ensemble average of O . Its time dependence comes from that of ρ , which is governed by Liouville's theorem.

There are three common ensembles in statistical mechanics namely the *canonical*, *micro-canonical* and *grand canonical* ensemble.

In the Microcanonical ensemble, the macrostate of a system is defined by the number of particles N , volume V , and energy E . This ensemble describes an isolated system that does not exchange energy/matter with its environment. For this ensemble, if we calculate the number of distinct microstates $\Omega(N, V, E)$, then the phase space density $\rho(q, p)$ can be obtained as

$$\rho(q, p) = \frac{1}{\Omega} \delta(\mathcal{H}(q, p) - E), \quad (2.11)$$

where δ is the Dirac delta function and \mathcal{H} is the Hamiltonian. The Microcanonical ensemble is the most natural ensemble for MD, since in integrating Newton's equation of motion the energy is conserved.

In physical experiments, controlling the energy E of a system is hard and we never deal with a completely isolated system. Therefore, the canonical ensemble is introduced which describes a closed system in contact with a heat bath at a constant temperature T , thus being able to exchange energy with the environment. The macrostate of the system is defined by the number of particles N , volume V , and temperature T . The equilibrium phase space density can be obtained as

$$\rho(q, p) = \frac{1}{Q_N(V, T)} \exp[-\mathcal{H}(q, p)/k_B T], \quad (2.12)$$

where $Q_N(V, T)$ is the canonical partition function, and k_B is the Boltzmann constant. The average value of a physical quantity $\langle O \rangle$ (averaged over the ensembles not over the time) is then

$$\langle O \rangle = \langle O \rangle_{\text{ens}} = \frac{\int O e^{-\mathcal{H}(q,p)/k_B T} d^{3N}q d^{3N}p}{\int e^{-\mathcal{H}(q,p)/k_B T} d^{3N}q d^{3N}p} \quad (2.13)$$

where O_{ens} is the ensemble average. The canonical ensemble is the natural ensemble for the *Monte Carlo* method since it generates ensembles according to a canonical distribution.

A physical system represented by a grand-canonical ensemble is in equilibrium with an external reservoir with respect to both particle and energy exchange. This is an extension of the canonical ensemble, but instead the grand canonical ensemble is allowed to exchange energy and particles with its environment. The chemical potential is introduced to specify the fluctuation of the number of particles, just as temperature is introduced into the canonical ensemble to specify the fluctuation of energy.

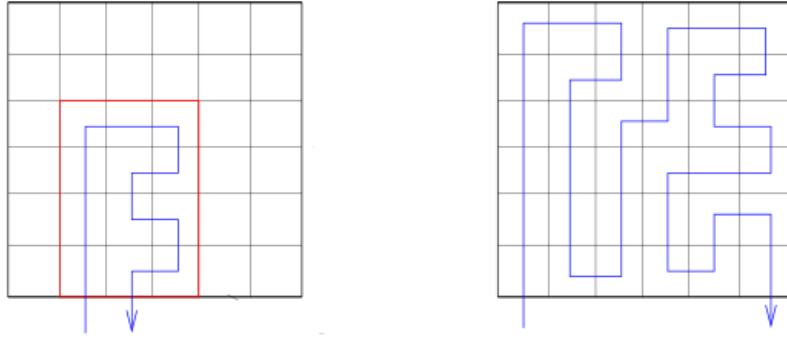


Figure 2.1: (*left*) A non-ergodic system in which the trajectory does not pass through all possible states in the phase space. The *red* line encases the sub-phase space of the non-ergodic trajectory. (*right*) An ergodic system in which the trajectory passes through all possible states in phase space.

2.1.3 Ergodicity

As mentioned in the previous section, any point in the $6N$ -dimensional classical phase space represents a microstate. Suppose the system is connected to a heat bath at temperature T . If we track the evolution of a point in such a system with time, i.e., if we look at its trajectory, any region in the phase space can be explored with a probability given by eq. 2.12. According to the *ergodic hypothesis*, if we wait for a very long time, the trajectory will eventually pass through all the regions of the phase space (Fig. 2.1). To understand the usefulness of ergodicity, suppose we would like to find the average of a quantity O under certain experimental conditions. Obviously, we cannot take advantage of eq. 2.13 as we do not have access to detailed information of microstates. When the ergodic hypothesis is true, we are allowed to make measurements for a long enough time and then calculate the average of A over that period of time. This is called the time average. In other words, according to ergodic hypothesis $\langle A \rangle_{\text{ens}}$ which is given by eq. 2.13 is equal to $\langle O \rangle_{\text{time}}$:

$$\langle O \rangle_{\text{ens}} =$$

$$\langle O \rangle_{\text{time}} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t O(t') dt'. \quad (2.14)$$

When the ergodic hypothesis is true, if one waits for a sufficiently long time, the trajectory of the representative point of a system will cover the entire accessible phase space. More precisely, the representative point comes arbitrarily close to any point in the accessible phase space.

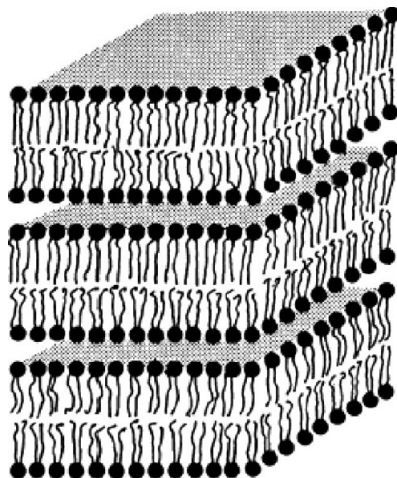


Figure 2.2: A stack of lipid bilayers forming a multi-bilayer. The picture is taken from Ref. [77] © Elsevier Limited.

2.2 Phase behavior of model membranes

In experiments, membranes are usually modeled *in vitro*. One of the most studied model membranes is lipid multi-bilayers. As the name indicates, and as shown in Fig. 2.2, the structure of these multi-bilayers is characterized as a stack of lipid bilayers. A lipid multi-bilayer is usually deposited on a solid substrate such as glass, silicon or mica. The hydration level is then controlled by exposing the sample to water [70]. By changing the hydration level and temperature, one can study the phase behavior of these systems. By utilizing different spectroscopic methods such as X-ray or neutron diffraction, structural properties of different phases can be explored.

There are several methods and techniques to study various physical and chemical changes of materials such as crystallization, melting, freezing, oxidation, etc. A well-known technique is *Differential Scanning Calorimetry* (DSC) which was developed in 1962 [158]. In this method, the difference in the amount of heat required to increase the temperature of a sample and reference is measured as a function of temperature. Both the sample and reference are maintained at nearly the same temperature throughout the experiment. The reference material should have a well-defined heat capacity over the range of temperatures to be scanned. As an example, when a solid sample melts to a liquid, it will require more heat flowing to the sample to increase its temperature at the same rate as the reference. This is due to the absorption of latent heat by the sample as it undergoes the phase transition from solid to liquid. By observing the difference in heat flow between the sample and reference, differential scanning calorimeters are able to measure the amount of heat absorbed during such a transition.

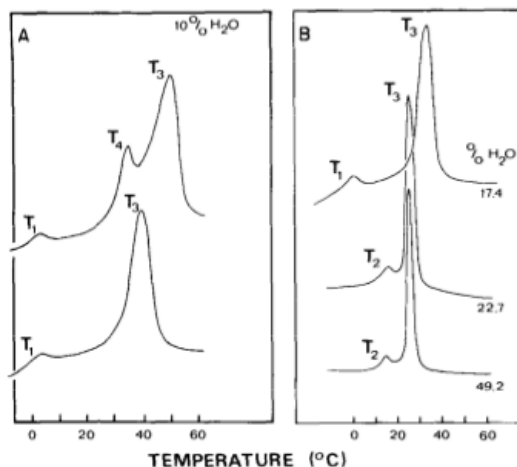


Figure 2.3: A typical DSC graph from an experiment on Dimyristoyl phosphatidylcholine (DMPC) lipids. Each curve corresponds to a different sample preparation and hydration. Temperatures at which the peaks occur correspond to different phase transitions. This research was originally published in Journal Name. Author(s). Title. *The Journal of Biological Chemistry*. 1979; 254:6068-6078. © the American Society for Biochemistry and Molecular Biology.

The result of a DSC experiment is a curve of heat flux versus temperature. In Fig. 2.3 such a curve is shown for Dimyristoyl phosphatidylcholine (DMPC) lipids as the sample material. The peaks show the temperatures at which a phase change happens.

Different phases categorized for DMPC lipids are shown in Fig. 2.4. At high temperatures and high hydration level, lipid multi-bilayers are in a *lamellar fluid phase* (L_{α}). The fluid phase is characterized by the rapid translational and rotational diffusion of lipids and highly disordered hydrocarbon core of the membranes. As temperature is lowered, a phase transition from the lamellar fluid phase (L_{α}) to a *pre-transition phase* ($P_{\beta'}$) occurs. This transition is called the *main transition* and for DPPC, DMPC and POPC it happens at $T_m = 41^{\circ}\text{C}$ [25], $T_m = 23.5^{\circ}\text{C}$ [77] and $T_m = -5^{\circ}\text{C}$ [138] respectively. The pre-transition phase, also known as the *ripple phase*, is characterized by nearly frozen hydrocarbon chains of lipids and the long-wavelength (10-20 nm) in-plane undulation of the bilayers (Fig. 2.4).

At much lower temperatures, the lipids are in a solid-like phase known as the *gel phase* ($L_{\beta'}$). For DPPC, this phase change happens at 35°C [25]. In the gel phase, the hydrocarbon chains of the lipids in the bilayer are frozen in the all-*trans* conformation and are highly ordered and the lipid chains make a certain angle with the bilayer normal (Fig. 2.4). This tilt angle, for example, for DPPC lipids at 25°C and 0% relative humidity is 21.5° [71]. In general, the multi-bilayers in the $L_{\beta'}$ and $P_{\beta'}$ phases usually develop hexagonal or distorted hexagonal symmetry depending on the amplitudes of the tilt angles of lipid tails.

In addition to the three major phases mentioned above, many more *sub-phases* have been

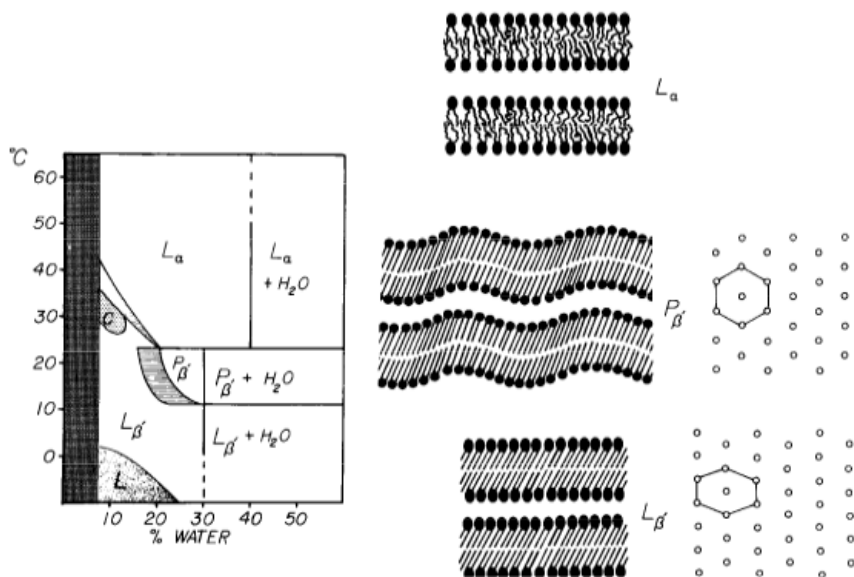


Figure 2.4: Temperature-composition phase diagram of hydrated DMPC multi-bilayers. The structures of different phases L_{α} , $L_{\beta'}$, and $P_{\beta'}$ are shown in the middle. The packing lattice (viewed from the top) is hexagonal for the $P_{\beta'}$ and distortedly hexagonal for the $L_{\beta'}$ phase. The picture is taken from Ref. [77] © Elsevier Limited.

identified. For DPPC, a phase change at 18°C has been revealed [25]. For DMPC, X-ray measurements have revealed the existence of three distinct sub-gel phases within the major $L_{\beta'}$ phase [141]. These sub-gel phases are distinguished by the direction of the chain tilt with respect to the position of the neighboring lipids within the same layer.

The phase transition temperature for single-lipid bilayers depends on several factors:

1- Number of hydration water molecules: Lipid head groups become hydrated when in contact with water. In this process, similar to any solvation or hydration process, they accommodate water molecules in the *hydration shell*. It is known that the phase transition temperature T_m decreases as the amount of water in bilayer-water systems (water concentration) increases. T_m , however, does not decrease indefinitely, but reaches a limiting value independent of the water concentration. At this water concentration, the bilayer is said to be fully hydrated.

2- Fatty acid length: Lipids with longer tails have a higher phase transition temperature. With longer lipid tails, van der Waals interactions are higher and the lipids are more condensed. This leads to a higher phase transition temperature. For instance, DPPC and DMPC have 16 and 14 carbon atoms in their hydrocarbon chains respectively, and their head groups are identical. The phase transition temperature for DPPC is $T_m = 41^{\circ}\text{C}$ [25], while for DMPC it is $T_m = 23.5^{\circ}\text{C}$ [77].

3- The presence of double bonds in the fatty acid chains of the lipids: It is now well known

that the presence of a *cis* double bond in the hydrocarbon chain of lipids can considerably affect the phase transition temperature. An increase in the number of double bonds in the chain causes a lowering of the lipid transition temperature. For example, POPC and DPPC have the same structure in their head groups. They only differ in their tails as DPPC has 16 carbon atoms in both tails, while POPC has 16 carbon atoms in one tail and 18 carbon atoms in the other. The fact that POPC is longer may lead to the wrong assumption that it has a higher T_m . But this is not the case. One of POPC's tails is not fully saturated and contains a double bond. This leads to a much lower, relative phase transition temperature ($T_m = -5^\circ\text{C}$ [138]) when compared to DPPC. The *cis* characteristic of the double bond reduces the inter-lipid contacts, and as a result, lowers the van der Waals interaction between adjacent lipid tails, leading to a more fluid lipid phase.

2.3 Protein conformation

When talking about protein or peptide structure, we often hear different terminologies. The structure of these molecules is frequently decomposed into primary structure, secondary structure, tertiary structure, and quaternary structure. These structures are shown in Fig. 2.5. Primary structure refers to the linear amino acid sequence of a polypeptide. Secondary structure is the local structure of linear segments of the polypeptide backbone atoms without regard to the conformation of the side chains. Super secondary structure (motif) is the associations of secondary structural elements through side chain interactions. Tertiary structure is the three-dimensional arrangement of all atoms in a single polypeptide chain. Quaternary structure is the arrangement of separate polypeptide chains (subunits) into the functional protein.

α -helix, *β -sheet*, and *turn* are three common secondary structures in proteins. Those which cannot be classified as one of the standard three classes are usually grouped into a category called "other" or "random coil" (*flexible chain*). One of these secondary structures, namely Alpha helix, is shown in Fig. 2.6. We briefly explain the helices as they are the most prevalent and also relevant to the thesis.

In a helical conformation, the relationship of one peptide unit to the next is the same for all α -carbons (See section 1.2.1 for the definition of C_α). This means that the dihedral angle pairs ϕ and ψ are the same for each residue in the helical conformation. For the geometrically ideal, right-handed, α -helix, these values are $\phi = -57.8^\circ$ and $\psi = -47.0^\circ$. Helices are often designated by the number of residues per helical turn and the number of atoms in one hydrogen-bonded ring.

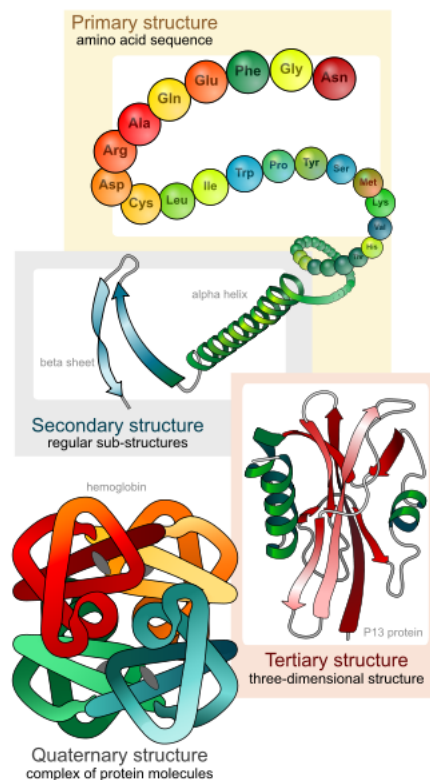


Figure 2.5: Primary, secondary, tertiary and quaternary structures of proteins. The picture has been released into the public domain. Source: wikipedia:protein structure.

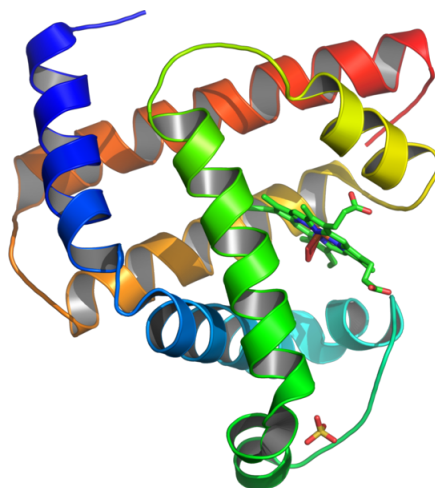


Figure 2.6: A representation of the 3D structure of the myoglobin protein. Alpha helices are shown in colour. The picture has been released into the public domain. Source: wikipedia:myoglobin.

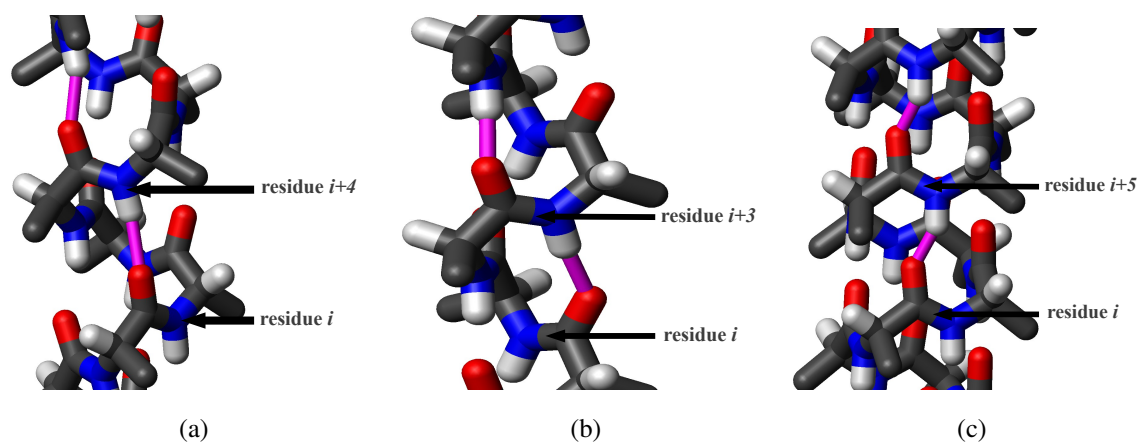


Figure 2.7: Side view of an α -helix (a), a 3_{10} -helix (b), and a π -helix of alanine. The hydrogen bonds are shown in *magenta*. Oxygens are in *red*, hydrogens in *white*, carbons in *gray* and nitrogens in *blue*. The pictures are licensed under the Creative Commons license.

2.3.1 α -helix

A common motif in the secondary structure of proteins, the α -helix, is a right-handed spiral conformation, in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier ($i + 4 \rightarrow i$). This structure is illustrated in Fig. 2.7a for alanine. Each amino acid residue corresponds to a 100° turn in the helix (i.e., the helix has 3.6 residues per turn), and a translation of 1.5 \AA along the helical axis. Therefore, the pitch of the α -helix (the vertical distance between one consecutive turn of the helix) is 5.4 \AA which is the product of 1.5 \AA and 3.6. α -helix has 10 atoms (including the hydrogen) in the closed loop formed by the hydrogen bond.

2.3.2 3_{10} -helix

The amino acids in a 3_{10} -helix are arranged in a right-handed helical structure (Fig. 2.7b). Each amino acid corresponds to a 120° turn in the helix (i.e., the helix has three residues per turn), and a translation of 2.0 \AA along the helical axis, and has 10 atoms in the ring formed by making the hydrogen bond. Most importantly, the N-H group of an amino acid forms a hydrogen bond with the C=O group of the amino acid three residues earlier; this repeated $i + 3 \rightarrow i$ hydrogen bonding defines a 3_{10} -helix.

2.3.3 π -helix

Similar to α -helix and 3_{10} -helix, the amino acids in a standard π -helix are also arranged in a right-handed helical structure. Each amino acid corresponds to a 87° turn in the helix (i.e., the helix has 4.1 residues per turn), and a translation of 1.15 \AA along the helical axis. In this structure which is shown in Fig. 2.7c, the N-H group of an amino acid forms a hydrogen bond with the C=O group of the amino acid five residues earlier ($i + 5 \rightarrow i$).

2.3.4 Helical wheel diagram representation

A helical wheel is a type of plot or visual representation used to illustrate the properties of alpha helices in proteins. The sequence of amino acids that make up a helical region of the protein's secondary structure are plotted in a rotating manner where the angle of rotation between consecutive amino acids is 100° (the turn angle in α -helix conformation), so that the final representation looks down the helical axis. The plot reveals whether hydrophobic/hydrophilic amino acids are concentrated on one side of the helix or not. As an example, the wheel diagram of transportan, which will be discussed in chapter 4, is presented in Fig. 4.3.

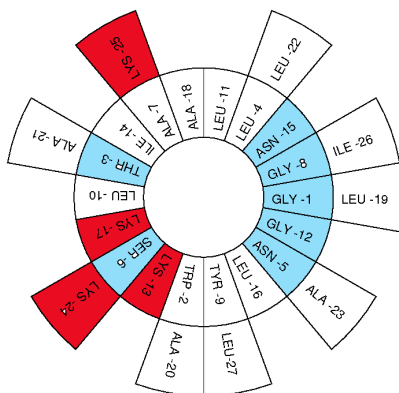


Figure 2.8: The helical wheel diagram of transportan. Each residue is categorized according to the following color scheme: *white* (hydrophobic in neutral pH), *cyan* (hydrophilic and neutral in neutral pH), *red* (hydrophilic and charged in neutral pH). There are 27 residues in total.

2.4 Electrostatics of peptide-membrane systems

Membranes can be cationic (being positively charged), anionic (being negatively charged) or zwitterionic (being neutral but with a positive and a negative charge group). Electrostatic interactions within membranes have great effect on a number of membrane properties, such as rigidity, structural stability and dynamics as well as phase behavior. Also, these interactions between membranes and other biological systems, such as peptides or proteins, can play very important roles for a biological process to happen. Therefore, studying these interactions in more detail is necessary.

2.4.1 Hydrogen bonding

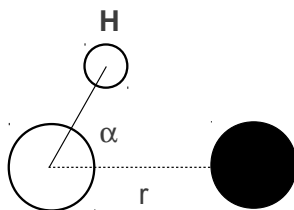


Figure 2.9: Hydrogen bond criterion. The donor/acceptor molecules are shown in white/black spheres. The cutoff values of α and r used in this simulation are 30° and 0.35 nm respectively. $r = 0.35$ nm corresponds to the first minimum of the RDF of SPC water molecules [11, 104, 50, 90].

Hydrogen bonds are responsible for many unusual properties of water such as high melting point, boiling point and heat of vaporization. Hydrogen bonds are a consequence of electrostatic interaction. They usually mediate interactions between molecules, with a typical strength of several $k_B T$ ($k_B T \approx 2.5$ kJ/mol). For comparison, the typical energy for a carbon-carbon double bond is $240 k_B T$ and for a single C-O bond is $136 k_B T$. For van der Waals interactions, this energy is approximately $1 k_B T$ when two molecules are 3 \AA apart. A hydrogen bond is an attractive interaction of a hydrogen, covalently attached to an electronegative atom, and another electronegative atom belonging to other molecules or chemical groups. To determine if a hydrogen bond exists, the following criteria (see Figure 2.9)

$$r_{\text{HB}} < 0.35\text{nm}, \quad (2.15)$$

and

$$\alpha < 30^\circ \quad (2.16)$$

are used. These are the default criteria of GROMACS package [84] for identification of a hy-

drogen bond. The value of 0.35 nm corresponds to the first minimum of the radial distribution function (RDF) of water molecules. The first minimum of the RDF separates the first and the second hydration shells. For any water molecule, the closest neighboring water molecules are those which are within the first hydration shell.

Hydrophobic effect: The hydrophobic effect can be understood if we define it in terms of free energies. Consider the Gibbs free energy when a solute is in its bulk state and when it is in water. If in transferring the solute from its bulk environment to water the free energy increases, i.e., if $\Delta G = G_{\text{water}} - G_{\text{bulk}}$ is positive, it means that the solute is hydrophobic and prefers to be surrounded by other solute molecules rather than by water molecules. In contrast, negative ΔG implies that the molecule is hydrophilic and likes to be surrounded by water molecules, and thus, is soluble in water.

Hydrogen bonding plays a crucial role in the hydrophobic effect: Because a purely hydrocarbon molecule or region is incapable of forming hydrogen bonds with water, introduction of the hydrocarbon into water causes disruption of the hydrogen bonding network between water molecules. It is energetically unfavorable for water molecules to lose hydrogen bonds within themselves, and therefore, they prefer rearrangements that provide as many hydrogen bonds as possible. The hydrogen bonds are partially reconstructed by building a water "cage" around the hydrocarbon molecule, but the water molecules that form the "cage" (or solvation shell) have substantially restricted mobilities. This leads to a significant loss in translational and rotational entropy of water molecules and makes the process unfavorable in terms of free energy of the system. By aggregating together, non-polar molecules reduce the surface area exposed to water and minimize their disruptive effect. Therefore, the hydrophobic effect is an entropic effect originating from the disruption of highly dynamic hydrogen bonds between molecules of liquid water by the non-polar solute.

2.4.2 Charge-pair interaction

In biological systems, beside forming hydrogen bonds with water molecules, lipids are also able to form charge pairs with themselves as well as with other charged systems such as charged amino acids of peptides or proteins. Figure 2.10 shows two cases of charge-pair interactions in lipid head groups. The methyl groups attached to nitrogen atoms have partial positive charges. Also, oxygen atoms attached to carbon or phosphorus atoms have partial negative charges. This gives rise to an electrostatic attraction between these groups. The definition of a charge pair is based on the respective radial distribution function, i.e., the RDF of charged components in membrane systems. Here, the RDF between the neighboring groups shown in Fig. 2.10 has



Figure 2.10: Charge pair interaction formed between adjacent lipids: charge-pair between methyl group and carbonyl group (left), and charge-pair between methyl group and phosphate group (right).

a minimum when the distance between the neighboring groups is 0.4 nm [112]. This means that at any distance less than 0.4 nm, the closeness of these groups can be attributed to the electrostatic interaction. For other systems comprising of different components, this value may no longer be 0.4 nm and the criterion for the charge pair interaction should be changed accordingly.

2.5 Methods

2.5.1 The molecular dynamics methods

The molecular dynamics (MD) method is an approach to find the equilibrium as well as dynamical properties of molecular systems by solving Newton's equation of motion numerically:

$$\frac{d^2 \mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i}. \quad (2.17)$$

where \mathbf{r}_i is the position vector of particle i . This equation describes the motion of a particle with the mass of m_i when the total force applied on it is \mathbf{F}_i . In a typical MD simulation, at each time the total force applied on a particle is computed. Having the force and using the position and velocity of the particle, the position and velocity after a certain *time step* can be computed. By repeating this cycle, the coordinates of the particle at all times can be obtained. This cycle is applied on all the particles in the system. Therefore, time development of the coordinates of all particles (the trajectory of the system) can be obtained. The forces on each particle originates from other particles present in the system which relates to the force through the following equation:

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V(\mathbf{r}_1, \dots, \mathbf{r}_N). \quad (2.18)$$

Hence, knowing the interactions of a particle with others is of great importance. Also, to simulate the statistical ensembles such as NPT (constant number of particles, constant pressure and constant temperature), special algorithms should be used to keep these quantities constant. Another important point is the method that is used to solve eq. 2.17. These aspects are further discussed in the next sections.

MD simulations can be divided into three main categories: *classical MD* [45], *ab initio MD* [61] and *coarse-grained MD*. At the very heart of any MD scheme is the question of how to describe - that is in practice how to approximate - the interatomic interactions. The traditional route followed in classical MD is to determine these potentials in advance. That is, before setting up the simulation, one should know these potentials to be able to incorporate them into the simulation. Typically, the full interaction is broken up into two-body, three-body and many-body contributions, long-range and short-range terms etc. The aforementioned term "body" in classical MD refers to nuclei of atoms which are much heavier than electrons. This does not mean, however, that the electrons are neglected. It means that the effect of the electronic structure of particles is incorporated into *force field* by which the interaction potentials of the particles are described. More detailed discussion about force fields in classical MD will be presented in section 2.5.3.

In *ab initio* MD, the electronic structure of atoms is taken into account. In other words, in this method, the forces acting on the nuclei are computed from electronic structures of atoms in the system. In this way, the electronic variables are not integrated out beforehand, but are considered as active degrees of freedom. Thus, *ab initio* MD is based on Newton's as well as Schrödinger equations. This method not only allows one to tackle chemically complex problems such as those involving breaking and formation of chemical bonds, but also it has a truly predictive power. Despite its obvious advantages, it is evident that a price has to be paid for putting MD on *ab initio* grounds: the correlation lengths and relaxation times that are accessible are much smaller than what is affordable via classical MD. *Ab initio* MD simulations are limited to time scales of a few tens of picoseconds and length scales of a few nanometers, i.e., they are limited to systems comprising of hundreds of atoms [96]. There are two main *ab initio* schemes which are in widespread use namely *Born-Oppenheimer* [18] and *Car-Parrinello* [24].

In Born-Oppenheimer MD [18], the electronic structure in MD simulations is included by straightforwardly solving the static electronic structure problem in each MD step given the set of fixed nuclear positions at that instance of time. Thus, the electronic structure part is reduced to solving a time-independent quantum problem, e.g. by solving the time-independent Schrödinger equation, concurrently to propagating the nuclei via classical MD. Thus, the time-dependence of the electronic structure is not intrinsic and is a consequence of nuclear motion [96].

The Car-Parrinello method [24] explicitly introduces the electronic degrees of freedom as (fictitious) dynamical variables, writing an extended Lagrangian for the system which leads to a system of coupled equations of motion for both ions and electrons. In this way, an explicit electronic minimization at each time step, as done in Born-Oppenheimer MD, is not needed: after an initial standard electronic minimization, the fictitious dynamics of the electrons keeps them on the electronic ground state corresponding to each new ionic configuration visited along the dynamics, thus yielding accurate ionic forces. In order to maintain this adiabaticity condition, it is necessary that the fictitious mass of the electrons is chosen small enough to avoid a significant energy transfer from the ionic to the electronic degrees of freedom. This small fictitious mass in turn requires that the equations of motion are integrated using a smaller time step than the one commonly used in Born-Oppenheimer MD.

By switching from *ab initio* to classical MD, we lose some detailed information related to electronic properties. One can go further ("lose more") and develop a method that captures the important atomistic aspects without the computational cost of a brute force classical or *ab initio* MD for the entire system. In this way, *coarse-grained* MD is introduced which was first proposed by Broughton *et al.* in 1998 [130]. In this method, instead of explicitly representing

every atom of the system, as in classical MD, one uses “pseudo-atoms” to represent groups of atoms [45]. For example, colloidal suspensions are dispersions of mesoscopic (10 nm - 1 μm) solid particles. These particles themselves consist of millions, or even billions, of atoms. Furthermore, the number of solvent molecules per colloid is comparable or even larger. Clearly, a MD simulation that follows the behavior of several thousands colloids over an experimentally relevant time interval (milliseconds to seconds) would be prohibitively expensive. This is why colloidal suspensions are always modeled using a coarse-grained model [45].

As another example, in simulations of biological membranes, one can reach time and length scales as large as microseconds and micrometers respectively by using a coarse-grained model [93]. In such simulations, the tails of lipids are represented by a few pseudo-atoms by gathering two to four methylene groups into each pseudo-atom. The parameterization of these coarse-grained models must be done empirically [93], by matching the behavior of the model to appropriate experimental data or all-atom simulations.

Next, we discuss the most important aspects of classical MD which was used throughout this thesis. From now on, we just use the term MD instead of classical MD.

2.5.2 MD integrators

Here, we introduce one of the most important algorithms for solving the equations of motion. This algorithm is called *Verlet* and was first proposed in 1967 [75]. Depending on the way in which the coordinates and velocities are produced, different types of Verlet algorithms can be categorized. In the so-called *coordinate Verlet* [75] the position at time $t + \Delta t$, $\mathbf{r}(t + \Delta t)$, can be obtained using the position and acceleration at time t and also the position at time $t - \Delta t$.

$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \Delta t \frac{d}{dt} \mathbf{r}_i(t) + \frac{\Delta t^2}{2} \frac{d^2}{dt^2} \mathbf{r}_i(t) - \frac{\Delta t^3}{6} \frac{d^3}{dt^3} \mathbf{r}_i(t) + O(\Delta t^4), \quad (2.19)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \frac{d}{dt} \mathbf{r}_i(t) + \frac{\Delta t^2}{2} \frac{d^2}{dt^2} \mathbf{r}_i(t) + \frac{\Delta t^3}{6} \frac{d^3}{dt^3} \mathbf{r}_i(t) + O(\Delta t^4). \quad (2.20)$$

Summing these two equations, one obtains

$$\mathbf{r}_i(t + \Delta t) \approx 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i(t). \quad (2.21)$$

As can be seen, the Verlet algorithm is time-reversible, because one can obtain $\mathbf{r}_i(t - \Delta t)$ from $\mathbf{r}_i(t + \Delta t)$ by interchanging $\mathbf{r}_i(t + \Delta t)$ and $\mathbf{r}_i(t - \Delta t)$ in eq. 2.21. The velocity at time t can be approximated as:

$$\mathbf{v}_i(t) \approx \frac{1}{2\Delta t} (\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)). \quad (2.22)$$

Combining eq. 2.21 and 2.22, the coordinate Verlet algorithm is generated. The truncation error introduced by this algorithm is Δt^4 for the new coordinates and Δt^2 for the new velocities. A practical problem with this approach is its sensitivity to numerical precision and round-off errors because the velocities are obtained as the difference of two coordinates of the same magnitude. In case this drawback is significant, one can use the *velocity Verlet* algorithm [144]. In velocity Verlet algorithm, new velocities are generated after computing new positions and new forces:

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{1}{2} \Delta t^2 \frac{\mathbf{F}_i(t)}{m_i}, \quad (2.23)$$

$$\mathbf{v}_i(t + \Delta t) \approx \mathbf{v}_i(t) + \Delta t \frac{\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)}{2m_i}. \quad (2.24)$$

It can be shown that this algorithm is equivalent to the coordinate Verlet algorithm [45].

In GROMACS [84], the so-called Leap-Frog algorithm [55] is used:

$$\mathbf{r}_i(t + \Delta t) \approx \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \frac{\Delta t}{2}) \quad (2.25)$$

$$\mathbf{v}_i(t + \frac{\Delta t}{2}) \approx \mathbf{v}_i(t - \frac{\Delta t}{2}) + \frac{\Delta t}{m_i} \mathbf{F}_i. \quad (2.26)$$

As can be seen, in this algorithm the coordinate and velocity are not calculated at the same time, but rather at $t + \Delta t$ and $t + \Delta t/2$ respectively. In GROMACS, however, the velocities at $t - \Delta t/2$ and $t + \Delta t/2$ are averaged to give a velocity at the time t just the same as the one in the Verlet algorithm [84, 13].

There are also some other Verlet-like algorithms such as the *Beeman algorithm* [10] and the velocity-corrected Verlet algorithm [45]. For further discussion on these algorithms, see Ref. [45].

The Verlet algorithm is a symplectic integrator. This means that it conserves the volume in phase space, i.e., it is area-preserving in the absence of round-off error. This property guarantees that there is no severe energy drift in the Verlet scheme even in long time scales. This is a very important feature since we have to live with the Lyapunov instability [45], which states that trajectories depend on initial coordinates very sensitively. Due to Verlet's area-preserving feature, we can use it safely since the propagation of the phase space vector evolves in the phase space with a constant volume. This means we can have truncating and round-off errors which can lead to different trajectories but without the changing ensemble behavior.

Three important considerations influence the choice of algorithm: (i) Time reversibility, inherent in the Newtonian equations of motion, should be conserved. (ii) The generated trajectories should conserve volume in phase space, i.e., the algorithm should be symplectic. Being

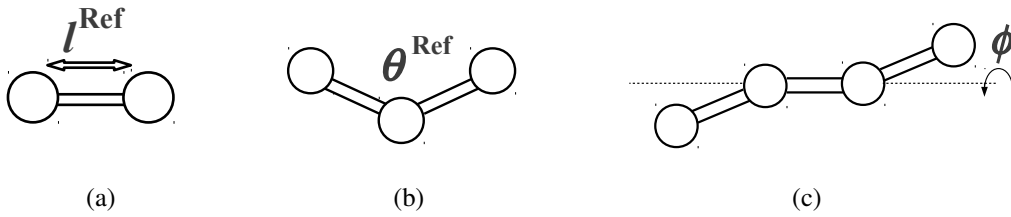


Figure 2.11: Illustration of bond stretching (left), bond bending (middle) and torsional terms appearing in eq. 2.27.

symplectic is important to conserve equilibrium distributions in phase space, because deviation from symplectic behavior will produce time-dependent weight factors in phase space. (iii) Since the computational effort is completely dominated by the force calculation, methods that use only one force evaluation per time step are to be preferred. This rules out the well-known *Runge-Kutta* methods, which moreover are also not symplectic and lead to erroneous behavior on longer time scales.

2.5.3 Force field

As mentioned before, the essence of MD is to integrate the Newton's equation of motion numerically. Therefore, one should have a complete knowledge about the interactions and physical properties of the particles. All this information is gathered in the *force field*. A force field accounts for the internal energy of the system. To better understand what information is included in a force field, we look at the potential part of a model Hamiltonian:

$$\begin{aligned}
 V = & \sum_{\text{bonds}} \frac{k_i^b}{2} (l_i - l_i^{\text{Ref}})^2 + \sum_{\text{angles}} \frac{k_i^a}{2} (\theta_i - \theta_i^{\text{Ref}})^2 \\
 & + \sum_{\text{torsions}} \frac{V_T}{2} \left[1 + \cos(n\phi - \gamma) \right] \\
 & + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left[4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\epsilon_0 r_{ij}} \right]. \quad (2.27)
 \end{aligned}$$

The first term in eq. 2.27 corresponds to the bonded atoms (Fig. 2.11a) modeled by a harmonic potential with a force constant of k_i^b , the bond length l_i and a reference bond length of l_i^{Ref} . It is important to remember that this is an approximation to the real bond stretching potential and that it no longer holds true for large deviations from the harmonic approximation. For situations where the bond lengths may deviate far, or to accurately calculate molecular structures and vibrational frequencies, it is necessary to go beyond the harmonic approximation

and include higher order terms usually up to $(l_i - l_i^{\text{Ref}})^4$. Another description of the potential function of the bonded atoms is the Morse function [100]:

$$V(r) = D[1 - e^{-a(l_i - l_i^{\text{Ref}})}]^2 \quad (2.28)$$

where D is the dissociation energy of the bond, and a is a constant related to the steepness of the potential. The Morse curve can be approximated near the minimum at $l_i = l_i^{\text{Ref}}$ by the harmonic potential with force constant $k = 2Da^2$. The Morse curve is only a convenient analytical expression that has some essential features of a diatomic potential, including a fairly good agreement with vibration spectra of diatomic molecules, but there is no theoretical justification for this particular form. In many instances we may not even have an analytical form for the potential, but know the potential at a number of discrete points, e.g., from quantum-chemical calculations. In that case a good way to proceed is to construct a potential function from *cubic spline interpolation* of the computed points. Because cubic splines have continuous second derivatives, the forces will behave smoothly as they will have continuous first derivatives everywhere.

The second term describes a harmonic potential for bending of three consecutive atoms (Fig. 2.11b) where k_i^a is the angle bending force constant, θ_i is the actual bond angle and θ_i^{Ref} is the reference bond angle. The energy needed to distort an angle away from equilibrium is much lower than that needed to distort a bond, so bond angle bending force constants tend to be proportionally smaller than those for bond-stretching. As with the bond stretching potential, the accuracy can be improved by including higher order terms.

The third term involves four consecutive atoms (Fig. 2.11c), representing torsional interactions, which describes how energy changes when bonds rotate, i.e., when dihedral angles change (see also section 1.2.2 for the definition of dihedral angles in peptides). There, V_T gives a quantitative measure of the energy change due to torsional rotation. Torsional angle interactions are different to stretching and bending interactions in two important ways. The first is that internal rotation barriers are low compared to other interactions, meaning that changes in dihedral angles can be large. Secondly, the torsional potential is periodic through a 360° rotation. This implies that it would be inappropriate to approximate torsional potential by a Taylor series. The phase angle γ is usually chosen so that terms with positive V_T have a minimum at 180° .

The last term corresponds to the non-bonded interactions which include electrostatic interactions for charged particles and *Lennard Jones* for van der Waals interactions. In Fig. 2.12, a plot of Lennard-Jones potential is provided. In this potential, ϵ is the depth of the potential, r is the distance between two particles, σ is the distance at which the inter-particle potential is

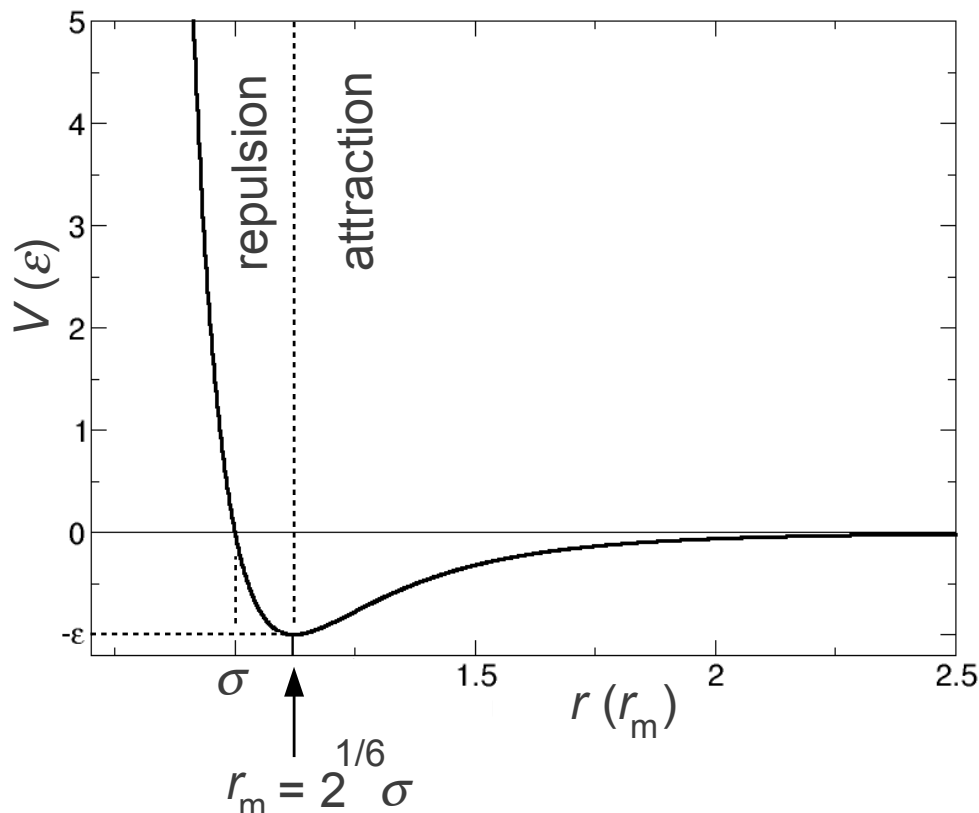


Figure 2.12: The Lennard-Jones interaction: ϵ is the depth of the potential, r is the distance between two particles, and σ is the distance at which the inter-particle potential is zero. The regions where the slope of the potential is negative (positive) corresponds to a repulsion (an attraction) interaction.

zero, and i and j account for different particles. These parameters can be fitted to reproduce experimental data or accurate quantum chemistry calculations. Due to its computational simplicity, the Lennard-Jones potential is used extensively in computer simulations even though more accurate potentials exist such as *Buckingham potential* in which the repulsive term r^{-12} is replaced with an exponential term, thus being more computationally expensive.

Choosing the force field

The precise form of a force field is strongly influenced by the way it is derived (e.g. from experiment or quantum computation) and also by the accuracy needed for its intended purpose. For example, some force fields are intended for simulations of bulk phases (e.g. the AMBER force field [29]). These generally have a simple form with harmonic terms for bond stretching and bending, and usually a Lennard-Jones term for the van der Waals interaction.

The force field of any molecule is defined not only by eq. 2.27, but also by values of

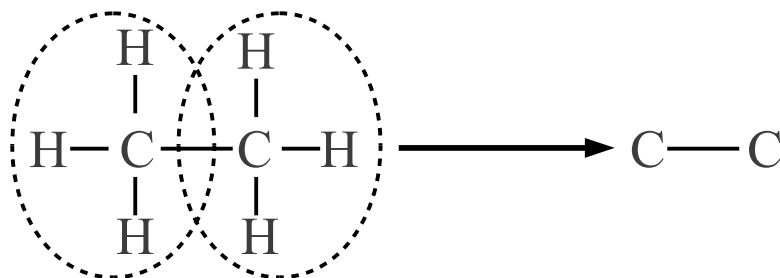


Figure 2.13: Schematic illustration of the united atom model.

parameters associated with this equation. These parameters have to be chosen against some model of reality. Traditionally, this is taken from experimental data. For some molecules, such as alkanes, experimental data is plentiful. For other classes of molecules, data is more scarce. Even when experimental data is available, some force field parameters can be hard to be determined from it. Bond lengths and angles can be determined from crystallography. Van der Waals parameters can be determined from crystal structures. However, torsional potentials can be hard to determine, and so are atomic charges. These can be found from *ab initio* simulations. For atomic charges, for example, the method of *Restrained ElectroStatic Potential (RESP)* [9] is commonly used. The basic idea of this method is that a least squares fitting algorithm is used to derive a set of atom-centered point charges which best reproduce the electrostatic potential of the molecule. In other words, one can derive atomic charges by fitting the charges to reproduce the electrostatic potential (ESP) calculated at a large number of grid points around the molecule.

The hardest parameters to determine, whether from experimental or *ab initio* simulations, are the van der Waals parameters. As these determine intermolecular interactions, they are highly important for determining condensed phase properties. Van der Waals parameters are often found from experimental crystal structures. They can also be calculated from *ab initio* simulations. However, van der Waals parameters calculated from gas phase *ab initio* simulations can neglect interactions involving three or more atoms (three-body interactions). These can have a large contribution to the energy of a system, so neglecting these leads to large errors in the calculated values of some quantities. Another method has been used by Jorgensen *et al.* in deriving the Optimized Parameters for Liquid Simulation (OPLS) parameter set [67]. Here, initial guesses were made for the van der Waals parameters. Monte Carlo simulations were then

performed on pure liquids and the van der Waals parameters were adjusted to reproduce experimental values of thermodynamic quantities such as density and heats of vaporization. As these parameters were fitted to condensed phase properties, they implicitly include many body interactions. They are often referred to as effective two-body potentials. As another example, in the case of Berger *et al.* [15] force field used for lipids, the standard parameters of the Optimized Potentials for Liquid Simulations (OPLS) are modified such that they reproduce the density and the heat of vaporization for pentadecene [15]. Then the partial charges are taken from the work by Chiu *et al.* [26], in which they derived the partial charges on DMPC phospholipids by *ab initio* self-consistent field calculations, and GROMOS parameters [153, 152, 92, 154] are used for bonded potentials. This combination of parameters reproduce the experimental value for the area per molecule of many lipid bilayers [15, 148, 104].

One type of commonly used force field that provides a large saving in computer time is called united atom force field [79]. In this force field, the hydrogen atoms connected to carbon atoms (such as those in methyl or methylene group) are incorporated in the carbon atoms. This is illustrated in Fig. 2.13. In this way, the methyl or methylene groups are considered as a "united atom". The van der Waals interactions in the force field should be adjusted to take care of this change: these neglected atoms are accounted for by increasing the size of the atoms they are bonded to, usually by increasing the van der Waals radius. This kind of treatment is only acceptable when the hydrogen atoms are not polarized. For example, this approach cannot be done for the hydroxyl group (-OH) since the highly charged hydrogen atom can have a big contribution to charge pair interactions or hydrogen bondings.

Simple Charge Model (SPC) for water molecules

Despite the small size and seemingly simple structure of water, proper modeling of water has turned out to be very challenging. In a liquid state, water molecules are all differing in various properties such as the precise geometry and molecular vibration, since, from a quantum mechanical point of view, the molecular orbitals of water molecules depend on the arrangement of their neighboring water molecules [20]. However, in molecular modeling of water, we are not able to build a water model that covers the whole range of physical behavior of real water molecules, but have to make various approximations to obtain a model that is not expensive in computation, and yet usable for specific cases.

In general, water models can be divided into three types. The first kind are characterized by rigid geometry and pairwise interactions including the electrostatic and Lennard-Jones interactions. The second class, so-called flexible models, permits internal conformational changes. In the third case, models have been developed to include the effects of polarization and many-body effects explicitly [79]. For a review of different models see Ref. [51].

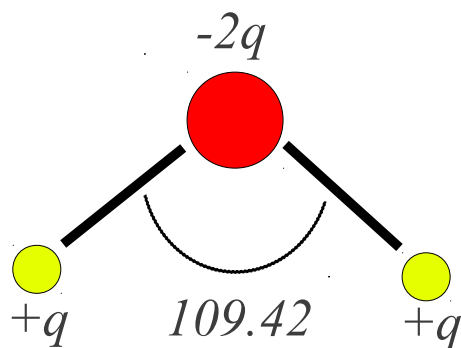


Figure 2.14: The SPC water model.

In practice, the rigid water models are the most commonly used. The use of rigid internal geometry is, of course, an approximation, and it means that some properties of water, such as the vibrational spectrum, cannot be studied. Introducing more parameters, describing the flexibility and polarization, can enhance the performance of water model. However, it is usually feasible to employ simple water models in simulations of large biological systems considering the computational cost.

The Simple Point Charge (SPC) model [11], is a rigid water model in which the water molecule has three centers of concentrated charge: a positive charge ($+0.41 e$) on the H atoms and excess negative charge ($0.82 e$) on the O atom (Fig. 2.14). The assumption that they are point charges is an approximation that leads to an incorrect value for the permanent dipole moment of the water molecule. To correct this, the H–O–H bond angle is changed to 109.42° in the model (compared with the experimentally-found H–O–H bond angle of 104.45°). The van der Waals interaction between two water molecules is modeled using a Lennard-Jones potential between the oxygens without consideration of hydrogen atoms, and $r_{\text{OH}} = 0.1 \text{ nm}$.

2.5.4 Temperature and pressure coupling

The direct integration of Newton's equation of motion leads to a constant energy of the system and this is not useful for studying the dissipative non-equilibrium systems such as lipid bilayers. Furthermore, since the experiments are conducted in constant temperature and pressure, it is reasonable to keep these quantities constant in simulation.

To keep the temperature constant, a usual approach is to couple the simulation box to a heat bath which is commonly referred to as a thermostat. Several algorithms have been developed for this purpose including the *Andersen* thermostat [4], the *Langevin* thermostat [64], the *Lowe-Andersen* thermostat [89], the *Nosé-Hoover* thermostat [45] and the *Nosé-Poincaré* thermostat

[17], and *Berendsen* thermostat [12], and *velocity rescaling* [21].

In the Nosé-Hoover thermostat [45], the system Hamiltonian is extended by introducing extra variables describing the thermal reservoir and a friction term in the equations of motion:

$$H_{\text{Nosé}} = \sum_{i=1}^N \frac{\mathbf{P}_i^2}{2m_i} + U(\mathbf{r}^N) + \frac{\xi^2 Q}{2} + L \frac{\ln s}{\beta} \quad (2.29)$$

where $H_{\text{Nosé}}$ is the extended Hamiltonian, \mathbf{P}_i is the momentum of particle i , Q is the parameter describing the coupling strength, and ξ is a thermodynamic friction coefficient related to an additional coordinate s . (For details of this method, see Ref. [45].) This method aims to generate a canonical ensemble and is widely used. However, one drawback of the Nosé-Hoover thermostat is that it is not a symplectic algorithm. This problem has been corrected in a modified version of the Nosé-Hoover thermostat called the Nosé-Poincaré method [17].

Next, we discuss the Berendsen temperature and pressure coupling as they are used in the simulations of this thesis.

Berendsen thermostat

We first consider the Langevin equation [64] which describes coupling of a system to a heat bath with fixed reference temperature T_0 :

$$m_i \dot{\mathbf{v}}_i = \mathbf{F}_i - m_i \gamma_i \mathbf{v}_i + R(t), \quad (2.30)$$

where R_i is a Gaussian stochastic variable with zero mean. In the Berendsen scheme [12], the equation of motion is changed to

$$m_i \dot{\mathbf{v}}_i = \mathbf{F}_i + m_i \gamma(t) \left[\frac{T_0}{T - T_0} \right] \mathbf{v}_i \quad (2.31)$$

without adding a local stochastic term, as in the Langevin thermostat. This equation represents a proportional scaling of the velocities per time step Δt in the algorithm from v to λv with

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T} - 1 \right) \right]^{\frac{1}{2}} \quad (2.32)$$

where τ_T is the temperature relaxation time. Equation 2.32 is equivalent to correcting the spontaneous temperature according to

$$\frac{dT}{dt} = 2\gamma[T_0 - T] = \frac{T_0 - T}{\tau_T}. \quad (2.33)$$

Therefore, in the Berendsen thermostat, the velocity is rescaled at each time step Δt to control the kinetic energy, and hence, the temperature. More detailed derivation of this can be found in [12].

Then, the velocities are rescaled and the positions are updated in leap-frog integration manner:

$$\mathbf{v}_i(t + \frac{\Delta t}{2}) \rightarrow \left[\mathbf{v}_i(t - \frac{\Delta t}{2}) + \frac{\Delta t}{m_i} \mathbf{F}_i \right] \lambda \quad (2.34)$$

$$\frac{1}{2} \left[\mathbf{v}_i(t - \frac{\Delta t}{2}) + \mathbf{v}_i(t + \frac{\Delta t}{2}) \right] \rightarrow \mathbf{v}_i(t) \quad (2.35)$$

$$\mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \frac{\Delta t}{2}) \rightarrow \mathbf{r}_i(t + \Delta t) \quad (2.36)$$

As can be seen by eq. 2.31 and 2.33, for very large τ_T , which corresponds to very small γ , eq. 2.31 becomes Newton's equation of motion. Direct integration of Newton's equation of motion without any extra term corresponding to the thermostat, leads to the sampling of an ensemble for which energy is conserved, thus a microcanonical ensemble. Therefore, eliminating random forces in a Langevin equation, as done in Berendsen thermostat, fails to produce the canonical ensemble [99]. On the contrary, for very small τ_T (i.e. $\tau_T = \Delta t$), the global coupling works nearly as the local disturbance (e.g. random forces) and the energy exchange between the system and the heat bath becomes large. Thus, the Berendsen thermostat approaches the canonical ensemble for very small τ_T . For intermediate values of τ_T , the statistical mechanical ensemble produced by the Berendsen thermostat has properties between a canonical and a microcanonical ensemble. One advantage of this method is that the coupling can be made as weak as desired to minimize the disturbance of the system, and the strength of the coupling can easily be varied to suit the needs of a given application. The other advantage is that the algorithm is numerically stable and truncation errors will not develop undesired deviations. This is of considerable practical value when conditions are adjusted to new values, as well as for long "unattended" runs.

With regard to generating a canonical ensemble, an improvement to the Berendsen thermostat is the *velocity rescaling* or the *Parrinello-Bussi* thermostat [21] which samples the correct canonical ensemble. In this approach the velocities of all the particles are rescaled by a properly chosen random factor, to enforce the correct distribution for the kinetic energy. Also, a quantity is defined which is constant and plays a similar role as that of the energy in the microcanonical ensemble. It can be used to verify how much the numerical procedure generates configurations that belong to the desired *NVT* ensemble and to provide a guideline for the choice of the integration time step.

Berendsen barostat

In cases where it is desired to control the pressure, the system can also be coupled to a "pressure bath". There are several algorithms to achieve the desired pressure such as the *Andersen barostat* [4], the *Parrinello-Rahman barostat* [111] and the *Berendsen barostat* [12].

In the Berendsen barostat [12], the coupling to a pressure bath can be accomplished similar to a heat bath in the Berendsen thermostat. An extra term is added to the equations of motion that effects a pressure change [12]

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_b}. \quad (2.37)$$

where P and P_0 are the spontaneous and reference pressure, respectively. For an isotropic system, the pressure is

$$P = \frac{2}{3V}(E_k - \Xi), \quad (2.38)$$

where V is the spontaneous volume of the system, and E_k is the kinetic energy and Ξ is the virial for the pair-additive potentials:

$$\Xi = -\frac{1}{2} \left[\sum_{i<j} \mathbf{F}_{ij} \cdot \mathbf{r}_{ij} \right], \quad (2.39)$$

where

$$\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j. \quad (2.40)$$

and \mathbf{F}_{ij} is the force on particle i exerted by particle j . A pressure change can be accomplished by changing the virial through scaling of interparticle distances. So an extra term proportional to x is added in the equation of $\dot{\mathbf{r}} = \mathbf{v}$ is added:

$$\dot{\mathbf{r}} = \mathbf{v} + \alpha \mathbf{r}. \quad (2.41)$$

It can be shown [12] that for having a pressure change given by eq. 2.37, the coefficient α should be

$$\alpha = -\frac{\beta(P_0 - P)}{3\tau_b}, \quad (2.42)$$

where β is the compressibility. Then, a proportional scaling of coordinates \mathbf{r}_i and the box length L per time step Δt will change the pressure. As the compressibility only enters the algorithm in conjunction with the time constant, its value need not be precisely known. Finally, the scale factor, μ , is given by

$$\mu = \left[1 + \frac{\Delta t}{\tau_b}(P_0 - P) \right]^{\frac{1}{3}}. \quad (2.43)$$

Therefore, by scaling the particle coordinates and box size, pressure control will be achieved:

$$\mathbf{r}_i \rightarrow \mu \mathbf{r}_i. \quad (2.44)$$

$$L \rightarrow \mu L \quad (2.45)$$

For systems with interfaces, such as lipid bilayers, semi-isotropic scaling can be useful. In this case, the xy -directions are scaled isotropically and the z direction is scaled independently. For anisotropic systems, equation 2.43 becomes a tensorial equation with similar form given by the pressure P (more details can be found in Ref. [12]).

In practice, the coupling constants, τ_T and τ_b , are used as an empirical parameter to adjust the strength of the coupling. Values of $\tau_T \approx 0.1$ ps are typically used in MD simulations of condensed-phase systems [60]. Note, however, that this choice generally leads to fluctuations close to those of the microcanonical ensemble. With this choice, the Berendsen thermostat merely removes energy drifts from a MD simulation (since in microcanonical ensemble energy is conserved), without significantly altering the ensemble sampled. It is recommended to use pressure coupling time constants of $\tau_b = 0.1$ ps or larger [12] since fluctuations of global properties are strongly influenced for time constants less than 0.1 ps; the intensity of such fluctuations cannot be used to derive thermodynamic properties. Also, reliable dynamic properties can be derived for both temperature and pressure coupling time constants above 0.1 ps [12].

2.5.5 Constraint dynamics

In MD simulations, the time step is limited by bond oscillations. These oscillations can be estimated by considering the highest phonon or vibrational frequencies which are of the order of $\approx 10^{-14}$ s or 0.01 ps [129]. Internal vibrations with frequencies higher than $k_B T/h$, where h is the *Planck constant*, exhibit essential quantum behavior. At very high frequencies the corresponding degrees of freedom are in the ground state and can be considered static. It seems logical, therefore, to treat such degrees of freedom as constraints. If these oscillations are to be included in a simulation, the time step of the simulation must be smaller than the period of these oscillations, otherwise the dynamics of the system are not obtained properly. Simulations with smaller time steps are computationally more expensive than those with larger time steps. By replacing the bond vibrations with constraints, the time step in MD can be increased by a factor of four [53], leading to a faster simulation. Once again, this replacement is possible, as in polyatomic molecules, the fast internal vibrations are usually decoupled from rotational and translational motions and can therefore be frozen by introducing a certain number of rigid bonds and angles in the skeleton of the molecule. For example, N_2 becomes a rod [8] and

H₂O a rigid triangle [120], etc. An example of a rigid model for water is SPC model which is explained in section 2.5.3.

Because bonds in molecules are coupled, resetting coupled constraints is a nonlinear problem. Many algorithms have been proposed for solving this problem. The most widely used algorithms are LINCS [53] and SHAKE [131] for large molecules and SETTLE [98] for smaller molecules such as water. In the SHAKE algorithm, the nonlinear equations are solved iteratively. But due to this iterative nature, it is difficult to parallelize it. SETTLE is the analytical version of SHAKE for three atoms which is easy to be parallelized and is widely used to constrain water molecules. In this thesis, the LINCS algorithm is applied to constrain all the bond lengths.

LINear Constraint Solver (LINCS)

Consider a system of N particles. The motion of a set of N particles can be described by a set of second-order ordinary differential equations, Newton's second law, which can be written in matrix form

$$\mathbf{f} = \mathbf{M} \cdot \frac{d^2 \mathbf{q}}{dt^2}, \quad (2.46)$$

where \mathbf{M} is a mass matrix and \mathbf{q} is the vector of generalized coordinates that describe the particles' positions. For example, the vector \mathbf{q} may be a $3N$ Cartesian coordinates of the particle positions \mathbf{r}_k , where k runs from 1 to N . In the absence of constraints, \mathbf{M} would be the $3N \times 3N$ diagonal square matrix of the particle masses. The vector \mathbf{f} represents the generalized forces. If n constraints are present, the coordinates must also satisfy n time-independent algebraic equations

$$\sigma_k(\mathbf{q}) = 0 \quad (2.47)$$

where the index k runs from 1 to n .

Suppose we want to constrain the length of the bond between two molecules. Then the constraints are

$$\sigma_k^{(t)} \equiv \|\mathbf{X}_{k\alpha}^{(t)} - \mathbf{X}_{k\beta}^{(t)}\|^2 - d_k^2 = 0, \quad k = 1 \cdots n \quad (2.48)$$

where $\mathbf{X}_{k\alpha}^{(t)}$ and $\mathbf{X}_{k\beta}^{(t)}$ are the positions of the two particles involved in the k th constraint at time t and d_k is the prescribed inter-particle distance. These constraint equations, are added to the potential energy function in the equations of motion, resulting in

$$\frac{\partial^2 \mathbf{X}_i^{(t)}}{\partial t^2} m_i = -\frac{\partial}{\partial \mathbf{X}_i} \left[V(\mathbf{X}_i^{(t)}) + \sum_{k=1}^n \lambda_k \sigma_k^{(t)} \right] \quad i = 1 \cdots n \quad (2.49)$$

for each of the N particles in the system. Adding the constraint equations to the potential does

not change it, since all $\sigma_k^{(t)}$ should ideally be zero. Integrating both sides of the equations of motion twice in time yields the constrained particle positions at the time $t + \Delta t$

$$\mathbf{X}_i^{(t+\Delta t)} = \hat{\mathbf{X}}_i^{(t+\Delta t)} + \sum_{k=1}^n \lambda_k \frac{\partial \sigma_k^{(t)}}{\partial \mathbf{X}_i} (\Delta t)^2 m_i^{-1}, \quad i = 1 \cdots N \quad (2.50)$$

where $\hat{\mathbf{X}}_i^{(t+\Delta t)}$ is the unconstrained (or uncorrected) position of the i th particle after integrating the unconstrained equations of motion. To satisfy the constraints $\sigma_k^{(t+\Delta)}$ in the next time step, the Lagrange multipliers must be chosen such that

$$\sigma_k^{(t+\Delta t)} \equiv \|\mathbf{X}_{k\alpha}^{(t+\Delta t)} - \mathbf{X}_{k\beta}^{(t+\Delta t)}\|^2 - d_k^2 = 0, \quad k = 1, \dots, n \quad (2.51)$$

This implies solving a system of n non-linear equations

$$\begin{aligned} \sigma_j^{(t+\Delta t)} \equiv \\ \|\hat{\mathbf{X}}_{j\alpha}^{(t+\Delta t)} - \hat{\mathbf{X}}_{j\beta}^{(t+\Delta t)} + \sum_{k=1}^n \lambda_k \left[\frac{\partial \sigma_k^{(t)}}{\partial \mathbf{X}_{j\alpha}} m_{j\alpha}^{-1} - \frac{\partial \sigma_k^{(t)}}{\partial \mathbf{X}_{j\beta}} m_{j\beta}^{-1} \right]\|^2 - d_j^2 = 0, \\ j = 1, \dots, n \end{aligned}$$

simultaneously for the n unknown Lagrange multipliers λ_k . The solution vector for this n non-linear equations in n unknowns is updated using

$$\bar{\lambda}^{(l)} - \mathbf{J}_\sigma^{-1} \bar{\sigma} \rightarrow \bar{\lambda}^{(l+1)} \quad (2.52)$$

where \mathbf{J}_σ is the Jacobian of the of the equations σ_k

$$\mathbf{J}_\sigma = \begin{pmatrix} \frac{\partial \sigma_1}{\partial \lambda_1} & \frac{\partial \sigma_1}{\partial \lambda_2} & \cdots & \frac{\partial \sigma_1}{\partial \lambda_n} \\ \frac{\partial \sigma_2}{\partial \lambda_1} & \frac{\partial \sigma_2}{\partial \lambda_2} & \cdots & \frac{\partial \sigma_2}{\partial \lambda_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \sigma_n}{\partial \lambda_1} & \frac{\partial \sigma_n}{\partial \lambda_2} & \cdots & \frac{\partial \sigma_n}{\partial \lambda_n} \end{pmatrix} \quad (2.53)$$

Since not all particles are involved in all constraints, \mathbf{J}_σ is blockwise-diagonal and can be solved blockwise, i.e. molecule for molecule.

The above framework is common for a number of constraint algorithm such as LINCS and SHAKE. These methods differ only in how they solve the system of equations. LINCS solves

for the multipliers by using a series expansion to approximate the inverse of the Jacobian \mathbf{J}_σ :

$$(\mathbf{I} - \mathbf{J}_\sigma)^{-1} = \mathbf{I} + \mathbf{J}_\sigma + \mathbf{J}_\sigma^2 + \mathbf{J}_\sigma^3 + \dots \quad (2.54)$$

The first power of \mathbf{J}_σ gives the coupling effects of neighboring bonds. The second power gives the coupling effects over a distance of two bonds, not only between bonds that are separated by one bond, but also the feedback of a bond on itself through neighboring bonds. The third power gives the third order coupling effects and so on. So, truncating the series after a specified number of terms means neglecting all higher order coupling effects.

The inversion through an expansion is efficient, because the inverse itself is not needed, only the product of the inverse with a vector. In one time step, the bonds only influence each other when they are separated by fewer bonds than the highest order in the expansion; with the correction for the rotation this number is doubled. Because of this local coupling a decomposition method can be applied.

The approximation in eq. 2.54 only works for matrices with eigenvalues smaller than 1, making the LINCS algorithm suitable only for molecules with low connectivity. In molecules with only bond constraints the connectivity is so low that this will always be true, even if ring structures are present. Problems can arise in angle-constrained molecules.

Figure 2.15 shows a schematic picture of how the algorithm works for one bond. At each time step, the LINCS constraints, which consist of two steps, are applied after an unconstrained update. First, the projections of new bonds on the old bonds are set to zero. This step is a velocity correction which means that the velocity along the bond's direction is zero. In the second step the bond's length is corrected. The second step is actually an iterative process, but in practice, one iteration is usually sufficient for the accuracy required in MD simulations. In energy minimization this might not be accurate enough, and usually four iterations are used. A detailed derivation of this algorithm can be found in Ref. [53]. LINCS is three to four times faster than SHAKE and has better convergence behavior [53]. More importantly, parallelization of LINCS is much easier to implement because of the inversion method shown in eq. 2.54.

2.5.6 Long-Range interactions

Simulation of biological systems often requires handling of electrostatic interactions. Unlike the van der Waals interactions which are usually represented using a Lennard-Jones potential, the electrostatic interactions, represented by Coulomb potential, are long-ranged and cannot be cut off at a certain distances, as is done with the Lennard-Jones interactions. Studies have shown that the cutoff treatments lead to severe unphysical behavior since it implies a delta-function in the force at the cutoffs. In modeling of biological systems, the Ewald [41] and P³M

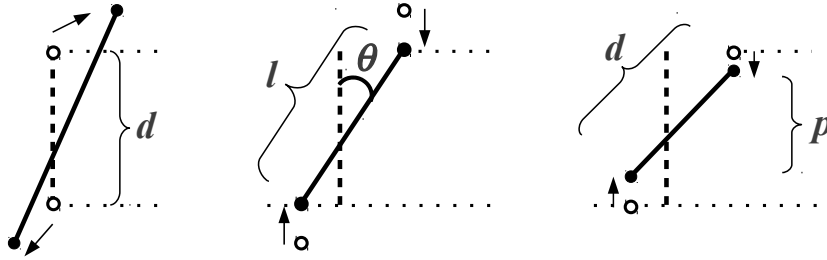


Figure 2.15: The three position-updates needed for one time step. The dashed line is the old bond of length d , the solid lines are the new bonds. In the left picture, the updated positions are not constrained yet. In the middle one, the projection of the new bond on the old bond is set to zero, and in the right one, a correction is applied to set the bond's length equal to d . $l = d / \cos \theta$ and $p = (2d^2 - l^2)^{\frac{1}{2}}$.

(particle-particle-particle-Mesh) [55] methods as well as their variants, such as PME (particle-Mesh-Ewald) [149], are the most commonly used methods. In these methods, the Coulomb potential is split into a short-range and a long-range contributions, chosen to be smooth functions of the distance. The short-range part converges at a certain distance and a cutoff can be applied. The computation of the long-range part involves solving the Poisson equation which describes the potential as a function of charge density. Next, we present the Ewald method, and then we explain the PME method which was used in our simulations. The PME algorithm scales as $N \log(N)$ which is much faster than the ordinary Ewald summation which scales as N^2 .

The Ewald summation of Coulomb energy

First, assume we have a collection of charged particles in a cube with side length L , with periodic boundary conditions. The collection is assumed neutral; there is an equal number of positive and negative charges. The total Coulomb energy in this system is given by

$$\mathcal{U}_{\text{Coul}} = \frac{1}{2} \sum_{i=1}^N q_i \phi(r_i), \quad (2.55)$$

where $\phi(r_i)$ is the electrostatic potential at position r_i ,

$$\phi(r_i) = \sum_{j=1}^{N'} \sum_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|}, \quad (2.56)$$

where \mathbf{n} is a three dimensional integer vector. The prime on the first summation indicates that we do not include terms for which $j = i$ if $\mathbf{n} = (0, 0, 0)$. That is, we allow each particle to interact with its periodic images, but not with itself.

To evaluate \mathcal{U} efficiently, we break it into two parts: a short-ranged potential treated with a simple cutoff, and long-ranged potential which is periodic and slowly varying, which can therefore be represented to an acceptable level of accuracy by a finite Fourier series. The idea of Ewald method is to do two things: first, screen each point charge using a diffuse cloud of opposite charge around each point charge, and then compensate for these screening charges using a smoothly varying, periodic charge density. The screening charge is constructed to make the electrostatic potential due to a charge at position \mathbf{r}_j decay rapidly at a prescribed distance. These interactions are treated in real space. The compensating charge density, which is the sum of all screening densities except with opposite charges, is treated using a Fourier series.

The standard choice for a screening density is Gaussian:

$$\rho_s(r) = -q_j(\alpha/\pi)^{\frac{3}{2}} \exp(-\alpha r^2). \quad (2.57)$$

So, for each charge we add such a screening charge cloud to the system. To evaluate \mathcal{U} , we have to evaluate the potential of a charge density that compensates for the screening charge densities at each particle. This is done in Fourier space. The potential of a given charge distribution is given by Poisson's equation

$$-\nabla^2 \phi(\mathbf{r}) = 4\pi\rho(\mathbf{r}). \quad (2.58)$$

The compensating charge distribution ρ_1 , can be written:

$$\rho_1 = \sum_{j=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} q_j(\alpha/\pi)^{\frac{3}{2}} \exp[-\alpha|\mathbf{r} - (\mathbf{r}_j + \mathbf{n}L)|]. \quad (2.59)$$

Notice that the sum over j includes the self-interaction when we include the potential due to this charge density in the calculation of the total Coulombic energy.

Now, consider the Fourier transform of the Poisson's equation

$$k^2 \tilde{\phi}(k) = 4\pi \tilde{\rho}(k). \quad (2.60)$$

The Fourier transform of ρ_1 is given by

$$\rho_1(k) = \sum_{j=1}^N q_j e^{-i\mathbf{k} \cdot \mathbf{r}_j} e^{-k^2/4\alpha}. \quad (2.61)$$

The \mathbf{k} -vectors are given by

$$\mathbf{k} = \frac{2\pi}{L} \mathbf{l} \quad \mathbf{l} \in \mathbb{Z}^3. \quad (2.62)$$

We can use eq. 2.60 to solve for $\tilde{\phi}(k)$:

$$\tilde{\phi}(k) = \frac{4\pi}{k^2} \sum_{j=1}^N q_j e^{-ik \cdot r_j} e^{-k^2/4\alpha}. \quad (2.63)$$

Note that this solution is not defined for $k = 0$. In fact, we have to assume that $\tilde{\phi}(0) = 0$, which is consistent with the notion that our system and all its periodic images is embedded in a medium of infinite dielectric constant.

Fourier inverting $\tilde{\phi}(k)$ gives

$$\phi_i(r) = \frac{1}{V} \sum_{k \neq 0} \tilde{\phi}(k) e^{ik \cdot r_i}. \quad (2.64)$$

After substitution for $\tilde{\phi}(k)$ from eq. 2.63 we obtain

$$\phi_i(r) = \sum_{k \neq 0} \sum_{j=1}^N \frac{4\pi}{V} \frac{q_j}{k^2} e^{ik \cdot (r_i - r_j)} e^{-k^2/4\alpha}. \quad (2.65)$$

So, the total Coulombic energy due to the compensating charge distribution is

$$\begin{aligned} \mathcal{U}_1 &= \frac{1}{2} \sum_i q_i \phi_1(\mathbf{r}_i) \\ &= \frac{1}{2} \sum_{k \neq 0} \sum_{j=1}^N \frac{4\pi}{V} \frac{q_i q_j}{k^2} e^{ik \cdot (r_i - r_j)} e^{-k^2/4\alpha} \\ &= \frac{1}{2V} \sum_{k \neq 0} \frac{4\pi}{k^2} |\rho(k)|^2 e^{-k^2/4\alpha}, \end{aligned} \quad (2.66)$$

where

$$\rho(k) = \sum_{i=1}^N q_i e^{ik \cdot r_i}. \quad (2.67)$$

Notice that this does indeed include a spurious self-self interaction, because the point charge at \mathbf{r} interacts with the compensating charge cloud also at \mathbf{r} . This self-interaction is the potential at the center of a Gaussian charge distribution, and can be obtained as follows

$$-\frac{1}{r} \frac{\partial^2 (r \phi_{\text{Gauss}}(r))}{\partial r^2} = 4\pi \rho_{\text{Gauss}}(r), \quad (2.68)$$

yielding

$$\phi_{\text{Gauss}}(r) = \frac{q_i}{r} \text{erf}(\sqrt{\alpha} r), \quad (2.69)$$

where erf is the error function:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-r^2} dr. \quad (2.70)$$

At $r = 0$, we have

$$\phi_{\text{self}} = \phi_{\text{Gauss}}(0) = 2\left(\frac{\alpha}{\pi}\right)^{1/2} q_i. \quad (2.71)$$

So the total self-interaction energy becomes

$$\mathcal{U}_{\text{self}} = \left(\frac{\alpha}{\pi}\right)^{1/2} \sum_{i=1}^N q_i^2, \quad (2.72)$$

which must be *subtracted* from the total Coulombic energy.

Finally, the real-space contribution of the point charge at \mathbf{r}_i is the screened potential:

$$\phi_{\text{short}}(r) = \frac{q_i}{r} - \frac{q_i}{r} \text{erf}(\sqrt{\alpha r}) \equiv \frac{q_i}{r} \text{erfc}(\sqrt{\alpha r}), \quad (2.73)$$

where erfc is the complementary error function. The total real-space Coulombic potential energy is therefore

$$\mathcal{U} = \frac{1}{2} \sum_{i \neq j}^N \frac{q_i q_j}{r_{ij}} \text{erfc}(\sqrt{\alpha r_{ij}}). \quad (2.74)$$

Putting it all together:

$$\begin{aligned} \mathcal{U}_{\text{Coul}} &= \frac{1}{2V} \sum_{k \neq 0} \frac{4\pi}{k^2} |\rho(k)|^2 e^{-k^2/4\alpha} \\ &- \left(\frac{\alpha}{\pi}\right)^{1/2} \sum_{i=1}^N q_i^2 \\ &+ \frac{1}{2} \sum_{i \neq j}^N \frac{q_i q_j}{r_{ij}} \text{erfc}(\sqrt{\alpha r_{ij}}), \end{aligned} \quad (2.75)$$

where

$$\rho(k) = \sum_{i=1}^N q_i e^{ik \cdot r}. \quad (2.76)$$

The choice of α is arbitrary. Clearly, very small α makes the Gaussians tighter and therefore the compensating charge distribution less smoothly varying. This means a Fourier series representation of \mathcal{U}_1 with a given number of terms is more accurate for large α .

Particle-Mesh-Ewald (PME)

The computational effort of the Ewald summation scales as N^2 with the number of charges N and becomes prohibitive for large systems. Fast Fourier transforms (FFT) are computationally attractive although they restrict the spatial solutions to lattice points. Interpolation is then needed to obtain the energies and forces acting on charges. They scale as $N \log N$ and are the long-range methods of choice, e.g., as implemented in the particle-mesh-Ewald (PME) method of Darden *et al.* [149] who use a Gaussian spread function and a Lagrange interpolation, or - preferably - the more accurate smooth particle-mesh-Ewald (SPME) method of Essmann *et al.* [40] who use a B-spline interpolation. The advantage of using B-spline interpolation is that the resulting potential function is twice continuously differentiable if the order of the spline is at least four; smooth forces can therefore be immediately obtained from the differentiated potential. With Lagrange interpolation the interpolated potential is only piecewise differentiable and cannot be used to derive the forces. In GROMACS package, which was used throughout this thesis, SPME is not implemented to this date, and therefore, we used PME for calculating the Coulomb potential energy.

Three parameters control the convergence of the sums in eq. 2.75: n_{\max} , an integer which defines the range of the real-space sum and controls its maximum number of vectors (i.e. image cells), similarly m_{\max} , an integer defining the summation range in the reciprocal-space and its number of vectors, and α , the Ewald convergence parameter, which determines the relative rate of convergence between the real and reciprocal sums. Note that a large value of α in eq. 2.57, i.e. a narrow Gaussian distribution, makes the real-space sum converge more quickly. In contrast, a small α causes the reciprocal-space sum to converge more quickly.

The choice of Ewald parameters should be based on several considerations:

1- *System size N* : larger systems may require a larger α and/or R_{cutoff} to limit the number of pair-wise interactions such that the real-space sum converges faster.

2- *Accuracy desired*: choosing a larger R_{cutoff} , n_{\max} , or m_{\max} will yield more accurate results, however it may be inefficient.

3- *Cutoff radius*: the smaller R_{cutoff} , the larger α needs to be for the real space sum to converge rapidly with a reasonable number of n -vectors.

Since PME is Fourier-based and utilizes fast Fourier transforms to evaluate the reciprocal sum, it is more efficient to “bias” the Ewald summation towards the reciprocal sum and limit the real-space sum within a small cutoff radius.

2.5.7 Free energy calculation

Free energy is one of the most important quantities in thermodynamics. The calculation of free energy of complex biological systems is one of the big challenges in computational physics and chemistry. There are several computational methods developed to tackle this task with the aim of better understanding of biological phenomena. These include *thermodynamic integration* [72], *slow growth* [143], Jarzynski-based [66] fast-growth technique and *umbrella sampling* [91] to name a few.

Among the methods based on nonequilibrium simulations, Jarzynski technique is very well-known. Jarzynski [66] demonstrated the equivalence of the free-energy change and an exponential average over the work W along *nonreversible* paths originating from a *canonic* ensemble:

$$\exp(-\beta\Delta A) = \langle \exp(-\beta W) \rangle \quad (2.77)$$

This is the remarkable Jarzynski equation, which at first sight is a counterintuitive expression, relating a thermodynamic quantity to a rather ill-defined and very much process-dependent amount of work. Equation 2.77 can be exploited in practical simulations by moving a constraint on the reaction coordinate relatively fast from an equilibrated system to the target system. This method became known as fast growth. The changes in the energy along these paths are averaged according to eq. 2.77. The computational tradeoff is that the faster the constraint is varied, the larger is the statistical spread, and thus, the more trajectories have to be calculated.

Umbrella sampling method

Umbrella sampling was developed by Torrie and Valleau [91]. The method is based on equilibrium properties of the system and is discussed here.

In statistical mechanics, the free energy A is related to the canonical partition function Q via $A = -1/\beta \ln Q$ where $\beta = 1/(k_B T)$, k_B being the Boltzmann's constant. The canonical partition function involves NVT ensembles. If the pressure, instead of the volume, is kept constant, the Gibbs free energy (usually denoted as G) is obtained. The following formalism is the same for both A (eq. 2.5) and G (eq. 2.6).

If the potential energy E is independent of the momentum, Q can be simplified as:

$$Q = \int \exp[-\beta E(r)] d^N r, \quad (2.78)$$

where N is number of degrees of freedom of the system. So, the integral is over the whole coordinate part of the configuration space.

In free energy calculations, one is generally interested in the free energy of the system when

it is in a certain state. These states can be described by a parameter, known as the reaction coordinate (ξ). The choice of ξ is usually based on geometrical bases. For example, in our study which involves peptide-membrane systems, ξ can be chosen as the distance between the center of mass of the peptide and the membrane. In simulations, the free energy is calculated for a set of values of reaction coordinate, known as a window. Therefore, several simulations corresponding to each window should be run in parallel. In the end, different free energy values associated with each window are combined together to give a free energy for the whole range of the reaction coordinate. With ξ defined, the probability of finding the system in a small interval $d\xi$ around ξ can be written as

$$Q(\xi) = \frac{\int \delta[\xi'(r) - \xi] \exp[-\beta E(r)] d^N r}{\int \exp[-\beta E(r)] d^N r}, \quad (2.79)$$

where the prime means that the integration is over all degrees of freedom except ξ . The free energy along the reaction coordinate, also known as potential of mean force (PMF), can be obtained as $A(\xi) = -1/\beta \ln Q(\xi)$.

In computer simulations, the direct phase-space integrals used in Eqs. 2.78 and 2.79 are impossible to calculate. However, if the system is ergodic, i.e., if every point in the phase space is visited after a long enough time, $Q(\xi)$ is equal to

$$P(\xi) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \rho[\xi(t')] dt', \quad (2.80)$$

where ρ counts the occurrence of ξ in an infinitesimal interval dt' . $P(\xi)$ is accessible in simulations, and therefore, provided the simulations are long enough to meet the ergodicity condition, the free energy $A(\xi)$ is in principle obtainable from $P(\xi)$. However, there are regions in the phase space which are very hard to reach in simulation times. In other words, for rare events, i.e., those with an energy barrier significantly larger than $k_B T$, hardly ever does the direct sampling of the corresponding phase space occur. Different free energy techniques have actually been developed to overcome this problem.

In umbrella sampling, the reaction coordinate is restrained at certain values by a bias potential to sample those rare events. For example, a harmonic potential can be added to the system's Hamiltonian to bias the system toward those regions inaccessible by unbiased simulations. This can be written as

$$E_b(r) = E_u(r) + \omega_i(\xi), \quad (2.81)$$

where $\omega_i(\xi)$ is the biased potential applied on the point ξ in the window i . Obviously, what we obtain from these biased simulations is not exactly what we are looking for, because by

adding a biased potential the system changes. But statistical mechanics help us to convert different properties from biased simulations to unbiased ones. In order to obtain the unbiased free energy $A(\xi_i)$ for the i th window, we need the unbiased distribution which is

$$P_i^u(\xi) = \frac{\int \delta[\xi'(r) - \xi] \exp[-\beta E(r)] d^N r}{\int \exp[-\beta E(r)] d^N r}. \quad (2.82)$$

MD simulation of the biased system provides the biased distribution along the reaction coordinate $P_i^b(\xi)$:

$$P_i^b(\xi) = \frac{\int \delta[\xi'(r) - \xi] \exp[-\beta(E(r) + \omega_i(\xi'(r)))] d^N r}{\int \exp[-\beta(E(r) + \omega_i(\xi'(r)))] d^N r}. \quad (2.83)$$

Because the bias depends only on ξ and the integration in the numerator is performed over all degrees of freedom but ξ ,

$$\begin{aligned} P_i^b(\xi) &= \exp[-\beta \omega_i(\xi)] \\ &\times \frac{\int \exp[-\beta E(r)] \delta[\xi'(r) - \xi] d^N r}{\int \exp[-\beta(E(r) + \omega_i(\xi'(r)))] d^N r}. \end{aligned} \quad (2.84)$$

Using eq. 2.82 results in

$$\begin{aligned} P_i^u(\xi) &= P_i^b(\xi) \exp[\beta \omega_i(\xi)] \\ &\times \frac{\int \exp[-\beta(E(r) + \omega_i(\xi(r)))] d^N r}{\int \exp[-\beta E(r)] d^N r} \\ &= P_i^b(\xi) \exp[\beta \omega_i(\xi)] \\ &\times \frac{\int \exp[-\beta(E(r))] \exp[-\beta \omega_i(\xi(r))] d^N r}{\int \exp[-\beta E(r)] d^N r} \\ &= P_i^b(\xi) \exp[\beta \omega_i(\xi)] \langle \exp[-\beta \omega_i(\xi)] \rangle. \end{aligned} \quad (2.85)$$

$P_i^b(\xi)$ is obtained from the biased MD simulation and $\omega_i(\xi)$ is given analytically, e.g. as a harmonic potential. Therefore, from the above equation $A_i(\xi)$ can be obtained. The average value denoted by $\langle \dots \rangle$ in the above equation is independent of ξ and we can define

$$F_i = -(1/\beta) \ln \langle \exp[-\beta \omega_i(\xi)] \rangle. \quad (2.86)$$

Hence, we obtain

$$A_i(\xi) = -(1/\beta) \ln P_i^b(\xi) - \omega_i(\xi) + F_i. \quad (2.87)$$

This gives the free energy for the i th window up to a constant F_i .

If we had only one window, i.e., the bias potential is applied only on one reaction coordinate point, we could neglect the constant. But because there is always more than one window, as the difference in free energy is interesting not just single value of it, F_i 's must be calculated for each window. There are several methods to analyze Umbrella Sampling simulations including the Weighted Histogram Analysis Method (WHAM) [73] and Umbrella Integration [69]. WHAM, which is used in this thesis, is an iterative method and leads to

$$\exp(-\beta F_i) = \int P^u(\xi) \exp[-\beta \omega_i(\xi)] d\xi. \quad (2.88)$$

In eq. 2.88, P^u is not known. This can be seen also in 2.85 in which F_i has appeared. Therefore, $P^u(\xi_i)$ and F_i should be calculated iteratively until convergence.

WHAM combines different free energy values of different windows together to give one global $A(\xi)$. Different windows have different constants F_i and WHAM uses the overlap between the probability distributions of adjacent windows to calculate F_i 's. One should always run the simulations in such a way that the probability distributions of adjacent simulations overlap. For example, when the restraining potential is harmonic, by choosing a low enough harmonic constant the restrained part of the system can sample more regions in the phase space leading to a good overlap between the adjacent simulations.

Chapter 3

Interaction of POPC bilayers with Cholesterol and Dehydroergosterol

3.1 Introduction

Cholesterol (CHOL) is an important constituent of cellular membranes playing a fundamental role in many biological processes. This sterol affects membrane permeability, lateral lipid organization, signal transduction and membrane trafficking [127]. Different cellular functions depend on the amount of CHOL in the cell membrane. Processes like endo- and exocytosis, signal transduction and the presence of a permeability barrier for water and ions require a large amount of CHOL in the plasma membrane. For example, the amount of CHOL in human neurons is 34% [87]. In contrast, the amount of CHOL is very low in the endoplasmic reticulum (ER) (7% in rats [87]) which is in fact responsible for the abundance of CHOL in different organelles. There are many questions about the distribution and behavior of CHOL in membranes. For example, how is the amount of CHOL maintained and regulated in intracellular regions, or what is the rate of transbilayer flip-flops of CHOL? Is cholesterol mainly in the inner leaflet of cell membranes or in the outer leaflet? To answer these questions, one can make use of a CHOL reporter, i.e. a molecule which reveals information about the sterol orientation, distribution and interaction with other cell components. A perfect CHOL reporter is obviously one that has the least perturbing effect on the system and is easy to be detected in experiment.

Cholesterol reporters can be divided into two classes: (1) CHOL binding molecules which form a complex with CHOL. For example, Filipin, is a fluorescent polyene antibiotic that binds to CHOL and is commonly used to visualize the cellular distribution of free CHOL [156, 23, 22]. (2) CHOL analogues which can be photoreactive, such as [³H]6-azi-5 α -cholesterol [49], spin-labeled, such as 3 β -doxyl-5 α -cholestane [103], or fluorescent, such as dehydroergosterol

[135]. CHOL analogues can be incorporated into membranes for studying different aspects of CHOL behavior. That is, due to their detectable properties, instead of CHOL, CHOL analogues can be studied.

The question arises here as to whether or not CHOL analogues are able to mimic the behavior of CHOL and which one has the least perturbing effect on the system. Several biophysical studies reveal that even slight changes in sterol structure have large effects on sterol properties in lipid membranes[101, 126]. Therefore, finding a good cholesterol analogue, which can mimic CHOL the best, is crucial for CHOL research.

Dehydroergosterol (DHE) is a CHOL analogue with intrinsic fluorescence which naturally occurs in yeasts and certain sponges[135]. This sterol was discovered over 80 years ago, but its complete structure was only confirmed in 1985[43]. It differs from CHOL in having two additional double bonds in the steroid ring system making it slightly fluorescent. The tail of DHE is identical to that of ergosterol with an extra double bond and methyl group compared to CHOL.

DHE has proved to faithfully mimic CHOL in many respects. For example, it co-distributes with CHOL in membranes, exhibits the same exchange kinetics as CHOL in membranes, is nontoxic to cultured cells or animals, and is accepted as substrate for esterification[135, 136, 46]. But it has a lower ability than CHOL to stiffen [47] and condensate lipids[134].

To our knowledge, no MD simulation has been reported about DHE in lipid bilayers. Obviously such a study can shed a great deal of light on the behavior of DHE in cell membranes. Here, we perform atomistic MD simulations to study the properties of DHE in lipid bilayers and compare them to those of CHOL. See Fig. 3.2 for the chemical structures of CHOL and DHE.

3.2 Methods

We have performed atomic-scale MD simulations of three different membrane systems. The first bilayer was composed of 128 palmitoyl-oleoyl phosphatidylcholine (POPC) molecules, the second of 128 POPC and 32 CHOL molecules, and the third of 128 POPCs and 32 DHE. All three bilayers were hydrated with 3500 water molecules. The initial structures of all bilayers were obtained by arranging the POPC molecules in a regular array in the bilayer (x,y) plane with an initial surface area of 0.67 nm^2 per POPC molecule. An equal number of sterol molecules were inserted into each leaflet. Before the actual MD simulations, the steepest-descent algorithm was used to minimize the energy of the initial structure. The simulations were performed using the GROMACS software package. The MD simulations of all bilayer systems were carried out over 150 ns. The first 50 ns was considered as an equilibration period, and the last 100 ns of the trajectory were analyzed. Figure 3.1 shows the chemical structure of POPC with the atom numbering used in this thesis. Fig. 3.2 illustrates the chemical structure of cholesterol and DHE. Figure 3.3 shows a snapshot of the simulation of the POPC-DHE system.

As shown in Fig. 3.2, both CHOL and DHE are composed of a ring system with a hydroxyl group and a short atom chain attached to carbon C_{17} . The ring system forms an asymmetric planar structure with two off-plane methyl substituents (C_{18} and C_{19}) defining the so-called rough side often called as the β -face. The smooth side which has no methyl group attached to it is called the α -face.

We used standard force-field parameters for POPC molecules [15] where the partial charges were taken from the underlying model description [148]. For water, we employed the SPC model [11]. For the sterol force field, we used the description of Holtje *et al.* [57]. For the additional double bond in DHE, standard GROMACS parameters were used.

Periodic boundary conditions with the usual minimum image convention were used in all three directions. The bond lengths of all molecules were preserved using the LINCS algorithm

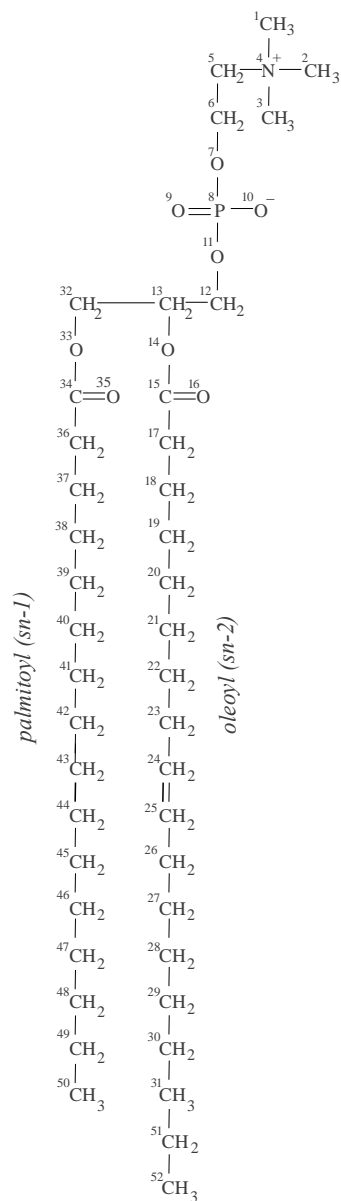


Figure 3.1: Chemical structure of POPC.

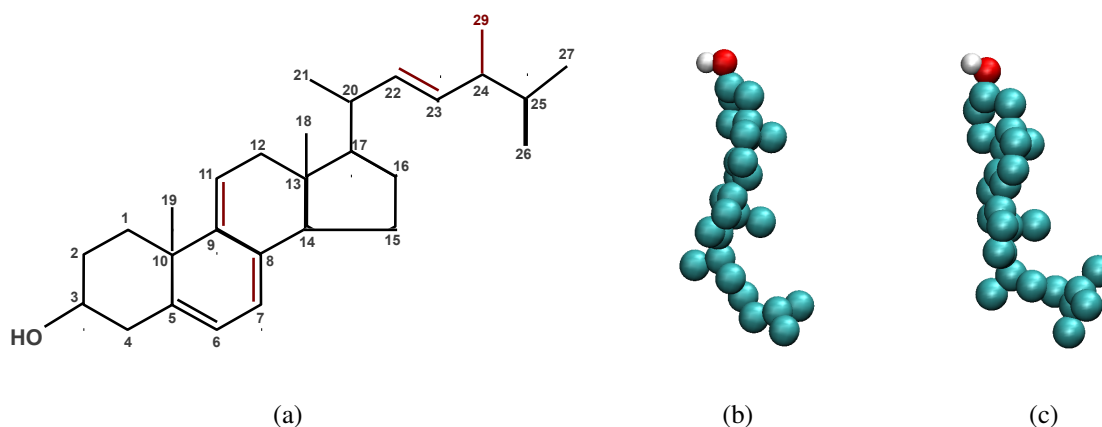


Figure 3.2: Chemical structure of CHOL and DHE together while differences are indicated in *red* (a), representation of CHOL (b) and DHE (c).

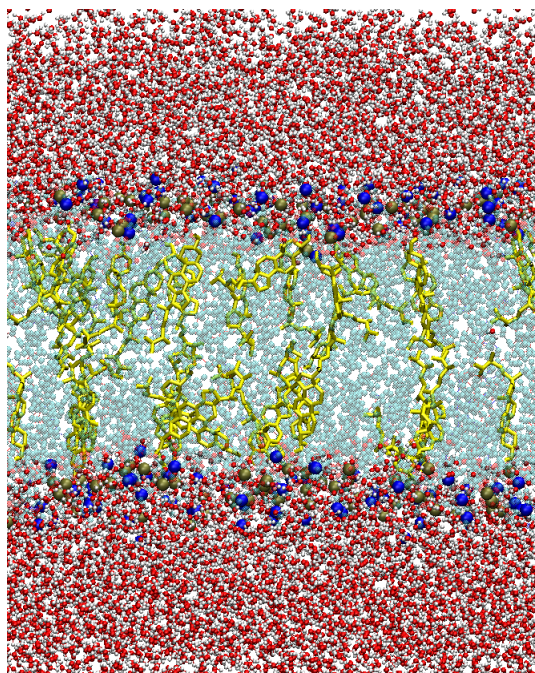


Figure 3.3: A snapshot of the simulation of the POPC-DHE system at 0 ns. Water is shown in *red*, DHE molecules are represented in *yellow* while POPC molecules are shown as transparent beads in *light blue*. Nitrogen and phosphorus atoms are shown as beads in *blue* and *light grey* respectively.

[53]. The time step was set to 2 fs, and the simulations were carried out at constant pressure (1 atm) and temperature (303 K), which is above the main phase transition temperature of POPC [138]. The temperature was controlled using the No \acute{s} e-Hoover method [17]. The temperature of the bilayer and water was controlled independently. The pressure was controlled semi-isotropically using Parrinello-Rahman barostat [111]. The Lennard-Jones interactions were cut

off at 1.0 nm. For the electrostatic interactions we employed the particle-mesh Ewald method [149] with a real space cutoff of 1.0 nm, b-spline interpolation (of order 4), and a direct sum tolerance of 10^{-5} . The list of nonbonded pairs was updated every 10 steps.

3.3 Results

3.3.1 Area per molecule and membrane thickness

The time development of the temperature (Fig. 3.4), the potential and total energy (Fig. 3.5) and the surface area per POPC (Fig. 3.6) for the three different systems was monitored throughout the 150-ns-long simulation. From the time profile of the surface area per POPC and the low value of drift compared to standard deviation (9×10^{-8} vs 0.01 nm^2), it was reasoned that after the first 50 ns, the system had equilibrated. Therefore, the first 50 ns were considered as the equilibration period and thus the last 100 ns of trajectories were analyzed.

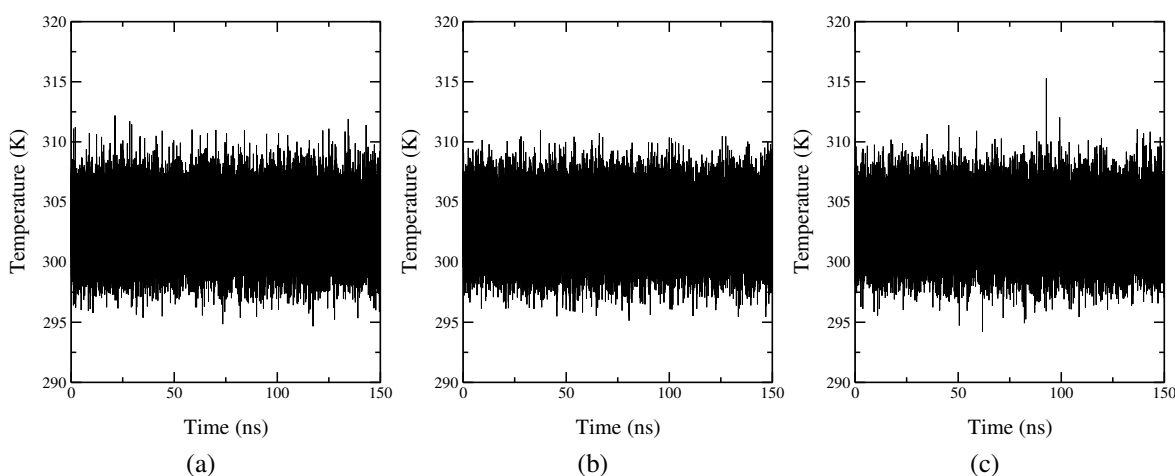


Figure 3.4: temperature vs time for pure POPC (a), POPC+CHOL (b) and POPC+DHE (c).

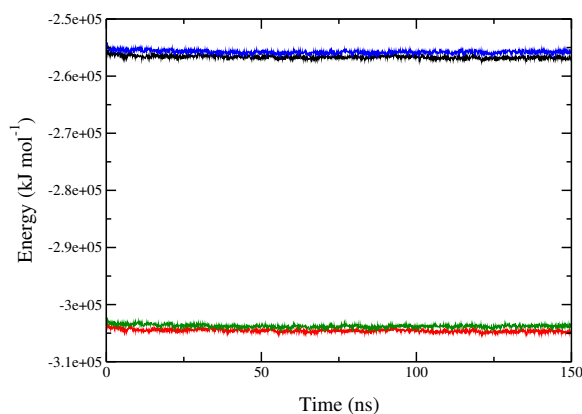


Figure 3.5: Total energy (top two graphs) and potential energy (bottom two graphs) for POPC+CHOL (black and red) and POPC+DHE (blue and green) systems.

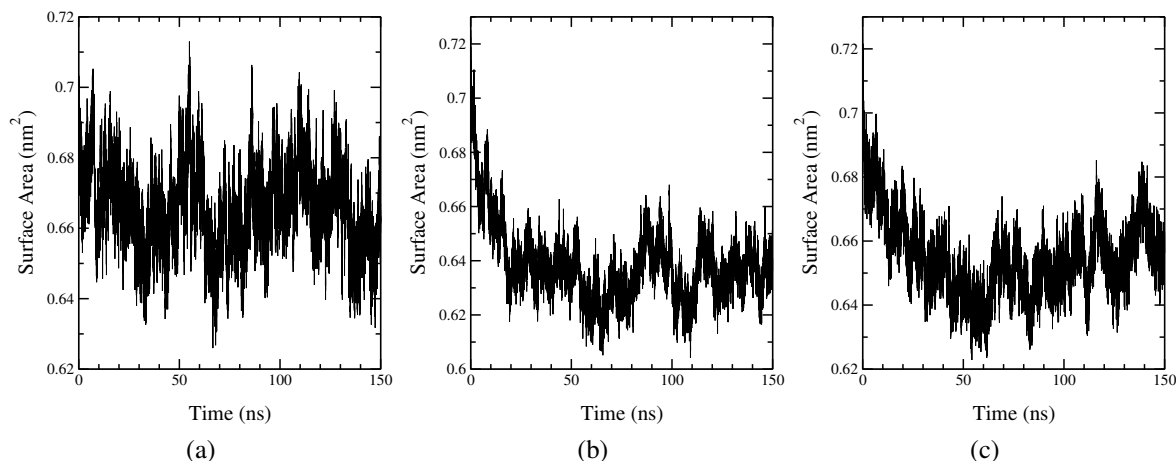


Figure 3.6: Area per lipid vs time for pure POPC (a), POPC+CHOL (b) and POPC+DHE(c).

The surface area per lipid is easy to calculate in a single component bilayer by dividing the total area of the bilayer by the number of lipids in a single leaflet. For binary mixtures and many component systems, this is no longer obvious, as has been discussed in recent works [125, 56, 126]. In this work the area per POPC is computed by considering the total area divided by the number of POPC molecules only (Table 3.1). For our purposes this is completely reasonable because our main objective is to compare the influence of DHE and CHOL on membrane systems. The area per POPC is comparable with other simulations. For example, Karttunen *et al.* have found it to be $(0.677 \pm 0.003) \text{ nm}^2$ at 323 K. Here we found it to be $(0.67 \pm 0.01) \text{ nm}^2$ at 303 K and the difference can be attributed to the difference in temperature. The surface areas given in Table 3.1 show that the presence of either sterol leads to membrane condensation, but the effect of cholesterol is a little stronger than its analog; for POPC-DHE the area per lipid is around 2% more than that of POPC-CHOL. The decrease of the surface area is associated with an increase of membrane thickness (Table 3.1).

3.3.2 Location and orientation of sterols in the bilayer

Figures 3.7 and 3.8 show the mass density profiles along the bilayer normal for the systems studied. Figure 3.8 has more details of the POPC-CHOL and POPC-DHE systems and gives insight to the local pressure profile of systems as will be discussed later. The density profiles indicate that CHOL and DHE have the same effect on the orientation of POPC molecules: The whole POPC molecule (Fig. 3.7) and its tails (Fig. 3.8) have the same density profiles. Regarding the sterols (CHOL and DHE), the position of sterol head groups is very similar. A

Table 3.1: Comparison of different properties of pure POPC, CHOL and DHE.

Membrane	POPC	POPC-CHOL	POPC-DHE
Area/POPC (nm ²)	0.67 ± 0.01	0.64 ± 0.01	0.65 ± 0.01
Thickness (nm)	2.63 ± 0.07	2.84 ± 0.06	2.73 ± 0.06
S_{mol}	0.3±0.01	0.5±0.01	0.5±0.01
P-N angle(°)	77.3±0.1	79.3±0.2	78.2±0.2
Tilt of palmitoyl(°)	32.6±0.1	20.9±0.1	22.8±0.1
Tilt of oleoyl(°)	34.9±0.1	25.0±0.1	26.0±0.1
Tilt of sterol ring structure(°)		23.2±0.1	27.1±0.1
Tilt of sterol tail(°)		33.6±0.1	46.3±0.1
Neighbors of sterol rings		52.1	51.4

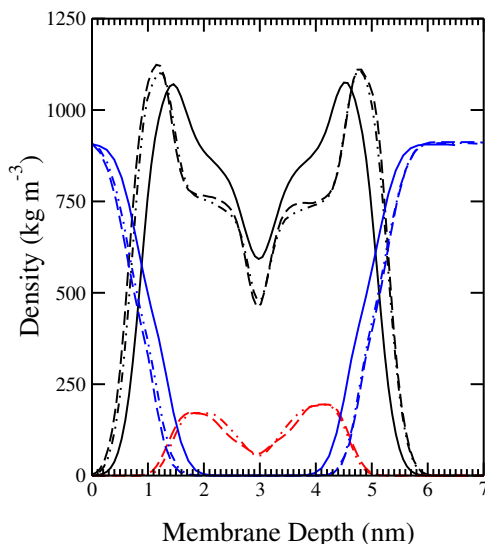


Figure 3.7: Partial density profiles along the bilayer normal for POPC (*black*), water (*blue*) and sterol molecules (*red*) for pure POPC (*solid line*), POPC+CHOL (*dashed line*) and POPC+DHE (*dash-dotted*)

slight change in the position of sterol tails and ring structures can be observed in Fig. 3.8 which is related to the different orientation of these components. That is, although the head groups of sterols are in the same place, the other parts of the body of sterols have different orientations. As will be shown later, the ring structure as well as the tails of sterols make different angles with the z -axis giving rise to a slight change in the density profiles of these components. The fact that the peak of both sterols' carboxyl group (OH) occurs at the same z coordinate as the peak of carbonyl group (C=O) of POPC suggests that sterols' head groups are mostly bound to the carbonyl group of POPC. This will be more discussed when analyzing the membrane-water interface (section 3.3.7).

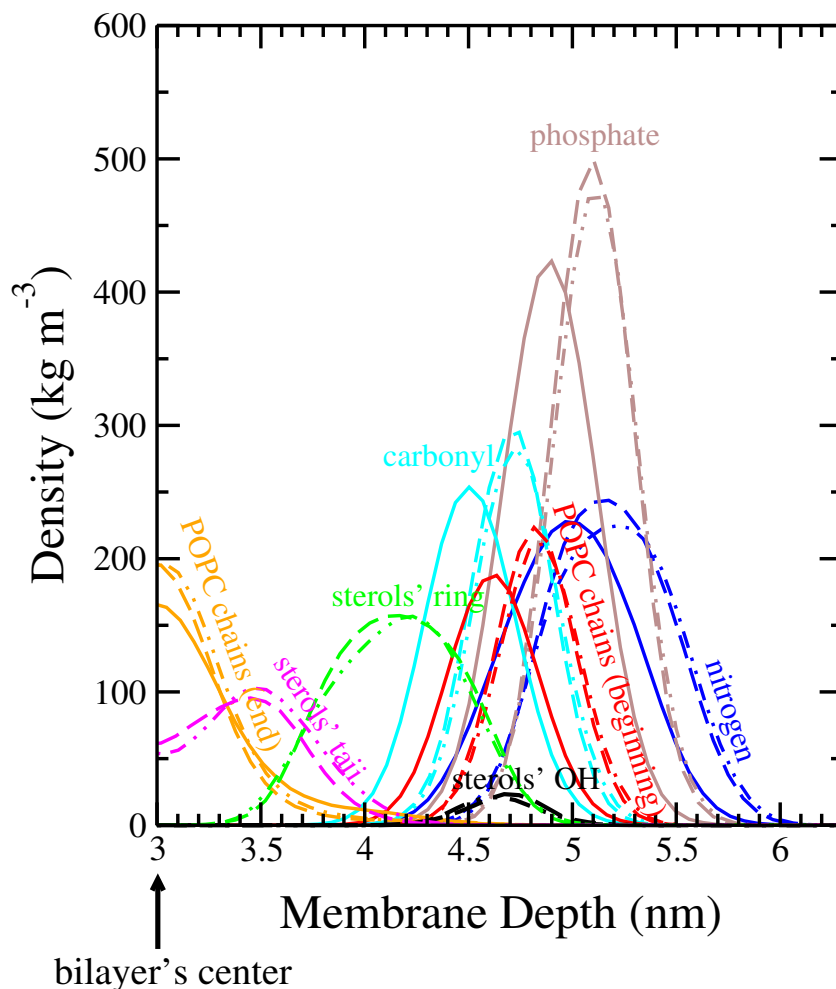


Figure 3.8: The density profiles of the lower leaflet of pure POPC (*solid line*), POPC-CHOL (*dashed line*) and POPC-DHE (*dashed-dotted*). The density profile represents the nitrogen group (NH_4^+), the phosphate group (PO_4^-), the beginning of POPC chains (C_{12} , C_{13} , and C_{32}), carbonyl groups ($\text{C}=\text{O}$), the end of POPC chains (C_{50} , and C_{52}), sterols' head groups (OH), sterols' ring system, and sterols' tails.

3.3.3 Order and conformation of acyl chains

Although the subtle difference in the structure of CHOL and DHE is not well reflected in the density profile, it does have an effect on the ordering of POPC acyl chains. This can be illustrated by the molecular order parameter, S_{mol} , whose profiles along the palmitoyl acyl chains of POPC are shown in Fig. 3.9. Both sterols increase S_{mol} of POPC acyl chains at all depths in the membrane. In table 3.1, mean values (average over segments 2-15) of S_{mol} for the palmitoyl acyl chains are given. The average values of S_{mol} over the whole chain in both POPC-CHOL and POPC-DHE systems are the same, but averaging over segments 4-10, where the difference is noticeable, leads to different result: ordering of POPC palmitoyl acyl chains is 8% higher in POPC-CHOL compared to POPC-DHE system. This is in line with

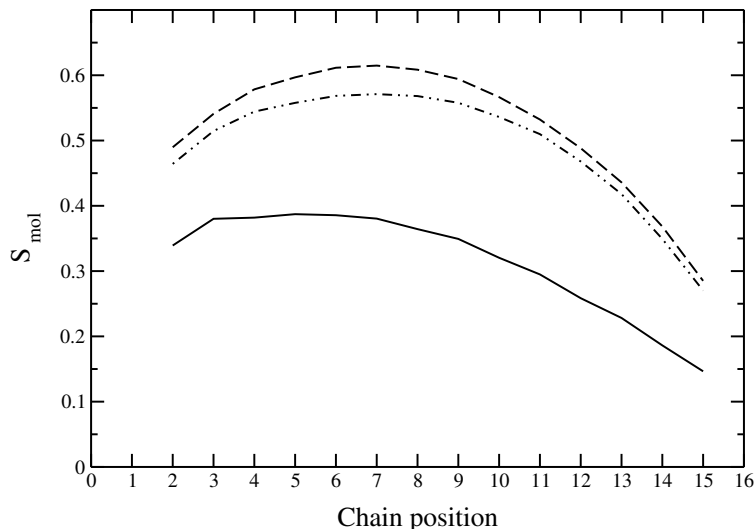


Figure 3.9: Order parameter S_{mol} of the palmitoyl acyl chains in pure POPC (solid line), POPC+CHOL (dashed line) and POPC+DHE (dash-dotted).

experimental results. For example, Scheidt *et al.* have experimentally shown that DHE has a lower acyl-chain-ordering effect in liposomes made of POPC compared to CHOL [134].

In addition to the order parameter for the acyl chains, we studied the ordering of head groups in a similar fashion. To this end, we have chosen the angle of the P-N vector with respect to the bilayer normal (z -axis) and have computed its distribution. The result is shown in Fig. 3.10 and the corresponding mean values are given in Table 3.1. The P-N vector has a significantly higher tendency of being in the bilayer plane ($\theta = 90^\circ$) than of sticking out of it. The presence of sterols makes the P-N vector a little more inclined to the bilayer plane (see Table 3.1) and this effect is just a bit more in the case of POPC-CHOL.

Distributions of the of palmitoyl and oleoyl acyl chains of POPC are shown in Fig. 3.11 and the corresponding mean values are given in Table 3.1. As seen from Fig. 3.11 and from the values of Table 3.1, in all the systems the tilt angle of the oleoyl acyl chain is higher than that of palmitoyl. This is because of the double-bond present in the oleoyl acyl chain. This double-bond, makes the dihedral angle of C-C=C-C to take a *cis* conformation and this leads to an increase in the tilt angle of oleoyl acyl chain. Figure 3.11 also shows that both CHOL and DHE decrease the tilt angles of both acyl chains. According to the Table 3.1, CHOL decreases the mean value of tilt of palmitoyl and oleoyl chains by $\sim 12^\circ$ and $\sim 10^\circ$, respectively. A similar effect is seen for DHE with a respective reduction of $\sim 10^\circ$ and $\sim 9^\circ$. Therefore, the effect of CHOL on the POPC acyl chains is a bit stronger than that of DHE on the tilt angle of POPC acyl chains.

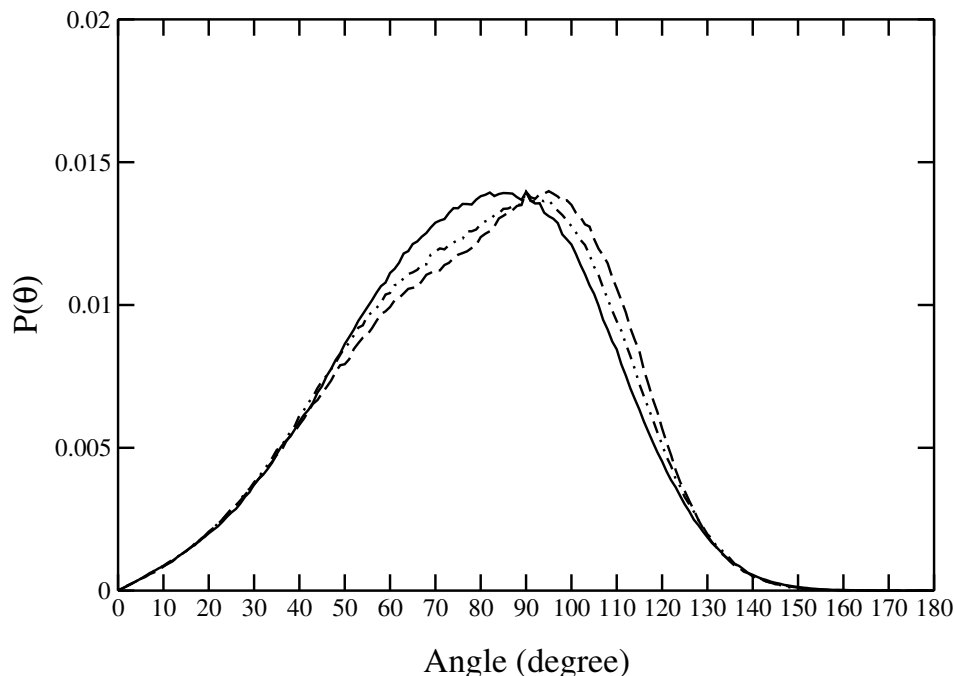


Figure 3.10: Normalized distribution of the angle θ between the P-N vector of POPC lipids and the z -axis in pure POPC (*solid line*), POPC+CHOL (*dashed line*) and POPC+DHE (*dash-dotted*).

3.3.4 Conformation of sterols

The key factor that is related to the effects of different sterols on bilayer properties is the sterol orientation in the bilayer [151, 161]. Studies indicate that the tilt of a sterol in a membrane correlates with its ordering and condensing ability. It should be kept in mind, though, that the microscopic origin of the tilt is in the atomic-level interactions, and thus the tilt of a sterol is a manifestation of its interactions with other molecules in the membrane. Nevertheless, the tilt angle seems to provide a physically meaningful and experimentally measurable quantity that can be used to compare the ordering properties of various sterols in membranes. A higher sterol tilt weakens the ordering effect on the sterols nearest neighbors [1]. Here, we measured the tilt angle of sterols' ring structure as well as their tails. For the ring structure, we measured the angle between the z -axis and the vector connecting the C_3 methyl group to C_{17} (See Fig. 3.2 for atom numbering). Distributions of such an angle are shown in Fig. 3.12 and the corresponding mean values are given in Table 3.1. The mean value of CHOL's tilt angle is 17% less than that of DHE. This indicates that CHOL should have a higher ordering effect than DHE. Indeed, as it is evident in Fig. 3.9 and was mentioned before, the palmitoyl order parameter is higher than that of DHE (8% difference for carbons 4 to 10). This is also in line with the more condensing effect of CHOL compared to DHE as can be seen from the less area per POPC in the POPC-CHOL system compared to POPC-DHE system (see Table 3.1). As a comparison

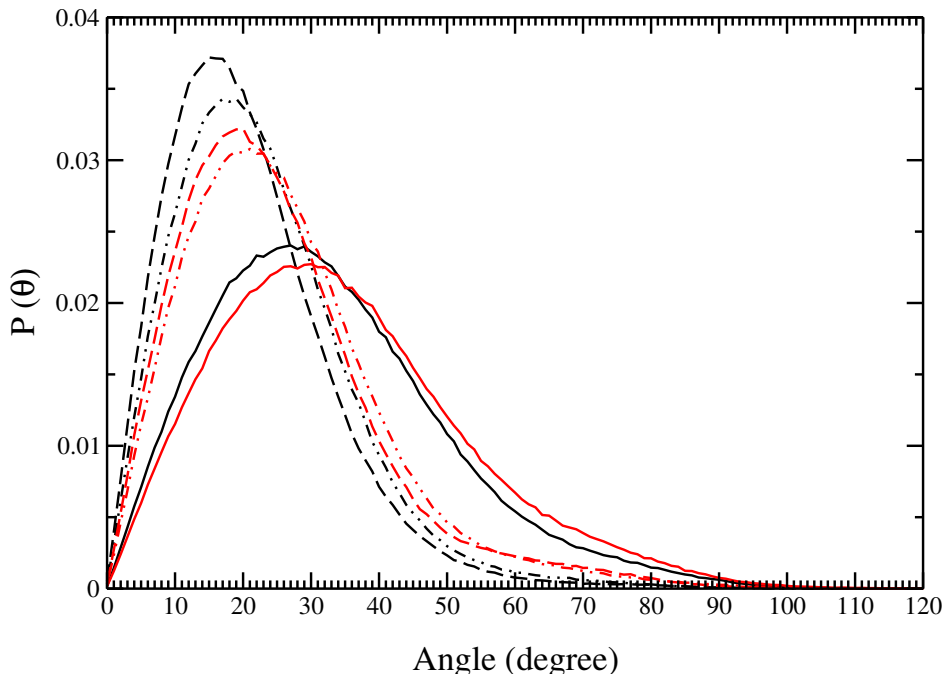


Figure 3.11: Normalized distribution of the tilt angle of palmitoyl (*black*) and oleoyl (*red*) acyl chains in pure POPC (*solid line*), POPC+CHOL (*dashed line*) and POPC+DHE (*dash-dotted*). Here, the angle is defined to be the angle between the z -axis and the vector connecting the two ends of POPC tails, namely C_{50} to C_{34} for palmitoyl, and C_{52} to C_{15} for oleoyl acyl chain (see Fig. 3.1 for atom numbering).

with other simulations, the average tilt angles of CHOL in CHOL+DOPC and CHOL+DPPC systems have been found to be 24.7° [127] and 19.8° [110], respectively. The corresponding value in CHOL-POPC (our simulation) is 23.2° . It is interesting that our value of tilt angle in POPC-CHOL system is between those in DPPC-CHOL and DOPC-CHOL. DPPC has two tails without any double-bond while DOPC has two tails with one double-bond in each of them. POPC, as depicted in Fig. 3.1, has one tail without double-bond and one tail with a double-bond. Therefore, it seems reasonable to say that the double bonds in lipid tails increase the tilt angle of CHOL.

To compare the orientation of CHOL and DHE in more detail, in addition to the sterol tilt, one can also study the angle between the z -axis and the vector connecting the two ends of sterols' tail, i.e. C_{25} to C_{17} (see Fig. 3.2 for atom numbering). This gives rise to the tilt angle of sterols' tails. The result is shown in Fig. 3.13 and the corresponding mean values are shown in Table 3.1. The mean value of the tail tilt of CHOL and DHE are $33.6 \pm 0.1^\circ$ and $46.3 \pm 0.1^\circ$, respectively. That is, the tail of CHOL has a higher tendency to stay close to the bilayer normal.

To shed more light on the structure of CHOL and DHE, and to understand how different

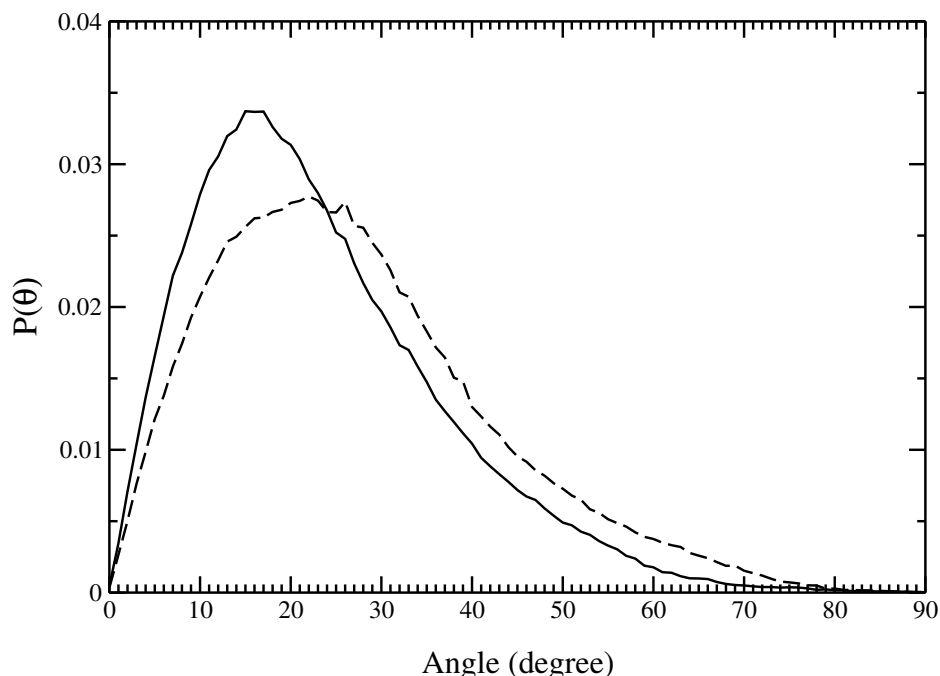


Figure 3.12: Normalized distribution of the tilt angle of CHOL ring structure (*solid line*) and DHE ring structure (*dashed line*). Here, the tilt angle is defined to be the angle between the z -axis and the vector connecting the two ends of the sterols' ring structure, i.e. C_3 and C_{17} (see Fig. 3.2 for atom numbering).

methyl groups locate in the structure of these sterols, we performed some analyses on the angles between different methyl groups. We wanted to know whether the C_{21} methyl group is on the same side as C_{18} and C_{19} or on the opposite side (For atom numbering see Fig. 3.2). To this end, we computed the dihedral angle between four methyl groups C_{18} - C_{13} - C_{20} - C_{21} . Notice that these groups are not consecutive, and the desired angle is the one between the vector connecting C_{20} to C_{21} and the plane formed by three methyl groups C_{18} , C_{13} and C_{20} . If this angle is less (greater) than 90° , C_{18} and C_{21} are (not) on the same side with respect to the ring structure. Similarly, we analyzed C_{29} in DHE to see if it is on the same side as C_{18} or not. Figure 3.14 shows the results. As evident from the figure, for C_{18} - C_{13} - C_{20} - C_{21} , these angles are greater than 90° in both CHOL and DHE, and for C_{18} - C_{13} - C_{24} - C_{29} , the angle is less than 90° . These results show that C_{18} and C_{21} are on the opposite side while C_{29} in DHE is on the same side as C_{18} .

Combining the results of this section, we are able to construct a schematic picture of CHOL and DHE in the POPC bilayer. In this picture, the sterols have two parts, namely a ring system and a tail. The methyl groups protrude from the body of this structure as shown in Fig. 3.15. In DHE, the C_{29} methyl group is on the same side as C_{18} and C_{19} . The C_{21} methyl group is on the opposite side in both CHOL and DHE.

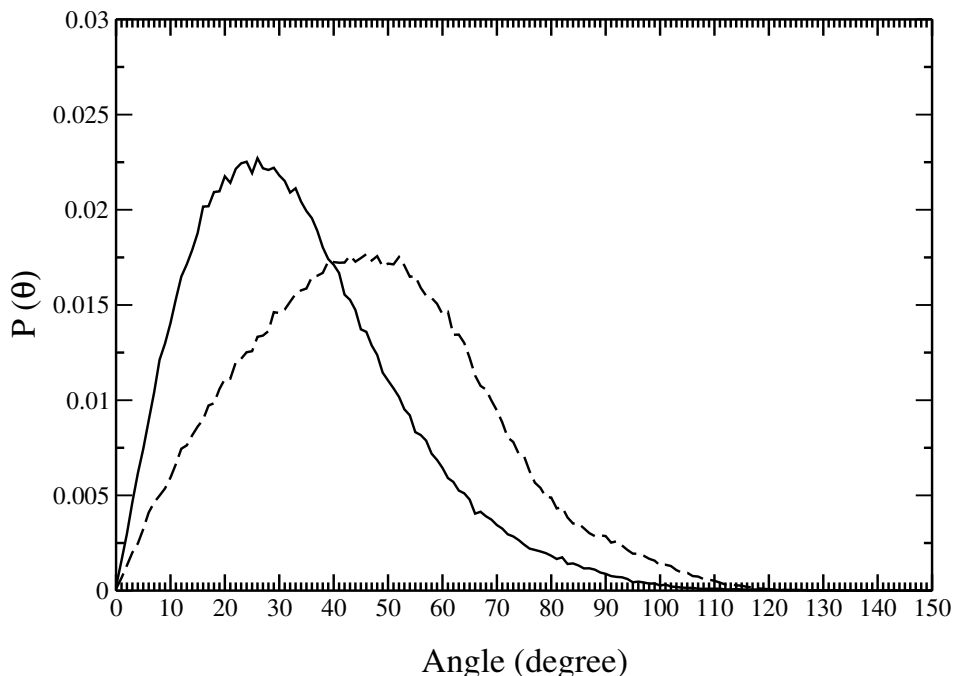


Figure 3.13: Normalized distribution of the tilt angle of CHOL tail (*solid line*) and DHE tail (*dashed line*). Here, the tilt angle is defined to be the angle between the z -axis and the vector connecting C_{17} to C_{25} (see Fig. 3.2 for atom numbering).

3.3.5 Packing of chains

Radial distribution function (RDF), or pair correlation function $g(r)$, of two sets of particles with respect to each other, describes how the average density of one set of particles varies as a function of the distance from the other one. This quantity is of fundamental importance in thermodynamics as macroscopic thermodynamic quantities can be calculated using $g(r)$. One can measure the three-dimensional (xyz) or two-dimensional (xy) distance of two sets of molecules to obtain a three-dimensional or two-dimensional RDF, respectively. As shown in Fig. 3.16a, we computed the two-dimensional RDFs of the centers of mass of sterols versus themselves (i.e. CHOL versus CHOL and DHE versus DHE). In Fig. 3.16b, the two-dimensional RDFs of sterol ring structures (carbon atoms C_1 to C_{17} as depicted in Fig. 3.2) versus different tails of POPC are also shown.

As seen in Fig. 3.16a, the peaks of the RDF of POPC+CHOL system is more pronounced compared to POPC-DHE. This shows a more ordered packing of CHOL molecules. Nevertheless, the position of peaks is approximately the same in both CHOL and DHE systems. Of all the peaks in the RDF graph of CHOL-CHOL and DHE-DHE, the first one is the strongest. This peak occurs at ~ 1 nm showing that the majority of sterols tend to maintain a distance of 1 nm with respect to each other.

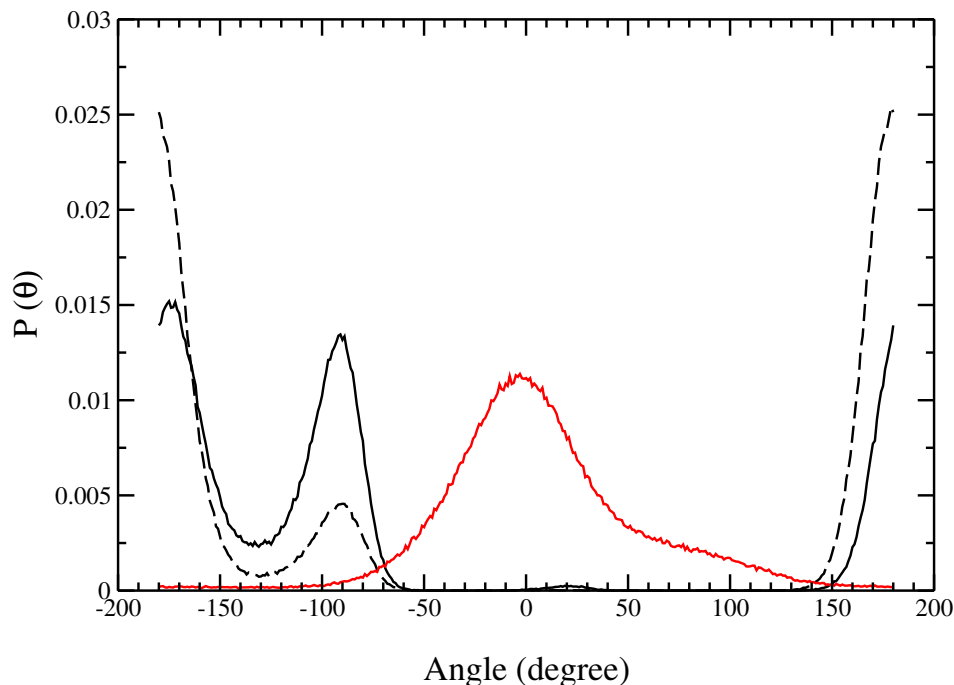


Figure 3.14: The angle between the vector connecting C_{20} to C_{21} and the plane formed by three methyl groups C_{18} , C_{13} and C_{20} in CHOL (*solid black line*) and DHE (*dashed black line*). The graph in *red* relates to the extra methyl group C_{29} and is the angle between the vector connecting C_{24} to C_{29} and the plane formed by three methyl groups C_{18} , C_{13} and C_{24} . For atom numbering see Fig. 3.2.

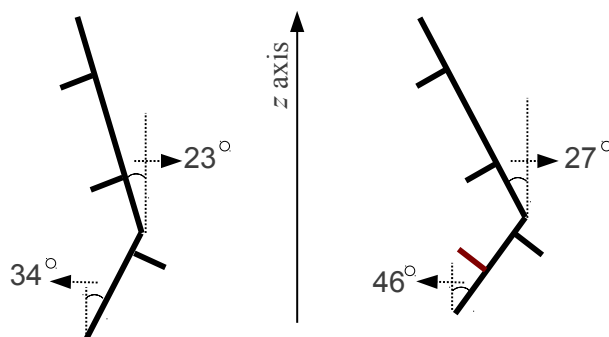


Figure 3.15: Schematic illustration of CHOL and DHE. The sterols are divided into two main parts, a ring system and a tail. The other methyl groups are budded out of the body of sterols in the way which is shown. The extra methyl group of DHE (C_{29}) is shown in *red*.

As seen in Fig. 3.16b, both palmitoyl and oleoyl acyl chains have similar RDFs with the same position of peaks. The only difference is that the peaks are stronger in the case of POPC-CHOL system compared to POPC-DHE system. The strongest peaks in all the RDF graphs are the first ones which occur at ~ 0.6 nm. By comparing the position of the strongest peaks (see Figs 3.16a and 3.16b) one can conclude that both CHOL and DHE avoid being located in

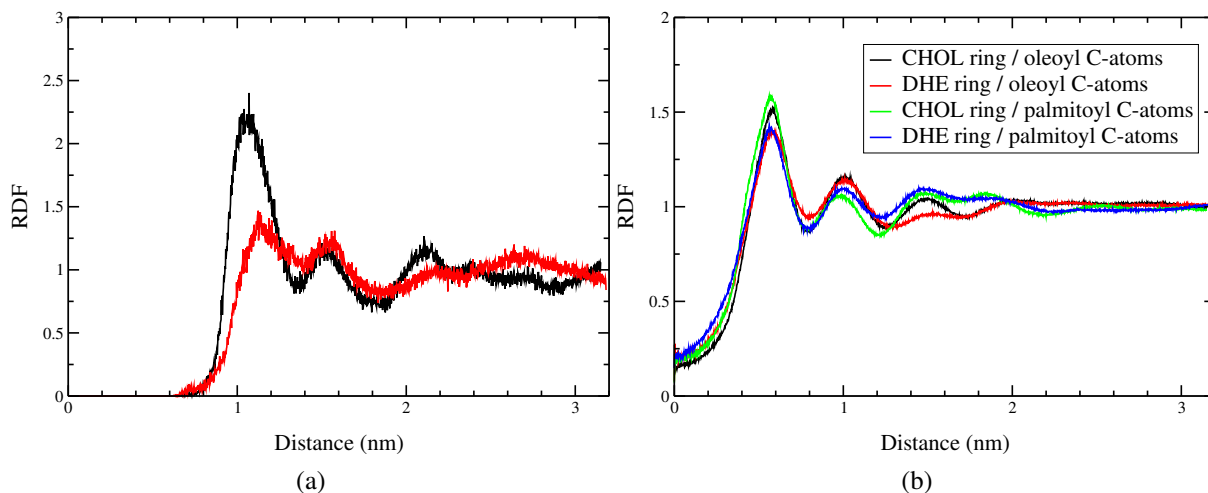


Figure 3.16: Two-dimensional RDFs of COM of CHOL- (*black*) and DHE-ring (*red*) versus themselves (a), and of COM of acyl chains (palmitoyl or oleoyl) versus sterols' ring structure (C₁ to C₁₇) (b).

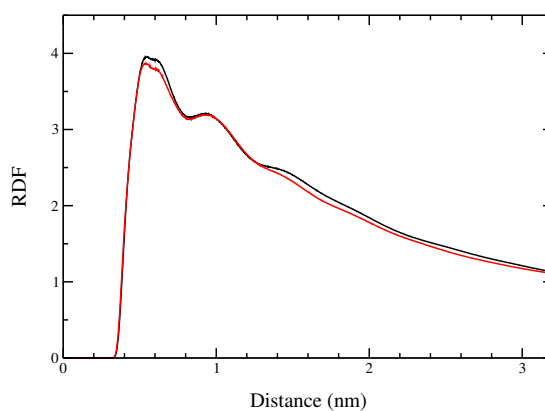


Figure 3.17: Three-dimensional RDFs of carbon atoms of POPC molecules versus sterol ring (C₁ to C₁₇).

adjacent positions. Rather, they prefer to be surrounded by POPC molecules. This result is in line with other simulations [108, 123, 95]. For example, Reigada *et al.* [95] showed that only at high CHOL concentrations (> 30 mol%) does the occurrence of close contacts (< 0.8 nm) become relevant since then the membrane is crowded with CHOL molecules. In our case, the CHOL concentration is 20% and they avoid being in close contact with each other.

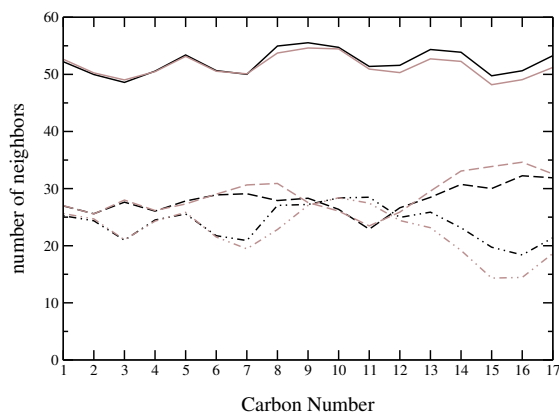


Figure 3.18: Number of neighbors of sterol ring methyl groups located on α -side (*dashed line*), β -side (*dashed-dotted*) and both sides (total) (*solid line*) for CHOL (*black*) and DHE (*gray*).

3.3.6 Packing of atoms relative to sterol ring atoms

To better understand how POPC carbon atoms locate around sterols, we first studied the three-dimensional RDF of POPC carbon atoms relative to carbon atoms belonging to sterol ring (the methyl groups C_{18} and C_{18} were not included) as shown in Fig. 3.17). The RDF profiles look alike in both cases showing the first peak at around 0.5 nm, which corresponds to the van der Waals distance of carbon atoms, and the minima at 0.8 nm. Therefore, the closest neighbor carbon atoms of sterols' ring are within an imaginary sphere of radius 0.8 nm. It is then interesting to know how these close neighbors are distributed around the sterol rings, or to know which side of the sterol, α -face or β -face, they like more. To quantify this, we studied the average number of neighbors of the sterol rings. The neighbor of an arbitrarily chosen carbon atom in a sterols' ring is defined to be an atom belonging to the POPC molecule and located no further than 0.8 nm (the position of the first minimum in the RDF) from the carbon atom in question. In addition to measuring the distance between a POPC carbon atom C_{POPC} and a sterol ring carbon atom C_{ring} (which should be < 0.8 nm), we need to determine whether the neighboring atom C_{POPC} is on the β -side of the sterol ring structure or on its α -side. To this end, we measured the angle $C_{\text{POPC}}-C_{19}-C_{\text{ring}}$. If this angle is less than 90° , the neighboring carbon atom is on the β -side of the sterol ring. Otherwise, it is on the α -side.

Figure 3.18 shows the average numbers of the so-defined neighbors. As for the average number of neighbors of the CHOL ring, we found 52.1, of which 28.1 are located in the α -side, and 24.0 in the β -side. For the DHE ring, this number turned out to be 51.4, of which 28.9 are located in the α -side, and 22.5 in the β -side. Except for C_{10} and C_{11} , for both CHOL and DHE systems, the number of α -side neighbors is always higher than those in β -side. For C_{10} and C_{11} which are close to the methyl group C_{19} , the number of neighbors is more in the

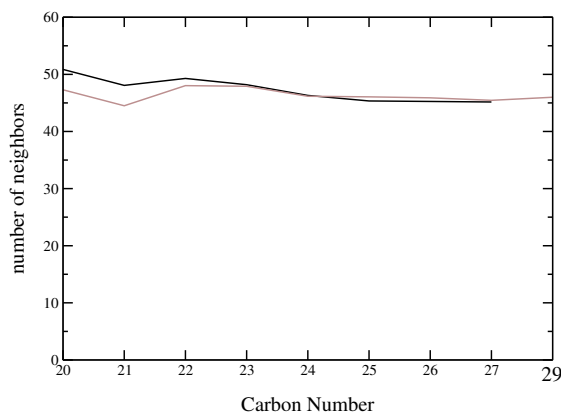


Figure 3.19: Number of neighbors of sterol tails of CHOL (*black*) and DHE (*gray*). Here, the tail of CHOL is numbered from methyl groups C₂₀ to C₂₇. DHE has one more methyl group, namely C₂₉, which is shown in a larger font size to emphasize that there is no C₂₈ in the numbering of DHE.

β -side than in the α -side. This is in line with other simulation on CHOL+DPPC system [126]. For methyl groups close to the OH-group (C₁ to C₆), the number of neighbors is essentially the same in CHOL and DHE. This is also in agreement with another simulation [126] on CHOL and a “smooth” CHOL called demethylated CHOL (with C₁₈ and C₁₉ removed). For methyl groups C₁₂ to C₁₇, the number of neighbors on the α -side is higher in DHE system than in CHOL system. This is understandable if we look at the schematic picture of these sterols in Fig. 3.15: DHE is more tilted and this makes it harder for the neighboring atoms to become close to DHE from the β -side. Instead, it is easier for them to reach the sterol ring from the α . The total number of neighbors, however, is more in CHOL system than in DHE system.

The same analysis can be performed on the sterols’ tails. As shown in Fig. 3.19, the number of neighbors is higher in CHOL system for methyl groups up to C₂₀. This is also expected since CHOL is less tilted than DHE and is more exposed to the neighboring methyl groups of POPC molecules.

3.3.7 Membrane-water interface

To elucidate the effect of DHE on the bilayer/water interface, we analyzed the atomic-level interactions of the hydroxyl groups (OH) of CHOL and DHE with POPC head groups. The OH group in DHE, like the OH group in CHOL, participates in hydrogen bonding with water and POPC oxygen atoms. The average numbers of OH-DHE and OH-CHOL H-bonds with water and the oxygen atoms belonging to the phosphate group of POPC are given in Table 3.2. The H-bond pattern is almost the same for both sterols: they make H-bonds predominantly with the carbonyl group of the *sn*-2 chain (46% of all H-bonds between CHOL and POPC, and 51% of

Table 3.2: Hydrogen bonds and charge-pair interactions per sterol in the membrane/water interface. Atom numbering is shown in Fig. 3.1.

	POPC+CHOL	POPC+DHE
Sterol-water H bonds	0.43	0.43
Sterol-PC H bonds	0.90	0.88
Sterol-POPC(O7)	0.14	0.12
Sterol-POPC(O9)	0.01	0.01
Sterol-POPC(O10)	0.02	0.01
Sterol-POPC(O11)	0.01	0.01
Sterol-POPC(O14)	0.16	0.08
Sterol-POPC(O16)	0.46	0.51
Sterol-POPC(O33)	0.03	0.04
Sterol-POPC(O35)	0.07	0.10
Sterol-POPC charge-pair	3.74	3.53
Sterol(O)-POPC(methyl) charge-pair	0.82	0.82

all H-bonds between DHE and POPC). This result is comparable with other simulations [126]. For example, in a DPPC-CHOL simulation with the same mole fraction of CHOL, Karttunen *et al.* found that, of all the possible H-bonds between OH-CHOL and DPPC molecules, the number of H-bonds between the carbonyl group of *sn*-2 chain and DPPC molecules is the highest (0.38 H-bond per CHOL) [126]. This is also in line with the density profile shown in Fig. 3.8 in which OH head groups of sterols are at the level of carbonyl groups of POPC chains.

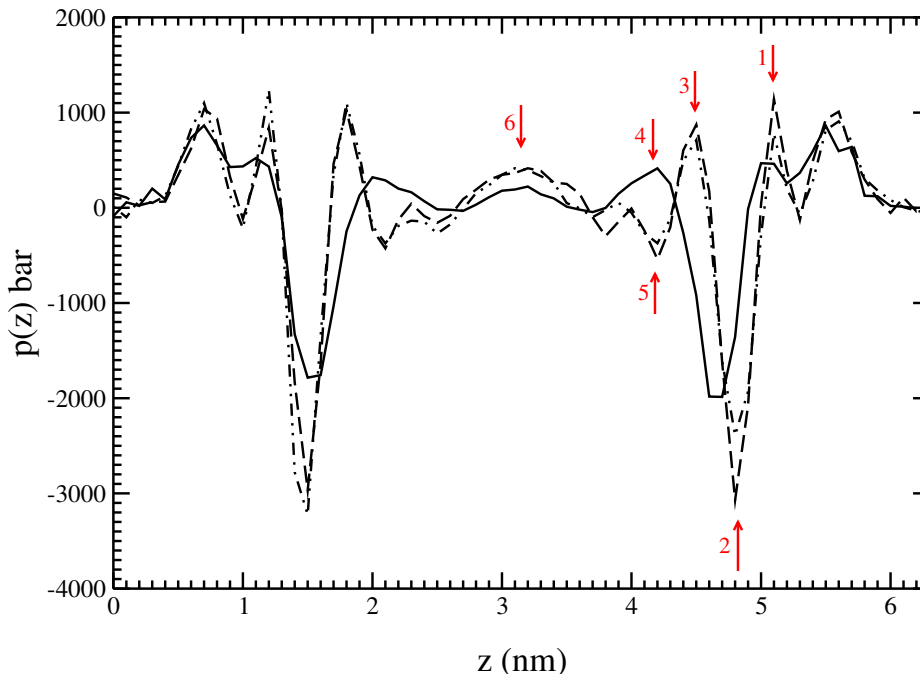


Figure 3.20: Lateral pressure for pure POPC (*solid line*), POPC+CHOL (*dashed line*), and POPC+DHE (*dashed-dotted*). The horizontal axis corresponds to the bilayer normal. Important maxima and minima of the profiles are indicated by *red* arrows and are explained in the text.

3.3.8 Lateral pressure

The distribution of local pressure inside the bilayer along the bilayer normal (z -axis) is commonly referred to as lateral pressure, $p(z)$. The lateral pressure profile is related to many important macroscopic and measurable quantities, such as surface tension, surface free energy, and spontaneous curvature [132]. $p(z)$ arises from local forces acting inside a lipid bilayer in the direction of the membrane plane. The condition for mechanical stability and equilibrium requires that the integrated lateral pressure profile across the membrane vanishes. Nevertheless, the profile may display different behavior in different regions in the membrane due to a variety of interactions whose relative importance varies across the membrane. Traditionally, three different regimes have been identified: (1) a repulsive contribution in the hydrophilic head group region due to electrostatic and steric interactions and hydration repulsion; (2) an attractive contribution at the membrane-water interface due to the interfacial energy between the water and the hydrocarbon phase, trying to minimize the surface area; and (3) a repulsive contribution inside the membrane due to steric interactions between hydrophobic chains [94]. These forces are assumed to create a non-uniform lateral component of local pressure inside a bilayer [94]. The details may vary considerably from one system to another. The local pressure for a system consisting of particles with many-body potentials can be defined using the local

stress tensor. The lateral pressure profile $p(z)$ is then defined as a difference between the normal (along the membrane normal direction) and the lateral components of the pressure tensor, that is, $P_N = P_{zz}$ and $P_L = (P_{xx} + P_{yy})/2$: $p_z = P_L - P_N$. Qualitatively, this means that a bilayer tends to expand along the membrane xy -plane with positive $p(z)$ and contract with negative $p(z)$.

Local pressure is difficult to measure experimentally. In experiments, probes should be employed in different parts of membrane components [146], thus the measured system is not a native one, as is preferred. Computational studies based on atomistic modeling of lipids, however, can shed light on this issue. Here, by means of MD simulations, we analyzed the dependence of lateral pressure profile on the sterol type. The results are presented in Fig. 3.20. The lateral pressure was determined using a customized version of Gromacs, Gromacs 4.0.2. The approach is to re-run the simulations using the customized version of Gromacs while the system is divided into 0.1 nm thick slabs for which the pressure is calculated. Pairwise forces during the pressure calculation are computed from the force field description and MD trajectory. Because pairwise electrostatic forces cannot be obtained using PME method, re-run simulations were carried out by truncating the long-range interactions. As Sonne *et al* [142] showed that the truncation distance for pressure calculation should be $r_{\text{cut}} > 1.8$ nm, we used $r_{\text{cut}} = 2.0$ nm as for electrostatic cut-off distance.

As shown in Fig. 3.20, the presence of either sterol affects the qualitative nature of the pressure profile of pure POPC. In CHOL-POPC and DHE-POPC systems the number of peaks and their heights are more than that of pure POPC. To analyze the pressure profile, one should compare it with the density profile presented in Fig. 3.8 and relate the maxima and minima in the lateral pressure profile to the maxima in the density plot. These maxima and minima are indicated in Fig. 3.20. The local maximum #1, which is common in pure POPC and sterol-POPC systems, corresponds to the repulsion of head groups of POPC lipids since the density profile of nitrogen and phosphate groups has its maximum at that region. The absolute minimum #2, which is common in pure POPC and sterol-POPC systems, is typical for all membranes and corresponds to the region where the tails start, namely C_{12} , C_{13} and C_{32} (see Fig. 3.8 for comparison with the density profile). It is an attraction contribution to the pressure profile and is because the membrane tries to minimize its surface area to prevent water from reaching to the hydrophobic region. The next local maximum, #3, which is solely present in sterol-POPC systems is because of the repulsion of OH group of sterols. Its occurs at the same position where the peak of OH density profile does. There is a local maximum, #4, in pure POPC which corresponds to the steric repulsion of POPC acyl chains: the acyl chains repel each other at this region because they are already very close at the parts near the water-membrane interface (peak #2). The next local minimum, which is also not present in the pure POPC system, #5, occurs approximately at $z = 4.3$ nm and corresponds to the ring structure of sterols. This

means that at this region the membrane components attract each other. In the middle region of the bilayer, where peak #6 occurs, the pressure is almost zero in pure POPC, thus this region can be regarded as a free zone. When sterols are present in the system, however, there are repulsive contribution to the pressure profile due to steric repulsion of lipid tails: because lipids are tightly packed at the sterols' ring area, here, where the ring is no longer present, they try to avoid each other.

Comparison of lateral pressure of CHOL-POPC and DHE-POPC reveals no major difference. However, studies show that the differences due to the sterols in unsaturated bilayers, such as POPC, are considerably smaller than in a saturated ones [110, 109]. Therefore, one should expect more differences to emerge when sterols interact with saturated bilayers. In other words, differences in effects of either sterol on bilayers' lateral pressure and elasticity are more pronounced when they interact with saturated bilayers such as DPPC.

3.3.9 Discussion

Both CHOL and its analogue increase the ordering and decrease the area per POPC, but the effect of CHOL is stronger. To understand this difference, in previous sections we performed several analyses on both sterols. All the differences in behaviors of sterols originate from the interactions between sterols and POPC molecules. As we showed, the interactions in the head group region is almost the same and both sterols similarly anchor to the carbonyl as well as nitrogen group of POPCs through H-bond and charge-pair interactions. Inside the bilayer, where the methyl groups of sterols and POPCs are present, van der Waals interactions are involved. Differences in van der Waals interactions in CHOL and DHE systems must be the reason why CHOL has a less tilt angle than DHE, thus stronger ordering and condensing effect. Differences in packing of carbon atoms of POPC relative to carbon atoms of sterols reveal the differences in van der Waals interactions in CHOL and DHE systems. A question remains as to why CHOL has more van de Waals interactions with POPCs than DHE. Undoubtedly, this difference is the consequence of structural differences in CHOL and DHE. DHE has two more double bonds in its ring structure, and one more double bond in its tail. In addition, DHE has an extra methyl group attached to its tail. To investigate the effect of these structural differences, we need to deal with each of them individually.

A lot of simulations as well as experiments have been performed on the CHOL-like molecules interacting with PC membranes [126, 30, 110, 151, 95, 128]. Here, we consider demethylated cholesterol (DCHOL), desmosterol (DESMO), 7-dehydrocholesterol (7-DHC), and ergosterol (ERG) the chemical structure of which are shown in Fig. 3.21. In DCHOL, the two methyl groups C_{18} and C_{19} , which are present in the β -face of CHOL, are removed. Thus, DCHOL

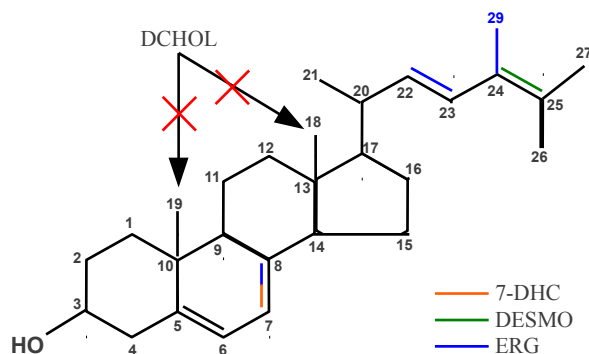


Figure 3.21: Different sterols that are similar to CHOL (*black*) in structure. Differences to CHOL are shown in different colors. DCHOL does not have the two methyl groups C₁₈ and C₁₉. DESMO has an extra double bond between C₂₄ and C₂₅. 7-DHC has an extra double bond between C₇ and C₈. ERG has a double bond between C₇ and C₈, a double bond between C₂₂ and C₂₃, and an extra methyl group attached to C₂₄.

is a “smooth” CHOL not having the rough β -face. DESMO is the same as CHOL in structure except it has a double bond in its tail between C₂₄ and C₂₅. 7-DHC is the same as CHOL in structure except it has one more double bond in the ring system between C₇ and C₈. ERG is also similar to CHOL in structure, but has one more double bond in the ring system between C₇ and C₈ (like 7-DHC), a double bond between C₂₂ and C₂₃ (like DHE), and an extra methyl group attached to C₂₄ (like DHE). Thus, ERG is very much similar to DHE in structure, except it does not have a double bond between C₉ and C₁₁.

Of all these cholesterol-like molecules, only ERG has a more ordering effect than CHOL [30]. In all others, the tilt angle of the sterols is more than that of CHOL leading to less of an ordering effect of sterols [126, 110, 151, 95, 128]. Structural differences in these molecules come from either the presence or absence of double bonds, whether in the ring structure or in the tail, or from the presence or absence of methyl groups, whether in the ring structure or in the tail.

To elucidate the importance of double bonds, we concentrate on 7-DHC, which differs from CHOL only by one more double bond in the ring structure, and also on DESMO, which has only one more double bond in the tail. Studies have shown that in systems comprised of PC molecules and these sterols, sterols’ tilt angle is greater than that of CHOL [110, 128]. Thus, the presence of double bond, whether in the ring structure or in the tail, increases the tilt angle of sterols. In the cholesterol force field developed by Holtje *et al.* [57], which is based on GROMACS, a carbon atom involved in a double bond (CR61) can interact with a PC carbon atom (LP2), described by Berger *et al.* [15] force field. The Lennard-Jones parameters for such an interaction leads to a van der Waals radius which is greater than the van der Waals radius of

a carbon atom not involved in the double bond interacting with a PC carbon atom. Moreover, the van der Waals potential minimum of CR61 and LP2 is also shallower than the one of a carbon atom not involved in the double bond interacting with a PC carbon atom. This means that the presence of double bond weakens van der Waals interactions. Therefore, it is sensible to expect that DESMO and 7-DHC are less tightly bounded to the neighboring PC atoms and their ordering ability is less effective than CHOL.

The importance of methyl groups in the ring structure of CHOL can be understood when we note that in DCHOL, which does not have the two C₁₈ and C₁₉ methyl groups, the tilt angle is more than that of CHOL. This should be due to the absence of van der Waals interactions between neighboring PC molecules and these methyl groups in the case of DCHOL. In other words, these methyl groups act as “handles” that are exposed to the neighboring molecules via which CHOL is kept in a more upright orientation (see Fig. 3.15). When we compare ERG and 7-DHC, we see that ERG has an extra double bond in its tail which should in turn decrease the van der Waals interactions. But as ERG has even higher ordering effect than CHOL, the presence of the extra methyl group attached to C₂₄ must be crucial in the almost upright orientation (13.2° [30]) of this sterol.

Based on the above discussion, one can distinguish two competing factors that affect the sterols orientation in a membrane, one being possible double bonds in the structure of sterols which can lower the van der Waals energies, and the other being the methyl groups that can act as “handles” via which the sterols can be kept in a more upright position, thus interact with neighboring atoms more. ERG and DHE seems to be two examples in each of which one of these factors is dominant. ERG has a lower tilt angle, and consequently, higher ordering effect than CHOL [30]. This molecule has two double bonds in the ring structure, i.e. one double bond more than CHOL, and also a double bond in its tail. These double bonds should reduce the van der Waals interactions in the way which was explained above. But the extra methyl group attached to C₂₄ seems to be the dominant factor. In DHE, however, there is one more double bond than ERG in the ring structure, i.e., there are three double bonds in the ring structure in total. This extra double bond seems to favor a more tilted orientation of the sterol leading to a higher tilt angle and less ordering effect of DHE as was shown in our study.

3.3.10 Conclusion

In this study, we compared the effect of CHOL and DHE on POPC membranes. By analyzing the order parameter and area per lipid, we showed that DHE is weaker in ordering and condensing of POPC membranes. We also showed that the orientation of these sterols in the membrane is different: CHOL has a more vertical orientation than DHE favoring more van der Waals interactions. This was evident from the number of packing carbon atoms around the sterols' ring and tail (Fig. 3.18 and 3.19).

Chapter 4

DPPC Bilayer Interacting With Transportan

4.1 Introduction

Cell-penetrating peptides (CPPs) are short peptides of less than 30 amino acids that are able to penetrate cell membranes and transport different cargo into cells without disrupting them [164]. Most of these peptides have a positive net charge and/or an amphipathic nature, but can otherwise have very different characteristics [39]. The mechanism of cell translocation is not known. As a matter of fact, because different amino acid motifs have been observed in different CPPs, it is believed that the translocation mechanism can differ for different CPPs[164]. Nevertheless, the mechanism seems to be receptor and energy independent. In certain cases translocation is partially mediated by endocytosis[38, 81]. A variety of possible mechanisms have been proposed, including the carpet model [145], the formation of transient pores [48, 52], the formation of inverted micelles [33], local electroporation [36], micropinocytosis [163] and direct insertion of an unfolded peptide into the membrane [137]. Cargoes that are successfully translocated by CPPs range from small molecules, such as small peptides or fluorescent labels, to proteins and supermolecular particles[164]. Most CPPs are inert or have insignificant side effects[164]. The half-time of penetration of these peptides ranges from 5 to 20 minutes[164]. CPPs are novel vehicles for delivering cargoes into cells, and thus, have promising applications in drug delivery.

Transportan (GWTLN SAGYL LGKIN LKALA ALAKK IL-amide), a 27 amino-acid-long member of a class of galanin-based chimeric peptides, is an amphipathic cell-penetrating peptide which was designed and synthesized in 1996 [78]. It is constructed from the 12 N-terminal residues of galanin, the 14 C-terminal residues of mastoparan and a connecting lysine. Galanin

is a neuropeptide that is widely expressed in the brain, spinal cord, and gut of mammals, and mastoparan is a peptide toxin from wasp venom.

Several experiments have been performed on transportan-lipid systems [6, 7, 119, 86, 27], but the penetration mechanism still remains unknown. It has been shown [119] that internalization of radiolabelled transportan is receptor/protein independent and that the peptide penetrates into every mammalian cell type (Bowes's melanoma, HeLa, HEK293, Jurkat, etc.) with the same localization on the plasma membrane, nuclear membrane and all other intracellular membrane structures. In the same study[119], it was also shown that endocytosis cannot be involved in the cellular uptake of the peptide since penetration is not blocked by the presence of inhibitors of endocytosis. The penetration was observed to be slightly concentration dependent with a higher relative uptake at lower concentrations[119]. In plant cells, the cellular uptake of transportan was observed to be tissue dependent being significantly weak in leaf tips, and strong in protoplast which is comprised of phospholipids[27]. Moreover, like the mammalian cells, in plant cells the penetration is concentration dependent and does not involve endocytosis[27]. The overall dynamics of the peptide is significantly slower in the partly negatively charged bicelle environment [7]. Transportan has a close-to-random coil secondary structure in water [86], but adopts a helical structure when bound to phospholipid bicelles as well as SDS micelles with the helix mostly localized to the mastoparan part[86].

Several MD simulations on cell-penetrating peptides (CPPs) and antimicrobial peptides (AMPs) can be found in literature [82, 37, 163, 139, 83, 52]. In some of them, but not all, translocation was observed within simulation times (hundreds of nanoseconds). For example, spontaneous pore formation was not observed for penetratin (a CPP) interacting with phospholipids such as palmitoyl-oleoyl phosphatidylcholine (POPC), palmitoyl-oleoyl phosphatidylglycerol (POPG), palmitoyl-oleoyl phosphatidic acid (POPA), and palmitoyl-oleoyl phosphatidylserine (POPS) membranes[82], and for the 21-amino-acid-long transportan analogue, transportan-10 (a CPP), interacting with POPC[37], and for penetratin and TAT peptide (a CPP which is obtained from the transcription transactivation protein from human HIV-1 virus) with dioleoyl phosphatidylcholine (DOPC) and dipalmitoyl phosphatidylcholine (DPPC) membranes[163]. Penetration was observed, however, for Melittin (an AMP) interacting with DPPC membranes[139], magainin-2 (an AMP) with DPPC membranes[83] and also TAT peptide with DOPC membranes[52], although the latter simulation has been argued to include artifacts and not to represent experimental conditions appropriately [163] as their system was strongly charged and not neutralized by counter ions leading to a high membrane expansion[52].

Computer simulations are particularly helpful in gaining detailed information about biophysical systems and can be regarded as complementary to experiments. The spontaneous translocation of CPPs is known from experiments to be a rare process [164], making it very

hard to simulate translocation within current simulation times accessible. However, simulating these peptides can shed light on the possible translocation mechanism. By studying the peptide/membrane interactions and behavior, some mechanisms can be ruled out or the peptide/membrane conformations can be studied in early stages of penetration. To our knowledge, no computer simulation has been performed on transportan interacting with a bilayer. Here, we present results of atomistic molecular dynamics simulations performed to enhance the knowledge of the nature of the interaction between transportan and a zwitterionic membrane (DPPC).

4.2 Simulation Details

All the simulations were performed at constant temperature and pressure with a constant number of particles (NpT ensemble). Simulations were initiated using GROMACS 3.1.4 [84], but higher versions up until 4.0.5 [54] were used in extending them. The simulation systems consisted of 2×64 dipalmitoyl phosphatidylcholine (DPPC) lipids hydrated with ≈ 7500 water molecules, and one transportan peptide. The initial size of the system was $6.5 \times 6.5 \times 9.5 \text{ nm}^3$. Figure 4.1 shows a schematic representation of DPPC with atom numbering used in this study.

We performed 10 unbiased simulations corresponding to 10 different initial configurations. As for the initial configuration of lipids, we used the coordinates of previously studied DPPC lipids in water [117] to which we added transportan. The initial conformation of transportan was obtained from the parameters developed by Maler *et al.* (PDB code: 1SMZ) [6] in which they utilized Nuclear Magnetic Resonance (NMR) spectroscopy to study the structure of transportan in phospholipid bicelles. As they showed, the peptide takes a more or less helical conformation when associated with the bilayer. We used that helical conformation as for the initial conformation of the peptide. The conformation at other times was not constrained and the peptide was able to change it. Figure 4.2 shows representative snapshots of the system.

The four positive charges carried by transportan were compensated by adding the corresponding number of Cl^- counterions. The systems were energetically minimized using the steepest descent method before the MD simulations. 8 simulations were run up to around 90 ns, 1 up to 320 ns, and 1 up to 780 ns.

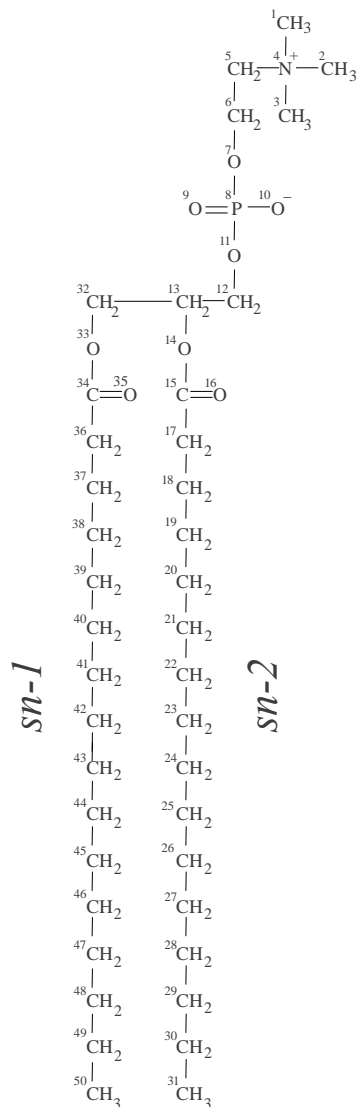


Figure 4.1: Schematic representation of DPPC.

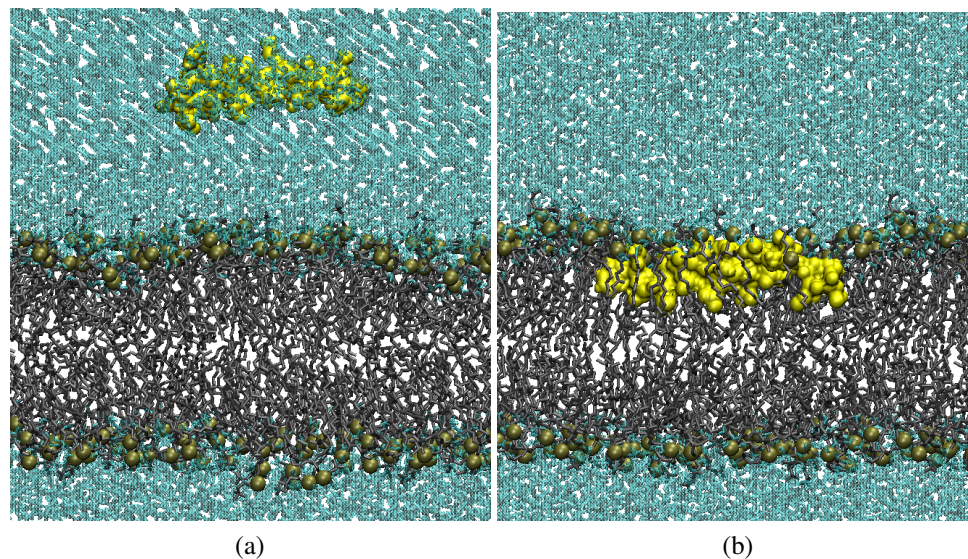


Figure 4.2: First configuration (a), and last configuration at 780 ns (b) of the longest simulation. The peptide is shown in *yellow*, water molecules in light *blue*, DPPC lipids in *gray* and phosphorus atoms are shown as balls in *tan*.

DPPC was modeled with the Berger *et al.* force field parameters [15]. Water was modeled using the simple point charge model (SPC) [11]. For transportan, GROMOS87 force field [153] with corrections as detailed in [152, 92] was used. For Cl^- , the same force field as for transportan was used, the validity of which has been evaluated in Ref. [114].

Temperature was kept constant at 323 K using the Berendsen weak coupling algorithm [12]. The heat bath coupling of lipids and the peptide was separated from that of water and counterions and both of them had a coupling constant of 0.1 ps. Pressure was set to 1 bar using the Berendsen algorithm [12] with the time constant of 0.1 ps. It was applied semi-isotropically, i.e., the extension of the simulation box in the bilayer normal direction (z) and its cross-section area in the bilayer plane (xy) could vary independently. Lennard-Jones and electrostatic interactions were cut off at 1.0 nm. The particle mesh Ewald (PME) method [149] was used to evaluate long-range interactions as it has been shown that truncated electrostatic methods lead to physical artifacts [117, 116, 115]. All bonds and angles were constrained to their equilibrium values using the Linear Constraint Solver (LINCS) algorithm [53]. The time step for all simulations was set to 2 fs.

To obtain the potential of mean force (PMF), the umbrella sampling method [91] was used in which a biasing potential acts on the center of mass of the peptide to restrain it at different distances from the center of mass of the bilayer. The free energy of the unbiased system was calculated from the probability distribution of the peptide when restrained at different positions

in biased simulations. We used a series of harmonic potentials with force constants of 3000 $\text{kJ mol}^{-1} \text{nm}^{-2}$ for the umbrella potential. A total of 41 simulation windows were used and in each window the center of mass of the peptide was restrained to a particular depth within the membrane, separated by 0.1 nm intervals. The starting configuration for each window was chosen as follows: For the first 10 windows, corresponding to distances from 0.0 nm to 1.0 nm between the peptide and the middle of the bilayer, the peptide was pulled from the center of the bilayer to assure that it did not stick to any of the charged head-groups which were about 2 nm away from the center of bilayer. We took a horizontal configuration as the starting point. For other windows, since we already had a 780 ns unbiased simulation in which peptide penetrated up to 1.5 nm away from the center of the bilayer, we could extract plausible configurations of the peptide at different distances from the bilayer, and picked those configurations as the starting point for each window. Finally, the PMF was constructed from the biased distributions of the center of mass of the peptide at each window using the weighted histogram analysis method (WHAM) [73] with a relative tolerance of 10^{-11} . Each window was equilibrated for 40 ns and production simulations were run for 60 ns. To see if the free energy profile changes with extension of simulations, the simulation of some windows, corresponding to the interior of the bilayer, were extended for additional 50 ns, but no change in the free energy profile was observed.

4.3 Analysis

The residues are named according to their three-letter abbreviation names (see table 1.2) followed by a number indicating their order in the peptide. Figure 4.3 shows the helical wheel diagram of the peptide (see section 2.3.4 for more information about the helical wheel representation).

For the calculation of the order parameters, the numbering shown in Fig. 4.1 was used.

We are interested in local properties of the membrane; that is, where lipids are close to the peptide. To measure the proximity of lipids and the peptide, we define a circle on the xy -plane with radius d to the center of all the components of the peptide. Lipids are defined as interacting with transportan if any component of them lies within any of the circles. The value of d was selected based on the lateral radial distribution function (RDF) of the peptide with respect to DPPC atoms. The first minimum of the RDF occurs at ≈ 0.5 nm (data not shown). Therefore, d was selected to be 0.6 nm. The local properties, such as the local thickness, the local P-N angle, and the local order parameter, were computed by means of our own codes. All other analyses, such as hydrogen bonds, the area per lipid, and the protein structure, were performed using the GROMACS package [54]. To elucidate the effect of the presence of transportan on

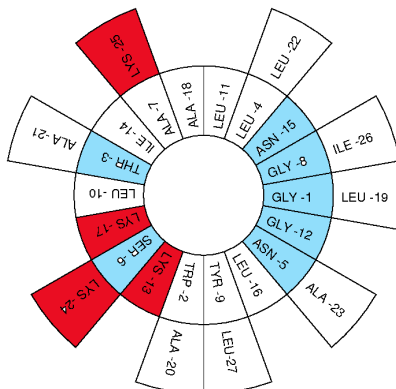


Figure 4.3: The helical wheel diagram of transportan. Each residue is categorized according to the following color scheme: *white* (hydrophobic in neutral pH), *cyan* (hydrophilic and neutral in neutral pH), *red* (hydrophilic and charged in neutral pH).

lipid bilayer, we compared our results with those of a previous 100-ns simulation of a bilayer consisting of pure POPC molecules [74].

4.4 Results

4.4.1 Unbiased Simulations

As outlined above, we performed 10 unbiased simulations. In all simulations except simulation #3 and #9, the initial configuration of the peptide was parallel to the membrane surface, but the peptide was slightly rotated around its helical axis (see Fig. 4.2a). In the other two simulations, the peptide had an angle less than 90° with the membrane surface. Table 4.1 shows a summary. Here, insertion depth is defined to be the distance between the center of mass of phosphorus atoms and the peptide. The longest simulation (780 ns) was used for all major analyses. Figure 4.2b shows the orientation of the peptide after 780 ns.

Distance between several system components and the middle of the bilayer

Having mass and position of all elements of the system, center of mass (COM) of each component (residues or lipid subgroups) can be calculated. Then, we can compute the vertical distance between COM of these components and COM of the whole bilayer as a function of time. This way, we can see how the peptide approaches the membrane surface. Figure 4.4a shows the time development of the distance of several system components from the bilayer's center. This figure corresponds to the simulation #2 (Table 4.1) and shows that in this particular

Table 4.1: Details of different simulations. Insertion depth is defined to be the distance between the center of mass of the peptide and phosphorus atoms. N-terminus corresponds to GLY1 residue and C-terminus corresponds to LEU27 residue. Insertion depth is defined to be the average distance between the center of mass of the peptide and phosphorus atoms within the indicated time interval.

Number	Simulation length	First Approaching Terminal	Insertion Depth
1	86 ns	C-terminus	0 ± 0.1 nm (70 ns-86 ns)
2	780 ns	N-terminus	0.7 ± 0.1 nm (400 ns-500 ns)
3	320 ns	C-terminus	0.5 ± 0.1 nm (220 ns-320 ns)
4	90 ns	N-terminus	0.6 ± 0.1 nm (80 ns-90 ns)
5	90 ns	N-terminus	0.4 ± 0.1 nm (80 ns-90 ns)
6	90 ns	C-terminus	0.4 ± 0.1 nm (80 ns-90 ns)
7	90 ns	N-terminus	0.1 ± 0.1 nm (80 ns-90 ns)
8	90 ns	C-terminus	0.5 ± 0.1 nm (80 ns-90 ns)
9	90 ns	C-terminus	0.2 ± 0.1 nm (80 ns-90 ns)
10	90 ns	N-terminus	0.7 ± 0.1 nm (80 ns-90 ns)

simulation, the peptide approaches the bilayer with the N-terminus (GLY1). It takes approximately 300 ns for the other end of the peptide, i.e., LEU27 residue, to completely associate with the bilayer's surface, so that the whole peptide is internalized.

Similar analysis can be done for the second longest simulation (#3) the plot of which is shown in Fig. 4.4b. Here, the peptide approaches with the C-terminus (LEU27), but despite of that, the N-terminus has a higher tendency to go into the bilayer core. In this case, it takes \approx 200 ns for the peptide to completely associate with the bilayer.

Figure 4.5 shows the same trend for all remaining simulations. No matter which end touches the bilayer surface first, the N-terminus of the peptide easily penetrates into the bilayer while the other end associates with the bilayer after a longer time. The exception is the simulation #1 in which the whole peptide associated with the bilayer weakly and no penetration was observed (see also table 4.1).

Minimal distance between lysine residues and bilayer's phosphate groups

To elucidate the importance of lysine residues on the translocation mechanism, and to investigate the possibility of salt bridges between these residues and the bilayer, we plotted the minimal distance between the NH_3^+ of the four positively charged transportan residues (lysine) and phosphate group for simulation #2. As presented in Fig. 4.6, after 100 ns all the lysine residues are associated with the phosphate group and maintain a minimum distance of less than 0.2 nm throughout the simulation. The same behavior was observed in another simulation on a transportan analogue, transportan-10 [37].

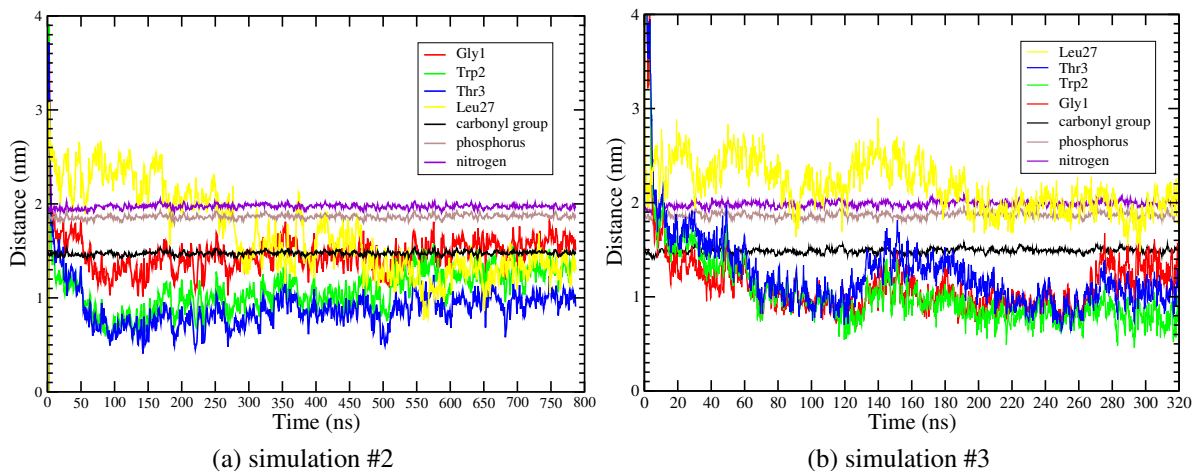


Figure 4.4: Distance of several components of the system from the middle of the bilayer. The COM of the whole bilayer is set to zero. Color coding: *yellow* (LEU27), *blue* (THR3), *green* (TRP2), *red* (GLY1), *black* (lipids' carbonyl group), *brown* (lipids' phosphorus atoms), *violet* (lipids' nitrogen atoms).

Hydrogen bonding between lysine residues and bilayer's phosphate groups

A hydrogen bond is an attractive interaction of a hydrogen, covalently attached to an electronegative atom, and another electronegative atom belonging to other molecules or chemical groups. To determine if a hydrogen bond exists, the following criterion is used [104, 50, 90] (see Fig. 4.7):

$$r_{\text{HB}} < 0.35 \text{ nm}$$

$$\alpha < 30^\circ$$

The value of 0.35 nm corresponds to the first minimum of the RDF of water.

In our system, the -NH or -OH groups of peptide's backbone can form hydrogen bonds with phosphate, glycerol or carbonyl groups of lipids.

In the previous section we showed that there is a charge-pair interaction between the lysine residues and the bilayer's phosphate groups. To see how tightly these side chains are bound to the membrane surface, we also computed the number of hydrogen bonds between lysine residues and bilayer's phosphate groups. As shown in Fig. 4.8, these side chains always maintain at least one hydrogen bond with the phosphate groups. Further analyses show that the main contribution to the hydrogen bonds come from the NH_3^+ of the lysine's side chain than from the NH group of these residues (data not shown). More information about the sudden change in the number of hydrogen bonds of LYS25 and the membrane surface is provided in the Discussion section.

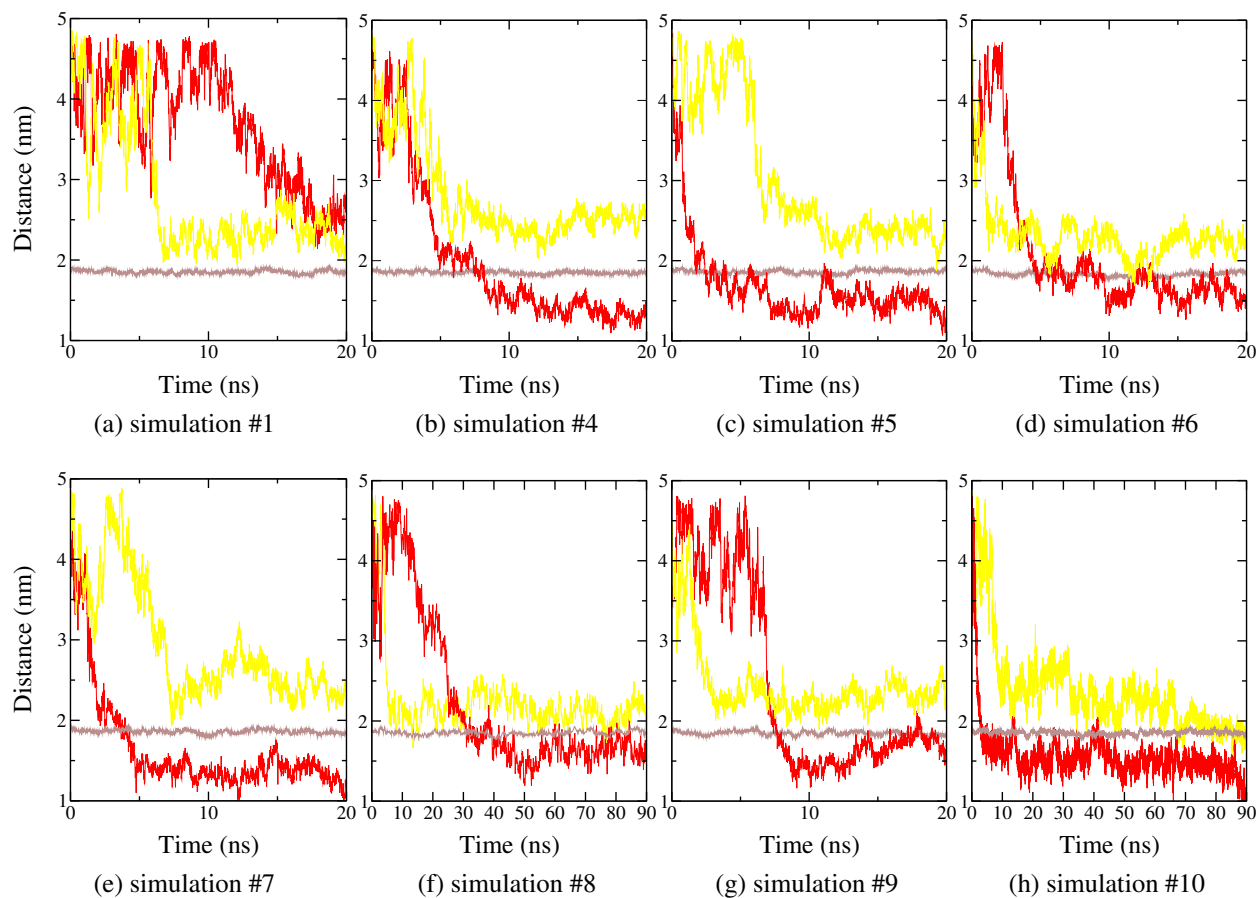


Figure 4.5: Distance of GLY1 (*red*), LEU27 (*yellow*) and phosphorus atoms (*brown*) from the middle of the bilayer. These graphs demonstrate which end of the peptide associates with the membrane first in each simulation. For more information regarding the insertion depth, see Table 4.1.

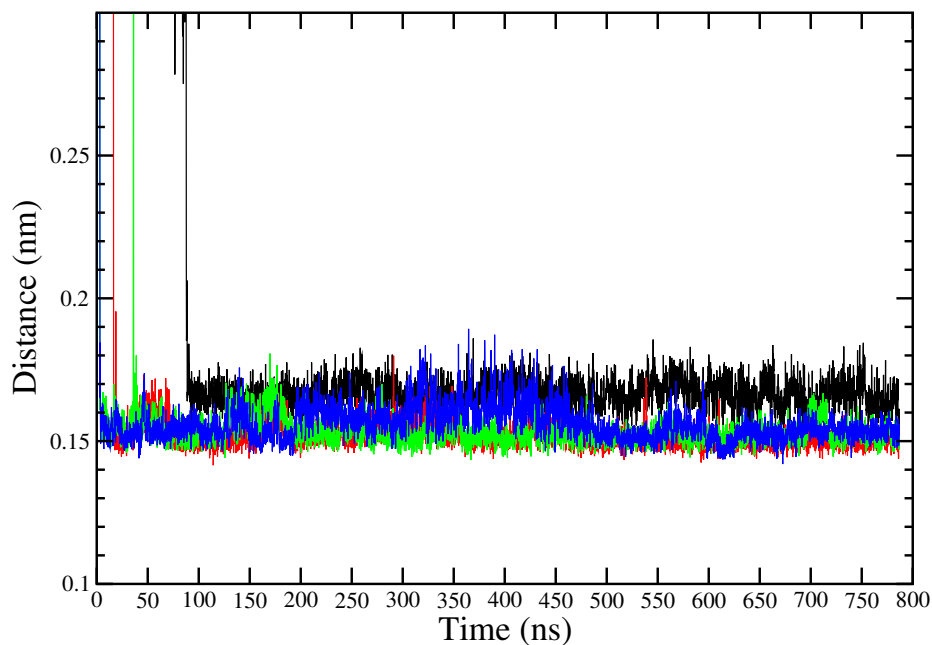


Figure 4.6: Minimal distance between NH_3^+ of lysine side chains and phosphate groups. Color coding: *black* (LYS13), *red* (LYS17), *green* (LYS24) and *blue* (LYS25).

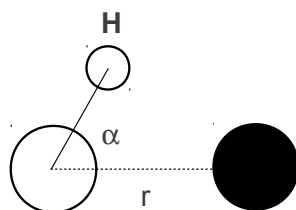


Figure 4.7: Hydrogen bond criterion. The donor/acceptor molecules are shown in white/black spheres. The values of α and r used in this simulation are 30° and 0.35 nm respectively.

Area per Lipid and Thickness

The area per lipid is defined as the surface area of the bilayer in the xy -plane (the membrane normal is the z direction) divided by the number of molecules per leaflet and is experimentally accessible by NMR. In a single component PC bilayer the average area per PC is calculated simply by dividing the total cross-sectional area of the bilayer by the number of PCs in one leaflet. But when the membrane consists of more than one molecule type, care must be taken in how to define the area per lipid as discussed in some works[125, 56]. In our system, although the bilayer consists of a single molecule type (DPPC), but at most of the times it is associated with transportan. Therefore, two molecule types are present on the xy -plane. One approach is to compute the average area per DPPC by extracting the area of transportan from the total area.

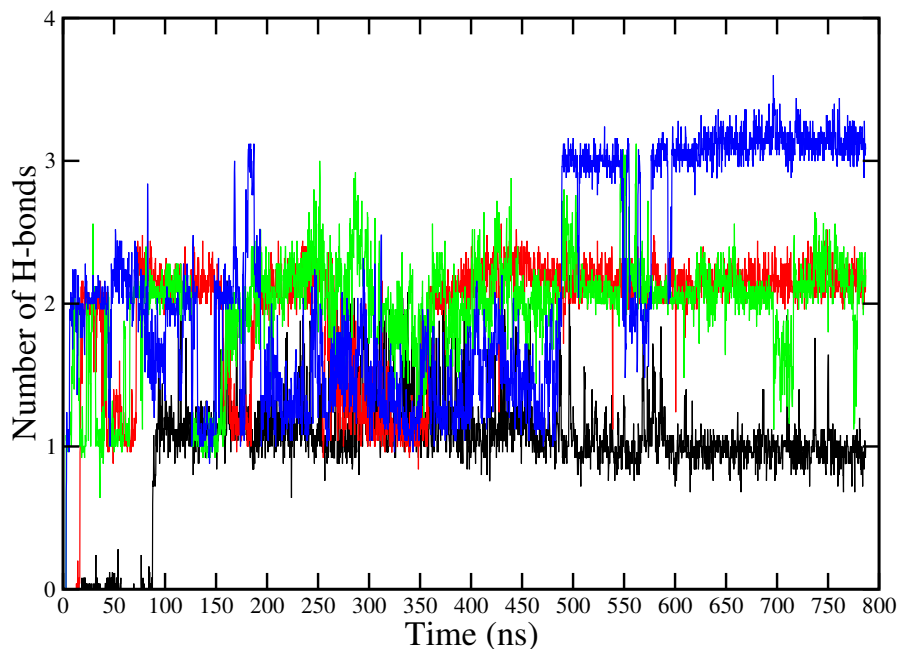


Figure 4.8: Hydrogen bonds between LYS13, LYS17, LYS24 and LYS25 with lipid phosphate groups. Color coding: *black* (LYS13), *red* (LYS14), *green* (LYS24) and *blue* (LYS25).

But using a value for the area of transportan, only an approximate value for the area of DPPC can be obtained. This is because transportan is mostly associated with the head-group region of the bilayer and hence the distribution of the free area per DPPC molecules must depend on the distance from bilayer center. In this work we divide the total area by the number of DPPC molecules to compute the area per lipid. Figure 4.9a shows the time development of the area per DPPC for DPPC-transportan system. A similar analysis has been performed on the pure DPPC system. We found that the average values of area per DPPC are $0.66 \pm 0.01 \text{ nm}^2$ and $0.65 \pm 0.01 \text{ nm}^2$ for DPPC-transportan and pure DPPC, respectively.

We defined the bilayer thickness as being the distance between the COM of nitrogen atoms of one leaflet from the other one's. Figure 4.9b shows the bilayer's thickness as a function of time. Comparison between the numerical values of the average thickness as well as the average area per lipid for DPPC-transportan and pure DPPC systems will be provided in discussion section.

Figure 4.10 shows the local and non-local positions of nitrogen atoms of both bilayer leaflets with respect to the bilayer's center of mass. In average, 19 ± 6 lipids were in neighborhood of transportan when the peptide was associated with the upper monolayer.

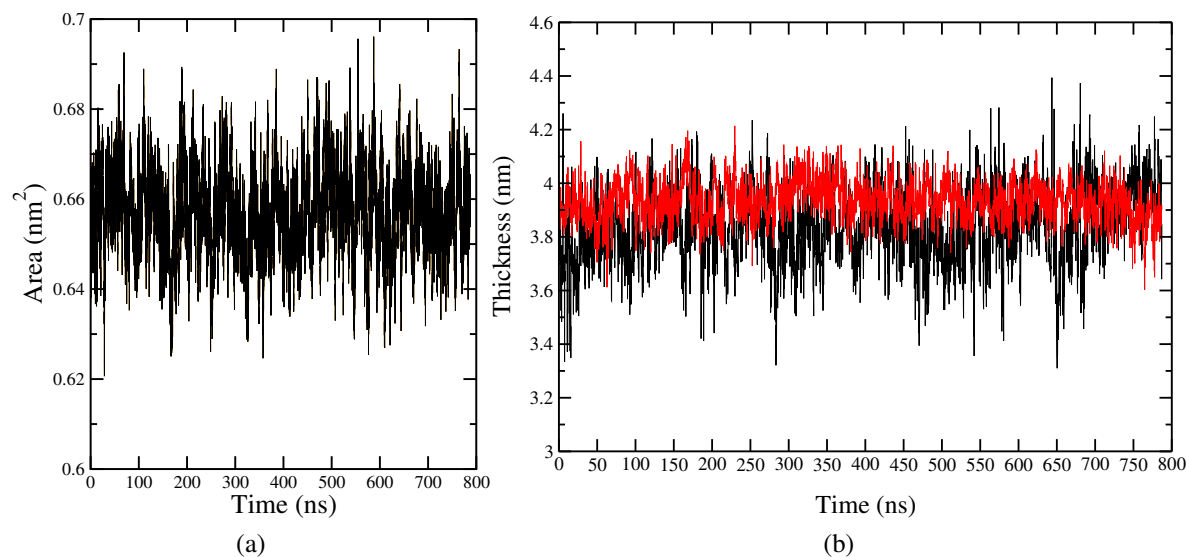


Figure 4.9: Area per lipid (a), and the thickness of neighboring (*black*) and distant (*red*) lipids (b). As defined in section 4.3, the neighboring lipids are no more than 0.6 nm away from the peptide.

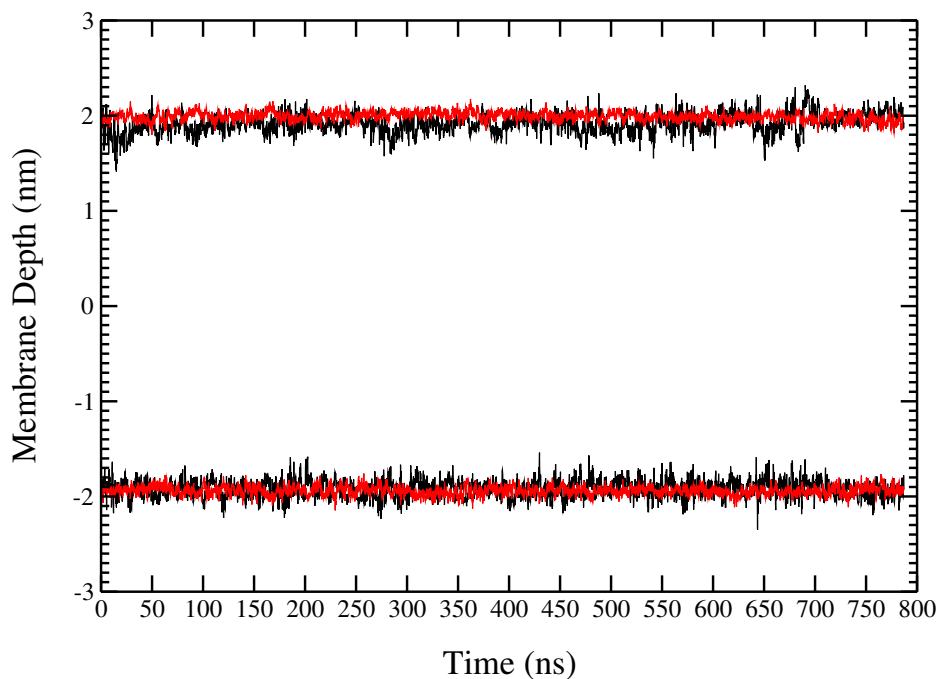


Figure 4.10: z -coordinates of nitrogen atoms of neighboring (*black*) and distant (*red*) lipids. The peptide is associated with the upper monolayer. As defined in section 4.3, the neighboring lipids are no more than 0.6 nm away from the peptide.

Order parameter

Robinson *et al.* [124] were the first ones to show the existence of the ordering effect in MD simulations. To quantitatively evaluate the order of lipid chains, one can use the molecular

order parameter, S_{mol} :

$$S_{\text{mol}} = 1/2 \langle 3 \cos^2(\theta_n) - 1 \rangle \quad (4.1)$$

where, as illustrated in Fig. 4.11, θ_n is the instantaneous angle between the bilayer normal and the n^{th} segmental vector, i.e., (C_{n-1}, C_{n+1}) the vector linking $n - 1$ and $n + 1$ carbon atoms in the hydrocarbon chain. The angular brackets, $\langle \dots \rangle$, denote both the ensemble and the time averages. S_{mol} is closely related to the deuterium order parameter (S_{cd}) and is measured in NMR spectroscopy: $S_{\text{mol}} = -2S_{\text{cd}}$ [31, 35] which holds for saturated chains. For regions close to a double bond, S_{cd} is the more appropriate order parameter. Fig. 4.12 shows the order parameter of the *sn-1* and *sn-12* acyl chains of neighboring and distant lipids as well as those of pure DPPC.

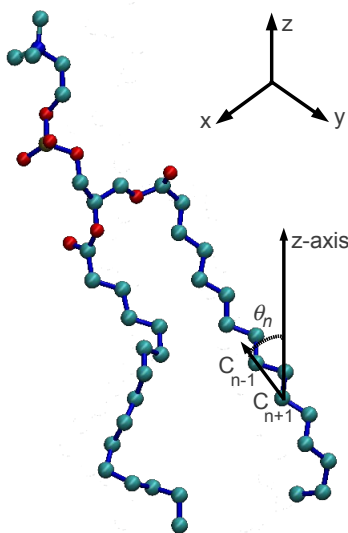


Figure 4.11: Representation of θ_n as used in computing S_{mol} . z -axis is defined to be in the direction of bilayer normal.

P-N angle

Having the position of phosphorus and nitrogen atoms (Fig. 4.1), one can calculate the angle between the bilayer normal and the vector linking phosphorus and nitrogen in each lipid at each time. Mathematically, this angle can be obtained using the following formula:

$$\cos(\theta) = \frac{z_N - z_P}{|\vec{R}_N - \vec{R}_P|}$$

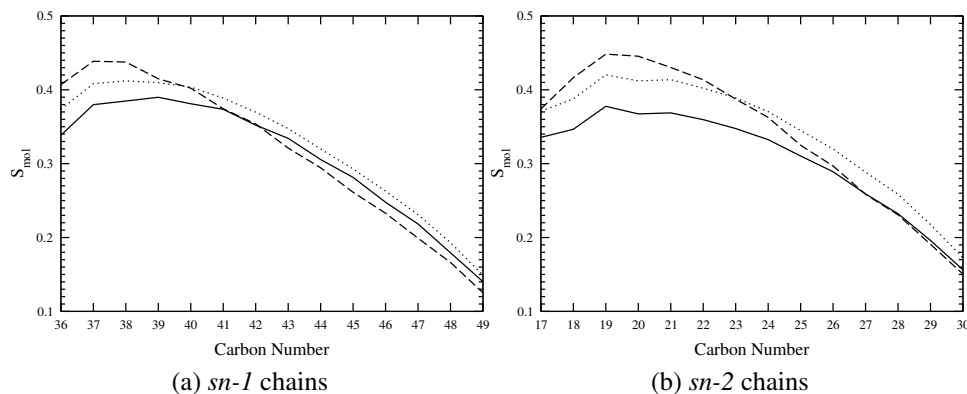


Figure 4.12: Order parameter S_{mol} of the *sn-1* (a) and *sn-2* (b) acyl chains of pure DPPC (*solid line*), peptide's neighboring (*dashed line*) and distant lipids (*dotted line*) obtained from the time interval of 600 ns to 780 ns of simulation #2.

where z_P and z_N are the z -coordinates of phosphorus and nitrogen respectively, and \vec{R}_P and \vec{R}_N are the position vectors of the respective atoms. It is assumed that the bilayer's normal is along the z -axis. The angle distribution of P-N vector around the z -axis is shown in Fig. 4.13. The distribution corresponds to 600 ns to 780 ns. The mean value of the P-N angle for pure DPPC, neighboring and distant lipids are $79.3 \pm 0.2^\circ$, $73.0 \pm 0.2^\circ$, and $77.9 \pm 0.2^\circ$, respectively.

Peptide structure

Upon peptide association, its helicity changes and it samples a variety of structures. Figure 4.14 shows the evolution of peptide's structure for the simulations of #2, #3, #8 and #10 in which the peptide approached the bilayer from GLY1, LEU27, LEU27 and GLY1 residue, respectively.

4.4.2 Insertion of transportan into bilayer

As outlined in the introduction, CPP translocation has proved experimentally to be a very slow process and we don't expect to see translocation under normal experimental conditions within hundreds of nanoseconds. That's why none of the simulations with the peptides initially placed in the water phase led to a configuration with transportan completely penetrated into the bilayer core. Looking for a possible pathway of peptide translocation, we performed four more simulations up to 172 ns in which the peptide was initially placed directly into the bilayer core horizontally (parallel to the bilayer surface). Figure 4.15 shows the first and the last configurations of these simulations. Fig. 4.15a shows the initial configuration of the first simulation. In other three simulations, the initial configuration of the peptide was obtained by rotating it around its axis by a certain angle.

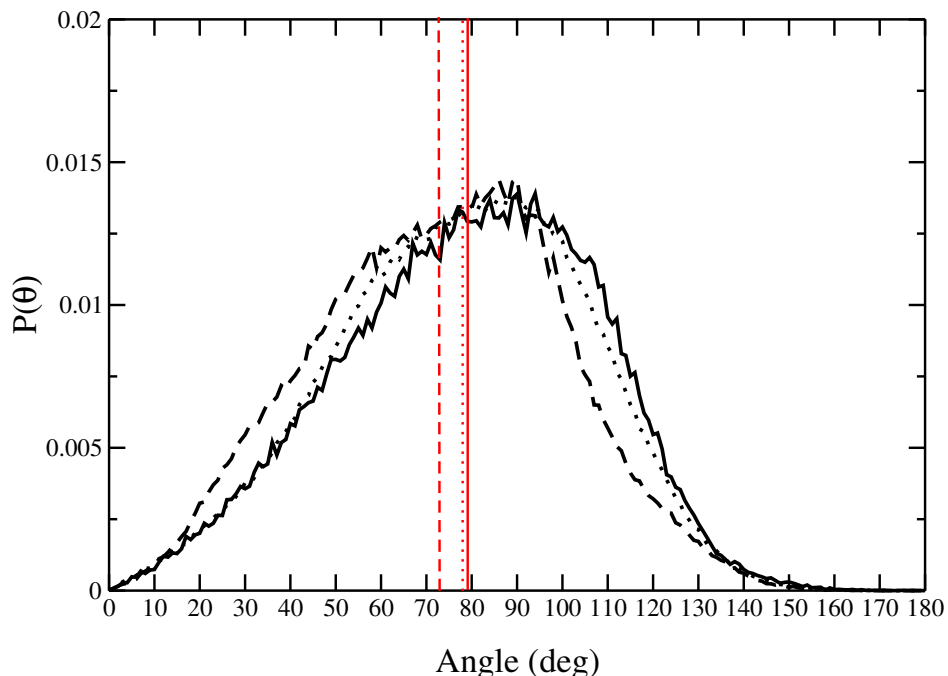


Figure 4.13: The distribution of the angle between P-N vector and the bilayer normal (z -axis) for pure DPPC (*solid line*), peptide's neighboring (*dashed line*) and distant (*dotted line*) lipids of the peptide-associated monolayer. The mean values are shown in *red* with the same line style.

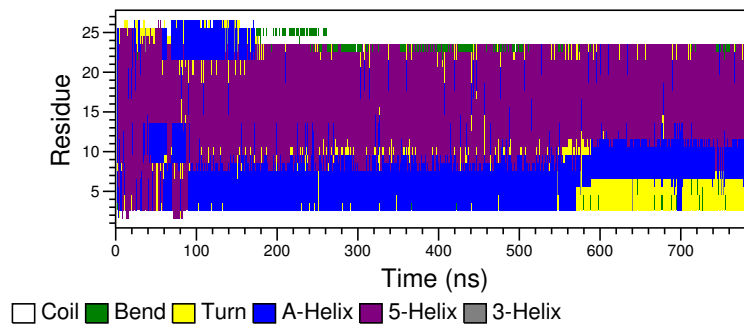
Table 4.2: Area per lipid and thickness of the bilayer when transportan is directly inserted in the bilayer's core.

Number	Simulation length	Area per lipid (nm^2)	Thickness (nm)
1	172 ns	$0.64 \pm 0.01 \text{ nm}^2$	$3.97 \pm 0.05 \text{ nm}$
2	172 ns	$0.65 \pm 0.01 \text{ nm}^2$	$3.93 \pm 0.04 \text{ nm}$
3	164 ns	$0.66 \pm 0.01 \text{ nm}^2$	$3.93 \pm 0.05 \text{ nm}$
4	134 ns	$0.65 \pm 0.01 \text{ nm}^2$	$3.93 \pm 0.05 \text{ nm}$

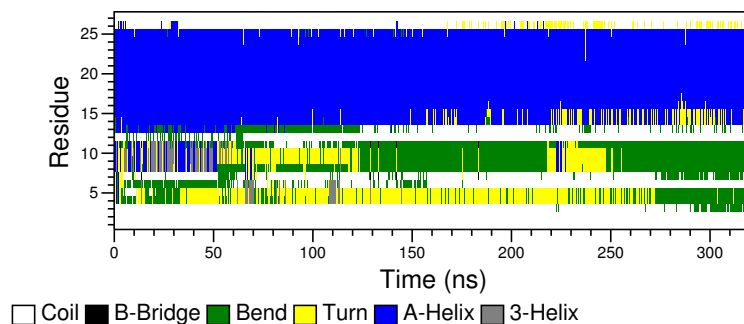
To see if the presence of transportan inside the bilayer core is destabilizing the membrane or not, we computed the order parameter for the $sn - 1$ acyl chain of DPPC lipids. As shown in Fig. 4.16, it is interesting to see that in all the simulations the order parameter for most of the carbon atoms is bigger than the order parameter in pure DPPC. This result will be discussed in more details in the Discussion section.

4.4.3 Biased Simulations

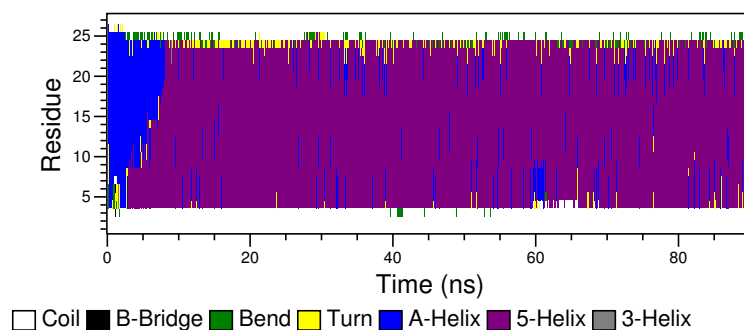
As mentioned in section 4.2, we computed the Gibbs free energy of the system when the peptide is at different distances from the middle of the bilayer along the z -axis (Fig. 4.17).



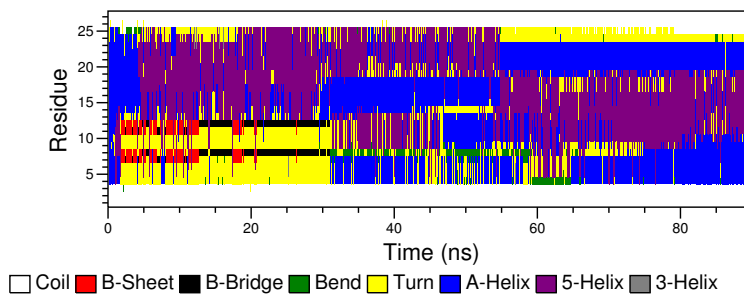
(a) simulation #2



(b) simulation #3



(c) simulation #8



(d) simulation #10

Figure 4.14: Evolution of secondary structures of transportan as defined by DSSP [68]. The residues start from GLY1 (N-terminus) to LEU27 (C-terminus). Different colors correspond to different secondary structures as shown in legends.

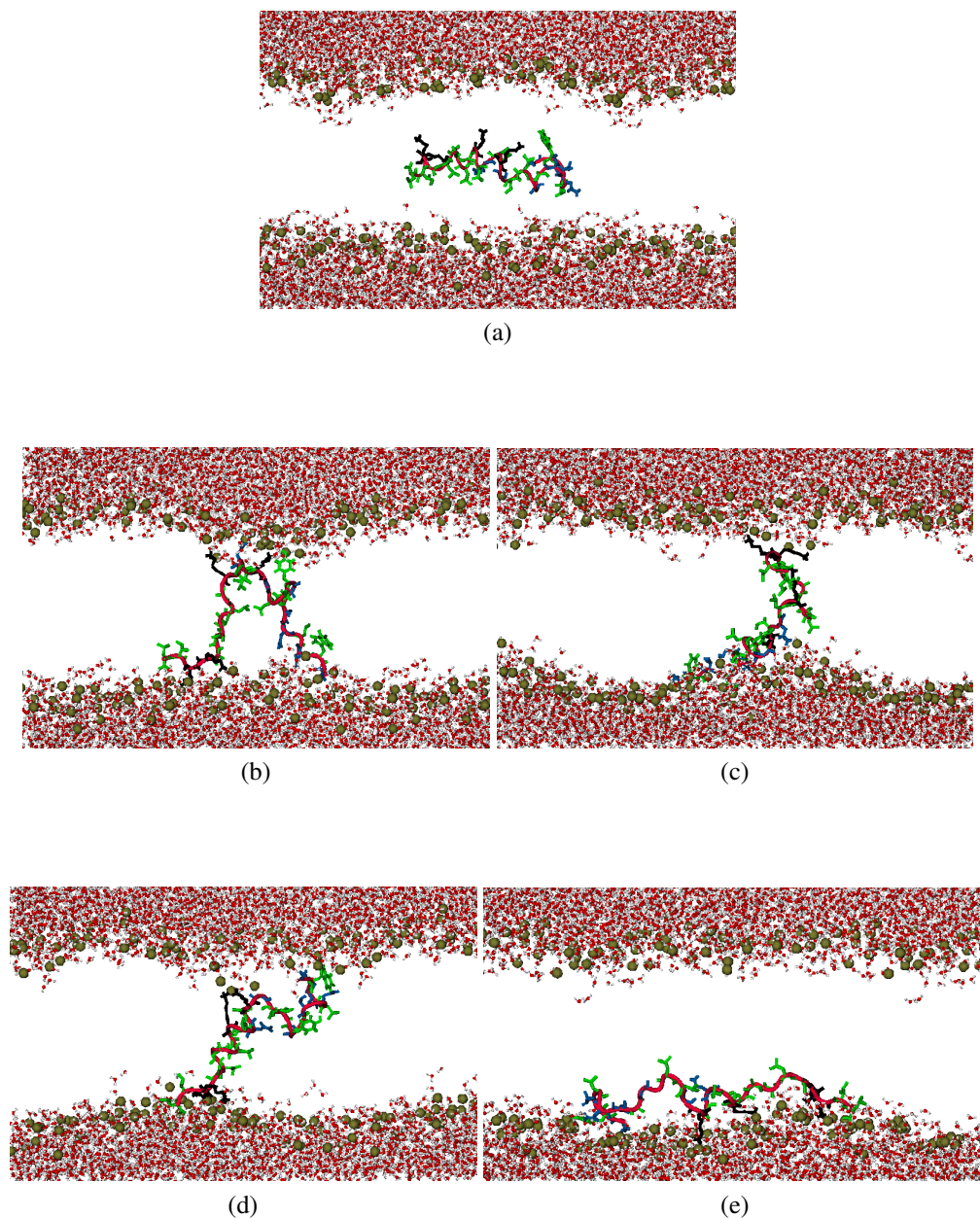


Figure 4.15: Transportan inserted into the bilayer's center: initial configuration of the system with transportan inside the bilayer (a), last configuration of the simulation #1 after 172 ns (b), last configuration of the simulation #2 after 172 ns (c), last configuration of the simulation #3 after 164 ns (d), and last configuration of the simulation #4 after 134 ns (e). Phosphorus atoms are shown as balls in *tan*, lysine residues in *black*, hydrophilic residues in *blue* and hydrophobic residues in *green*. The overall conformation of the peptide is shown as a tube in *red*.

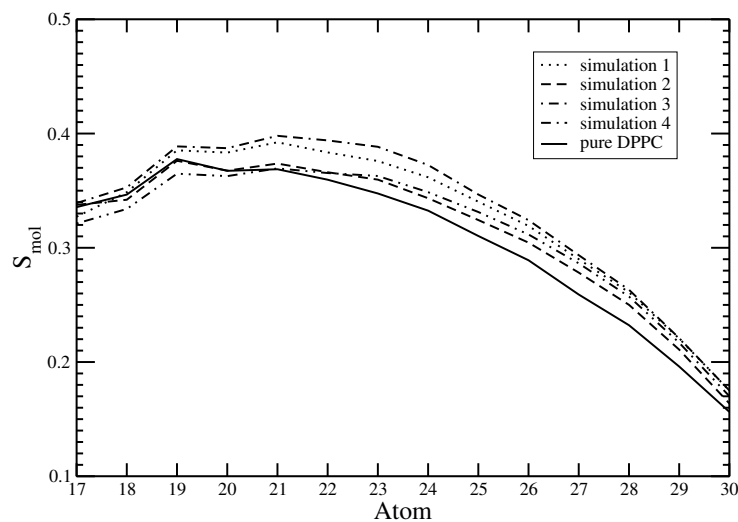


Figure 4.16: Order parameter, $S_{mboxmol}$ of *sn*-1 acyl chains of DPPC in the inserted-transporan simulations obtained from the last 30 ns of the simulations.

Figure 4.17 shows that there are three important energy minima inside the bilayer around 0.4 nm, 0.8 nm and 1.2 nm from the bilayer's center. Figure 4.18 shows four snapshots of biased simulations when the peptide is restrained at $z = 0$ nm, $z = 0.4$ nm, $z = 0.8$ nm and $z = 1.2$ nm, where z is the distance of the peptide from the bilayer's center.

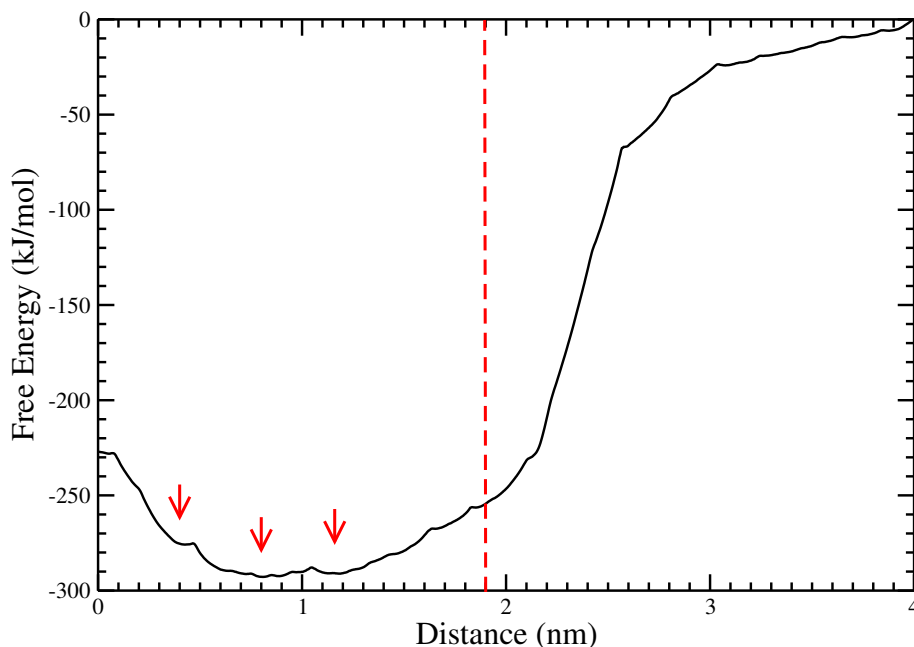


Figure 4.17: Free energy as a function of the peptide’s distance from the bilayer’s center. The bilayer’s center is set to zero. The water-membrane interface, defined as the average position of phosphorus atoms in the unbiased simulations, is shown as a *dashed line in red*. There are three important energy minima at $z = 0.4$ nm, $z = 0.8$ nm and $z = 1.2$ nm as indicated by the arrows.

4.5 Discussion

Fig. 4.4 and 4.5 shows a rapid transportan-membrane association: in all simulations, it took less than 10 ns for the peptide to reach the membrane’s nitrogen atoms. In half of the simulations, transportan approached the bilayer surface with the N-terminus (GLY1) and in the other half, it did with the C-terminus (LEU27). As it can be seen from Fig. 4.4, of all the residues, threonine (Thr) and tryptophan (Trp) are those which penetrate the most into the hydrophobic core of the bilayer. From Fig. 4.4a it is also noticeable that after 500 ns, these two residues as well as their neighbors start returning to the bilayer head-groups.

Fig. 4.6 shows that lysine residues are always in the vicinity of the bilayer’s phosphate groups. This is partly because of the electrostatic interaction between these side chains, which are positively charged, and the bilayer’s phosphate group, which are negatively charged. The other reason, as indicated by Fig. 4.8 is the hydrogen bonding between these residues and the membrane head-groups. These side chains always maintain at least one hydrogen bond with the phosphate groups. After 500 ns, the peptide re-orient itself in such a way that LYS25 triples its number of H-bonds (see also Fig. 4.14a). So the combination of Coulomb interaction and hydrogen bonds between lysine residues and bilayer’s phosphate groups are important factors

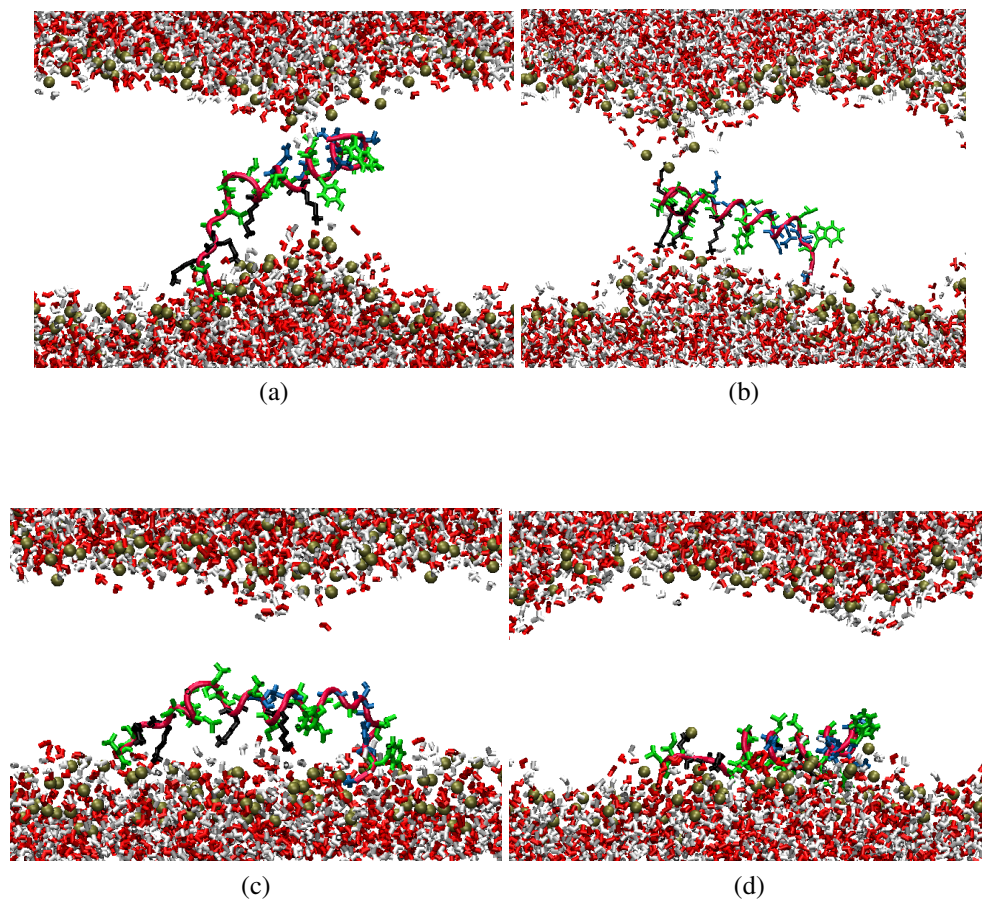


Figure 4.18: Snapshots of the biased simulations when the peptide is restrained in the middle of the bilayer (a), 0.4 nm from the bilayer's center (b), 0.8 nm from the bilayer's center (c), and 1.2 nm from the bilayer's center (d). Phosphorus atoms are shown as balls in *tan*, lysine residues in *black*, hydrophilic residues in *blue* and hydrophobic residues in *green*. The overall conformation of the peptide is shown as a tube in *red*.

in the transportan orientation in the bilayer. This is in line with the result obtained by Lee *et al.* [37] who studied the interaction of POPC membranes and Tp10, the shorter version of transportan consisting of 16 amino acids. They showed that the lysine-phosphate salt bridge is very important in the peptide orientation as well as the peptide-membrane interaction. In another study, the importance of salt bridge has been widely discussed and was predicted to modulate peptide's translocation[3].

The average area per DPPC for DPPC-transportan was found to be $0.66 \pm 0.01 \text{ nm}^2$. As a control, the average area per DPPC for pure DPPC was also computed and it was $0.65 \pm 0.01 \text{ nm}^2$ which is in agreement with other simulations as well as experiments. For example, Faller *et al.* found it to be $0.65 \pm 0.01 \text{ nm}^2$ at 325 K [80], and in an NMR experiment, Nagle *et al.* found it to be $0.629 \pm 0.013 \text{ nm}^2$ [105]. Therefore, simulations are compatible with the interpretation that there is no change in area per lipid within measurement accuracy. Consistently, there was no significant bilayer thinning: as can be seen in Fig. 4.10, the neighboring lipids of the peptide (in the upper monolayer) are slightly dragged toward the bilayer core, but this is very small. The thickness of peptide's neighboring lipids is $3.84 \pm 0.14 \text{ nm}$ and of distant ones is $3.93 \pm 0.08 \text{ nm}$. Therefore, the local thickness is 2.6% less than the non-local thickness. In pure DPPC, the average thickness was found to be $3.86 \pm 0.04 \text{ nm}$. Comparing the local and non-local thickness in transportan-DPPC system with the average thickness of pure DPPC shows that for all these three values, the error bars overlap and no particular decrease or increase can be attributed to the bilayer due to the presence of transportan. The same argument applies to the area per lipid for simulations with transportan directly inserted into the bilayer's hydrophobic core. But interestingly, according to table 4.2, thickening of the bilayer occurs when transportan is inside the membrane. This means transportan adjusts itself inside the bilayer core by increasing the bilayer's thickness.

As demonstrated in Fig. 4.12, the shape of the order parameter curves resembles those obtained from experiments, with a plateau near the carbonyl groups and increasing random structure toward the bilayer core. A significant increase in the ordering of the upper carbons of lipid tails can be observed due to the presence of transportan. This behavior is different from penetratin/POPC systems. Lensink *et al.* [82] showed that the presence of penetratin introduces disorder over the entire length of the lipids. Also, MD simulation of melittin, an antimicrobial peptide, shows that lipid molecules bound to the peptides are less ordered while those further away more ordered[5, 139]. But in the case of transportan, we see a rather different behavior: local carbon atoms close to head-groups (C_{17} to C_{23} in *sn-2* and C_{36} to C_{39} of *sn-1* acyl chains) are more ordered than non-local ones. Electrostatic interactions between the peptide and head-groups of local lipids constrain the carbon atoms that are close enough to head-groups leading to higher order parameter than the average. Because the temperature is constant and the peptide

should have a certain amount of kinetic energy, the rest of the carbon atoms move more to compensate the lack of thermal motion of the whole peptide leading to lower order parameter values than the average. This result may rule out the possibility of pore formation: The more ordered the lipids are, the less likely they can re-orient themselves to make a pore.

According to Fig. 4.13 the P-N vector has a significant tendency of being in the bilayer plane ($\theta = 90^\circ$). However, this tendency is influenced by the presence of transportan: In pure DPPC, the mean value of this angle is $79.3 \pm 0.2^\circ$ consistent with other simulations[118] while for neighboring lipids it decreases to $73.0 \pm 0.2^\circ$. This shows an ordering in the lipids head-groups and is consistent with the ordering behavior of carbons close to the head-groups (cf. Fig. 4.12). The mean value of P-N angle for distant lipids is $77.6 \pm 0.2^\circ$. close to the value for pure DPPC.

Figures 4.14 shows the conformation of transportan as a function of time for four simulations. There is no resemblance in transportan conformation when associated the membrane in the four cases shown. For every simulation, the conformation of transportan is different in different part of the peptide. Whether or not transportan should take a specific conformation prior to translocation remains as a question. In some antimicrobial peptides it has been shown that a stable helical secondary structure of the peptide is not a prerequisite to pore formation[139, 121]. It is interesting to note the change in the structure of peptide in Fig. 4.14a at around $t = 500$ ns. This change correlates with the sudden change in the number of hydrogen bonds between LYS25 and the membrane surface as shown in Fig. 4.8.

At the beginning of the inserted-transportan simulations, the peptide was attracted to the phosphate groups of the bilayer's surface from below via lysine and glycine. Then the rest of the peptide had to rearrange itself by bending at certain residues. As shown in figure 4.15, in three of the four simulations, the peptide bent itself in the bilayer core. In all these three cases, bending occurred at the position of leucine. In simulation 1, the peptide bent at LEU11 and LEU16, in simulation 2, it bent at LEU19 and LEU22, and in simulation 3, it bent at LEU10 and LEU11. Therefore, lysine, glycine and leucine play an important role in the peptide conformation inside the bilayer. It is also noticeable that, as opposed to other three simulations, in simulation 4 the peptide stays close to one of the monolayers and does not bound to the other one (Fig. 4.15e). At this state, the average distance of the peptide from the bilayer's COM was found to be 1.00 nm and this distance corresponds to one of the local minima of the free energy (cf. 4.17). In other words, in this case the peptide has been trapped in a local minimum. Another important thing is the water defect in the inserted-transportan simulations. Of all the four simulations, the second simulation, corresponding to Fig. 4.15c, shows the most noticeable water defect, although it did not lead to any pore formation. This can be seen from two-dimensional density plot of water and phosphate groups as illustrated in Fig. 4.19. The

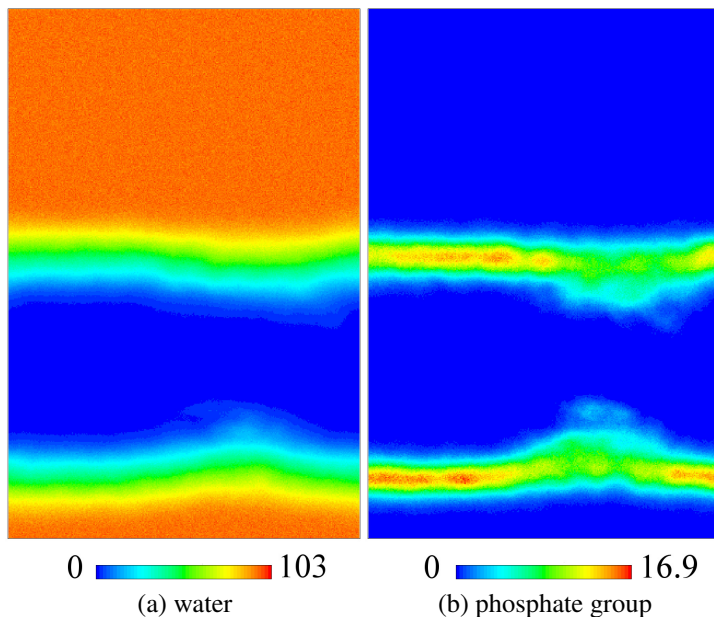


Figure 4.19: Two-dimensional density map of water (a), and phosphate-group (PO_4^{-1}) (b) when the peptide is inserted into the bilayer's center. The figures correspond to simulation #2 of the inserted-transportan simulation (Fig. 4.15c). The unit is nm^{-3} , and the vertical axis is along the bilayer normal.

figures show that water molecules penetrate into the bilayer core to keep the phosphate groups as well as the hydrophilic residues of the peptide solvated.

Figure 4.16 shows the ordering of the lipid tails when transportan is in the bilayer core. For most of the carbon atoms (C_{21} to C_{30}), the order parameter is higher for transportan-DPPC systems compared to pure DPPC. It is interesting to note that in the normal simulations in which the peptide was associated with the bilayer's surface but not completely inside the bilayer core, the order parameter of carbon atoms close to head-groups was higher in transportan-DPPC systems compared to pure DPPC (cf. Fig. 4.12). So based on these observations, we can conclude that wherever transportan is present, it induces order in its neighboring lipids.

The fact that the free energy is higher when the peptide is in the water than when it is inside the bilayer shows the tendency of the peptide for a passive translocation. However, for going from one monolayer to the other one, the peptide should pass some major energy barriers. The first bump is located between 1.0 nm to 1.1 nm from the bilayer's center, the next one is around 0.5 nm from the bilayer's center, and the last one is in the middle of the bilayer. There are three minima corresponding to these bumps: one around 0.4, one around 0.8 nm and the other around 1.2 nm from the bilayer's center. When the peptide is around 1.2 nm from the bilayer's center, it is highly associated with the bilayer's surface (Fig. 4.18d).

When it is around 0.8 nm from the bilayer's center, some of its hydrophobic residues have freed themselves from the bilayer's hydrophilic center (Fig. 4.18c). When it is around 0.4 nm from the bilayer's center, the lysine residue LYS25 frees itself from one monolayer and is attracted to the other one (Fig. 4.18b). This is another indication of how important lysine residues are. Comparable to experiments [159, 162], the presence of lysine is slowing down the penetration process by forming salt bridges with the bilayer. To complete the translocation, all remaining residues should free themselves, and to accomplish this, they should pass a potential barrier located in the middle of the bilayer. This is the main potential barrier in the free energy profile and its height is approximately 66 kJ/mol. According to the Eyring theory [42], this barrier is in principle accessible at 323 K on time scales of thousands of microseconds:

$$\tau = \frac{\hbar}{kT} e^{\Delta G/kT} = 0.007 \text{ s}$$

The height of this barrier is less than the barrier corresponding to the flip-flop of a single DPPC lipid which is ≈ 80 kJ/mol [133]. As another comparison, the potential barrier for penetratin (another CPP) translocation through DPPC lipids is around 75 kJ/mol [163]. Although in their study Mark *et al.* [163] argued that the translocation of penetratin through DPPC membranes is pore-mediated, in our simulations no formation of a pore was observed. However, water defects were observable in our bias simulations. Figure 4.20 shows the water defect when the peptide is restrained at $z=0$ nm and $z=0.4$ nm. It is worth noting that in general water defects are not guaranteed to lead to a successful pore formation. For example, in a study on the phospholipid flip-flop, Tieleman *et al.* [133] showed that when moving a single lipid towards the bilayer core, water defects were noticeable in POPC and DOPC lipids (more than what we observed in our simulations), but in neither of them a pore was formed. On the other hand, in the same study, water defect as well as pore formation were observed in DLPC, DMPC and DPPC lipids.

Overall, in our opinion there are two facts that make the formation of trans-membrane pores by single transportan in DPPC bilayers very unlikely: 1) the increase of the membrane's ordering due to the presence of the transportan, and 2) the fact that in the inserted-transportan simulations, the average thickness increased or remained unchanged, and in the normal simulations no significant thinning of the bilayer was observed. Regarding the penetration mechanism, either a collective behavior is involved or a direct translocation of the peptide is a possible pathway. As proposed by Mark *et al.* [163], a collective behavior can be such that multiple peptides induce curvature in the membrane, ultimately leading to the formation of small vesicles within the cell that encapsulate the peptides (micropinocytosis). To simulate this behavior, large enough bilayer is needed so that elastic effects of the bilayer can be seen more readily.

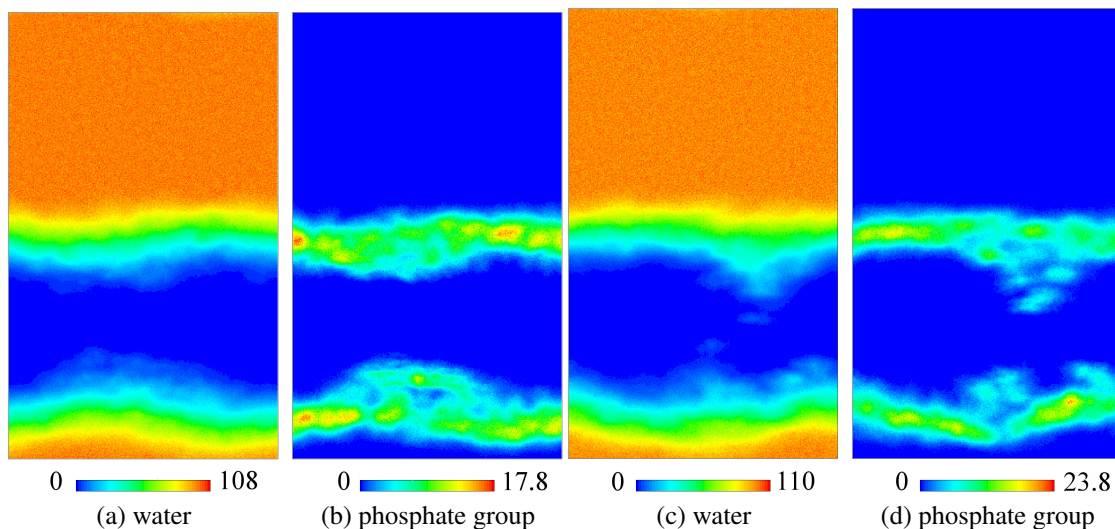


Figure 4.20: Two-dimensional density map of water ((a) and (c)) and phosphate-group (PO_4^{-1}) ((b) and (d)). In (a) and (b) the peptide is restrained at $z=0$ nm, and in (c) and (d) the peptide is restrained at $z=0.4$ nm. The unit is nm^{-3} , and the vertical axis is along the bilayer normal.

But there is also another possible mechanism in which a single peptide passes the potential barrier through a stochastic mechanism and reach the other side of the bilayer by using lysine residues. These residues may play an important role when transportan is inside the bilayer since, as seen in Fig. 4.18, they point towards the water-membrane interface and can drag the peptide from one interface to the other.

4.6 Conclusion

In conclusion, transportan seems to be adaptable to the membrane interior and is unlikely to be able to form a pore in the membrane when a single peptide is involved. This can be inferred from the ordering effect of transportan on the lipids, and also from the thickening of the bilayer when transportan is in its interior. Moreover, we highlighted the importance of lysine residues in the peptide orientation when they form salt-bridges with the membrane head-groups, whether they are associated with the membrane surface or they are in the bilayer core. We still cannot rule out the possibility of the pore formation when multiple peptides interact with the membrane and the next step will be assessing this possibility by considering multiple peptides interacting with a larger bilayer.

Chapter 5

DPPC Bilayer Interacting With Penetratin

5.1 Introduction

Penetratin, also called the pAntp peptide, is one of the most thoroughly studied cell-penetrating-peptides which was discovered in 1994 [32] and can translocate across cellular membranes without membrane damage [147]. It is a 16-residue long fragment derived from the homeodomain of the *Drosophila* (amino acids 43-58, RQIKIWFQNRRMKWKK-amide) transcription factor Antennapedia and has proven to be an excellent tool for cargo delivery [34]. In penetratin, 7 of the 16 residues are positively charged at neutral pH (four lysines and three arginines). The conformation of the penetratin peptide in the parent protein is helical. However, when isolated in solution, the peptide is largely unstructured [32].

To our knowledge, only two simulation studies have been reported. Lensink *et al.* [82] simulated the behavior of penetratin in water and the spontaneous binding of a single penetratin peptide to neutral and charged lipid bilayers. The next simulation of penetratin was reported in 2009 by Mark *et al.* [163]. Their simulations suggest that micropinocytosis may explain how CPPs facilitate the transport of cargo molecules into cells. In this mechanism, multiple peptides induce curvature in the membrane, ultimately leading to the formation of small vesicles within the cell that encases the peptides.

Here, we present the results of simulation of penetratin interacting with dipalmitoyl phosphatidylcholine (DPPC) bilayers in an attempt to better understand the nature of penetratin-membrane interaction.

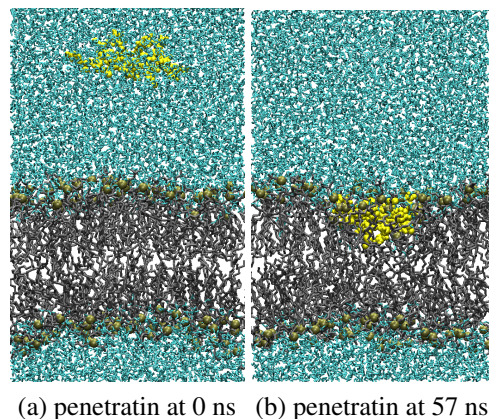


Figure 5.1: First configuration (a), and the configuration at 57 ns (b) of the the simulation #4. The peptide is shown in *yellow*, water molecules in light *blue*, DPPC lipids in *gray* and phosphorus atoms are shown as balls in *tan*.

5.2 Simulation Details

All the simulations were performed in constant temperature and pressure with constant number of particles (NpT ensemble), and were run using GROMACS 4.0 [54]. The simulation systems consisted of 2×64 DPPC lipids hydrated with ≈ 9000 water molecules, and one penetratin peptide. The initial size of the system was $6.4 \times 6.4 \times 10.8 \text{ nm}^3$. In chapter four, the chemical representation of DPPC is shown in Fig. 4.1. Figure 5.1 shows representative snapshots of the system.

We performed eight simulations corresponding to eight different initial configurations. As for the initial configuration of lipids, we used the coordinates of previously studied DPPC lipids in water [117] to which we added penetratin. The initial conformation of penetratin was obtained from the parameters developed by Maler *et al.* (PDB code: 10MQ) [85] in which they utilized Nuclear Magnetic Resonance (NMR) spectroscopy to study the structure of penetratin in negatively charged bicelles. The initial conformation of peptide was more or less helical and the conformation at other times was not constrained and the peptide was able to change it.

The seven positive charges carried by penetratin were compensated by the introduction of the corresponding number of Cl^- counterions. The systems were energetically minimized using steepest descent method before MD simulations. DPPC was modeled with the Berger *et al.* force field parameters [15]. Water was modeled using the simple point charge model (SPC) [11]. For penetratin, gromos87 force field [153] with corrections as detailed in [152, 92] was used. For Cl^- , the same force field as for penetratin was used, the validity of which has been evaluated in Ref. [114].

Temperature was kept constant at 323 K using the Berendsen weak coupling algorithm [12]. Heat bath coupling of lipids and protein was separated from that of solution and both of them had a coupling constant of 0.1 ps. Pressure was set to 1 bar using the Berendsen algorithm [12] with the time constant of 0.1 ps. It was applied semi-isotropically, i.e., the extension of the simulation box in the bilayer normal direction (z) and its cross-section area in the bilayer plane (xy) could vary independently. Lennard-Jones and electrostatic interactions were cut off at 1.0 nm. The particle mesh Ewald (PME) method [149] was used to evaluate long-range interactions as it has been shown that truncated electrostatic method leads to physical artifacts [117, 116, 115]. All bonds and angles were constrained to their equilibrium values using the Linear Constraint Solver (LINCS) algorithm [53]. The time step for all simulations was set to 2 fs.

5.3 Analysis

The residues are named according to their three-letter abbreviation names (see Table 1.2) followed by a number indicating their order in the peptide. Figure 5.2 shows the helical wheel diagram of the peptide (see section 2.3.4 for more information about the helical wheel representation).

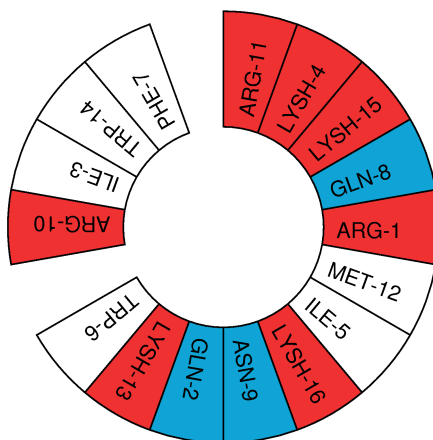


Figure 5.2: The helical wheel diagram of penetratin as obtained from Gromacs. Each residue is categorized according to the following color scheme: *white* (hydrophobic in neutral pH), *cyan* (hydrophilic and neutral in neutral pH), *red* (hydrophilic and charged in neutral pH).

We are interested in local properties of the membrane, that is, where lipids are close enough to the peptide. To measure the proximity of lipids and the peptide, we draw a circle on the xy plane with radius d to the center of all the components of the peptide. Lipids are defined as interacting with penetratin if any component of them lies within any of the circles. The value

of d was selected based on the lateral radial distribution function (RDF) of the peptide with respect to DPPC atoms. The first minimum of the RDF occurs at ≈ 0.5 nm (data not shown). Therefore, d was selected to be 0.6 nm. The local thickness was computed by means of our own codes. All other analyses were performed using GROMACS 4.0.7 [54].

To determine if a hydrogen bond exists, the following criterion is used [104, 50, 90] (see Figure 4.7 in chapter four.):

$$r_{\text{HB}} < 0.35 \text{ nm} ,$$

and

$$\alpha < 30^\circ .$$

The value of 0.35 nm corresponds to the first minimum of the RDF of water.

For order parameter, atom numbering of Fig. 4.1 was used.

To elucidate the effect of the presence of penetratin on lipid bilayer, we compared our results with those of a previous 100-ns simulation [74].

5.3.1 Results

Distance between several system components and the middle of the bilayer

In Figs. 5.3 and 5.4, we have presented the distance between several components of the system from the center of mass of the lipids (the middle of the bilayer). In this way, one can see how residues approach and penetrate into the bilayer. As presented in Table 5.1, the deepest penetration was for simulations #4 (0.3 nm). In simulation #1, which was the longest one, no penetration was observed. In the table, the terminals which first touch the membrane are presented. In five of the simulations, the C-terminus, corresponding to LYSH16 residue, touches the membrane first. The graphs show that not only the hydrophobic residues, but also the hydrophilic ones, including ARG, LYSH, and GLN, are able to penetrate the bilayer.

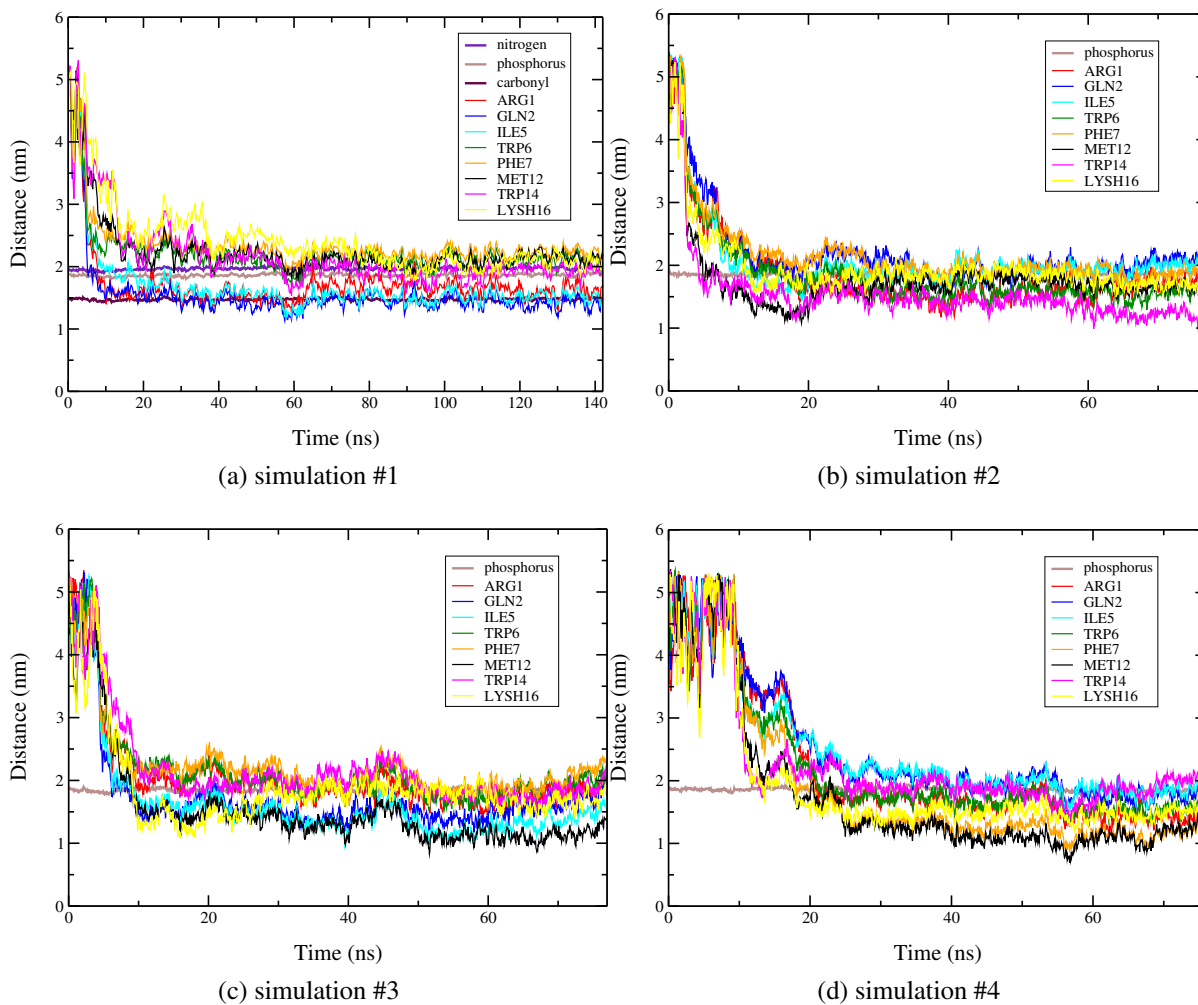


Figure 5.3: Distance of several components of the system from the middle of the bilayer for eight simulations. From these figures one can see which terminal touches the membrane first and which residues go deeper into the bilayer. For more information, see Table 5.1.

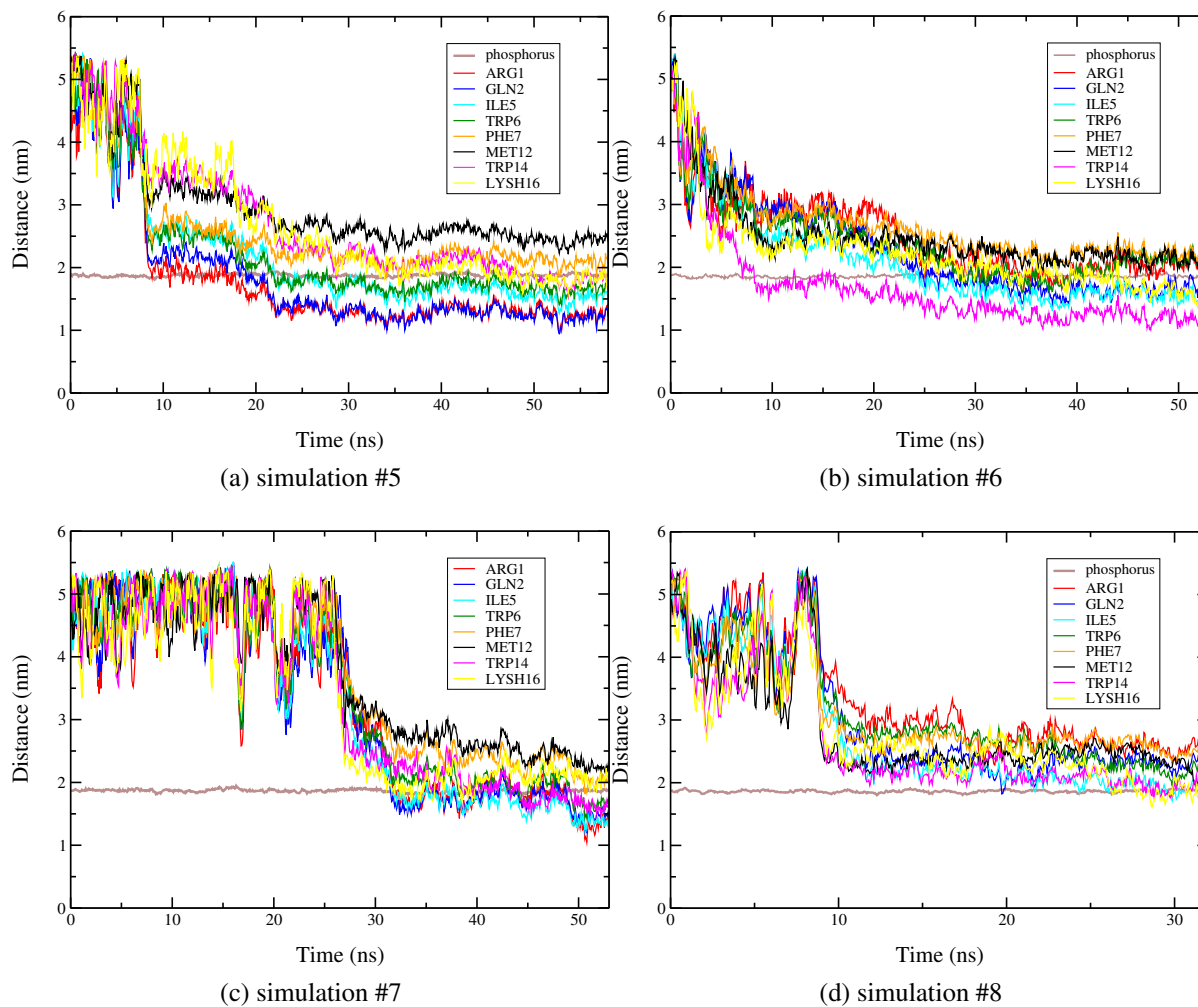


Figure 5.4: Distance between several components of the system from the middle of the bilayer for eight simulations. From these figures one can see which residue touches the membrane first and which residues go deeper into the bilayer. For more information, see Table 5.1.

Table 5.1: Details of different simulations. N-terminus corresponds to ARG1 residue and C-terminus corresponds to LYSH16 residue. Insertion depth is defined to be the average distance between the center of mass of the peptide and phosphorus atoms within the indicated time interval. A negative insertion depth means that the peptide has not penetrated. When it is zero, it means that the the penetration is about to occur. A positive insertion depth means that the peptide has penetrated.

Number	Simulation length	First Approaching Terminal	Insertion Depth
1	142 ns	N-terminus	-0.1 ± 0.1 nm (120 ns- 142ns)
2	77 ns	C-terminus	0.2 ± 0.1 nm (60 ns- 77 ns)
3	77 ns	N-terminus	0.2 ± 0.1 nm (60 ns-77 ns)
4	76 ns	C-terminus	0.3 ± 0.1 nm (60 ns- 76 ns)
5	58 ns	N-terminus	0.0 ± 0.1 nm (50 ns- 58ns)
6	53 ns	C-terminus	0.1 ± 0.1 nm (45 ns-53 ns)
7	53 ns	C-terminus	0.1 ± 0.1 nm (50 ns-53 ns)
8	33 ns	C-terminus	-0.2 ± 0.1 nm (28 ns-33 ns)

Hydrogen bonds

As was shown in previous section (Figs. 5.3 and 5.4), hydrophobic residues (TRP, ILE and MET) are playing an important role in dragging the peptide to the bilayer core. Interestingly, some hydrophilic residues, such as ARG, are able to penetrate into the bilayer. Arginine is positively charged in neutral pH, but it is able to act as a donor and make hydrogen bonds (H-bonds) with phosphate and carbonyl groups of DPPC. In addition to ARG, LYSH, GLN, ASN and TRP are also able to make H-bonds because of the presence of N-H groups in their structure. To investigate the possibility of H-bonds between these residues and the lipid bilayer, we computed the number of H-bonds between these residues and the phosphate and carbonyl groups of the lipids. As presented in Fig. 5.5, all these residues are able to make H-bonds with phosphate groups, but only ARG and LYSH make H-bond with carbonyl groups. This is an indication of the stability of these charged residues even when they are below the phosphate groups (close to carbonyl groups).

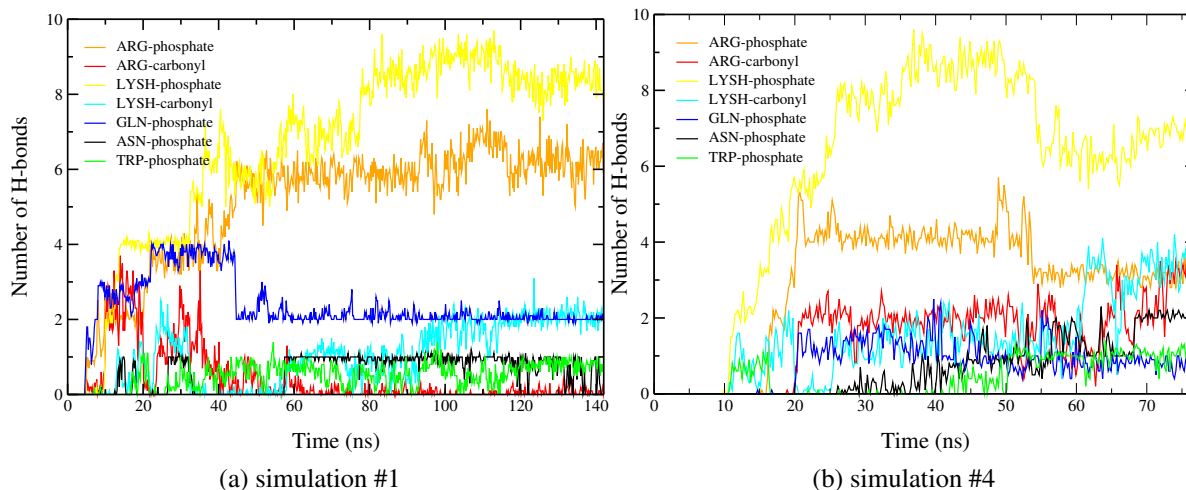


Figure 5.5: Number of H-bonds as a function of time for simulations #1 and #4. The donor groups are ARG, LYSH, GLN, ASN, and TRP residues. The acceptor groups are the C=O and PO_4^- groups of DPPC lipids.

Membrane thickness

To shed light on how the membrane reacts to the presence of penetratin, we computed the z coordinate of the nitrogen atoms of bilayer's head groups for both neighboring and distant lipids. As presented in Fig. 5.6, the neighboring lipids of the peptide are dragged towards the bilayer center. To quantify this, we computed the membrane thickness for neighboring and distant lipids. The thickness is defined to be the z -distance between the center of mass of nitrogen atoms of one monolayer from the other monolayer. As presented in Fig. 5.7, the membrane thins wherever the peptide is associated with it. The average values of the local thickness in simulations #1 and #4 are 3.84 ± 0.17 nm and 3.73 ± 0.14 nm, respectively. For distant lipids, the average values of thickness are 3.95 ± 0.63 nm and 4.00 ± 0.64 nm in simulations #1 and #4, respectively. These average values were obtained after 20 ns of simulations within which the peptide is completely associated with the bilayer (see Figs. 5.3a and 5.3d). The average thickness of a pure DPPC, as was presented in previous chapter, is 3.86 ± 0.04 nm. The high standard deviation of the values of thickness in penetratin-DPPC systems compared with the pure DPPC system, shows the high oscillation of membrane when interacting with penetratin. However, the average values are almost the same since they overlap within the given error bar. Nevertheless, comparing the local and non-local thickness in simulation #4 shows that the local thickness is 7.5% less than the non-local thickness. In transportan (previous chapter), the difference was found to be 2.6%.

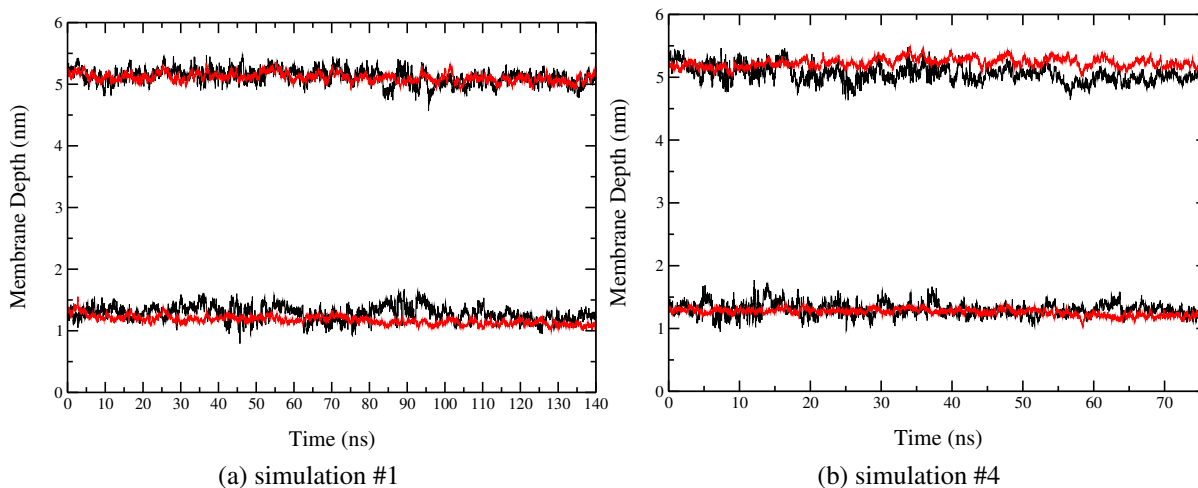


Figure 5.6: z -coordinates of nitrogen atoms of neighboring (*black*) and distant (*red*) lipids in simulations #1 (a), and #4 (b). In (a) the peptide is associated with the lower monolayer, and in (b) the peptide is associated with the upper monolayer.

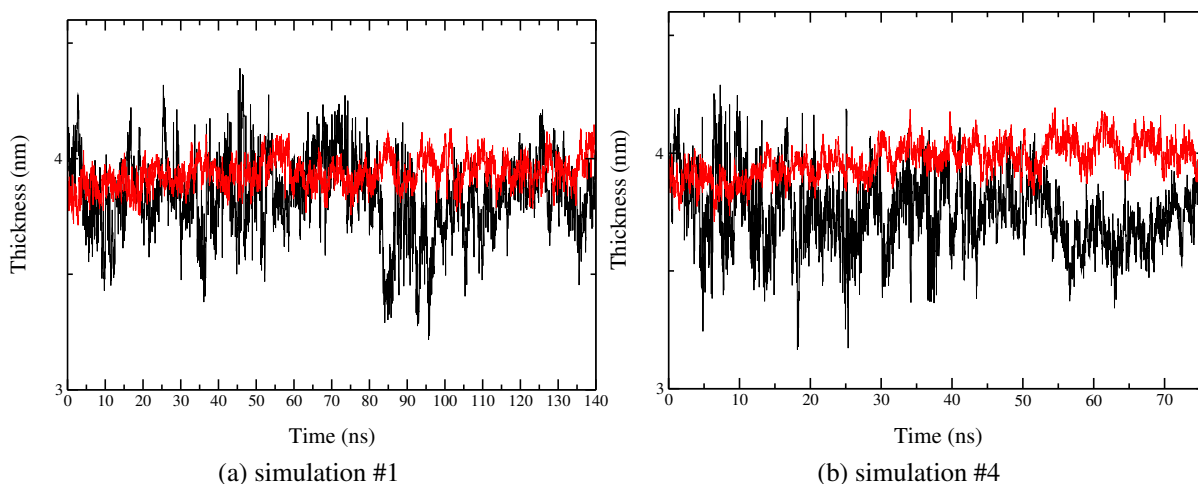


Figure 5.7: Membrane thickness for simulations #1 (a), and #4 (b), computed for the neighbor (*black*) and distant (*red*) lipids.

Order Parameter

Figure 5.8 shows the molecular order parameter, S_{mol} , as defined in previous chapter. The associated monolayer shows more disorder than the free monolayer, and this effect is more clear for *sn-1* acyl chains. For *sn-2* acyl chains, the values of order parameter is almost the same. The average value of S_{mol} is 0.31 ± 0.08 and 0.32 ± 0.08 for *sn-1* chains in the associated and free monolayer, respectively. For *sn-2* chains, this value is 0.32 ± 0.08 for both associated and free monolayer. This effect, which was also observed by Lensink *et al.* [82] in penetratin-

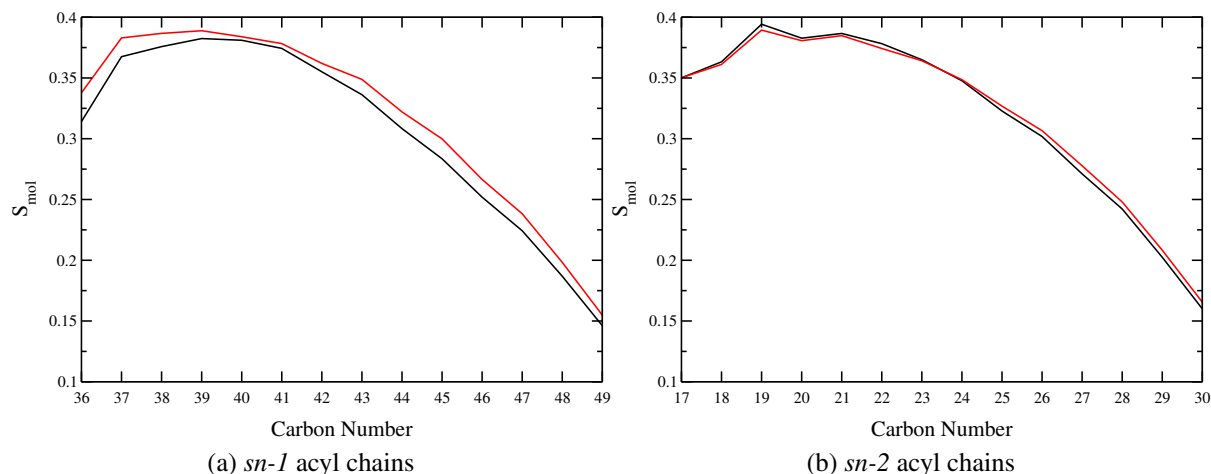


Figure 5.8: Molecular order parameter of *sn-1* (a) and *sn-2* (b) acyl chains. The associated monolayer is shown in *black* and the free monolayer is shown in *red*. The presence of penetratin induces disorder in *sn-1* acyl chains.

PC systems, is the opposite of what we observed in transportan-DPPC system (chapter four). There, we showed that the presence of transportan induces order in the bilayer.

Peptide conformation

Figure 5.9 shows the evolution of the peptide's conformation for simulations #1, #3, and #4. These results, together with other analyses (data not shown), show that there is a slight resemblance in penetratin conformation in the residues from one to six. The peptide almost takes an α -helical structure in these residues. Beyond these residues, the conformation of penetratin is different in different parts of the peptide, but a predominant motif seems to be the turn and the bend. This finding is a good match with experimental data obtained by Maler *et al.* [85] in which they found the helicity between the residues LYSH4 to MET12.

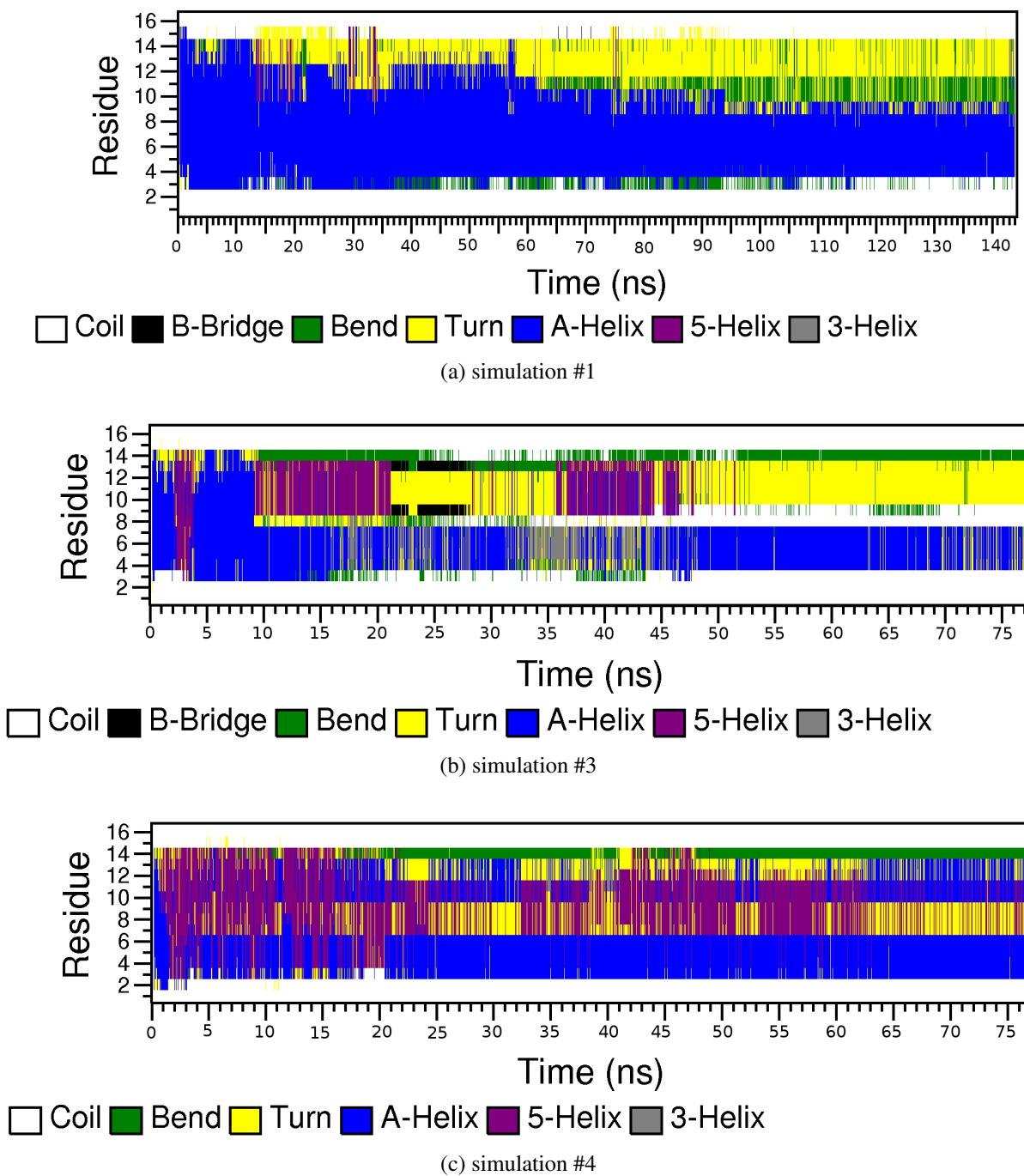


Figure 5.9: Evolution of secondary structures of transportan as defined by DSSP [68]. The residues start from ARG1 (N-terminus) to LYSH16 (C-terminus). Different colors correspond to different secondary structures as shown in legends.

5.4 Discussion and Conclusion

In this study, we showed that penetratin disorders the DPPC lipids (Fig. 5.8). Also, we showed that neighboring lipids of penetratin drag down the lipid head groups towards the bilayer core. This result was inferred from the fact that the local thickness (due to neighboring lipids) is 7.5% less than the non-local thickness (due to distant lipids). These two results are in complete contrast with those in transportan-DPPC system which was studied in previous chapter, and may be indications of different penetration mechanisms in these systems. Also, we showed the importance of charged residues, namely ARG and LYSH, in the peptide localization within the membrane. These residues can make H-bonds with carbonyl groups of lipids. Therefore, although they are charged, they can locate below the phosphate groups as indicated in Figs. 5.3a, 5.3d, 5.4a, 5.4b, and 5.4c. More analyses of this study are in progress.

Chapter 6

Conclusion and Proposal for Future Work

6.1 Conclusion

The theme of this thesis is computer simulation of cell membrane and peptides. We employed classical molecular dynamics to study the effect of dehydroergosterol (DHE) on cell membranes and compared it with cholesterol (CHOL). Then, using the same method, we studied the interaction of a cell-penetrating peptide, transportan, with cell membranes.

In chapter three, DHE was studied when interacting with POPC membrane and was shown to mimic CHOL to a certain extent: the condensing effect of CHOL was found to be 2% more than DHE. The average ordering effect of both sterols were found to be the same. The interaction of head groups of both sterols with other membrane components was also found to be the same. However, some tiny differences were found due to structural differences of sterols. For example, although the average order parameter of POPC acyl chains in both cases are the same, but the average order parameter of middle parts of acyl chains is higher in POPC-CHOL system than POPC-DHE system. This, in turn, is due to different orientation of sterols within the bilayer: CHOL is more vertical than DHE. In addition, we showed that both sterols have the same effect on the local pressure, and thus, on the elasticity of POPC membranes.

In chapter four, translocation of transportan through DPPC membranes was studied. It was shown that the free energy of transportan is the lowest when it resides within the bilayer core below membrane head groups. It was also shown that a single molecule of transportan is unlikely to make a pore in membranes. Apart from the fact that simulations showed no pore formation, this claim was supported by several indications. First, no significant thinning was observed when the peptide was associated with the membrane. Surprisingly, when the peptide was directly inserted into the bilayer core, thickening of the bilayer was observed. Secondly, not only is the peptide unable to perturb the membrane when it locates below the head groups, but also it induces order in its neighboring acyl chains. These two facts show that the peptide

adapts itself inside the bilayer after the penetration. Besides, we showed that peptides make different conformations when bound to the membrane. Another important result was the crucial role of lysine residues in translocation. These residues form salt bridges with the membrane head groups and slow down the translocation.

In chapter five, the interaction of penetratin with DPPC membranes was studied. Through detailed analyses, we showed that not only hydrophobic residues, but also charged ones, namely ARG and LYSH, are able to penetrate into the bilayer. These charged residues can make H-bonds with the carbonyl groups of the lipids and locate below the phosphate groups. As opposed to transportan, penetratin has a thinning and disordering effects on the lipids which may be an indication to a different penetration mechanism.

6.2 Proposal for Future Work

POPC lipids are not fully saturated and have a double bond in the oleoyl acyl chain. There is a general understanding that the role of CHOL is most conspicuous in bilayers composed of saturated chains such as DPPC [150, 110, 109]. That is, the distinction between the behavior of CHOL and other sterols is most clear when sterols interact with fully saturated bilayers. As shown in chapter three, there are some differences in the behavior of DHE compared to CHOL when they interact with POPC bilayers. However, one can expect a broader range of differences when these sterols interact with fully saturated lipids. Simulations on CHOL interacting with saturated lipids are abundant [125, 95, 30, 126, 151]. To our knowledge, there is no simulation on DHE interacting with saturated lipids. Such a study can give a better answer to the question of how good DHE can mimic CHOL in biological systems.

In chapter four, the interaction of a single transportan with membranes was studied. Although a lot of insights were gained by studying the behavior of system when the peptide is associated with the membrane, no complete translocation was observed within the accessible simulation time. The possibility of pore formation by a single peptide was ruled out, but the mechanism of translocation still remains unknown. One possibility is a single-peptide mechanism in which the peptide translocates the bilayer through a stochastic process. In fact, we proposed a time scale for such a process being thousands of microseconds. Another possibility is that multiple peptides are involved in the translocation process. That is, several peptides can aggregate and facilitate the translocation. Experiments as well as MD simulations support this idea for most of antimicrobial peptides [97, 76, 160, 88, 139, 62]. According to these studies, when the peptide/lipid concentration increases above a certain threshold value, translocation occurs. Simulation of several molecules of penetratin with membranes, has also provided a lot of insights into the translocation process [163]. As proposed by Mark *et al.* [163], multiple

penetratin peptides can form small vesicles within the cell that encapsulate the peptides, thus facilitate the translocation. Such a simulation should be performed with large enough membranes. Imposition of periodic boundary conditions constrain the bilayer and makes it difficult for the membrane to deform unless the membrane is large enough. In their simulation [163], 512 lipid molecules were used to simulate the bilayer making the simulation box twice the size of ours. Therefore, to investigate the possibility of a collective behavior of transportan peptides in translocation, the same procedure can be followed. One can perform several simulations in large boxes in each of which several peptides interact with the bilayer. This work is in progress.

Bibliography

- [1] J. Aittoniemi, T. Róg, P. Niemelä, M. Pasenkiewicz-Gierula, M. Karttunen, and I. Vattulainen. Tilt: major factor in sterols ordering capability in membranes. *J. Phys. Chem. B.*, 110, 2006.
- [2] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell*. Garland Science, New York, US, fourth edition, 2002.
- [3] P. F. Almeida and A. Pokorny. Mechanisms of antimicrobial, cytolytic, and cell-penetrating peptides: From kinetics to thermodynamics. *Biochem.*, 48, 2009.
- [4] H. C. Anderson. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72, 1980.
- [5] M. Bachar and O. M. Becker. Protein-induced membrane disorder: a molecular dynamics study of melittin in a dipalmitoylphosphatidylcholine bilayer. *Biophys. J.*, 78, 2000.
- [6] E. Barany-Wallje, A. Andersson, A. Graslund, and L. Maler. NMR solution structure and position of transportan in neutral phospholipid bicelles. *Febs Lett.*, 567, 2004.
- [7] E. Barany-Wallje, A. Andersson, A. Graslund, and L. Maler. Dynamics of transportan in bicelles is surface charge dependent. *J. Biomol. NMR*, 35, 2006.
- [8] J. Barojas, D. Levesque, and B. Quentrec. Simulation of diatomic homonuclear liquids. *Phys. Rev. A.*, 7, 1973.
- [9] C. I. Bayly, C. Piotr, W. D. Cornell, and P. A. Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.*, 97, 1993.
- [10] D. Beeman. Some multistep methods for use in molecular dynamics calculations. *J. Comp. Phys.*, 20, 1976.

- [11] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. (1981). Interaction models for water in relation to protein hydration. In: *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry*, ed. B. Pullman. Dordrecht, Reidel Publ. Cy., pp. 331-42.
- [12] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81, 1984.
- [13] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. Gromacs: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, 91, 1995.
- [14] J. M. Berg, J. L. Tymoczko, and L. Stryer. *Biochemistry*. W. H. Freeman and Company, 5th edition.
- [15] O. Berger, O. Edholm, and F. Jähnig. Molecular dynamics simulations of a fluid bilayer of Dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.*, 72, 1997.
- [16] D. Boal. *Mechanics of the Cell*. Cambridge University Press, UK, 2002.
- [17] S. D. Bond, B. J. Leimkuhler, and B. B. Laird. The Nosé-Poincaré method for constant temperature molecular dynamics. *J. Comp. Phys.*, 151, 1999.
- [18] M. Born and J. R. Oppenheimer. Zur quantentheorie der molekeln. *Ann. Phys.*, 389, 1927.
- [19] L. Brocchieri and S. Karlin. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res.*, 33, 2005.
- [20] C. J. Burnham and S. S. Xantheas. Development of transferable interaction models for water. II. Accurate energetics of the first few water clusters from first principles. *J. Chem. Phys.*, 116, 2002.
- [21] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126, 2007.
- [22] J. D. Butler, J. Blanchette-Mackie, E. Goldin, R. R. O'Neill, G. Carstea, C. F. Roff, M. C. Patterson, S. Patel, M. E. Comly, and A. Cooney. Progesterone blocks cholesterol translocation from lysosomes. *J. Biol. Chem.*, 267, 1992.

- [23] J. D. Butler, M. E. Comly, H. S. Kruth, M. Vanier, M. Filling-Katz, J. Fink, N. Barton, H. Weintroub, J. M. Quirk, and T. Tokoro. Niemann-pick variant disorders: comparison of errors of cellular cholesterol homeostasis in group D and group C fibroblasts. *Proc. Nat. Aca. Sci.*, 84, 1987.
- [24] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55, 1985.
- [25] S. C. Chen, J. M. Sturtevant, and B. J. Gaffney. Scanning calorimetric evidence for a third phase transition in phosphatidylcholine bilayers. *Proc. Natl. Acad. Sci.*, 77, 1980.
- [26] S. Chiu, M. Clark, V. Balaji, S. Subramaniam, H. Scott, and E. Jakobsson. Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. *Biophys. J.*, 69, 1995.
- [27] A. Chugh and F. Eudes. Cellular uptake of cell-penetrating peptides pVEC and transportan in plants. *J. Pept. Sci.*, 14, 2008.
- [28] T. H. Creighton. *Proteins: structures and molecular properties*. W. H. Freeman, San Francisco, 1993.
- [29] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117, 1995.
- [30] J. Czub and M. Baginski. Comparative molecular dynamics study of lipid membranes containing cholesterol and ergosterol. *Biophys. J.*, 90, 2006.
- [31] J. H. Davis. The description of membrane conformation, order and dynamics by ²H-NMR. *Biochim. Biophys. Acta*, 737, 1983.
- [32] D. Derossi, A. H. Joliot, G. Chassaing, and A. Prochiantz. The third helix of the antennapedia homeodomain translocates through biological membrane. *J. Biol. Chem.*, 269, 1994.
- [33] D. Derossi, S. Calvet, A. Trembleau, A. Brunissen, G. Chassaing, and A. Prochiantz. Cell internalization of the third helix of the Antennapedia homeodomain is receptor independent. *J. Bio. Chem.*, 271, 1996.
- [34] D. Derssi, G. Chassaing, and A. Prochiantz. Trojan peptides: the penetratin system for intracellular delivery. *Trends Cell Biol.*, 8, 1998.

- [35] J. P. Douliez, A. Leonard, and E. J. Dufourc. Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splitting. *Biophys. J.*, 68, 1995.
- [36] G Drin, S. Cottin, E. Blanc, A. R. Rees, and J. Temsamani. Studies on the internalization mechanism of cationic cell-penetrating peptides. *J. Biol. Chem.*, 278, 2003.
- [37] C. M. Dunkin, A. Pokorny, L. Hee-Seung, and P. F. Almeida. Molecular dynamics studies of (Tp10) interacting with a POPC lipid bilayer. *J. Phys. Chem. B*, 115, 2011.
- [38] V. E., Richard J. P., C. Rispal, and B. Lebleu. TAT peptide internalization: seeking the mechanism of entry. *Curr. Protein. Pept. Sci.*, 4, 2003.
- [39] E. Ei-Andaloussi, T. Holm, and U. Langel. Cell-penetrating peptides: mechanism and applications. *Curr. Phar. Des.*, 11, 2005.
- [40] J. U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee, and L. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103, 1995.
- [41] P. P. Ewald. Die berchnung optischer und elektrostatischer gitterpotentiale. *Ann. Phys.*, 64, 1921.
- [42] H. Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3, 1935.
- [43] R. T. Fischer, F. A. Stephenson, A. Shafiee, and F. Schroeder. Structure and dynamic properties of Dehydroergosterol, $\Delta^{5,7,9,(11),22}$ Ergostateraen- 3β -o1. *J. Bio. Phys.*, 13, 1985.
- [44] P. L. Freddolino, C. B. Harrison, Y. Liu, and K. Schulten. Challenges in protein-folding simulations. *Nature Phys.*, 6, 2010.
- [45] D. Frenkel and B. Smit. *Understanding molecular simulation: From algorithms to applications*. International Series in Natural Philosophy. Academic Press, San Diego, 2nd edition, 2002.
- [46] A. Frolov, J. K. Woodford, E. J. Murphy, J. T. Billheimer, and F. Schroeder. Spontaneous and protein-mediated sterol transfer between intracellular membranes. *J. biol. Chem.*, 271, 1996.
- [47] O. Garvik, P. Benediktson, A. C. Simonsen, J. H. Ipsen, and D. Wustner. The fluorescent cholesterol analog dehydroergosterol induces liquid-ordered domains in model membranes. *Chem. Phys. Lipids*, 159, 2009.

- [48] E. Gazit, W. J. Lee, P. T. Brey, and Y. Shai. Mode of action of the antibacterial cecropin B2: a spectrofluorometric study. *Biochem.*, 33, 1994.
- [49] G. Gimpl, U. Klein, H. Reilander, and F. Fahrenholz. Expression of the human oxytocin receptor in baculovirus-infected insect cells: high-affinity binding is induced by a cholesterolcyclodextrin complex. *Biochem.*, 34, 1995.
- [50] E. Guardia, J. Marti, L. Garcia-Tarres, and D. Laria. A molecular dynamics simulation study of hydrogen bonding in aqueous ionic solutions. *J. Mol. Liquids*, 117, 2005.
- [51] B. Guillot. A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.*, 101, 2002.
- [52] H. D. Herce and A. E. Garcia. Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes. *Proc. Nat. Aca. Sci.*, 104, 2007.
- [53] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaij. LINCS, a linear constraint solver for molecular simulations. *J. Comp. Chem.*, 18, 1997.
- [54] B. Hess, C. Kutznerand, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4, 2008.
- [55] R. Hockney and J. Eastwood. *Computer simulation using particles*. Institute of Physics publishing, London, 1988.
- [56] C. Hofsaß, E. Lindhal, and O. Edholm. Molecular dynamics simulations of phospholipid bilayers with cholesterol. *Biophys. J.*, 84, 2003.
- [57] M. Holtje, T. Forster, B. Brandt, T. Engels, W. von Rybinski, and H. D. Holtje. Molecular dynamics simulations of stratum corneum lipid models: fatty acids and cholesterol. *Biochim. Biophys. Acta*, 1511, 2001.
- [58] R. Hooke. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses*. John Martyn and James Allestry, London, 1665.
- [59] K. Huang. *Statistical Mechanics*. John Wiley & Sons, 2nd edition, 1987.
- [60] P. H. Hünenberger. Thermostat algorithms for molecular dynamics simulations. *Adv. Polym. Sci.*, 173, 2005.

- [61] R. Iftimie, P. Minary, and E. Tuckerman. Ab initio molecular dynamics: Concepts, recent developments, and future trends. *Proc. Natl. Acad. Sci.*, 102, 2005.
- [62] G. Illya and M. Deserno. Coarse-grained simulation studies of peptide-induced pore formation. *Biophys. J.*, 95, 2008.
- [63] J. N. Israellachvili. *Intermolecular and surface forces*. Academic Press, San Diego, CA, 2nd edition, 1991.
- [64] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel. Langevin stabilization of molecular dynamics. *Chem. Phys.*, 114, 2001.
- [65] M. J. Janiak, D. M. Small, and G. G. Shipley. Temperature and compositional dependence of the structure of hydrated dimyristoyl lecithin. *J. Bio. Chem.*, 254, 1979.
- [66] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78, 1997.
- [67] W. Jorgensen and J. Tirado-Rives. The opls potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Ame. Chem. Soc.*, 110, 1988.
- [68] W. Kabsch and C. Sandra. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymer*, 22, 1983.
- [69] J. Kastner and W. Thiel. Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: Umbrella integration. *J. Chem. Phys.*, 123, 2005.
- [70] J. Katsaras. Adsorbed to a rigid substrate, Dimyristoylphosphatidylcholine multibilayers attain full hydration in all mesophases. *Biophys. J.*, 75, 1998.
- [71] J. Katsaras, D. S. Yang, and R. M. Epanand. Fatty-acid chain tilt angles and directions in dipalmitoyl phosphatidylcholine bilayers. *Biophys. J.*, 63, 1992.
- [72] J. G. Kirkwood. Statistical mechanics of fluid mixtures. *J. Chem. Phys.*, 3, 1935.
- [73] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comp. Chem.*, 13, 1992.

- [74] M. Kupiainen, E. Flack, S. Ollila, P. S. Niemelä, A. A. Gurtovenko, M. T. Hyvönen, M. Patra, M. Karttunen, and I. Vattulainen. Free volume properties of sphingomyelin, DMPC, DPPC, and POPC bilayers. *J. Comp. Theo. Nan. Sci.*, 2, 2005.
- [75] L. Verlet. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159, 1967.
- [76] L. Yang, T. A. Harroun, T. M. Weiss, L. Ding, and H. W. Huang. Barrel-stave model or toroidal model? a case study on melittin pores. *Biophys. J.*, 81, 2001.
- [77] B. D. Ladbrooke and D. Chapman. Thermal analysis of lipids, proteins and biological membranes. a review and summary of some recent studies. *Chem. Phys. Lipids*, 3, 1969.
- [78] U. Langel, M. Pooga, K. Czeslava, M. Zilmer, and T. Bartfai. A galanin-mastoparan chimeric peptide activates the Na⁺,K⁺-ATPase and reverses its inhibition by ouabain. *Elsevier Sci. B.V.*, 62, 1996.
- [79] A. R. Leach. *Molecular modeling: principles and applications*. Longman, Harlow, UK, 1996.
- [80] B. W. Lee, R. Faller, A. K. Sum, I. Vattulainen, M. Patra, and M. Karttunen. Structural effects of small molecules on phospholipid bilayers investigated by molecular simulations. *Fluid Phase Equilibria*, 225, 2004.
- [81] J. A. Leifert and J. Lindsay Whitton. Translocatory proteins and protein transduction domains: a critical analysis of their biological effects and the underlying mechanisms. *Molec. Ther.*, 8, 2003.
- [82] M. F. Lensink, B. Christiaens, J. Vandekerckhove, A. Prochiantz, and M. Rosseneu. Penetratin-membrane association: W48/R52/W56 shield the peptide from the aqueous phase. *Biophys. J.*, 88, 2005.
- [83] H. Leontiadou, A. E. Mark, and S. J. Marrink. Antimicrobial peptides in action. *J. Ame. Chem. Soc.*, 128, 2006.
- [84] E. Lindahl, B. Hess, and D. van der Spoel. Gromacs 3.0: a package for molecular simulation and trajectory analysis. *Mol. Mod.*, 7, 2001.
- [85] M. Lindberg, H. Biverstahl, A. Graslund, and L. Maler. Structure and positioning comparison of two variants of penetratin in two different membrane mimicking systems by NMR. *Eur. J. Biochem.*, 270, 2003.

- [86] M. Lindberg, J. Jarvet, U. Langel, and A. Graslund. Secondary structure and position of the cell-penetrating peptide transportan in SDS micelles as determined by nmr. *Biochem*, 40, 2001.
- [87] H. Lodish, A. Berk, S. Lawrence Zipursky, P. Matsudaira, D. Baltimore, and J. E. Darnell. *Molecular Cell Biology*. W. H. Freeman and Company, 4th edition, 2000.
- [88] M. L. Longo, A. J. Waring, L. M. Gordon, and D. A. Hammer. Area expansion and permeation of phospholipid membrane bilayers by influenza fusion peptides and melittin. *Langmuir*, 14, 1998.
- [89] C. P. Lowe. An alternative approach to dissipative particle dynamics. *Europhys. Lett.*, 47, 1999.
- [90] A. Luzar and D. Chandler. Structure and hydrogen bond dynamics of water-dimethyl sulfoxide mixtures by computer simulations. *J. Chem. Phys.*, 98, 1993.
- [91] T. G. M. and J. P. Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J. Comp. Phys.*, 23, 1977.
- [92] A. E. Mark, S. P. van Helden, Smith. P. E., L. H. M. Janssen, and W. F. van Gunsteren. Convergence properties of free energy calculations: α -cyclodextrin complexes as a case study. *J. Am. Chem. Soc.*, 116, 1994.
- [93] S. J. Marrink, A. H. de Vries, and A. E. Mark. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B.*, 108, 2004.
- [94] D. Marsh. Lateral pressure in membranes. *Biochim. Biophys. Acta*, 1286, 1996.
- [95] H. Martinez-Seara, T. Róg, M. Karttunen, I. Vattulainen, and R. Reigada. Cholesterol induces specific spatial and orientational order in cholesterol/phospholipid membranes. *PLoS One*, 5, 2010.
- [96] D. Marx and J. Hutter. Ab initio molecular dynamics: Theory and implementation. *Modern methods and algorithms of quantum chemistry*, J. Grotendorst (Ed.), John von Neumann Institute for Computing, Jülich, NIC Series, Vol. 1, ISBN 3-00-005618-1, pp. 301-449, 2000.
- [97] K. Matsuzaki, K. Sugishita, N. Ishibe, M. Ueha, S. Nakata, K Miyajima, and R. M. Epand. Relationship of membrane curvature to the formation of pores by magainin 2. *Bio. Chem.*, 37, 1998.

- [98] S. Miyamoto and P. A. Kollman. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comp. Chem.*, 13, 1992.
- [99] T. Morishita. Fluctuation formulas in molecular-dynamics simulations with the weak coupling heat bath. *J. Chem. Phys.*, 113, 2000.
- [100] P. M. Morse. Diatomic molecules according to the wave mechanics. II. vibrational levels. *Phys. Rev.*, 34, 1929.
- [101] O. G. Mouritsen and M. J. Zuekermann. What's so special about cholesterol? *Lipids*, 39, 2004.
- [102] G. J. Mulder. On the composition of some animal substances. *Journal für praktische Chemie*, 16, 1839.
- [103] P. Muller and A. Herrmann. Rapid transbilayer movement of spin-labeled steroids in human erythrocytes and in liposomes. *Biophys. J.*, 82, 2002.
- [104] K. Murzyn, W. Zhao, M. Karttunen, and T. Róg. Dynamics of water at membrane surfaces: Effect of headgroup structure. *Biointerphases I*, 1, 2006.
- [105] J. F. Nagle, R. Zhang, S. Tristram-Nagle, W. Sun, H. I. Petrache, and R. M. Suter. X-ray structure determination of fully hydrated l_{α} phase dipalmitoylphosphatidylcholine bilayers. *Biophys. J.*, 70, 1996.
- [106] D. L. Nelson and M. M. Cox. In *Lehninger principles of biochemistry*. 4rd ed., (W. H. Freeman, New York, 2004).
- [107] D. L. Nelson and M. M. Cox. In *Lehninger principles of biochemistry*. 3rd ed., (Worth Publishers, New York, 2000).
- [108] P. S. Niemelä, M. T. Hyvönen, S. Ollila, M. Karttunen, and I. Vattulinen. Assessing the nature of lipid raft membranes. *PLoS Comput Biol*, 3, 2007.
- [109] S. Ollila, M. T. Hyvönen, and I. Vattulainen. Polyunsaturation in lipid membranes: Dynamic properties and lateral pressure profile. *J. Phys. Chem. B*, 111, 2007.
- [110] S. O. H. Ollila, T. Róg, M. Karttunen, and I. Vattulainen. Role of sterol type on lateral pressure profiles of lipid membranes affecting membrane protein functionality: comparison between cholesterol, desmosterol, 7-dehydrocholesterol and ketosterol. *J. Struc. Bio.*, 159, 2007.

- [111] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52, 1981.
- [112] M. Pasenkiewicz-Gierula, Y. Takaoka, H. Miyagawa, K. Kitamura, and A. Kusymi. Charge pairing of headgroups in phosphatidylcholine membranes: a molecular dynamics simulation study. *Biophys. J.*, 76, 1999.
- [113] R. K. Pathria. *Statistical mechanics*. International Series in Natural Philosophy. Pergman, Oxford, 1985.
- [114] M. Patra and M. Karttunen. Systematic comparison of force fields for microscopic simulations of NaCl in aqueous solutions: Diffusion, free energy of hydration and structural properties. *J. Com. Chem.*, 25, 2004.
- [115] M. Patra, M. Karttunen, M. T. Hyvonen, E. Flack, P. Lindqvist, and I. Vattulainen. Md simulations of lipid bilayers: Artifacts due to truncating electrostatic interactions. *Eur. Biophys. J.*, 32, 2003.
- [116] M. Patra, M. Karttunen, M. T. Hyvonen, E. Flack, P. Lindqvist, and I. Vattulainen. Molecular dynamics simulations of lipid bilayers: Major artifacts due to truncating electrostatic interactions. *Biophys. J.*, 84, 2003.
- [117] M. Patra, M. Karttunen, M. T. Hyvonen, E. Flack, and I. Vattulainen. Lipid bilayers driven to a wrong lane in molecular dynamics simulations by subtle changes in long-range electrostatic interactions. *J. Phys. Chem. B*, 108, 2004.
- [118] M. Patra, E. Salonen, E. Terama, I. Vattulainen, R. Faller, B. W. Lee, J. Holopainen, and M. Karttunen. Under the influence of alcohol: The effect of ethanol and methanol on lipid bilayers. *Biophys. J.*, 90, 2006.
- [119] M. Pooga, M. Hallbrink, M. Zorko, and U. Langel. Cell penetration by transportan. *FASEB*, 12, 1998.
- [120] A. Rahman and F. H. Stillinger. Molecular dynamics study of liquid water. *J. Chem. Phys.*, 55, 1971.
- [121] A. Ramamoorthy, S. Thennarasu, D. K. Lee, A. Tan, and L. Maloy. Solid-state NMR investigation of the membrane-disrupting mechanism of antimicrobial peptides MSI-78 and MSI-594 derived from magainin 2 and melittin. *Biophys. J.*, 91, 2006.
- [122] F. Reif. *Statistical thermal physics*. McGRAW-HILL Book Company. International student edition, 1965.

- [123] H. J. Risselada and J. S. Marrink. The molecular face of lipid rafts in model membranes. *Proc. Nat. Aca. Sci.*, 105, 2008.
- [124] A. Robinson, W. G. Richards, P. J. Thomas, and M. M. Hann. Behaviour of cholesterol and its effect on head group and chain conformations in lipid bilayers: a molecular dynamics study. *Biophys. J.*, 68, 1995.
- [125] T. Róg and M. Pasenkiewicz-Gierula. Non-polar interactions between cholesterol and phospholipids: a molecular dynamics simulation study. *Biophys. Chem.*, 107, 2004.
- [126] T. Róg, M. Pasenkiewicz-Gierula, I. Vattulainen, and M. Karttunen. What happens if cholesterol is made smoother: Importance of methyl substituents in cholesterol ring structure on phosphatidylcholinesterol interaction. *Biophys. J.*, 92, 2007.
- [127] T. Róg, M. Pasenkiewicz-Gierula, I. Vattulainen, and M. Karttunen. Ordering effects of cholesterol and its analogues. *Biochim. Biophys. Acta*, 1788, 2009.
- [128] T. Róg, I. Vattulainen, M. Jansen, E. Ikonen, and M. Karttunen. Comparison of cholesterol and its direct precursors along the biosynthetic pathway: Effects of cholesterol, desmosterol and 7-dehydrocholesterol on saturated and unsaturated lipid bilayers. *J. Chem. Phys.*, 129, 2008.
- [129] J. W. Rohlf. *Modern physics from α to Z $^\circ$* . John Wiley & Sons, 1st edition, 1994.
- [130] R. E. Rudd and J. Q. Broughton. Coarse grained molecular dynamics and the atomic limit of finite elements. *Phys. Rev. B.*, 58, 1998.
- [131] J.-P. Ryckaert, G. Ciccotti, and H. J. C.. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.*, 23, 1977.
- [132] S. Safran. *Statistical thermodynamics of surfaces, interfaces, and membranes*. Addison-Wesley, 1994.
- [133] N. Sapay, W. F. D. Bennett, and D. P. Tieleman. Thermodynamics of ip-op and desorption for a systematic series of phosphatidylcholine lipids. *Soft Matt.*, 5, 2009.
- [134] H. A. Scheidt, P. Muller, A. Herrmann, and D. Huster. The potential of fluorescent and spin-labeled steroid analogs to mimic natural cholesterol. *J. Biol. Chem.*, 278, 2003.
- [135] F. Schroeder. Fluorescent sterols: probe molecules of membrane structure and function. *Prog. Lipid. Res.*, 23, 1984.

- [136] F. Schroeder, A. A. Frolov, E. J. Murphy, B. P. Atshaves, J. R. Jefferson, L. Pu, G. W. Wood, W. B. Foxworth, and A. B. Kier. Recent advances in membrane cholesterol domain dynamics and intracellular cholesterol trafficking. *Proc. Soc. Exp. Biol. Med.*, 213, 1996.
- [137] S. R. Schwarze, K. A. Hruska, and S. F. Dowdy. Protein transduction: unrestricted delivery into cells? *Trends. Cell Bio.*, 10, 2000.
- [138] J. Seelig and N. Waespe-Sarcevic. Molecular order in cis and trans unsaturated phospholipid bilayers. *Biochem.*, 17, 1978.
- [139] D. Sengupta, H. Leontiadou, A. E. Mark, and S. J. Marrink. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochim. Biophys. Acta*, 1778, 2008.
- [140] S. J. Singer and G. L. Nicolson. The fluid mosaic model of the structure of cell membranes. *Science*, 175, 1972.
- [141] G. S. Smith, E. B. Sirota, C. R. Safinya, and N. A. Clark. Structure of the L_{β} phases in a hydrated phosphatidylcholine multimembrane. *Phys. Rev. Lett.*, 60, 1988.
- [142] J. Sonne, F. Y. Hansen, and G. H. Peters. Methodological problems in pressure profile calculations for lipid bilayers. *J. Chem. Phys.*, 122, 2005.
- [143] T. P. Straatsma, H. J. C. Berendsen, and J. P. M. Postma. Free energy of hydrophobic hydration: A molecular dynamics study of noble gases in water. *J. Chem. Phys.*, 85, 1985.
- [144] W. C. Swope, H. C. Anderson, P. H. Berens, and K. R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.*, 76, 1982.
- [145] C. T. Taylor, G. T. Furuta, K. Synnestvedt, and S. P. Colgan. Phosphorylation-dependent targeting of cAMP response element binding protein to the ubiquitin/proteasome pathway in hypoxia. *Proc. Natl. Acad. Sci.*, 97, 2000.
- [146] R. H. Templer, J. S. Castel, A. R. Curran, G. Rumbles, and D. R. Klug. Sensing isothermal changes in the lateral pressure in model membranes using di-pyrenyl phosphatidylcholine. *Faraday Discuss.*, 111, 1999.
- [147] D. Terrone, S. L. Sang, L. Roudaia, and J. R. Silvius. Penetratin and related cell-penetrating cationic peptides can translocate across lipid bilayers in the presence of a transbilayer potential. *Biochem.*, 42, 2003.

- [148] D. P. Tieleman and H. J. C. Berendsen. Molecular dynamic simulation of fully hydrated dipalmitoylphosphatidylcholine bilayer with different microscopic boundary conditions and parameters. *J. Chem. Phys.*, 105, 1996.
- [149] D. Tom, Y. Darrin, and P. Lee. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, 98, 1993.
- [150] T. Róg and M. Pasenkiewicz-Gierula. Cholesterol effects on a mixed-chain phosphatidylcholine bilayer: a molecular dynamics simulation study. *Biochimie*, 88, 2006.
- [151] S. Vainio, M. Jansen, M. Koivusalo, T. Róg, M. Karttunen, I. Vattulainen, and E. Ikonen. Desmosterol cannot replace cholesterol in lipid rafts. *J. Biol. Chem.*, 281, 2006.
- [152] A. R. van Buuren, S. J. Marrink, and H. J. C. Berendsen. A molecular dynamics study of the decane/water interface. *J. Phys. Chem.*, 97, 1993.
- [153] H. J. C., Berendsen and W. F., van Gunsteren. Gromos-87 manual. Biomos BV Nijenborgh 4, 9747 AG Groningen, The Netherlands 1987.
- [154] S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi and W. F van Gunsteren. Biomolecular simulation: The GROMOS96 manual and user guide. Zürich, Switzerland: Hochschulverlag AG an der ETH Zürich. 1996.
- [155] G. Vereb, J. Szöllösi, J. Matkó, P. Nagy, T. Farkas, L. Vigh, Mátyus L., T. A. Waldmann, and S. Damjanovich. Dynamic, yet structured: The cell membrane three decades after the SingerNicolson model. *Proc. Nat. Aca. Sci.*, 100, 2003.
- [156] A. J. Verkleij, B. de Kruijff, W. F. Gerritsen, R. A. Demel, L. L. van Deenen, and P. H. Ververgaert. Freeze-etch electron microscopy of erythrocytes, *Acholeplasma laidlawii* cells and liposomal membranes after the action of filipin and amphotericin B. *Biochim. Biophys. Acta*, 291, 1973.
- [157] K. W. von Nägeli and C. E. Cramer. Pflanzenphysiologische untersuchungen. 1 Heft. Zurich, F. Schultess, 1855.
- [158] E. S. Watson, Ridgefield, and M. J. O'Neill. Differential microcalorimeter. U.S. Patent 3263484.
- [159] P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Sterinman, and J. B. Rothbard. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *Proc. Nat. Aca. Sci.*, 97, 2000.

- [160] T. Wieprecht, O. Apostolov, M. Beyermann, and J. Seelig. Membrane binding and pore formation of the antibacterial peptide pglA: thermodynamic and mechanistic aspects. *Biochem*, 39, 2000.
- [161] X. Xu and E. London. The effect of sterol structure on membrane lipid domains reveals how cholesterol can induce lipid domain formation. *Biochem.*, 39, 2000.
- [162] L. Yang, V. D. Gordon, D. R. Trinkle, N. W. Schmidt, M. A. Davis, C. DeVries, A. Som, J. E. Cronan, Jr., G. N. Tew, and G. C. L. Wong. Mechanism of a prototypical synthetic membrane-active antimicrobial: Efficient hole-punching via interaction with negative intrinsic curvature lipids. *Proc. Nat. Aca. Sci.*, 105, 2008.
- [163] S. Yesylevskyy, S. J. Marrink, and A. E. Mark. Alternative mechanisms for the interaction of the cell-penetrating peptides penetratin and the TAT peptide with lipid bilayers. *Biophys. J.*, 97, 2009.
- [164] M. Zorko and U. Langel. Cell-penetrating peptides: mechanism and kinetics of cargo delivery. *Adv. Drug Delivery Rev.*, 57, 2005.

Curriculum Vitae

Name: Amir Mohsen Pourmousa Abkenar

Post-Secondary Education and Degrees: Sharif University of Technology
Tehran, Iran
1999 - 2004 B.Sc.

Institute for Advanced Studies in Basic Sciences
Zanjan, Iran
2004 - 2007 M.Sc.

University of Western Ontario
London, ON
2008 - 2012 Ph.D.

Honours and Awards: Western Graduate Research Scholarship
2008-2012

Best student poster presentation in HPC Innovation for Research
York University, Toronto
May 6, 2010

Full score (990/990) on the Graduate Record Examination (physics)
Dec. 2007

Related Work Experience: Teaching Assistant
The University of Western Ontario
2008 - 2012

Publications:

1. Pourmousa, M., Wong-ekkabut, J., Patra, M., and Karttunen, M. Molecular Dynamic Studies of Transportan Interacting with a DPPC Lipid Bilayer, *to be submitted*

2. Pourmousa, M., Róg, T., and Karttunen, M. Interaction of POPC bilayers with Cholesterol and Dehydroergosterol, *to be submitted*
3. Pourmousa, M., Patra, M., and Karttunen, M. Molecular Dynamic Studies of Penetratin Interacting with a DPPC Lipid Bilayer, *to be submitted*