TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

JUHO LIEKKINEN

COMPUTATIONAL STUDY OF PULMONARY SURFACTANT PROTEIN B USING MOLECULAR DYNAMICS SIMULATIONS

Master of Science thesis

Examiner: Prof. Ilpo Vattulainen
Examiner: Asst. Prof. Ville Santala
Examiner and topic approved by the
Faculty Council of the Faculty of
Natural Sciences
on February 1st 2017

ABSTRACT

JUHO LIEKKINEN: Computational Study of Pulmonary Surfactant Protein B Using Molecular Dynamics Simulations
Tampere University of Technology
Master of Science thesis, 83 pages, 10 Appendix pages
March 2017
Master's Degree Programme in Environmental and Energy Technology
Major: Bioengineering
Supervisors: Dr. Giray Enkavi, M.Sc. Matti Javanainen
Examiners: Prof. Ilpo Vattulainen, Asst. Prof. (tenure track) Ville Santala
Keywords: SP-B, lung surfactant, monolayer, self-assembly, protein-lipid interactions

Pulmonary surfactant is a surface active lipid–protein complex covering the air–liquid interface at the surface of the alveoli in the lungs. Its main function is to reduce the surface tension at the interface, and thus to minimize the work of breathing and prevent the alveoli from collapsing at the end of the breathing cycle. Pulmonary surfactant protein B (SP-B) is an essential protein, necessary for the formation and maintaining of the film at the interface. Despite its importance, there has been no structural model for SP-B, or information about its molecular mechanism of function.

In this thesis, we study the specific lipid interactions and membrane binding conformations of our new refined model for the SP-B hexamer of dimers. We use molecular dynamics simulations with the coarse-grained MARTINI force field for spontaneous lipid self-assembly and monolayer studies with SP-B. We concentrate on specific protein-lipid interactions, lateral lipid reorganization, and perturbations caused by SP-B in membranes.

The results show specific lipid interaction sites in the structure of SP-B. We found that the protein causes lateral reorganization of lipids in monolayers and shows specificity towards phosphatidylglycerol and cholesterol. We further found that SP-B as a hexamer of dimers has specific membrane binding residues that orient the protein parallel to the surface of the membrane. SP-B causes lipid protrusions in monolayers and membranes. These results imply a molecular mechanism for lipid transfer through a SP-B oligomer ring in the surfactant.

TIIVISTELMÄ

JUHO LIEKKINEN: Keuhkosurfaktanttiproteiini B:n laskennallinen tutkimus käyttänen molekyylidynamiikkasimulaatioita
Tampereen teknillinen yliopisto
Diplomityö, 83 sivua, 10 liitesivua
Maaliskuu 2017
Ympäristö- ja energiatekniikan koulutusohjelma
Pääaine: Bioengineering
Ohjaajat: Dr. Giray Enkavi, M.Sc. Matti Javanainen
Tarkastajat: Prof. Ilpo Vattulainen, Asst. Prof. (tenure track) Ville Santala
Avainsanat: SP-B, keuhkosurfaktantti, solukalvo, proteiini–lipidi -vuorovaikutukset

Keuhkosurfaktantti on pinta-aktiivinen lipoproteiinikompleksi keuhkorakkuloiden ilma–vesi-rajapinnassa. Sen tärkein tehtävä on alentaa keuhkorakkuloiden pintajännitystä ja siten vähentää hengittämiseen tarvittavaa työtä sekä estää keuhkorakkuloiden kasaanpainuminen uloshengityksen aikana. Keuhkosurfaktanttiproteiini B (SP-B) on elintärkeä proteiini, jota tarvitaan surfaktanttikalvon muodostumiseen ja ylläpitämiseen keuhkorakkuloiden pinnalle. Huolimatta SP-B:n välttämättömyydestä, sille ei ole ollut rakenteellista mallia tai molekyylitason toimintamekanismia.

Tässä työssä tutkittiin uuden kuudesta dimeeristä koostuvan oligomeerisen SP-B proteiinimallin lipidi-vuorovaikutuksia sekä proteiinin sitoutumista lipidikalvoihin. Tähän tarkoitukseen sovelsimme molekyylidynamiikkasimulaatioita käyttäen karkeistettua MARTINI-voimakenttää, jonka avulla tutkimme lipidien spontaania aggregoitumista ja monomolekulaarisia lilidikalvoja SP-B:n kanssa. Keskityimme pääasiassa SP-B:lle ominaisiin proteiini–lipidi -vuorovaikutuksiin, lipidien uudelleenjärjestymiseen kalvoissa sekä SP-B:n aiheuttamiin häiriöhin lipidikalvoissa.

Tutkimuksessa havaittiin SP-B:n rakenteessa olevan erityisiä lipidivuorovaikutuspaikkoja. Proteiini aiheutti lipidien uudelleenjärjestymistä lipidikalvoissa, ja se osoitti sitoutumisspesifisyyttä erityisesti fosfatidyyliglyserolia ja kolesterolia kohtaan. SP-B-oligomeerissa on membraaneihin sitoutumisen kannalta tärkeitä aminohappoja, jotka suuntaavat proteiinin yhdensuuntaisesti kalvon pinnan suhteen. SP-B aiheutti rakenteellisia muutoksia lipidikalvoihin. Nämä tulokset viittaavat molekyylitason toimintamalliin, jossa lipidit siirtyvät kalvolta toiselle kulkien SP-B-oligomeerirenkaan läpi.

PREFACE

This Master of Science Thesis was carried out in the Biological Physics and Soft Matter (BIO) research group of the Laboratory of Physics of Tampere University of Technology between September of 2016 and February of 2017. The molecular dynamics simulations were conducted using the computing services of the Finnish IT Center for Scientific Computing (CSC).

First of all, I would like to thank my examiner Prof. Ilpo Vattulainen for this one of a kind opportunity to work with the renowned Biological physics group and this interesting project. I would also like to thank Prof. Ville Santala for representing the Laboratory of Chemistry and Bioengineering as my second examiner.

Special thanks to my supervisors Matti Javanainen and Giray Enkavi for their advice and support during the various stages of this project. I would also like to thank all my colleagues in Tampere and Helsinki, especially Hanna and Outi for all the help and the inspiring discussions ranging from protein dynamics to the depths of the Bermuda Triangle during the past months.

Last but not least, I would like to thank my family and friends for their support during my years at TUT.

Tampere, March 2017

Juho Liekkinen

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

A	Area
AA	All atom
APL	Area per lipid
ARDS	Acute respiratory distress syndrome
b_{ij}	Equilibrium bond length
C_n	Proper dihedral force constant for Ryckaert-Bellemans potential function
Ca^+	Calcium ion
CATH	Class, Architecture, Topology, Homology -database
CO_2	Carbon dioxide
CG	Coarse-grained
CHOL	Cholesterol
CB	Composite body
DC	Lipid composition with DPPC and cholesterol
$\mathrm{Dim_c}$	Disulfide-bridged or covalent dimer
$\mathrm{Dim_f}$	Functional dimer
DPPC	Dipalmitoylphosphatidylcholine
EQ	Lipid composition with equal molar fractions of lipids
$\mathbf{F}_i$	Force acting on particle i
GROMACS	Groningen Machine for Chemical Simulations
H^+	Hydrogen ion / Proton
LB	Lamellar body
LA	Large surfactant aggregate
$\mathrm{L_c}$	Liquid condensed
$\mathrm{L_d}$	Liquid disordered
$\mathrm{L_e}$	Liquid expanded
$\mathrm{L_o}$	Liquid ordered
L_β	Fluidic gel
K	Kinetic energy
k_B	Boltzmann's constant
k^b_{ij}	Force constant
k^θ_{ijk}	Angular force constant
k_ϕ	Periodic proper dihedral force constant
k_ε	Improper dihedral force constant
m_i	Mass of a particle i

MD	Molecular dynamics
MM	Molecular mechanics
N_f	Number of degrees of freedom
n_{PC}	Number of steps between pressure scaling
NPT	Isothermal–isobaric ensemble
NVE	Microcanonical ensemble
NVT	Canonical ensemble
O_2	Dioxygen molecule
P	Pressure
P_0	Reference pressure
PC	Phosphatidylcholine
PDB	Protein Data Bank
PE	Phosphatidylethanolamine
PG	Phosphatidylglyserol
PHYS	Physiological lipid composition
PI	Phosphatidylinositol
POPC	Palmitoyloleoylphosphatidylcholine
POPG	Palmitoyloleoylphosphatidylglyserol
PS	Phosphatidylserine
$p(v_i)$	Probability function of a velocity v_i
$\mathbf{P}_{0ij}$	Target pressure
PME	Particle Mesh Ewald
q	Charge of a particle
QM	Quantum mechanics
RDS	Respiratory distress syndrome
$\mathbf{r}_i$	Position of particle i
r_{ij}	Distance between particles i and j
SAPLIP	Saposin like protein
SA	Small surfactant aggregate
SA	Self-assembly
SCOP	Structural Classification of Proteins database
SM	Sphingomyelin
SP	Surfactant protein
SP-A	Pulmonary Surfactant Protein A
SP-B	Pulmonary Surfactant Protein B
SP-C	Pulmonary Surfactant Protein C
SP-D	Pulmonary Surfactant Protein D

SRT	Surfactant replacement therapy
t	Time
T	Temperature
T_o	Reference temperature
T_m	Main transition temperature
TM	Tubular myelin
UA	United atom
V	Potential energy
V_{C}	Coulomb potential
V_{LJ}	Lennard-Jones Potential
WHO	World Health Organization
β	Isothermal compressibility
Δt	Time step
ϵ_0	Permettivity in vacuum
ϵ_r	Relative permettivity
θ_{ijk}	Angle between particles i, j, and k
θ^0_{ijk}	Equilibrium angle between particles i, j, and k
ϕ	Torsion angle
ϕ_s	Reference torsion angle
ξ_{ijkl}	Improper dihedral angle
ξ^0_{ijkl}	Equilibrium improper dihedral angle
Π–A	Pressure area isotherm
τ_p	Pressure time constant

1. INTRODUCTION

The respiratory surface of the lungs is covered by the pulmonary surfactant, a surface active film composed of membrane structures and surfactant proteins [1]. The main function of the surfactant is to optimize breathing by reducing the surface tension at the alveolar air–water interface, thus preventing the collapse of the alveoli during the compression–expansion cycles [2,3]. Pulmonary surfactant has also been shown to accelerate the diffusion of oxygen through the air–liquid interface [4].

As air is inhaled through the nose or the mouth, it first travels through the conducting zone of the respiratory system consisting of the pharynx, larynx, trachea, and the bronchi and bronchioles in the lungs (see Fig. 2.1). After many steps of bifurcation into smaller and smaller branches of the respiratory tree, the air ends up into the respiratory zone of the lungs. Alveoli are located at the end of the respiratory tree. They are the main gas exchange zone in the lungs, where the oxygen in the air is passed through the pulmonary surfactant covering the surface of the thin water phase, which in turn covers the respiratory epithelium. Together with the surfactant specific proteins, the pulmonary surfactant has a biologically unique phospholipid composition that makes it able to sustain high surface pressures without collapsing [2].

Pulmonary surfactant protein B (SP-B) is an essential part of the pulmonary surfactant. It participates in the interfacial adsorption of phospholipids, film stability, and surfactant re-spreading [3]. SP-B is a 79-residue polypeptide that forms a covalent homodimer. Due to the high hydrophobicity and difficulty of extraction, there has been no model for the structure or knowledge of the mechanism of function of SP-B [5]. Previous structural models for monomers and dimers of SP-B have been based on homology modeling with different saposin-like proteins (SAPLIPs) [5], but they have not provided unambiguous models that would explain all experimental results [6]. In this study, we have refined a recently published model for the new oligomeric 3D structure of SP-B [6].

Lung immaturity caused by premature birth or mutations in the genes coding for SP-B leads to a condition called respiratory distress syndrome (RDS) [7,8], which is one of the leading causes of neonatal deaths in the world [9]. RDS can be effectively

treated with surfactant replacement therapy (SRT), but its relatively high cost limits its use in low-income countries. Common clinical surfactants used for the treatment of SRT are derived from animal sources [10]. The development of a more affordable treatment based on synthetic surfactants would therefore be of interest. SP-B or its structurally analogous synthetic counterparts like Mini-B [11], have been shown to have a key role in the functionality of clinical surfactants [10].

Understanding the molecular mechanism of function of SP-B could provide vital information of this essential protein and give us ways to develop new drugs and treatments for conditions like RDS or acute respiratory distress syndrome (ARDS) in adults. The possible indirect involvement of SP-B in the transfer of oxygen across the pulmonary surfactant [4] could change our current view on the details of gas exchange in the alveoli. Traditional experimental methods can not achieve the molecular level of detail needed for the study of nanometer scale of size systems or reactions that occur within nanosecond timescales.

In this study, we utilized coarse-grained molecular dynamics (MD) simulations to study the molecular mechanism of function of our new model for the SP-B hexamer of dimers. MD simulations enable atomistic or near atomistic level of detail of biological systems, and insight to phenomena occurring therein. We concentrated on the specific lipid interactions of the protein to find preferential lipid binding conformations for SP-B. We also compared different versions of the SP-B dimer to the SP-B hexamer of dimers to validate the possibility of the higher oligomer structures, and see how their possible molecular mechanism of function differs from each other. We built a total of 9×24 lipid self-assembly (SA) systems starting from a random initial configuration with three lipid compositions, and 48 monolayer systems with a lipid composition resembling the physiological structure of pulmonary surfactant at the air–liquid interface.

This thesis is divided into seven chapters. Chapter 2 focuses on the biological background relevant to understanding the human respiratory system and SP-B. The general concepts of the classical molecular dynamics simulations and the coarse-grained MARTINI force field used in this study are discussed in chapter 3. Chapter 4 gives a short introduction to protein structure prediction methods that can be used to model the structures of novel proteins. Simulation models and analysis methods used in this study are introduced in chapter 5. Chapter 6 covers the most relevant results. Finally, the conclusions in chapter 7 summarize the key results and future prospects.

2. BIOLOGICAL BACKGROUND

In this chapter the biological background relevant to this study is discussed. First the overview of the human respiratory system describes the general structural and functional features characteristic to the mammalian lung. Then a short insight to the main structures making up the pulmonary surfactant is given. The chapter is finalized with a discussion about the most important surfactant protein, and the emphasis of this study, the pulmonary surfactant protein B.

2.1 Overview of the Human Respiratory System

The cardiovascular and respiratory systems are responsible for the supply of necessary oxygen (O_2) for aerobic cell respiration, as well as the disposal of waste carbon dioxide (CO_2) from the body [12, 13]. The cardiovascular system consists of the heart, blood, and the blood vessels while the respiratory system consists of the upper and the lower respiratory tracts (see Fig. 2.1). The continuous transport of O_2 and CO_2 between the lungs and the tissues is carried out by the cardiovascular system which pumps the oxygenated blood from the lungs to the rest of the body. In addition to the exchange of gases in the lungs the respiratory system has an important role in the immune system, heat exchange, fluid balance, and pH control of the whole body [14].

The human respiratory system can be divided into the conducting zone and the respiratory zone based the function of the structures. The main function of the conducting zone is to humidify, warm, and filter the inhaled air before it reaches the respiratory zone that is responsible for the gas exchange [15]. Pretreatment of the inhaled air before the alveoli prevents infections, the transport of harmful substances and particles into the respiratory zone, and the drying of the gas exchange interface. Inhaled air is relatively dry and cool, while exhaled air is heated up to the body temperature and saturated with water vapor [14]. The most important role of the lungs is the exchange of gases between the air and the transporter liquid (blood), which takes place in the respiratory zone.

Air is inhaled through the nose or the mouth, from where it travels through the pharynx, larynx, and the trachea (windpipe). The trachea divides into two primary

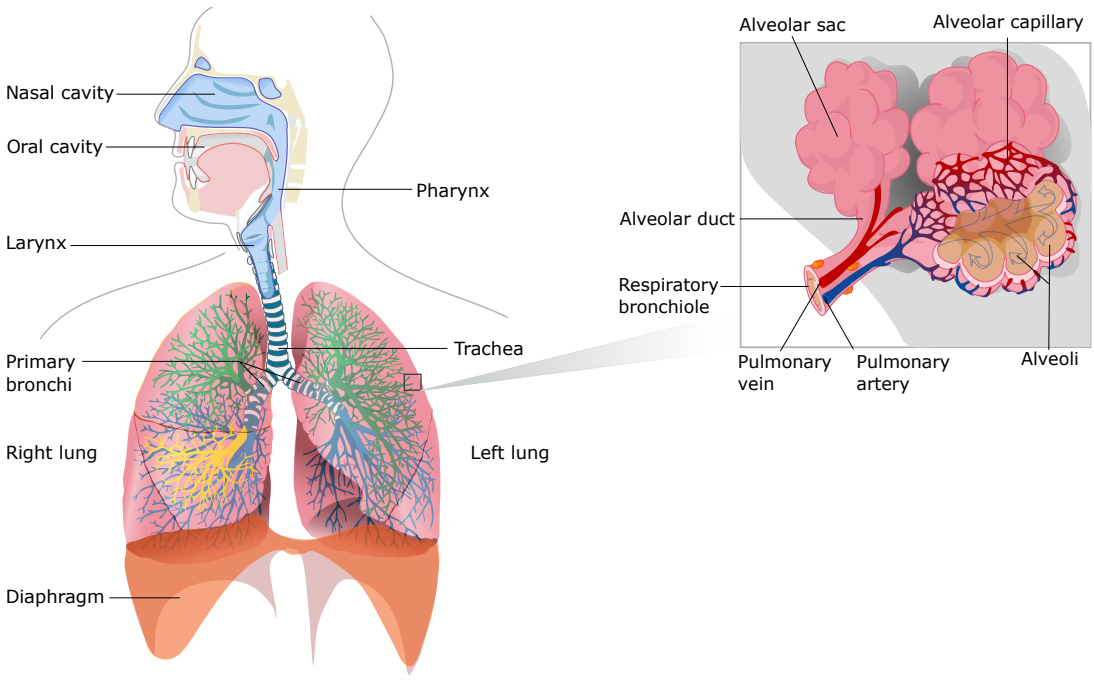


Figure 2.1 Structure of the human lungs. Left: the conducting zone of the respiratory system. Right: a close-up of the respiratory zone including a respiratory bronchiole, alveolar ducts, and alveoli. Modified from [16].

bronchi that continue to bifurcate systematically into smaller and smaller bronchi and bronchioles [14]. After many steps of bifurcation, at the end of the conducting zone the air reaches the terminal bronchioles that divide into respiratory bronchioles that finally lead to the alveolar ducts and the main gas exchange zone in the alveoli (see Fig. 2.1) [13, 14]. The process of gas flow between the atmosphere and the lung alveoli is called pulmonary ventilation [12]. In inhalation air is drawn into the respiratory zone by expanding the volume of the lungs and creating a negative pressure in respect to the atmospheric pressure. This can be achieved by contracting the muscles of inhalation, the diaphragm and external intercostal muscles between the ribs [13]. Exhalation is mostly passive in normal stress and occurs by relaxing of the respiratory muscles and elastic fibers in the lungs.

Efficient gas exchange in the alveoli requires a large air–liquid interface, and the total surface area of the approximately 500 million alveoli in the adult human lung can be up to $80\,\mathrm{m}^2$ [15, 17]. The volume of an alveolus is $\sim 4.2 \times 10^6\,\mu\mathrm{m}^3$ ($\sim 200\,\mu\mathrm{m}$ in diameter) and its limited by a $0.4\ \mu\mathrm{m}$ thick respiratory membrane wall [17]. The walls of the alveoli are composed of a single, thin layer of two types of alveolar

epithelial cells (see Fig. 2.2) [12]. The thin type I alveolar cells make up 95% of the wall and are the main site for gas transfer. The rest of the wall consists mostly of type II alveolar cells [18]. Their important function is to produce and secrete a thin layer of lipids and proteins called the pulmonary surfactant that covers the fluid lining inside the alveoli and prevents the lungs from collapsing at the end of expiration by lowering the surface tension at the interface [2]. The alveoli also contain alveolar macrophages that destroy pathogens that reach the respiratory zone, and participate in the recycling of used pulmonary surfactant with type II alveolar cells [19].

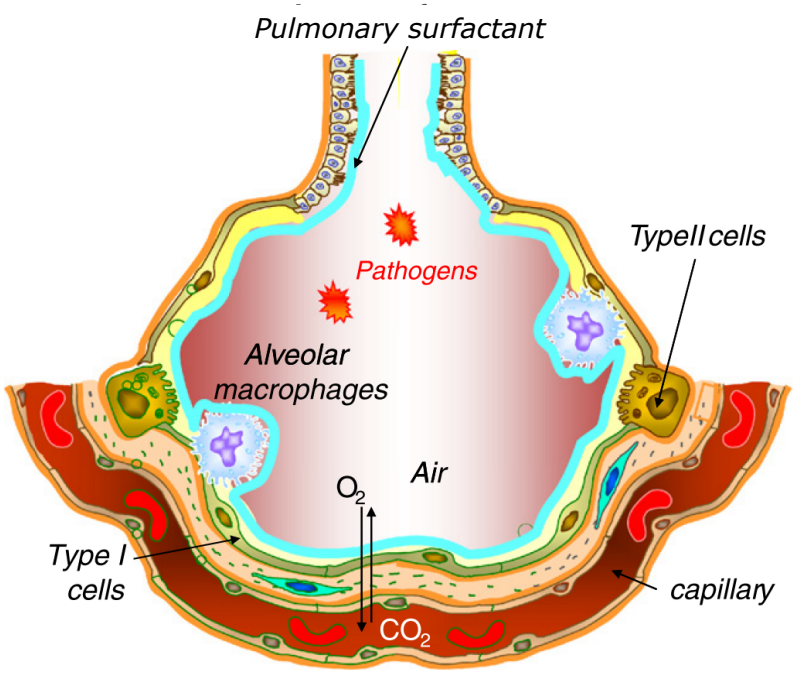


Figure 2.2 Gas exchange at the alveoli occurs through the respiratory epithelium. The walls of the alveoli is covered by type I alveolar cells (95%) and type II alveolar cells (5%). Type II cells secrete surface-active lung surfactant that covers the air–liquid interface. Alveolar macrophages participate in innate defense against pathogens and in the recycling of the surfactant. Modified from [20].

To oxygenate the blood, O_2 has to diffuse from the alveolar cavity through the respiratory membrane, into the blood in the pulmonary capillaries that cover about 70% of the outer alveolar surface area, while CO_2 mostly dissolved in the blood diffuses into the opposite direction [12]. Dissolved CO_2 is an essential part of blood, and it is part of the carbonate buffer system that controls the pH of blood [21]. The gas exchange at the respiratory zone is driven by the difference in the partial pressures of the gases in the alveolar air and the capillary blood [13]. As a non-polar molecule, O_2 has a low solubility in water. Without an efficient O_2 carrier

in the blood, the transfer of dissolved oxygen from the alveoli to the blood could not support the need of O_2 for aerobic respiration [12, 15]. Hemoglobin is a protein in the oxygen carrying red blood cells involved in the process of O_2 uptake in the capillaries in alveoli, and O_2 transport from the respiratory organs to the rest of the body.

2.2 Pulmonary Surfactant

Pulmonary surfactant is a lipid–protein complex produced and secreted onto the respiratory surface by type II alveolar cells [22]. It covers the whole air–liquid interface of the alveoli, and some pulmonary surfactant can be found also in other parts of the respiratory zone, which has been considered to be overflown from the alveoli [2, 23]. The main function of the pulmonary surfactant is to reduce the surface tension of alveoli from approximately 70 mN/m measured at the interface of pure water and air at physiological temperature, to close to 0 mN/m in the compressed state of the surfactant, thus preventing the alveoli from collapsing during exhalation [24]. Without the surfactant, the difference in the internal pressures between different sizes of interconnected air bubbles that are basically the alveoli would cause the collapse of the smaller bubbles due to the air flowing from the smaller bubbles with higher internal pressures into the larger bubbles [14]. The surface tension of the pulmonary surfactant at equilibrium, without compression is 25 mN/m [24]. The surfactant also lowers the force needed to be supplied by the diaphragm during inhalation [14], and significantly accelerates the diffusion of O_2 through the air–liquid interface at the surface of the alveoli [4]. In addition to the essential biophysical advantages of pulmonary surfactant, it also participates in innate defense mechanisms [25].

The pulmonary surfactant consists of about 90% of lipids and 10% of surfactant associated proteins (SPs) by mass (see Fig. 2.4) [24, 26]. The surfactant is assembled, stored, and secreted by type II alveolar cells in organelles called lamellar bodies (LBs) in a form of tightly packed membrane structures (see Fig. 2.3) [2]. LBs are secreted across the type II alveolar cell wall via exocytosis to the alveolar hypophase where they unravel and form tubular myelin (TM) and other smaller intermediate lipid–protein structures before associating with the monolayer at the air–liquid interface [2, 27]. TMs are not obligatory for the delivery and exchange of lipids at the pulmonary surfactant and thus, LBs are believed to be the primary form of lipid transport in the alveolar fluid [22]. The monolayer is at the air–water interface, connected to a network of membranes between the interfacial film and other surface-associated structures that act as reservoirs of surface active molecules [26, 28]. The

surfactant abilities of the interfacial monolayer are a consequence of the amphipathic nature of phospholipids.

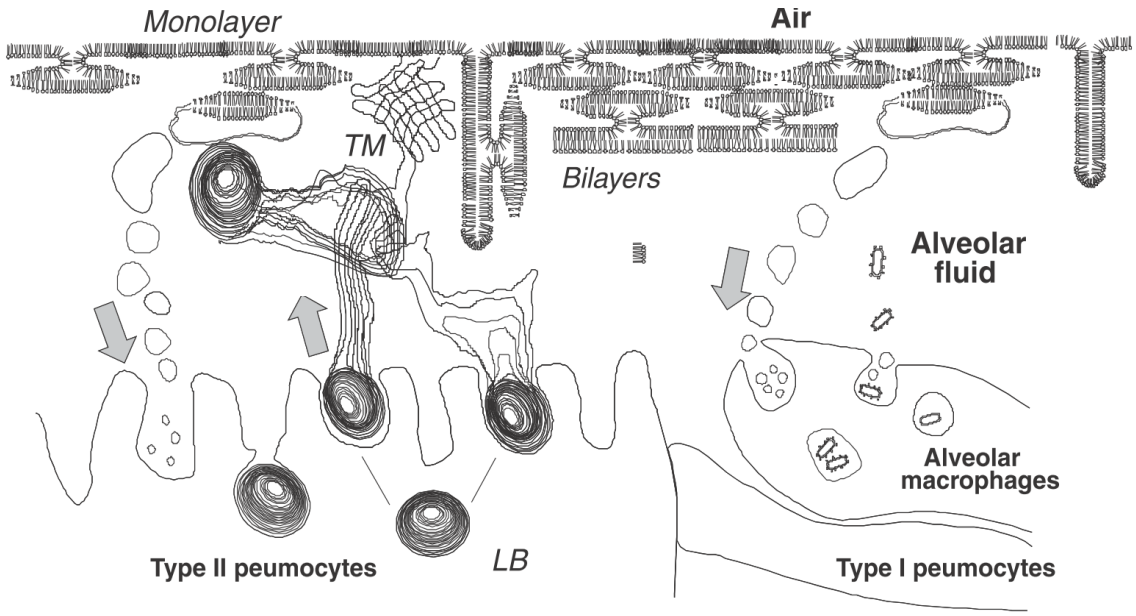


Figure 2.3 Pulmonary surfactant structures in the alveoli. The pulmonary surfactant is synthesized and packed into lamellar bodies (LB) in the type II alveolar cells (pneumocytes), and secreted into the alveolar fluid. The LBs unravel in the hypophase and form multilamellar vesicles, tubular myelin (TM), and other bilayer structures. The adsorption of phospholipids is mediated by the hydrophobic surfactant proteins SP-B and SP-C at the air–liquid interface, where the surfactant consisting of the phospholipids and the proteins forms a surface active monolayer. The surfactant is recycled by alveolar macrophages and type II alveolar cells. Modified from [20].

2.2.1 Lipids in the Pulmonary Surfactant

The pulmonary surfactant has a unique lipid composition compared to most membranes in mammalian tissues (see Fig. 2.4) [19]. Different lipids comprise approximately 90% of the pulmonary surfactant by mass and they are mainly responsible for forming the surface active monolayer films at the air–liquid interface of the alveoli [19, 24, 26]. By mass the pulmonary surfactant consists of 80% of phospholipids and about 10% of neutral lipids, most of which is cholesterol (CHOL) [2]. Zwitterionic phosphaditylcholines (PCs) comprise ∼65% of pulmonary surfactant by mass, over half of which is dipalmitoylphosphatidylcholine (DPPC, 16:0/16:0) [24]. Anionic phospholipids, like phosphatidylglycerol (PG) and phosphatidylinositol (PI)

comprise ∼10–15% of surfactant by mass [2]. Phosphatidylethanolamine (PE), sphingomyelin (SM), and phosphatidylserine (PS) appear in minor amounts in extracted pulmonary surfactant but they are considered to originate from other sources [19,29]. Structures of the most important pulmonary surfactant lipids are shown in Fig. 2.5.

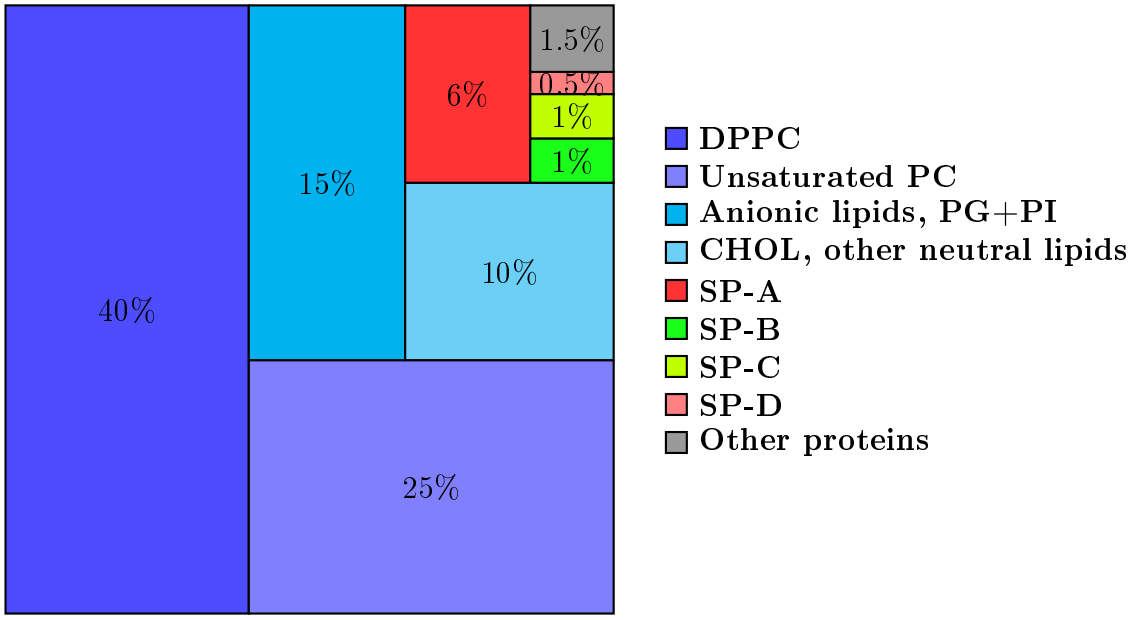


Figure 2.4 Composition of pulmonary surfactant. By mass 90% of pulmonary surfactant consists of lipids and 10% of surfactant proteins. DPPC is the most abundant phospholipid, and it is mainly responsible for the surface active properties of the surfactant. Modified from [1].

The amount of DPPC and PG lipids is unusually high in pulmonary surfactant compared to other tissues [19]. Together with cholesterol these three lipid types are responsible for most of the biophysical activity of the pulmonary surfactant [23]. The ratio of disaturated and unsaturated phospholipids has been observed to be constant (∼1:1) among adult mammalian lung surfactants [19, 23], DPPC being responsible for the disaturated phospholipid fraction in the pulmonary surfactant [2]. Anionic surfactant lipids like PG are important for the function and binding of the hydrophobic surfactant proteins [24, 30], and cholesterol affects the fluidity of the pulmonary surfactant at physiological temperatures [31]. Equal amount of saturated and unsaturated lipids is required for the stability of surfactant layers at high compression rates while at the same time having enough fluidity and dynamics [29].

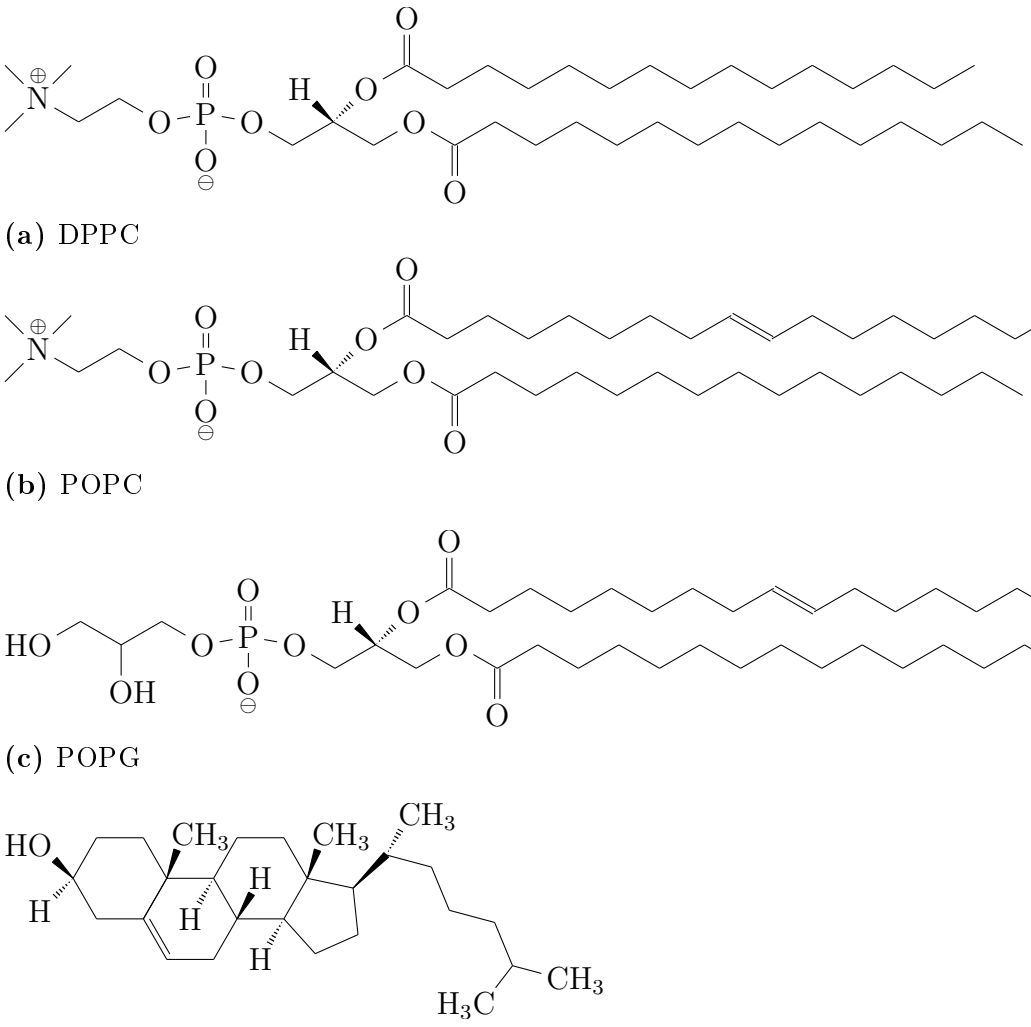
DPPC is considered to be the most important lipid in the surfactant as it is essential for lowering the surface tension of the surfactant close to 0 mN/m during compression at physiological temperatures [32]. This low surface tension prevents

the alveoli from collapsing. The disaturated acyl chains of DPPC can be packed to a very high density at the air–liquid interface, thus reducing the surface tension required to stabilize the interface at the end of expiration [32]. DPPC has a main transition temperature (T_m) of 41°C in bilayers, which is higher than the physiological temperature, thus pure DPPC membranes in the alveolar hypophase would not be fully fluid at the body temperature of 37°C [33]. This would inevitably hinder the spreading of surfactant throughout the compression–expansion cycles [34]. The unsaturated phospholipids and cholesterol lower the T_m of pulmonary surfactant membranes below the values of pure DPPC membranes and make them fluid at physiological temperatures [31]. The most common unsaturated phospholipid and the second most common PC in pulmonary surfactant is palmitoyloleoylphoshatidylcholine (POPC, 16:0/18:1), with T_m of -3°C [35]. Most of the surfactant anionic lipids are also unsaturated [20]. The unsaturated phospholipids lower the T_m of the pulmonary surfactant and together with the saturated lipids, provide the structural scaffold for the connected membrane structures that make up the pulmonary surfactant system [20].

Phosphatidylglycerol (POPG, 16:0/18:1) and other anionic phospholipids have been suggested to be important for the proper function and binding of the two hydrophobic pulmononary surfactant proteins (SP-B and SP-C) to surfactant membranes and monolayers [24, 30]. Both SP-B and SP-C have positively charged residues that could interact with the anionic lipids [3]. PG phospholipids are one of the minimal components required for the construction of TM in the type II alveolar cells along with DPPC, SP-A, SP-B and Ca^+ [26], which also implies specific interaction between PG and SP-B. Electrostatic interactions with the hydrophobic surfactant proteins would justify unusually high concentrations of anionic lipids (PG and PI) in the pulmonary surfactant compared to other tissues [3]. POPG is also an indicator of lung maturity in neonates, concentration of which can be tested from the amniotic fluid before delivery [19].

Cholesterol in general affects the packing of lipid membranes and modulates their thermodynamic properties, which can also be seen in pulmonary surfactant [1]. Two distinctive fluid phases, liquid ordered (L_o) and liquid disordered (L_d) phase, can be seen in native pulmonary surfactant preparations at physiological temperatures [20, 31]. Without cholesterol and the hydrophobic surfactant proteins only a solid and a fluid phase would coexist [20]. The L_o and L_d phases have small differences in their lipid composition. The L_o phase is enriched in DPPC and cholesterol [36], which enhances the fluidity of the L_o phase at physiological temperatures

by decreasing the packing of the lipids by intercalating between the phospholipid molecules [37]. Without cholesterol, the L_o phase would be in a more condensed and ordered but less fluidic gel (L_β) phase below the T_m [38,39]. On the other hand, the L_d phase transforms to more ordered L_o in the presence of cholesterol. The L_d phase contains a higher proportion of unsaturated lipids, which makes it more fluid at lower temperatures compared to the L_o phase. By increasing the fluidity of the membrane, cholesterol also limits the minimum surface tension attainable during compression of the film [20].



(a) DPPC

(b) POPC

(c) POPG

(d) Cholesterol

Figure 2.5 Structures of the main lipids in the pulmonary surfactant (a) DPPC, (b) POPC, (c) POPG, and (d) cholesterol.

Once reaching the air–liquid interface, the pulmonary surfactant organizes into

ordered and/or disordered phases, the proportions of which depend on the lipid composition, temperature, and compression state of the film [29]. These domains are analogous to the L_o and L_d phases in surfactant bilayers. The lipids arrange into a liquid condensed (L_c) phase at low molecular areas at high surface pressures, and into a liquid expanded (L_e) phase at high molecular areas and low surface pressures. Between these two there is a coexistence L_c/L_e phase where domains of both phases can be observed. When surface pressure is gradually applied to the film during compression, the L_e phase transitions to the L_c/L_e phase, and eventually a phase transition from L_c/L_e to L_c occurs [1]. If the surfactant is further compressed after the limit of surface pressure has been exceeded, the monolayer collapses and a 2D to 3D transition of surfactant monolayers to membrane-associated disk-like structures occurs (see Fig. 2.3) [40]. Cholesterol enhances the adsorption and respreading of surfactant lipids from the collapsed phases to the monolayer [41].

The surfactant phospholipids are directly associated with the surface tension reduction of the pulmonary surfactant, however none of the phospholipids alone or in conjunction with each other can maintain the function of the surfactant [3]. The tight packing of DPPC upon compression is responsible for the most of the surface tension reducing effect of the surfactant [32, 42], but the rate of adsorption of pure DPPC vesicles from the hypophase to the air–liquid interface is extremely slow [42, 43]. The unsaturated and anionic surfactant lipids support faster adsorption rates, but can not achieve low enough surface tension values without irreversible collapse and squeeze-out of the lipids [44]. The presence of hydrophobic surfactant proteins, discussed more in the next section, surpasses the shortcomings of surfactants consisting of only phospholipids [3]. These proteins are involved in both surface tension reduction and reversible phospholipid transfer at the pulmonary surfactant. Moreover they accelerate the adsorption of the lipids almost to the maximum extent possible [45, 46].

2.2.2 Pulmonary Surfactant Proteins

Pulmonary surfactant contains 8–10% surfactant specific proteins by mass [26]. There are four major surfactant proteins, the hydrophilic SP-A and SP-D, and the hydrophobic SP-B and SP-C [47]. All four surfactant proteins have been detected also from the human tear fluid, the ocular surface, and tissues of the lacrimal apparatus [48, 49], where they might have similar functions as in the pulmonary surfactant. This is not surprising as the alveoli and the eye are the only two sites in the human body with an air–water interface. SP-A and SP-D belong to the col-

lectin family of proteins and make up 5–6% and 0.5% of the surfactant by mass, respectively [3]. Both hydrophilic SPs participate in innate defense mechanisms by binding specific ligands on the surface of pathogens including bacteria, viruses, and fungi [50]. SP-A is also associated with surfactant phospholipids and is needed for the formation of TM in the pulmonary surfactant hypophase [51]. Regardless of their functions in the pulmonary surfactant neither is essential for the surface active functions of the surfactant and lack of these proteins does not cause any respiratory dysfunction [26, 52].

SP-B and SP-C are very hydrophobic and always associated with surfactant membranes [53]. These proteins are essential for the proper surfactant activity as they promote efficient interfacial adsorption, film stability, and re-spreading of the lipids during the course of the breathing cycle [42, 54, 55]. Lack or deficiency of SP-B or SP-C are common causes of different severe respiratory pathologies [7, 56]. Proper expression of SP-B is essential for the viability of a newborn and deficiencies in the expression of the functional protein leads to lethal respiratory distress syndrome shortly after birth [57–59]. SP-C deficiency is associated with the development of chronic lung disease in humans but the protein is not as critical for surfactant activity and breathing as SP-B [60].

SP-C is a small 4.5 kDa and 35 residue transmembrane peptide. It has a primarily α-helical secondary structure with an N-terminal with two palmitoylated cysteine residues and a positive net charge of $+3\,\mathrm{e}$ at pH 7. The C-terminal α-helix is enriched in aliphatic residues [61,62]. The positive charge at the N-terminus causes a preferential interaction with anionic surfactant lipids [63]. The palmitoylated chains are required for the tight association of the N-terminus with bilayers and interfacial films, which prevents its exclusion from the surfactant interface during compression–expansion cycles [64]. SP-C functions by promoting and stabilizing membrane–membrane and membrane–monolayer contacts [46, 65] and SP-C has been shown to perturb lipid packing in membranes [66]. One of the most important functions for the protein could thus be to promote lipid transfer between bilayers or between bilayers and monolayers independently, or in conjunction with SP-B [55, 65, 67].

Hereafter this thesis will concentrate on the structure and function of the most important surfactant protein SP-B, and further discussion on the other surfactant proteins is thus omitted unless related to the function of SP-B.

2.2.3 Structure and Function of Surfactant Protein B

Surfactant protein B belongs to the diverse lipid-interacting family of saposin like proteins (SAPLIPs) [5]. The SAPLIP family contains over 200 proteins in organisms ranging from the most primitive eukaryotes, like amoebozoans, to all mammals [68]. Thus, the SAPLIP family is assumed to have originated very early in the eukaryotic life. SAPLIPs are very diverse in protein sequence, but they share a few key features for the saposin fold: conserved hydrophobic amino acids in the core of the protein, six strictly conserved cysteines that form three stabilizing intramolecular disulphide bridges (see Fig. 2.6), and a high helical content of four or five α-helices in common positions [68]. The human genome codes for a total of 11 proteins with the SAPLIP fold in their structure. Three universal functions have been suggested for the SAPLIPs: membrane binding with local disordering of the lipids, membrane perturbation without permeabilization, and permeabilization by perturbation of the lipids [68], the latter being an assumed function of oligomeric proteins.

SP-B is a 8.7 kDa and 79-residue polypeptide that contains a high proportion of hydrophobic amino acids (approximately 52%) and adopts a mainly α-helical (30–45%) secondary structure [53,69,70]. SP-B contains the six conserved cysteines of SAPLIPs, forming three intramolecular cysteine bridges at positions C8–C77, C11–C71 and C35–C46, and also a SP-B specific seventh cystein C48 that forms a covalent homodimer of 19 kDa (see Fig. 2.6) [69,71]. SP-B is also the only SAPLIP protein that is permanently membrane-associated due to its high hydrophobicity [5]. SP-B monomer has a positive net charge of $+7\,e$, which leads to its observed preferential interaction with anionic surfactant lipids, especially PG [63,72,73]. SP-B has been shown to distribute preferentially into the disordered phase in bilayers and interfacial films [31,74].

SP-B is synthesized in the endoplasmic reticulum of the type II alveolar cells as a 381-residue precursor peptide proSP-B [75]. This precursor is processed post-translationally into the functional protein by proteolytic cleavage in at least three distinctive steps before it is packed into the mature LBs shown in Fig. 2.3 [76–78]. Dimerization and final permanent lipid association of SP-B does not occur until both flanking propeptides are cleaved from the precursor in the last step of SP-B processing in the intermediate composite bodies (CB), which are a precursor form of the LBs [71,78]. Both flanking regions of the precursor peptide contain functional saposin-like structures, but these peptides have not been observed in the alveolar fluid [53,79].

Mutations in the SP-B precursor can interfere with the processing of both SP-B

and SP-C, which indicates important proteolytic and chaperone functions of the propeptides of SP-B [76]. One of the important functions of mature SP-B is to mediate the packing of the surfactant lipids into LBs after the last step of post-translational processing [58,80]. Defects in SP-B expression cause significant reductions in the number and size of LBs secreted by type II alveolar cells [58,80]. Severe mutations or lack of expression of SP-B leads to lethal RDS shortly after birth due to the total lack of pulmonary surfactant at the air–liquid interface [75,81].

Two seemingly contradictory models for the orientation of membrane binding for SP-B have been detected in experiments (see Fig. 2.7) [73]. The first model suggests a superficial binding to the surface of bilayers and surfactant monolayers with the amphipathic helical structures as well as the positively charged residues in the C- and N-terminal regions of SP-B, causing little or no perturbations in the acyl chains of surfactant phospholipids [67,83,84]. The second model suggests much more significant perturbations in acyl chain packing and order, and deeper penetration of SP-B into lipid membranes [63,85,86].

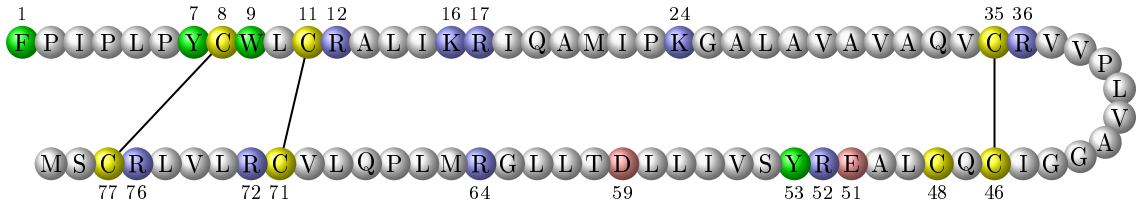
The observed orientation and level of membrane penetration of SP-B is highly dependent on the method used to prepare the samples and study the lipid–protein interactions [43, 73, 87]. Samples prepared with the SP-B in the organic solvent – isolated lipid fraction show deeper penetration of the protein inside bilayers whereas addition of the protein to preformed phospholipid vesicles leads to more superficial binding [43,87]. These two levels of membrane binding could indicate different functional stages of the protein: a superficial inactive membrane associated form, and a functional, more deeply inserted form of SP-B [87]. Altogether more experimental and molecular dynamics studies of the protein support a model of superficial or intermediate binding where the N- and C-terminal regions of the protein play a crucial functional role [73, 88–90]. The importance of the role of the N- and C-termini is further validated by the significant activity of the termini-including SP-B fragment called mini-B [91,92].

A possible model suggested for the interaction of SP-B with lipid membranes includes both electrostatic interactions by the cationic residues mainly in the N- and C-terminal and helical regions of the protein, and hydrophobic contacts by essential membrane binding residues, such as W9 in the N-terminus [3, 93, 94]. Mutation or oxidative modification of W9 has been shown to affect the surface activity and membrane association of SP-B, which could be caused by the improper binding of the protein to the surfactant membranes [93,95]. The aromatic side chain of W9 is believed to anchor SP-B firmly in the correct orientation to the surface of surfactant

```
                    5        10        15        20        25                  30        35        40
SP-B_h   FPIPLPYCWLCRALIKRIQAMIPKG.....ALAVAVAQVCRVVPL
SP-B_b   FPIPIPYCWLCRTLIKRIQAVIPKG.....VLAMTVAQVCHVVPL
SP-B_p   FPIPLPFCWLCRTLIKRIQAVVPKG.....VLLKAVAQVCHVVPL
SAP_A    ....SLPCDICKDVVTAAGDMLKDNATE.EEILVYLEKTCDWLPK
SAP_B    ....GDVCQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGP
SAP_C    ...SDVYCEVCEFLVKEVTKLIDNNKTE.KEILDAFDKMCSKLPK
SAP_D    ...DGGFCEVCKKLVGYLDRNLEKNSTK.QEILAALEKGCSFLPD

                    45        50        55        60        65        70        75
SP-B_h   VA.GGICQCLAERYSVILLDTLLGRM.LPQLVCRLVLRCSM..
SP-B_b   LV.GGICQCLVERYSVILLDTLLGRM.LPQLVCGLVLRCSS..
SP-B_p   PV.GGICQCLAERYIVICLNMLLDRT.LPQLVCGLVLRCSS..
SAP_A    PNMSASCKEIVDSYLPVILDIIKGEMSRPGEVCSALNLCESLQ
SAP_B    GM.ADICKNYISQYSEIAIQMMMH.M.QPKEICALVGFCDE..
SAP_C    SL.SEECQEVVDTYGSSILSILLEEV.SPELVCSMLHLCSG..
SAP_D    PY.QKQCDQFVAEYEPVLIEILVEVM.DPSFVCLKIGACPSAH
```

(a) Sequence alignment



(b) Primary structure of SP-B

Figure 2.6 (a) Sequence alignment of SP-B and other selected SAPLIPs using CLUSTAL [82]. SP-B_h, SP-B_b, and SP-B_p are the human, bovine, and porcine SP-B sequences, respectively, and SAP_A, SAP_B, SAP_C, and SAP_D are the four human saposin proteins, from A to D, respectively. Numbering of residues based on the human SP-B sequence. (b) A representation of some of the most important features of the primary structure of SP-B, especially the intra-chain cysteine bonds (yellow), positive residues (blue), negative residues (red), aromatic (green), and other conserved hydrophobic residues (brown in the sequence alignment) important for membrane binding and protein function.

phospholipid layers [96,97]. The positioning of these membrane interacting residues in many of the proposed SAPLIP homology models for SP-B could imply a specific orientation for the SP-B in surfactant membranes [5].

Functionally SP-B is indispensable for the three most important biophysical processes needed for proper pulmonary surfactant behavior during normal compression–expansion breathing cycles. These include: very rapid adsorption of surfactant

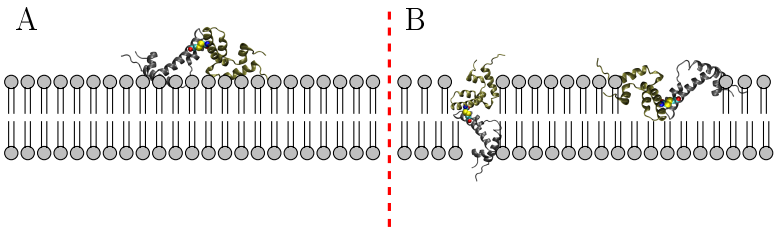


Figure 2.7 Suggested orientations for the membrane binding of SP-B dimer. A: The first model, SP-B dimer exhibits a superficial interaction with the membrane surface. B: The second model, SP-B attains a transmembrane orientation or inserts deeper into one leaflet of the membrane [73].

material from the aqueous hypophase into the air–liquid interface, good compressibility while still upholding the low surface tension and stability during expiration, and very efficient re-extension and spreading of the surfactant during inspiration [3]. These functions are all related to SP-B's ability to cause lipid acyl chain perturbations [63, 85], lipid reordering and domain formation [98], pore formation [99, 100], leakage of aqueous content from lipid vesicles [93, 101], reduction of pulmonary surfactant surface tension [54], the exchange of lipids between surfactant layers and reservoirs [42, 93, 102], and total fusion of lipid membrane structures [103, 104], only to name a few.

SP-B has been shown to mediate the interfacial adsorption and transfer of surface active material from the aqueous hypophase into the surfactant to form the surface active film, alone and in conjunction with SP-C [55, 105]. SP-B is the main protein responsible for promoting surface activity and stability in the surfactant [105]. Together with SP-B, SP-C was observed to accelerate the effect of initial surface film formation and re-extension after compression–expansion cycles compared to only SP-B in lipid extracts, which could indicate a synergetic effect of the two hydrophobic surfactant proteins in the lipid exchange between surfactant layers and reservoirs [105]. Both hydrophobic surfactant proteins have been observed to be present at the surface of the LBs that are responsible for the transport of lipids to the air–liquid interface [55]. The adsorption and spreading of lipids from the LBs to the interface is prevented in the presence of SP-B inhibitors. Thus SP-B functions as a gate for lipid transfer during adsorption to the air–liquid interface [55]. A conformational change of SP-B from an inactive closed to an active open form is believed to occur when the protein reaches the air–liquid interface, which could be related to the depth of insertion of the protein into membranes at different active states [6, 55].

The addition of SP-B into surfactant lipid extract improves the surface activity, and higher surface pressures without collapse can be achieved compared to pure lipid fraction [32, 105]. The presence of SP-B is needed for the proper compressibility and low surface tension of the pulmonary surfactant during compression–expansion cycles [32]. SP-B disrupts the L_c lipid microdomains in the coexistence L_c/L_e phase and creates a wider plateau range in surface pressure area isotherms ($\Pi - A$) *i.e.* it creates more disorder in the surfactant while optimizing the behavior in the physiological surface pressure range [42, 106]. Higher surface pressures cause a reversible 2D–3D structural transition in the surfactant in which part of the monolayer is squeezed out of the surface [3, 105]. SP-B has been suggested to contribute to the stabilization of this transition during both compression and expansion phases of the cycle and thus contribute also to the efficient re-extension and spreading of the surfactant during inhalation [3, 105].

The traditional model for O_2 permeation in lipid membranes, and especially in the respiratory epithelium, assumes that the concentration gradient over the membrane, or between the alveoli and the capillaries is the driving force for gas transfer [12]. As a small molecule O_2 has been assumed to diffuse freely over the phospholipid membrane based on the higher permeability compared to pure water surfaces [107, 108]. On the other hand, other studies have shown that pure lipid monolayers and surfactant lipid extracts, although significantly accelerating O_2 diffusion at the air–liquid interface, are not alone enough to maintain high-rate oxygen transfer [4, 109]. The permeation of O_2 has been found to be three to four orders of magnitude lower than expected by the traditional model, and thus the phospholipid membrane has been considered a barrier to free O_2 diffusion [109]. Based on their results with O_2 diffusion through phospholipid membranes, Ivanov et al. [109] suggested that protein-mediated O_2–channels through the membranes of aerobic organisms could be needed to further improve the flow of O_2.

The hydrophobic surfactant proteins in the pulmonary surfactant improve the rate of O_2 diffusion compared to a water interface, pure DPPC membranes, or surfactant extracts devoid of SP-B and SP-C [4]. The pulmonary surfactant is necessary for its surface tension lowering function in the air–liquid interface, but it could be also closely related to the high transfer rate of O_2 in the lungs. SP-B and SP-C promote connections and lipid transfer between the surfactant and the interconnected bilayer structures below the surface [3, 22]. Oxygen partitions preferably into the hydrophobic core region of the phospholipids [110]. The molecular connection of the surfactant lipid structures mediated by SP-B and SP-C could allow O_2 to partition

into the tail region and diffuse rapidly through the alveolar hypophase using the connected membrane structures as a scaffolding (see Fig. 2.8) [4].

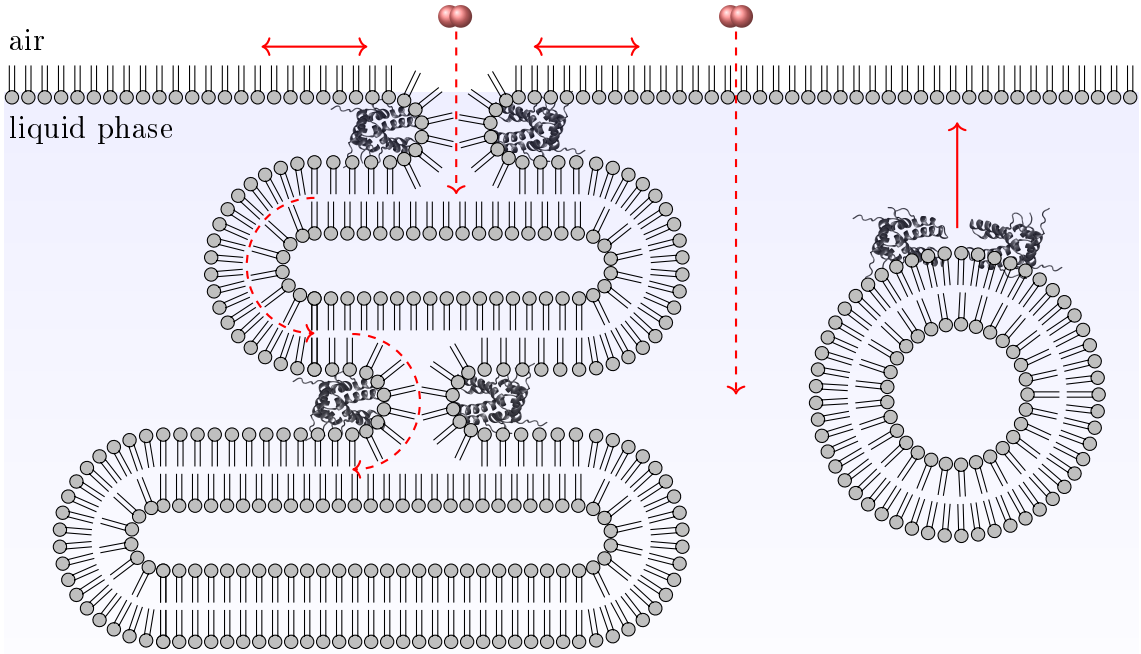


Figure 2.8 Main functions of SP-B in pulmonary surfactant during breathing. SP-B facilitates surfactant adsorption (solid red arrow on the right) and lipid spreading at the air–liquid interface, it preserves functional surfactant by maintaining low surface tensions, and mediates lipid transfer between the surfactant monolayer and the connected bilayer structures during compression–expansion cycles. Interconnected lipid structures below the surfactant could be associated with improved oxygen permeation needed for efficient breathing (left pathway marked with dashed arrows) compared to oxygen diffusion through the hypophase (right pathway). Modified from [4] and [3].

Despite the essential role of SP-B in the biophysical activity of the pulmonary surfactant system, a structural model and a molecular mechanism for the protein have been lacking. Determination of the 3D-structure of SP-B has been experimentally limited by its small size, the difficulty of extraction due to the extreme hydrophobicity, and the lack of methods to produce recombinant version of SP-B [5, 111]. Thus most protein models used in molecular dynamics studies so far have been based on sequence alignment and homology modeling of the SP-B and other SAPLIPs [112, 113]. There exist higher resolution models for the N- and C-terminal amphipathic α-helical sections of SP-B, which are the basis for a func-

tional 34-residue SP-B fragment called mini-B [11]. Mini-B contains two of the four intramolecular disulfide bonds but lacks the C48 residue needed for covalent dimerization. Mini-B has the same net charge of +7 e as native SP-B, and has been shown to retain most of the activity of native SP-B [92].

SP-B extracted and purified in organic solvents from animal lungs has been observed to exist as a covalently bound homodimer [114]. The dimerization occurs at the last step of post-translational modification, and has been suggested to be related to its permanent membrane association in agreement with other oligomerizing SAPLIPs [5]. SP-B has been shown to have a strong tendency to form larger non-covalent oligomers of dimers depending on the animal source, the purification method, solvents used, and the protein-to-lipid ratio. SP-B trimers have been found in studies with bovine and ovine lung extracts [70, 101, 115], but higher oligomers of up to octamers have been found with better purification techniques and analysis methods [116]. SP-B has a tendency to form oligomers even in the absence of C48-dependent intermolecular disulfide bonds [117]. This could be due to the E51–R52' and R52–E51' ion pairs that are believed to further stabilize the disulfide in the native state.

A recent study showed a significant improvement in the quality of the suggested structural model for SP-B [6]. The group was able to purify porcine SP-B using detergent-solubilized pulmonary surfactant, and visualize large ring-like oligomeric SP-B structures. Two homology models for non-covalent SP-B dimers were built based on analogous structural variants of the human saposin B protein. These so-called open and closed conformations for the SP-B dimers were further used to construct larger oligomeric structures by optimizing the C48–C48' as well as E51–R52' and R52–E51' bond lengths between the dimers in docking calculations and subsequent ranking of the models. This resulted in two highest ranking structures for a pentamer of dimers and a hexamer of dimers (see Fig. 2.9). These novel structural models have been used as a basis for the SP-B models built for this study, as described in section 5.

2.3 Diseases Related to SP-B

According to World Health Organization (WHO) [9, 118], every year approximately 15 million (over 10%) babies are born prematurely (before week 38 of gestation) worldwide, and the number is rising. Over 1 million children die annually because of complications caused by preterm birth [9]. Premature birth is a global problem, but it is more common in low-income countries, in which the mortality of babies born

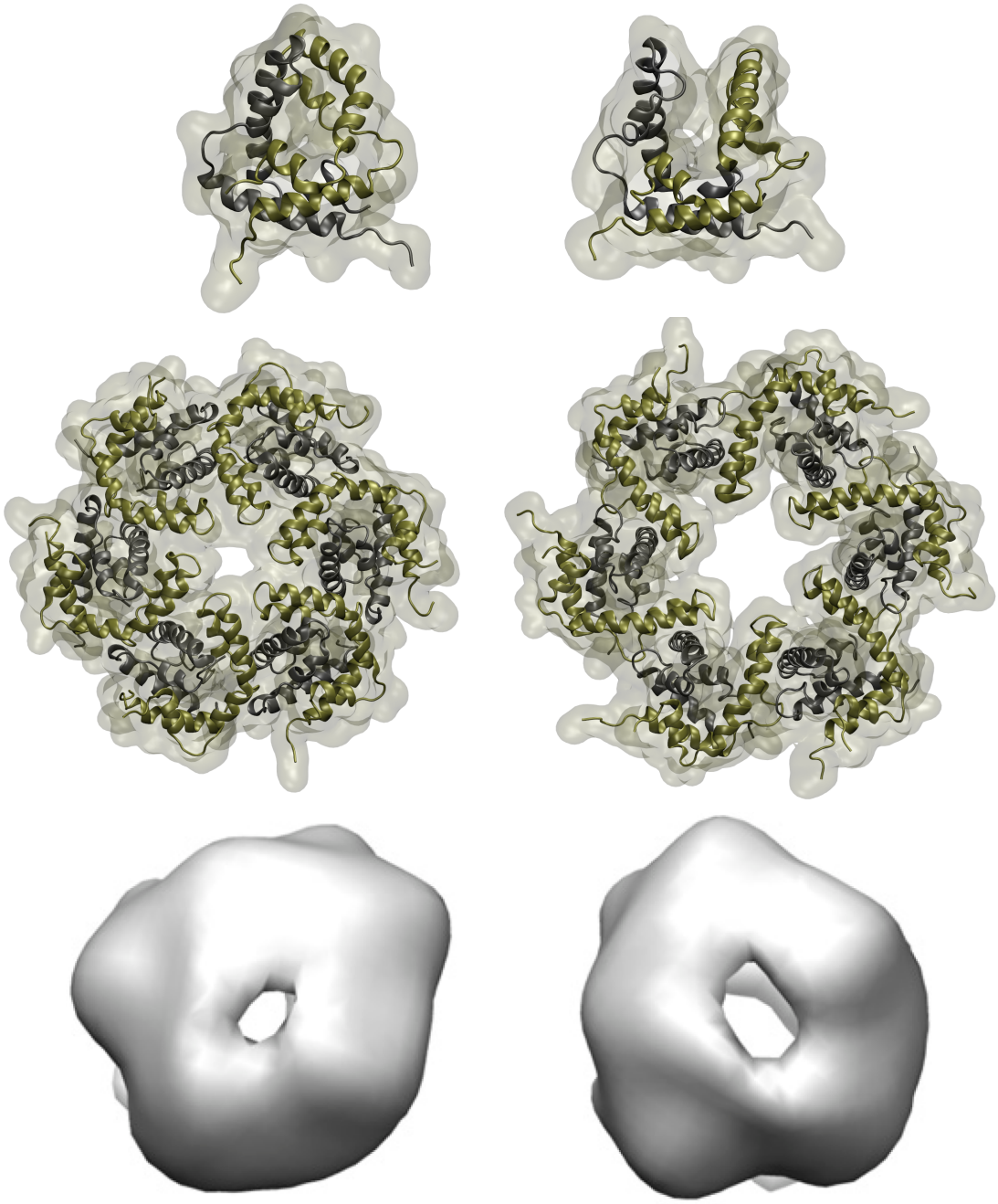


Figure 2.9 A model for the 3D structure of the closed (left) and open (right) conformation of the SP-B dimers, and in the middle closed (left) and open (right) conformation of the SP-B hexamer of dimers. The figures below represent the experimentally obtained front views of the 3D models built from negative stained oligomeric SP-B complexes [6].

before week 32 of gestation is over 50% due to the lack of feasible and cost-effective basic care for infections and breathing difficulties [118]. In addition to neonatal mortality, complications caused by breathing difficulties with preterm children can

result in lifelong disabilities, including learning disabilities and visual and hearing problems [9, 119].

The formation of the gas-exchange interface and the maturation of the lung is controlled by a number of genes during the stages of sacculation (before birth) and alveolarization (after birth) [120, 121]. The most important steps in the pulmonary maturation happen at the end of gestation before week 35 [121], thus majority of preterm babies born below 32 weeks of gestation develop the condition of respiratory distress syndrome (RDS) [121]. RDS is caused by lung immaturity and the deficiency of surfactant lipids in the lungs, and the condition is fatal without correct treatment [1,81,121]. RDS can be treated by administration of a clinical surfactant, in a treatment called surfactant replacement therapy (SRT) [29]. The replacement surfactant reduces the surface tension in the lungs and prevents the immature alveoli from collapsing [122]. Preterm babies of only 25 weeks of gestation can survive with SRT [122], but the treatment is still unattainable in low-income countries due to its high cost [9].

Both synthetic and natural pulmonary surfactants are effective in treating RDS, but the function of therapeutic surfactants has been shown to be highly dependent on SP-B content, or the content of SP-B and SP-C analogues [81,123]. Synthetic SP-B analogs, such as mini-B [11,92,124], are used in replacement surfactants, but they are not as effective as natural SP-B [124]. Mini-B consist of N- and C-terminal fragments of SP-B covalently bound together [11]. The production of functional recombinant SP-B has not been successful [29]. Understanding the molecular mechanism behind the function of SP-B can help in the development of an affordable and effective synthetic replacement surfactant that could save hundreds of thousands of prenates annually in low-income countries by lowering the price of SRT.

Mutations in the gene encoding for SP-B, *SFTPB*, are often fatal in neonates [29, 120, 121] without lung transplantation [81, 120]. The 381 amino acid preproprotein encoded by this gene is expressed and processed to the functional 79 amino acid protein in type II epithelial cells [120]. The SP-B preproprotein is associated with the proteolytic processing of the two hydrophobic surfactant proteins [58, 120]. Mutations in *SFTPB* gene interrupt the post-translational modification of SP-B and SP-C, which leads to the accumulation of partially processed, inactive forms of SP-B and SP-C in the alveoli [120, 121]. Neonates with mutated SP-B show similar symptoms to RDS, but SRT is not sufficient in the treatment of the condition [29]. Less severe mutations in *SFTPB* are associated with chronic lung diseases, similar to those caused by defects in SP-C expression [29, 125].

3. MOLECULAR DYNAMICS

In this thesis novel dimer and oligomer structures of SP-B, their interactions with lipid monolayers and bilayers, and their possible role in oxygen transfer at the air–liquid interface of lung alveoli are studied using classical molecular dynamics (MD) simulations. All simulations were performed using the GROMACS simulation package [126–128]. This chapter provides a short overview of MD simulations, based on the user manual of GROMACS [129], as well as Refs. [130–132].

Molecular modeling and simulation, especially that of large biological polymers such as proteins, lipids, and nucleic acids, have been rapidly growing multidisciplinary fields that have brought together scientists from a multitude of engineering and natural science fields. Biologists, chemists, and physicists are needed to describe the system of interest at an appropriate level, from the cellular level of the system down to the electronic level as well as the underlying forces behind the interaction between separate atoms. Mathematicians, computer scientists, and engineers are needed to formulate and refine algorithms and numerical models for the MD software and to efficiently implement them to a variety of different computational platforms, from a simple desktop computer to a state-of-the-art supercomputer.

The study of single biological molecules like proteins and lipids has been restricted by the size and timescale limitations of modern *in vitro* experimental methods. Often in the study of biomolecules, suitable experiments can be close to impossible to perform due to the nanometer size of the systems of interest and the nanosecond timescales of the reactions and interactions. MD simulations are a useful method in conjunction with experiments to explain the results at a molecular level and to build new models of biological systems. Simulations are often cheaper than experiments and they can be used to test many different sets of experimental conditions before actual laboratory experiments are conducted. Results from MD can thus be used to plan the tests and focus the limited experimental resources more efficiently.

Computational models are always highly-simplified versions of real systems and reactions that occur in reality. The goal of biomolecular modeling and MD is to describe a complex system or reaction occurring in nature with as realistic as possible atomistic or coarse-grained (CG) models that can be used to understand and

predict macroscopic properties of the system [129]. Classical MD simulations used in this thesis can be used to study systems governed by molecular mechanics (MM), or in other words systems that can be studied by using Newton's laws of motion and disregarding possible chemical reactions that might occur. Classical MD does not take into account quantum mechanics (QM), which includes *e.g.* chemical transformations governed by strong forces, such as chemical reactions, enzyme catalysis and photochemical reactions. The exclusion of QM in classical MD simplifies the computational process so that much larger systems and longer time scales can be achieved compared to computational quantum mechanics or the combination of QM and classical MD in mixed QM/MM simulations [130]. In QM/MM simulations part of the system can be treated by taking into account quantum mechanics, while majority of the system is treated as an atomistic MM simulation.

A schematic of the general algorithm for MD simulations is given in Fig. 3.1. MD simulation always starts with the preparation of the system of N particles, relaxation of the system by minimizing the initial forces affecting the system, equilibration simulations with possible restrains on some molecules, until the system is finally ready for actual production simulations.

3.1 Initial Conditions

The initial conditions of an MD simulation system are comprised of the simulation box size, the number, position, and starting velocities of all the particles in the system, their chemical identity, and the possible bonds each particle shares with others. Selecting the initial condition should be done with proper care, as the starting configuration often determines the success or the failure of a simulation. The goal is to build a simulation system that models the properties of a real physical system as realistically as the method enables. The system needs to be defined and its boundaries set so that only relevant features are studied. This might require simplifications in the system by *e.g* coarse-graining or by otherwise omitting unnecessary precision and limiting the size of the system.

Starting velocities of all particles in the system are required after the relevant macromolecules (proteins, lipids, carbohydrates, nucleic acids) and solvent are placed in a simulation box. The shape and size of the simplest cubic simulation box is defined by three basis vectors of the periodic box. Other shapes of the box exist with different number of basis vectors. In some applications the proper selection of the simulation box shape can accelerate the simulation as less solvent is required to fill the smaller volume. If the initial velocities of the atoms at the beginning of a sim-

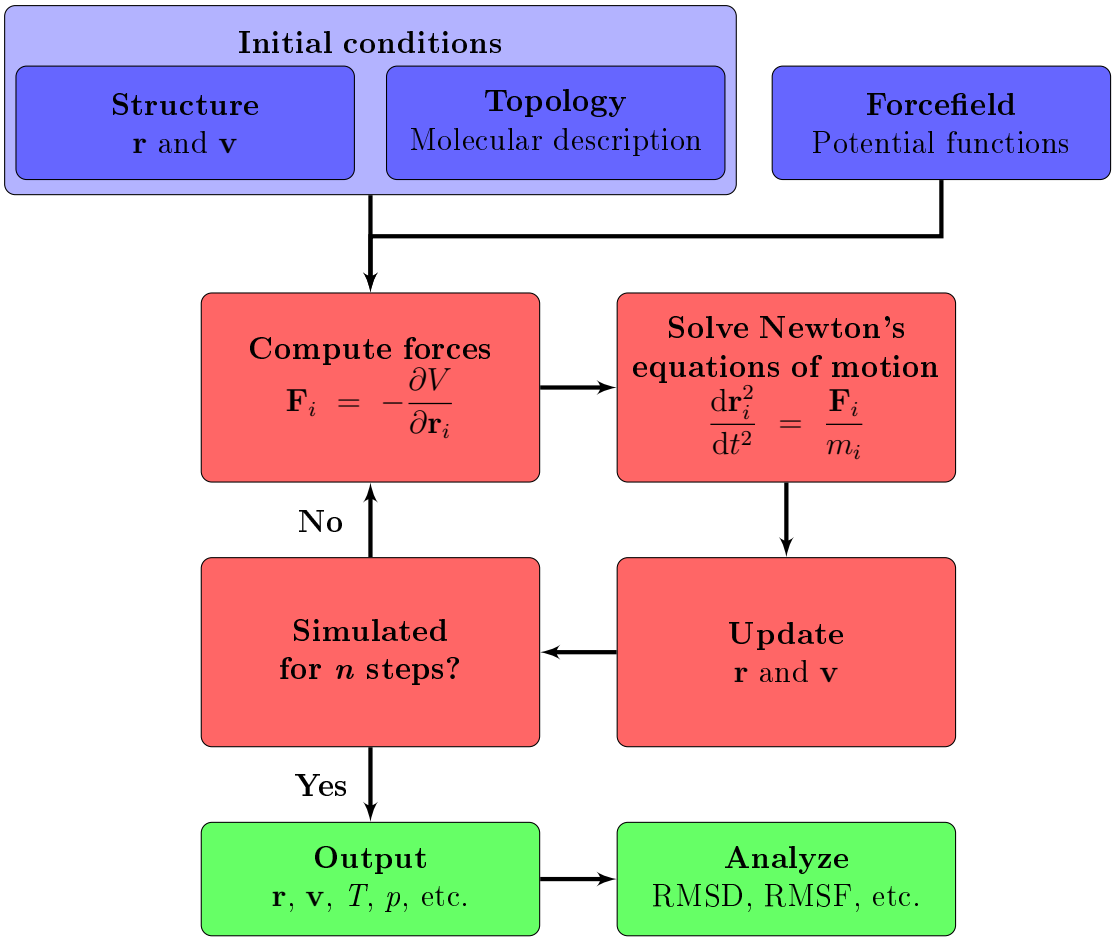


Figure 3.1 A general MD algorithm.

ulation are not known, they can be set by the user or they can be generated from the Maxwell–Boltzmann distribution at a given absolute temperature T:

$$p(v_i) = \sqrt{\frac{m_i}{\pi k_\mathrm{B} T}} \times \exp\left(-\frac{m_i v_i^2}{2 k_\mathrm{B} T}\right), \tag{3.1}$$

where p is the probability density as a function of particle velocity v_i, m_i is the mass of the particle, k_B is the Boltzmann's constant, and T is the temperature. A schematic picture of the velocity distribution generated for particles of three different masses is shown in Fig. 3.2.

MD simulations require a molecular description of the particles to be simulated, which is generally called a topology. The topology of the system contains parameters that describe the physical properties of all the particles and their interactions with each other. The properties of individual particles contain information of their atom

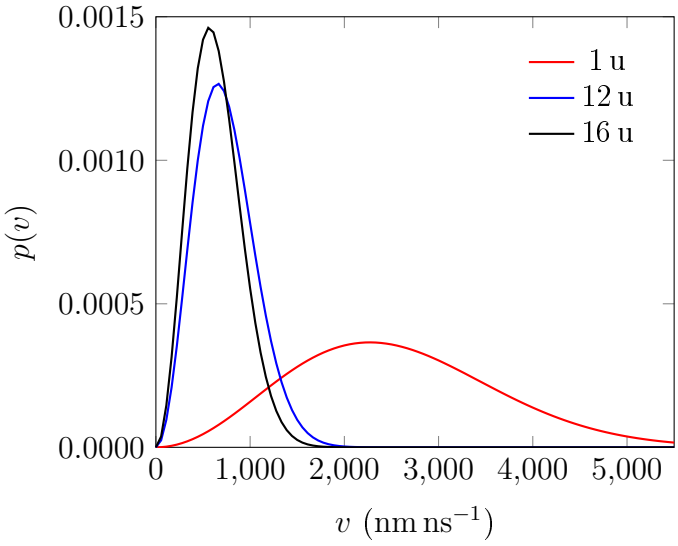


Figure 3.2 Maxwell-Boltzmann velocity distribution calculated at a physiological temperature of 310 K for three different particle masses representing hydrogen (1 u), carbon (12 u), and oxygen (16 u).

type, mass, and charge, while the inter-particle relations contain the bonds, angles and dihedrals between atoms. One must note that in classical MD simulations the topology is a fixed list where all of the bonded and special interactions are defined at the beginning of the simulation and they are not changed during the course of the simulation run. Thus, covalent bonds can not typically be formed or broken in classical MD simulations, and no chemical reactions can occur. In biological applications of MD, one has to consider the protonation state of a protein at a relevant pH as it remains constant during the simulation.

The initial coordinates of atoms in a structure can be attained from an experimental crystal structure, or if that is not available, then from otherwise built theoretical models, or the combination of both. These coordinates do not often represent the minimum potential energy attainable by the system. The total potential energy of the system needs to be minimized before simulation to prevent the system from collapsing or exploding due to the large forces caused by clashing atoms. Energy minimization relaxes the possible tensions or atomic clashes and prepares the simulation system for equilibration simulations. Thermodynamic and structural properties of the simulation system are monitored in the minimization and equilibration steps until stability of the system is achieved. This can require a varying amount of simulation time and precision, but these procedures are necessary to prevent any unusual behavior in the production phase.

3.2 Force Fields

The force field is the description of the interactions between atoms or coarse-grained particles in the simulation system as a set of potential functions and their associated parameters. Force fields are used to calculate the forces acting on the atoms at each step during the simulation. Equation 3.2 gives the total classical potential function which is a combination of the bonded and non-bonded interactions. The bonded terms include interactions of atoms linked by covalent bonds, and the non-bonded terms include electrostatic and van der Waals forces:

$$V_{\mathrm{TOT}} = V_{\mathrm{bonded}} + V_{\mathrm{non\text{-}bonded}}. \tag{3.2}$$

The functional forms of the potential and the way the empirical parameters used in the force field are defined differ between force fields. Each force field has its own strengths and weaknesses so the user has to have some experience and judgment when choosing the best fitting force fields [133]. Due to their suitability for different types of simulation setups, the choice of a force field has to be taken into account also when analyzing and interpreting the results.

The force field determines the level of detail and the time and length scales achievable with current computational methods. Most common force fields used with the GROMACS simulation package contain terms and parameters of either all atoms (AA) in the system or ones that include united atom (UA) groups for carbon atoms bonded with non-polar hydrogens. UA models offer faster computation compared to AA, but lack some accuracy that might be desired from an AA simulation. If even more detail with smaller systems and shorter time scales are desired, a QM/MM or QM force field can be used. QM simulations are useful in creating and refining parameters for AA force fields when combined with empirical data. Reactive force fields, such as ReaxFF [134,135], can be used for simpler chemical reactions of hydrocarbons, in which bonds can break and form during the simulation. Nevertheless QM/MM simulations of biomolecular systems require excessive computational capacity to be implemented in large scale with modern technology.

Common AA and UA force fields used with GROMACS include *e.g.* OPLS/AA [136], AMBER [137, 138], CHARMM [139, 140], and GROMOS [141, 142]. The most universal force field considering current state of parameters for all types of biomolecules has been CHARMM [131], but the others have also been actively developed to include parameters for *e.g.* nucleic acids, carbohydrates, and drug molecules [139,143].

MARTINI Coarse-grained models provide longer time and length scales at the expense of molecular detail achieved with AA and UA force fields by reducing the number of degrees of freedom in the system and allowing larger interaction time steps. In these CG models the system is divided into simplified beads that represent different sets of atoms or functional groups and the interaction parameters are then averaged over the grouped atoms. These so called effective potentials are chosen so that they reproduce the average interactions of the more detailed system. There has been significant development in the accuracy and efficiency of CG models in recent years, and the parameters have been extended from lipids and proteins to include also compounds such as carbohydrates, glycolipids, nucleic acids, and nanoparticles [144].

CG models are a useful method to probe a large number of different simulation setups when the molecular mechanism of function is yet unknown. For example when the level of detail of the protein model of interest is not reliable due to the assumptions and uncertainties in the modeling phase, or there is no previous model for the function of the protein, many different types of simulation system setups can be tested. The validity of many different conformational states of the same protein can be evaluated efficiently. It is possible to move between atomistic and CG levels of detail using methods such as protein coarse-graining [145], CG lipid membrane building [146], and reverse coarse-graining of CG systems [147]. These methods also provide a computationally efficient method to build and equilibrate complex membrane and membrane–protein structures before backmapping (fine-graining) them to AA detail.

One of the most used CG force fields is the MARTINI model also used in this study [145,148–150]. MARTINI has been used successfully to simulate a wide range of biomolecular systems and it has been extended to include parameters for most biologically relevant molecules [151–153]. Compared to atomistic models, the MARTINI force field enables two orders of magnitude larger system sizes or longer simulations to be done while still maintaining enough chemical specificity for many applications.

The MARTINI model is based on a four-to-one mapping method meaning that on average four heavy atoms and associated hydrogens are represented by a single interaction center. To simplify the model, there are only four types of interaction centers: polar (P), non-polar (N), apolar (C), and charged (Q). Each particle type is further divided into a number of subtypes based on their hydrogen-bonding capabilities or the degree of polarity, giving a total of 18 basic particle types [154]. Ring-like structures are mapped with a higher resolution of up to two heavy atoms

and associated hydrogens per MARTINI particle because the four-to-one mapping has been observed to be too coarse for ring geometry and behavior of molecules such as cholesterol [149]. Thus also the size of the ring type beads is smaller compared to the normal corresponding bead type. For computational efficiency the mass of the CG beads is set to 72 amu (4 water molecules) for all beads, except for the smaller beads in ring structures with a mass of 45 amu. One MARTINI water bead corresponds to four atomistic waters, but each ion is represented by a single CG bead that includes the ion and its first hydration shell [154].

The MARTINI force field's bonded and non-bonded interactions have been parameterized in a systematic way. The bonded parameters are derived from reference atomistic simulations, while the non-bonded parameters are based on the reproduction of experimental partitioning free energies between polar and apolar phases, mostly water and a hydrocarbon, for a large number of chemical compounds [154]. The MARTINI model aims for a broad range of applications without the need to reparameterize the model each time for a different system [144].

3.2.1 Bonded Interactions

Bonded interactions involve atoms or particles connected by chemical bonds within a molecule. They include 2-, 3- and 4-body interactions. The three types of interactions are shown in Fig. 3.3, and include the 2-body bond stretching, the 3-body bond angle bending, and the 4-body dihedral angle torsional terms.

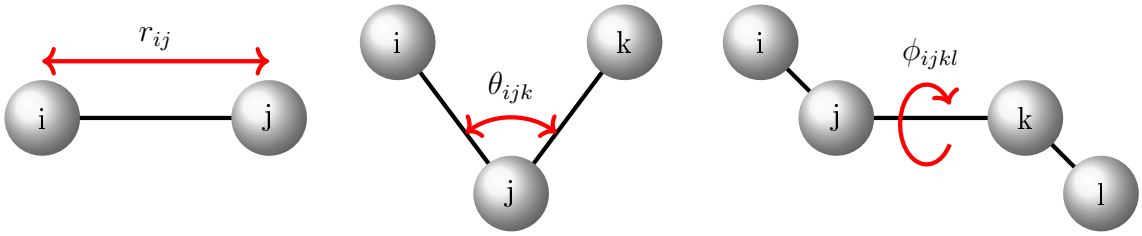


Figure 3.3 Schematic representation of the three bonded interactions: bond stretching, angle bending, and torsional terms

The total potential energy function of the bonded interactions can be written as a sum of the three aforementioned factors as:

$$V_{\text{bonded}} = V_{\text{bonds}} + V_{\text{angles}} + V_{\text{dihedrals}} \tag{3.3}$$

For example in the MARTINI force field the bond stretching term between two

covalently bonded atoms i and j (see Fig. 3.3) is represented by a harmonic potential modeled after Hooke's law as:

$$V_{\text{bonds}} = \frac{1}{2} k_{ij}^{b} \left(r_{ij} - b_{ij}\right)^2 , \tag{3.4}$$

where k_{ij}^{b} is the force constant of the bond, r_{ij} is the distance between the particles, and b_{ij} is the reference bond length value [149]. The bond-angle vibration between a triplet of atoms ijk can also be represented by a harmonic potential on the angle θ_{ijk} as:

$$V_{\text{angles}} = \frac{1}{2} k_{ijk}^{\theta} \left(\theta_{ijk} - \theta_{ijk}^{0}\right)^2 , \tag{3.5}$$

where k_{ijk}^{θ} is the angular force constant, θ_{ijk} is the angle between the three particles, and θ_{ijk}^{0} is the reference angle. For aliphatic chains, *cis* double bonds, and *trans*–unsaturated bonds the force constants and equilibrium bonds angles are $k_{ijk}^{\theta} = 25\,\text{kJ}\,\text{mol}^{-1}$ and $\theta_{ijk}^{0} = 180°$, $k_{ijk}^{\theta} = 45\,\text{kJ}\,\text{mol}^{-1}$ and $\theta_{ijk}^{0} = 120°$, and $k_{ijk}^{\theta} = 45\,\text{kJ}\,\text{mol}^{-1}$ and $\theta_{ijk}^{0} = 180°$, respectively.

The 4-body dihedral potential function can be divided again into the proper and the improper dihedral terms. The dihedral terms describe the torsional angle between the two planes formed by four successive bonded particles. The total dihedral potential function can be written as:

$$V_{\text{dihedrals}} = V_{\text{proper}} + V_{\text{improper}} \tag{3.6}$$

Proper dihedrals describe rotation of the ijk and jkl planes around the jk-axis shown in Fig. 3.3. In GROMACS proper dihedrals are defined according to the IUPAC/IUB convention, where ϕ is the angle between the ijk and the jkl planes, where 0° corresponds to the *cis* configuration and 180° to the *trans* configuration accordingly. There are two different potential functions used for proper dihedrals: periodic and Ryckaert–Bellemans potential. The former is used in the parametrization of the MARTINI force field and can be written as:

$$V_{\text{proper,periodic}} = k_{\phi} \left(1 + \cos\left(n\phi - \phi_s\right)\right) , \tag{3.7}$$

where k_{ϕ} is a force constant, n is a constant, ϕ is the torsion angle, and ϕ_s is the reference torsion angle. The Ryckaert–Bellemans potential function for proper dihedrals can be written as:

$$V_{\text{proper,RB}} = \sum_{n=0}^{5} C_n \left(\cos\left(\phi - 180^\circ\right)\right), \tag{3.8}$$

where C_n are constants for Ryckaert–Bellemans potential defined in the force field, and ϕ is the torsion angle.

Improper dihedrals are used to force planar groups remain planar and to prevent changes in the chirality of a molecule. The simplest improper dihedral potential is a harmonic potential defined as:

$$V_{\text{improper}} = \frac{1}{2} k_\xi \left(\xi_{ijkl} - \xi^0_{ijkl}\right)^2, \tag{3.9}$$

where k_ξ is a force constant, ξ_{ijkl} is the improper dihedral angle, and ξ^0_{ijkl} is the reference angle value defined in the force field.

3.2.2 Non-bonded Interactions

Non-bonded interactions contain the interactions of particles that are not covalently bonded but remain within a cut-off distance defined in the force field for van der Waals and electrostatic interactions. The cut-off distance can be implemented to speed up calculations, but especially in AA simulations the electrostatic interactions are calculated efficiently with the Particle mesh Ewald-method (PME) [155], which takes into account all interactions, also beyond the cut-off. Van der Waals interactions involve short-range repulsion and long-range attraction terms between a pair of particles. The non-bonded potential function is thus divided into two potential function terms, the Lennard-Jones (LJ) interaction describes the van der Waals dispersion and steric repulsion terms and the Coulomb (C) term the electrostatic interaction. The non-bonded interactions can be calculated as:

$$V_{\text{non-bonded}} = V_{\text{LJ}} + V_{\text{C}} \tag{3.10}$$

The non-bonded interactions are computed for each particle on the basis of a neighbor list, which keeps efficiently track of the particles that are within the cut-off radius. The LJ interactions contain the repulsion and dispersion terms and is defined as:

$$V_{\text{LJ}} = V_{\text{repulsion}} + V_{\text{attraction}} = \frac{C^{(12)}_{ij}}{r^{12}_{ij}} - \frac{C^{(6)}_{ij}}{r^{6}_{ij}}, \tag{3.11}$$

where $C^{(12)}_{ij}$ and $C^{(6)}_{ij}$ are constants depending on the pairs of particle types i and j,

and the r_{ij} term is their relative distance respectively. The electrostatic interaction between two charged particles can be calculated as a classical Coulombic potential as:

$$V_{\mathrm{C}} = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{\varepsilon_r r_{ij}}, \tag{3.12}$$

where ε_0 is the permittivity of vacuum, ε_r is the relative permittivity, q_i and q_j are the charges of the particles, and r_{ij} is the distance between the two particles.

Special potential functions such as absolute or relative position restrains can be applied to the simulation. The MARTINI model utilizes such relative position restrains by creating a network of pseudo-bonds between non-bonded particles in a protein structure. Due to the hereditary weaknesses of MARTINI to handle changes in the secondary structure of proteins as well as the overemphasized forces in protein–protein interactions, an elastic network of harmonic bonds can be used to stabilize and constrain a protein structure closer to its observed native state [154]. The standard MARTINI model has a built-in option to add such harmonic bonds between protein backbone beads to the topology based on a predefined cut-off distance. The behavior of a protein needs to be validated based on atomistic simulations and the cut-off and the force constants chosen accordingly. One way is to use an optimized method for combining an elastic network to a CG system, called ElNeDyn [156].

3.3 Equations of Motion

The standard MD sampling algorithm is a numerical solution of the classical Newtonian equations of motion based on the interactions defined in the force field. To compute the trajectory *i.e.* the evolution of positions and velocities of the N particles in the system, the forces acting on each particle at each time step have to be evaluated. The force $\mathbf{F}_i$ acting on a particle is the negative gradient of the total sum of the potential functions $V(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ as:

$$\mathbf{F}_i = -\frac{\partial V_i}{\partial \mathbf{r}_i} \tag{3.13}$$

where $\mathbf{r}_i$ is the position of the particle i. The motion of the particles can be calculated based on the forces acting on them as:

$$\mathbf{F}_i = m_i \frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2} \tag{3.14}$$

where m_i is the mass of the particle i, and t is time.

These two equations are solved simultaneously during each time step of the simulation. Several different numerical MD integrators can be used in GROMACS to calculate the new coordinates and velocities [157]. The two most commonly used are the leap-frog integrator, which also the default MD integrator, and the Verlet integrator. In the leap-frog algorithm the positions and velocities are calculated as:

$$\mathbf{r}_i\left(t+\Delta t\right) = \mathbf{r}_i(t) + \mathbf{v}_i\left(t+\frac{\Delta t}{2}\right)\Delta t \tag{3.15}$$

$$\mathbf{v}_i\left(t+\frac{\Delta t}{2}\right) = \mathbf{v}_i\left(t-\frac{\Delta t}{2}\right) + \frac{\mathbf{F}_i(t)}{2m_i}\Delta t \tag{3.16}$$

where $\mathbf{r}_i$ and $\mathbf{v}_i$ are the position and the velocity vectors of the particle i respectively, t is the point in time at which the system is studied at, Δt is the time step used in the simulation, and m_i is the mass of the particle. At representative predefined intervals these data are saved in an output trajectory file.

3.4 Temperature and Pressure Coupling

Without a specified temperature and pressure coupling scheme, MD simulations are performed in constant number (N), constant volume (V), and constant energy (E) environment, *i.e.* microcanonical ensemble (NVE). However, real experiments on biological systems take place at a constant temperature (T) and pressure (p) (NPT), also called the isothermal–isobaric ensemble, or constant temperature and volume (NVT) conditions, also called the canonical ensemble. These conditions are thus also desired in MD simulations, and achieved by selecting a suitable thermostat for the temperature and a barostat for the pressure.

One of the most commonly used temperature coupling methods is the Berendsen weak coupling algorithm [158]. Also the Nosé–Hoover thermostat [159, 160] and the velocity-rescaling scheme [161] are often utilized in MD simulations with GROMACS. The purpose of these algorithms is to slowly correct the deviations in the temperature of the system. In the Berendsen coupling the deviations decay exponentially with a time constant τ as:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{T_0 - T}{\tau_t}, \tag{3.17}$$

where T_0 is the reference temperature, and T is the temperature of the simulation system. The other temperature coupling method utilized in this study is the velocity-rescaling thermostat, which differs from the Berendsen by an additional stochastic

term to correct the kinetic energy distribution

$$\mathrm{d}K = (K_0 - K)\frac{\mathrm{d}t}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}}\frac{\mathrm{d}W}{\sqrt{\tau_T}}, \tag{3.18}$$

where K is the kinetic energy, N_f is the number of degrees of freedom, and $\mathrm{d}W$ a Wiener process. All the particle velocities are corrected by a properly chosen random factor. Velocity-rescaling produces a correct canonical ensemble while preserving the advantages of the Berendsen thermostat [161].

NPT simulation conditions can be achieved with a pressure coupling algorithm similar to the ones used in temperature coupling. The Berendsen pressure coupling [158] rescales the simulation box vectors and the coordinates of the particles to maintain the reference pressure P_0 as

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{P_0 - P}{\tau_p}, \tag{3.19}$$

A scaling matrix μ given as

$$\mu_{ij} = \delta_{ij} - \frac{n_{\mathrm{PC}}\Delta t}{3\tau_p}\beta_{ij}\{\mathbf{P}_{0ij} - \mathbf{P}_{ij}(t)\}, \tag{3.20}$$

is used in the method, where δ_{ij} is Kronecker delta ($0 \,\mathrm{if}\, i \neq j, 1 \,\mathrm{if}\, i = j$), n_{PC} the number of steps between pressure rescaling, τ_p the pressure time constant, β the isothermal compressibility of the system, and $\mathbf{P}_{0ij}$ the target pressure. The Berendsen barostat is useful especially in the system equilibration phase if other methods, such as the Parrinello–Rahman, would produce unwanted oscillations of the system.

The Parrinello–Rahman pressure coupling [162, 163] is used when accurate calculations on the thermodynamic properties of the system are of interest. It is similar to the Nosé–Hoover temperature coupling, and in theory gives the true NPT ensemble. The barostat is calculated as:

$$\frac{\mathrm{d}\mathbf{b}^2}{\mathrm{d}t^2} = V\mathbf{W}^{-1}\mathbf{b}'^{-1}(\mathbf{P} - \mathbf{P}_{\mathrm{ref}}), \tag{3.21}$$

where $\mathbf{b}$ is a matrix representing the box vectors, V is the volume of the simulation box, W is a matrix parameter that determines the strength of the coupling and how the box can be deformed, and the matrices $\mathbf{P}$ and $\mathbf{P}_{\mathrm{ref}}$ are the current and reference pressures, respectively.

3.5 Periodic Boundary Conditions

The simulation box is of finite size that depends on the fluctuations caused by the barostat. Hence without periodic boundary conditions (PBCs) this would mean that at the edge of the box the system would end at a wall or a vacuum. However, in MD simulations the simulation box is commonly made infinite by imposing uniform periodic images in all dimensions around the simulation box. Thus in reality there are no boundaries and a particle that passes through one side of the box re-enters in from the opposite side at the same time. The idea of PBCs is illustrated in Fig. 3.4.

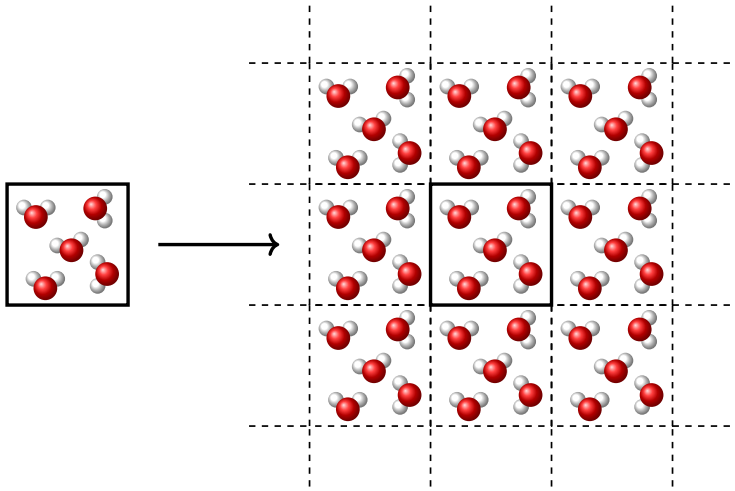


Figure 3.4 Periodic boundary conditions

PBCs minimize the edge effects of a finite system and the artifacts caused by unwanted system boundaries, but also creates a potential new artifact of periodic conditions. If the chosen simulation box size is too small, a particle can experience the forces caused by other particles more than once. Similarly a large biological macromolecule can experience forces caused by its own periodic image. Thus the size of the simulation box needs to be defined large enough to overcome these possible periodic artifacts from long-range interactions. The PBCs are combined in GROMACS with the minimum image convention, so that only one, the nearest image of each particle is considered for short-range non-bonded interaction terms.

3.6 Limitations of MD simulations

MD simulations are relatively simple in nature and rely on a well understood concept like classical mechanics. Still, the user should be aware of the limitations of MD

simulations and always check and validate their accuracy by comparing the behavior of the model to experimental results. As described earlier, the MD method is based on many approximations and assumptions that need to be taken into account when preparing and simulating a new system.

The use of only classical mechanics in describing the motion of atoms is suitable for most atoms in normal temperatures, but there are exceptions. For example, the motions of lightweight atoms like hydrogen and helium, and the vibrations of oscillating bonds can not always be described without quantum mechanics. Corrections to the force field parameters can at least partly overcome these issues. The exclusion of QM and bond formation and breaking also means that no chemical reactions like enzymatic catalysis or pH dependent change in the protonation state of a protein can take place.

MD force fields are inherently approximations of real life mechanics. All simulation results depend on the selection of an accurate force field for the application in question. The force fields are pair-additive, which means that all non-bonded forces result from the sum of non-bonded pairwise interactions. The PBCs may also cause problems if not addressed as mentioned earlier. The sizes and time scales achievable by MD simulations still remain very limited compared to experiment. This is both one of the advantages and downsides of MD as it offers resolutions that can not be achieved otherwise, but as new systems get bigger and the simulations longer, the costs of running a simulation also increase.

The MARTINI model has a very wide range of potential applications in MD simulations, but the force field has some important limitations that should be kept in mind. First of all, the model is parameterized for the fluid phase, and works best in such systems. It was originally parameterized especially for lipid and surfactant systems [148], and only later extended for other molecules. The behavior of solid and gas phase appears too stable with respect to the fluid phase, and thus the thermodynamic behavior of these phases at liquid interfaces should be interpreted with care [144].

The use of the CG model accelerates diffusion of molecules compared to experimental data by an average conversion factor of 4 [144], which has to be taken into account when comparing the time scales of reactions with AA or experimental data. The time step should be chosen between 20 and 40 fs to use the model as effectively as possible. The CG energy landscape is highly simplified compared to AA, which leads to increased sampling speed at the cost of a loss of detail [144]. Too short time step undermines the efficiency of the model and is waste of computing time.

When considering the simulation of proteins and peptides, the biggest shortcoming of the MARTINI model is that the secondary structure transformations are not possible. The secondary structure is fixed by the use of a dihedral potential energy function. On the other hand, the changes in the tertiary structure of proteins have been shown to be quite realistic [145]. The original parameterization of the MARTINI model has been shown to significantly overestimate the strength of protein–protein interactions [164]. This can be compensated by down-scaling the non-bonded interaction parameters between protein particle types.

4. PROTEIN STRUCTURE PREDICTION AND MODELING

This chapter gives a very short introduction to the basics of protein structure and the most common methods used in protein structure prediction. MD simulations are a method that can be used to predict and verify novel protein structures [130]. Protein models can be based on experimental data or amino acid sequence homogeneity. The first part of this chapter is based mostly on references [21, 165, 166], and the protein structure prediction and protein modeling on reference [130].

4.1 Basics of Protein Structure

The structure of a protein defines its function in biological systems. The protein structure is often described at four levels: primary, secondary, tertiary, and quaternary structure. The shape of a protein is defined by its primary structure, *i.e.* the amino acid sequence. The primary structure describes all the covalent bonds linking amino acid residues in polypeptide chains, which includes mainly the peptide bonds in the polypeptide backbone and the disulfide bonds between pairs of cysteine residues. The sequence of the 20 naturally occurring amino acids is essential for the three-dimensional structure and correct functioning of a protein. Single point mutation in the sequence can alter the structure and make the protein partly or completely inactive.

The secondary structure refers to commonly occurring stable local arrangements of amino acids forming regular structural patterns. The two most commonly occurring types of secondary structures are α-helix and β-sheets. The formation of these structures is driven by hydrogen bonding in the protein backbone. Secondary structure prediction is one of the challenges in defining the correct tertiary structure. Homology modeling and multiple sequence alignment can be used to identify common patterns in the protein sequence resulting in these local arrangements [167].

The tertiary structure of a folded protein describes the whole three-dimensional arrangement of all the residues in a polypeptide. Folding of a protein to its native conformation is driven by many types of covalent and noncovalent forces and in-

teractions, and may involve other proteins called chaperones. The forces include, in addition to the ones involved in the primary and secondary structure, hydrogen bonding of the protein side chains, clustering and burying of nonpolar residues to the inside of the protein by hydrophobic interactions, and electrostatic interactions between oppositely charged groups.

The quaternary structure describes three-dimensional arrangement of a protein consisting of two or more polypeptide subunits. The subunits of an oligomer are often held together by noncovalent interactions and they can be easily dissociated by treatment with denaturing agents. The subunits can be identical or different depending on the protein and thus similar or varying tertiary structures and conformational states.

4.2 Protein Structure Prediction

Protein structure prediction attempts to define a realistic model for the 3D structure of the protein by comparing it to pre-existing protein structures, or alternatively by computationally solving the folding of smaller secondary and tertiary structures in a protein. The challenge is to explore the whole conformational space of the molecule in order to identify the most appropriate, *i.e.* lowest energy, structures from the immense number of possibilities. The global minimum in the energy function is assumed to represent the native state of the structure of the molecule.

Protein structure prediction methods can be roughly divided into two main groups: comparative modeling and computational modeling. The first step in building a new model for a protein is to check if any protein structure databases, *e.g.* the Protein Data Bank (PDB) [168], contain known homologous sequences of similar proteins that can be used for homology modeling. If there are no known homologous proteins, but a structurally similar fold is recognized, a method called protein threading is an option. If neither comparative modeling method is suitable for the structure prediction or the structure needs to be refined, computational modeling can solve relatively small protein structures. *Ab initio* protein structure prediction methods try to solve the tertiary structure of a protein using only the primary sequence of amino acids.

Homology Modeling Proteins in the same family share similarities in the 3D structure as well as function, originating mostly from the homology of their primary structures [169]. Comparative modeling uses these structural similarities between proteins to construct a 3D-model based on the known structures of one or more

related template proteins [170]. There are four steps in constructing a homology model for an unknown protein structure: template selection, sequence alignment, model construction, and model assessment [170]. The template selection is the most critical step in homology modeling, and careful assessment needs to be done to find the best fitting structure. The simplest method is to use an automated program like FASTA [171] or BLAST [172] to do pairwise sequence alignment over known entries in protein databases. A sequence homology of over 30% is a requirement for correct homology modeling, but values of over 40% sequence identity have been found to produce more reliable structural models for new proteins [170]. Pairwise comparison might overlook some trends occurring in the structures. If more than one protein in the same family are known, a more accurate way for template searching is to use multiple sequence alignment tools, like PSI-BLAST [173], that can produce improved models from larger clusters of homologous proteins [167].

After template selection the sequences need to be properly aligned before model construction. One of the most commonly used programs for multiple sequence alignment is CLUSTAL [82, 174]. The 3D-model for the target protein is constructed based on the alignment using one of the three model generation methods: rigid body assembly, segment matching, and satisfaction of spacial restrains [170]. The first method assembles together rigid bodies or conserved structural fragments taken from the template proteins found in the alignment phase. In the second method the target protein sequence is aligned in shorter segments to several templates from PDB, and each segment is modeled based on the best segment homology. The third method generates a series of geometrical criteria from the templates, which are used in conjunction with an optimization procedure to determine a structure of the target. Accuracies of the methods are relatively similar, and template selection and sequence alignment have a much larger impact on the model accuracy. Noteworthy software used for protein structure homology modeling include *e.g.* MODELLER [175], SWISS-MODEL [176] and ROSETTA [177].

The initial structures obtained from comparative modeling can be of high energy and require energy minimization to refine the final structure. After construction the model needs to be evaluated. Most of the analysis can be done automatically using computer programs that examine the structure for common flaws in the fold and free energies. If discrepancies between the developed model and experimental data occur, the model needs to be refined until the best structure has been found. As experimental resolution and evaluation methods have improved over the years, thousands of submitted protein structures in the PDB have become obsolete [178].

Protein Threading Protein threading or fold recognition is a template-based method that is used if there are no known homologous proteins or the sequence homology is below 30%. Homology modeling aligns the target sequence to one or multiple template sequences whereas protein threading aligns the target sequence to template structures to find similar protein folds. Protein threading is based on the assumption that there is a limited number of possible protein folds in nature, and the amino acid sequence implies a preference for different structural environments in the folds. According to the Structural Classification of Proteins (SCOP) [179] and the Class, Architecture, Topology, Homology (CATH) [180] databases there are approximately 1400 known protein folds, and majority of new proteins submitted to PDB are structurally related to the ones already in the database with defined folds.

In protein threading the sequence is compared to a library of all known secondary and tertiary structures to get the best alignment of the sequence to a template. An objective scoring function is used to sort out the best sequence–structure alignment from all the possibilities. The function takes into account factors like amino acid preference for solvent (hydrophobicity), amino acid preference for a particular secondary structures and preferred interactions between residues, like ionic bonds and disulfides. Even after careful setup and consideration of the scoring function, protein threading has a high possibility of producing only incorrect structures.

***Ab initio* Protein Structure Prediction** The use of computational protein structure prediction methods is limited by the requirement for vast computational resources, and has been restricted to small proteins of a few dozen residues and targeted small sections of proteins observed to fold rapidly [181]. MD simulations are one of the computational methods used to predict and study new protein structures, but their ultimate shortcomings limit their use. Correct protein folding simulations require time-scales ranging from microseconds to several milliseconds, with a force field that describes correctly all the intermediate unfolded and misfolded conformations, reproduces the true potential energy surfaces of proteins, and minimizes all cumulative errors of long simulations [181]. Representative sampling of protein folding calls for numerous repetitions to characterize the heterogeneous folding processes.

5. SIMULATION MODELS AND ANALYSIS METHODS

In this thesis the interactions between SP-B and the main pulmonary surfactant phospholipids were studied using MD simulations described in chapter 3. The preparation of the simulation systems from protein model refining to simulation system construction, and the basics of the analysis methods used are covered in this chapter.

5.1 System Preparation

A good structural model is a requirement for a successful simulation. Based on initial simulation results with the original homology model for the SP-B multidimeric oligomers from [6] we decided to further refine the protein structure to better match experimental results. The systems were built to study SP-B's interactions with the pulmonary surfactant phospholipids. Due to their amphipathic nature, phospholipids assemble spontaneously into lipid aggregates, like micelles and liposomes. The best way to study any lipid-specific interaction sites in the structure of SP-B is to conduct lipid self-assembly (SA) simulations. These self-assembly studies work as a starting point for any other protein–lipid interaction simulations as they offer a potential starting conformation for the permanently membrane-bound SP-B in monolayers or bilayers.

5.1.1 Modeling

Neither high resolution structural model nor molecular mechanism of SP-B function have existed until recently. An experimental study by Olmeda *et al.* [6] revealed the presence of 10 nm ring-shaped multidimeric oligomer states of SP-B, based on atomic force and electron microscopy. A multiple sequence alignment with SP-B and saposin B proteins was done, and two 3D-models for dimers of SP-B were constructed based on the experimentally observed states of saposin B [6]. This resulted in the open and closed conformations for the dimers of SP-B (see Fig. 2.9), and finally the pentamers and hexamers of dimers. These oligomeric states were evaluated to best comply both to the observed size of the assemblies and the intra-chain C48

disulfide bonds and adjacent E51–R52' and E52–R51' salt bridges stabilizing the supradimeric complex.

For clarity a second distinction between two overlapping structural units of SP-B dimers needs to be done: a *functional dimer* ($\mathrm{Dim_f}$) and a *disulfide-bridged dimer* ($\mathrm{Dim_c}$). The functional dimer refers to a dimer formed by two SP-B monomers in a conformation akin to that of saposin B dimer. An inner hydrophobic cavity possibly responsible for lipid interaction is located between the monomers alike to the lipid binding pocket found in saposin B [182]. The disulfide-bridged dimer, or covalent dimer, refers to the structure commonly extracted in previous experiments, where the monomers are connected by a disulfide bridge at C48. Thus the covalent SP-B dimer contains one monomer from each adjacent functional dimers of the supradimeric structure.

A series of model refinement steps were conducted based on extensive initial MD simulation results with the original structures from [6]. We aimed to generate all-atom supradimeric oligomer models of various sizes that would:

(i) contain the SAPLIP-like functional dimer structure as a basic structural unit in either open or closed conformation
(ii) have the interface conductive to formation of the inter-chain disulfide and salt bridges
(iii) have the putative lipid binding site facing towards the central pore of the ring.

We used Rosetta's [183, 184] symmetric docking protocol with restraints on the distances between the interface forming residues to generate decoys that satisfy (ii). We later sorted these models based on their total score, clustered them, and extracted cluster centers that satisfy (iii) as final models. A hexamer of dimers (6mer) was selected as the main multi-dimeric oligomer to be used in the simulations based on initial results and reference [6]. The SP-B structures used in this study are shown in Fig. 5.1.

5.1.2 SP-B Systems

Three phospholipid compositions were selected for the lipid self-assembly (SA) and monolayer simulations, the characteristics of which can be seen in Table 5.1. The physiological lipid composition (PHYS) was derived from previous studies [2, 24], while the simplest DPPC and CHOL (DC) composition contained the same molar fraction of CHOL as the PHYS composition. A total of three different conformations of the modeled SP-B were used to study the protein's interactions with different

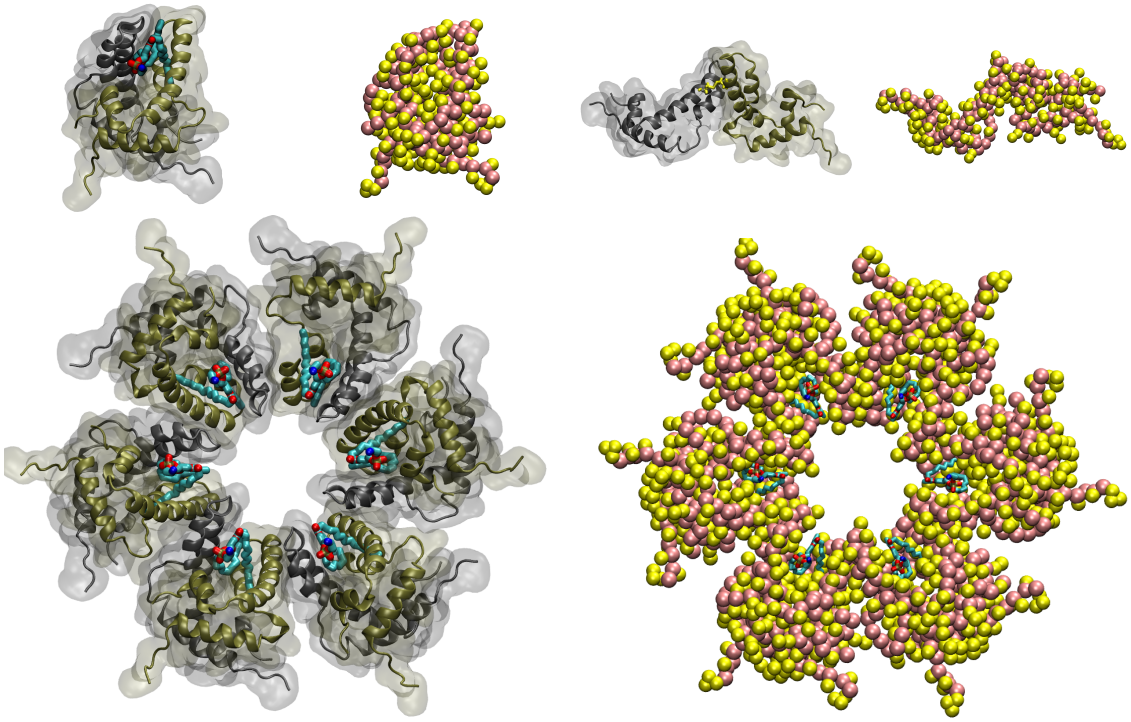


Figure 5.1 SP-B structures used in the study. Up left: the atomistic and the coarse-grained Dim_f. Up right: the atomistic and coarse-grained Dim_c. Below: the structure of the SP-B hexamer of dimers with phospholipids inside the lipid binding pockets. The SP-B oligomer forms a ring with a hyrdophobic central pore in the middle. The membrane binding residues in the N- and the C-termini are located at the outer rim of the initial structure. The lipid binding pockets of the functional dimers are facing the central pore.

phospholids. These structures were the Dim_c, the Dim_f, and the SP-B hexamer of dimers (6mer) (see Fig. 5.1 and Table 5.2).

Table 5.1 Lipid compositions in simulation systems. DC refers to a lipid composition of DPPC and cholesterol [185], EQ to an equal mole fraction of all four lipids, and PHYS represents an approximation of the physiological lipid composition of the pulmonary surfactant [2,24].

Name	**DPPC (mol-%)**	**POPC (mol-%)**	**POPG (mol-%)**	**CHOL (mol-%)**
DC	90			10
EQ	25	25	25	25
PHYS	50	25	15	10

SP-B lipid self-assembly In the preparation of the lipid self-assembly simulation systems, 24 independent starting conformations with each protein structure and

lipid composition were built by inserting the lipids randomly into the simulation box with the `gmx insert-molecules` GROMACS program. All dimer systems had a total of 300 lipids and 12000 CG water beads, and 6mer systems were built with 1100 lipids and 33000 water beads. In all systems the NaCl concentration was kept at a physiological value of 150 mM. A total of 216 self-assembly simulations were performed, 144 with the functional dimer ($\mathrm{Dim_f}$) and covalent dimer ($\mathrm{Dim_c}$) structures, and 72 simulations with the SP-B hexamer of dimers (6mer) structure (see Table 5.2). Each system was simulated for 2.5 µs.

SP-B monolayer systems The monolayers were constructed with the INSANE script [146]. The area per lipid (APL) was varied and four values were chosen to represent different compression states of the pulmonary surfactant. These values were 52.5, 55, 57.5, and 60 $\mathrm{Å}^2$. APL values below 52.5 $\mathrm{Å}^2$ resulted in folding (initial state of collapse) of the monolayer before protein contact, and values of over 60 $\mathrm{Å}^2$ resulted in spontaneous formation of holes in the monolayers. The protein was inserted between the monolayers into the water phase, after which the water beads were removed in small batches (500 water beads) from the systems followed by a short 10 ns simulation step with the protein position restrained. NaCl concentration was kept at physiological value of 150 mM. This resulted in systems where SP-B is in contact with two monolayers, one on each side of the protein, with a varying distance between the monolayers. Four inter-monolayer distances were chosen and selected from each set of systems with different APL. Three repetitions of each constructed monolayer system were simulated each for a total of 25.0 µs (see Table 5.2).

5.2 Simulation Parameters

All simulations were performed using the GROMACS package version 5.1.2 [128] and the coarse-grained MARTINI force field [145,149,150]. An energy minimization step with the steepest descent algorithm was performed before running the production simulations. The self-assembly systems were equilibrated in two steps before production runs: first at 310 K under NVT conditions with an integration time step of 25 fs for 200 ps, then under NPT conditions at a constant 1 bar pressure controlled with the isotropic Berendsen barostat [158] for 200 ps. Other conditions were kept the same. The production runs were simulated using an integration time step of 25 fs at a reference temperature of 310 K controlled with the v-rescale method [161] with a time constant of 1.0 ps. Separate heat baths for the protein, the lipids, and the solvent were used. A reference pressure of 1 bar was used and controlled with

Table 5.2 A list of the simulation systems in this study. Dim_f is the functional SP-B dimer, Dim_c is the covalently bound SP-B dimer, 6mer is the hexamer of SP-B dimers, SA is lipid self-assembly simulation, mono is monolayer interaction simulation, and the lipid compositions are as in Table 5.1. * Four APLs, with four monolayer distances, each with three repetitions.

Protein	SA/mono	Lipid composition	Repetitions	Time (µs)
Dim_f	SA	DC	24	2.5
Dim_f	SA	EQ	24	2.5
Dim_f	SA	PHYS	24	2.5
Dim_c	SA	DC	24	2.5
Dim_c	SA	EQ	24	2.5
Dim_c	SA	PHYS	24	2.5
6mer	SA	DC	24	2.5
6mer	SA	EQ	24	2.5
6mer	SA	PHYS	24	2.5
6mer	mono	PHYS	48 *	25.0

the isotropic Parrinello–Rahman barostat [162] with a time constant of 12.0 ps. All bonds were constrained using the LINCS algorithm [186]. Production simulations of the monolayer systems were performed at 310 K under NVT conditions with an integration time step of 25 fs.

5.3 Analysis Methods

Lipid Contact Frequency The number of frames in which any bead of the lipid is within a 0.6 nm cut-off distance of any bead of each protein residue is counted and divided by the total number of frames to get the lipid contact frequency for each protein residue. This analysis was performed for each repetition, 24 per self-assembly systems, and 3 for monolayer simulations, and normalized by dividing by the total number of frames in all repetitions. The mean of the results and the standard errors (SE) are presented. An additional normalization based on the relative concentration of each lipid type was performed in PHYS system. The analysis is implemented in a TCL script to be used with VMD [187].

Inter-Monolayer Distance The distance between the two monolayers on both sides of the 6mer was measured to select corresponding systems from different APL simulations. The distance of the monolayers was determined as the average distance of PO4 beads in the CG phospholipids, representing the phosphate group of the lipids, over the course of the simulation. The GROMACS program `gmx density`

was used to compute the partial densities of the particles across the box, and the inter-monolayer distance was calculated from the plotted data.

Area Per Lipid APL was determined for the monolayers and bilayers with

$$\mathrm{APL} = \frac{A_{\mathrm{membrane}}}{N_{\mathrm{lipids}}/2}, \tag{5.1}$$

where A_{membrane} is the area of the monolayer or bilayer, and N_{lipids} is the total number of lipids in both monolayers or leaflets of the bilayer.

Density Maps The lateral 2D number density maps of each lipid species were determined using the GROMACS program `gmx densmap` to show the average density of lipids in the monolayers around the protein. The density maps were calculated separately for each monolayer to compare the protein's ability to reorganize surfactant lipids upon adsorption to the surface, and to find specific interaction sites for the lipids in the protein structure.

The density maps were determined as a average over the repetitions. Before calculating the density maps, a proper translation and rotation was applied to each frame to superimpose the protein to the same reference structure. The heatmaps were normalized per lipid type to allow for comparison between different lipids. Four APL values with four monolayer distances allowed the evaluation of their effect in the lateral distribution of lipids around the protein.

6. RESULTS AND DISCUSSION

In this thesis, the specific lipid interactions and membrane binding of our new model for the structure of pulmonary surfactant protein B (SP-B) were studied. The functional investigation of pulmonary surfactant requires a delineation of lipid-protein interactions, which can be achieved by using molecular dynamics simulations. SP-B is an essential part of the human pulmonary surfactant and its functionality, yet its structure and molecular mechanism of function have remained elusive until recently. Thus, the study of the lipid-protein interactions is necessary for understanding the function of SP-B in the pulmonary surfactant.

To evaluate the effect of different oligomer structures of SP-B, we performed a total of 216 lipid self-assembly MD simulations using three SP-B structures in three lipid compositions. Two forms of SP-B dimers, the covalent dimer ($\mathrm{Dim_c}$) and the functional dimer ($\mathrm{Dim_f}$), and a SP-B hexamer of dimers (6mer) were studied (see Fig. 5.1). We chose lipid compositions, which contained DPPC and cholesterol (DC), equal mole fractions of the four main surfactant lipids (EQ), and a composition similar to the physiological structure of pulmonary surfactant (PHYS) (see Table 5.1). Additionally, the effect of SP-B on the lateral lipid reorganization of surfactant lipids in SP-B coupled monolayers at different compression states were studied by performing a total of 48 MD simulations (see Table 5.2 for a list of all simulation systems). The nomenclature used in this chapter is summed up in Fig. 6.1.

This chapter focuses on the possible lipid binding conformations and orientations of SP-B dimers and hexamer of dimers in membranes. In the latter part of the chapter, SP-B's observed membrane lipid perturbing effect is discussed as a possible molecular mechanism of function for the protein. All molecular images and analyses were prepared using VMD [187] and MATLAB®.

6.1 Lipid Self-Assembly Simulations

The MARTINI force field was originally parameterized for lipid and surfactant systems [148] and has been since extended to include other biomolecules, such as proteins [145]. As amphiphilic molecules, lipids have a tendency to spontaneously form

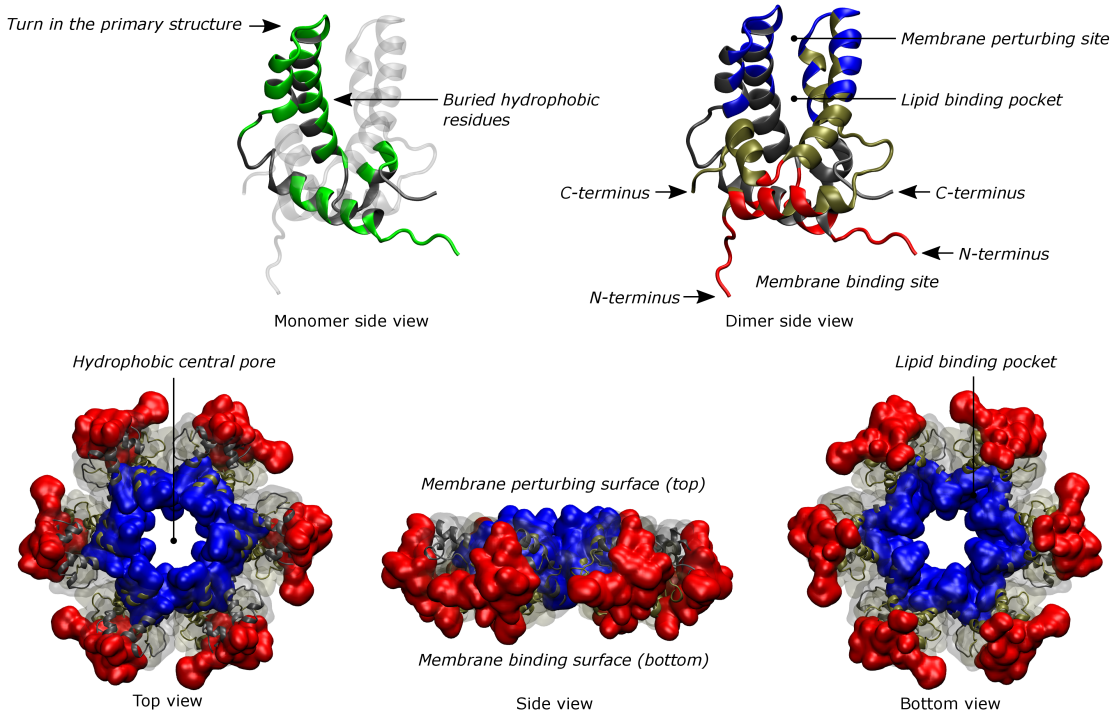


Figure 6.1 Reference structures. The buried hydrophobic residues are located on the inner face of the monomers in the functional dimer, and are shielded from water by the opposing monomer of the Dim_f. The residues in the membrane binding sites of the individual Dim_f structures are shown in red, and the residues in the membrane perturbing sites are shown in blue. The residues of the membrane binding site of the Dim_f are located in the N- and C-terminal helical structures, and the membrane perturbing site is located near the turn in the primary structure, approximately from residue V31 to Y53. The lipid binding pocket of the Dim_f is located between the SP-B monomers and is lined with the buried hydrophobic residues. The membrane binding surface and the membrane perturbing surface of the SP-B hexamer of dimers are directed towards opposing membranes or monolayers in the orientation of SP-B membrane binding suggested by our results. The hydrophobic central pore of the SP-B hexamer of dimers is enclosed by the lipid perturbing sites of the individual dimers.

different lamellar and non-lamellar phases in water [188]. Thus, we performed lipid self-assembly simulations to study the lipid interaction of the membrane associated SP-B. The formation of micellar aggregates, mostly spherical micelles and bilayer discs, was observed to be very rapid in the self-assembly systems, with SP-B partitioning exclusively into the lipid phase.

6.1.1 SP-B Dimers Have Two Distinctly Different Binding Configurations

We studied the lipid contact frequency of each residue in the SP-B structure to find specific lipid interaction sites in the dimers. The *overall* lipid contact frequency represents the average interaction time of a residue with any lipid in the system during the simulation time, while the *lipid-specific* contact frequency differentiates lipids by type. One protein residue can be in contact with more than one type of lipid in each frame, while contact with two lipids of the same type is counted as one. Thus the lipid-specific contact frequency results can be interpreted as a preference towards a specific type of a lipid, but not directly as a proportion of all contacts. Also the results of the contact frequency analysis show lipid-specific interaction sites in the structure of the protein. On the other hand, the overall lipid contact frequency results show preferences for the membrane binding orientation of the dimers.

Two distinctly different lipid binding configurations were observed between the covalent (Dim_c) and the functional (Dim_f) variations of the SP-B dimers (see Figs. 6.2 and 6.4). All three lipid compositions showed similar overall lipid contact frequencies with individual residues (see Figs. 6.3 and 6.5). The differences seen in lipid contact frequencies between the Dim_c and Dim_f structures arise from the orientation of the protein at the surface of the lipid interface.

Covalent dimers In the self-assembly simulations, the Dim_c structure is predominantly oriented towards the lipids in a binding mode shown in Fig. 6.2. SP-B consists of a large fraction of hydrophobic residues, which reside on both faces of the extended SP-B monomer. In the Dim_f structure most of the hydrophobic residues are buried within the cavity between the monomers while the rest of the hydrophobic residues in the helical N- and C-terminal portions of the protein are involved in membrane binding (see Fig. 6.1). In the beginning of the simulation of the Dim_c structure, these buried hydrophobic residues are exposed to the water phase. Our results show that with all tested lipid compositions, the Dim_c is at the surface of a bilayer disc oriented with these buried hydrophobic residues towards the lipids (see Fig. 6.3).

In our simulations Dim_c partitions to the edge of the bilayer discs formed during the lipid self-assembly. The covalently linked dimer attains an orientation in which it interacts with both leaflets of the bilayer disc (see Fig. 6.2). As previously mentioned, there are two proposed theoretical molecular models for the topology of SP-B in phospholipid membranes [73]. A transmembrane orientation of SP-B

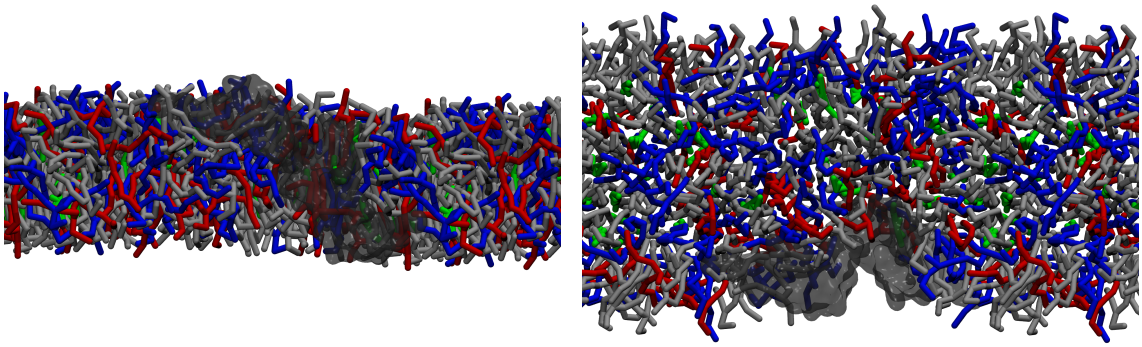


Figure 6.2 Lipid binding configuration of the Dim_c shows a possible orientation of binding with lipids for the SP-B dimer based on the self-assembly simulations. Dim_c binds to the edge of a bicelle in a transmembrane orientation. PHYS lipid composition, DPPC in grey, POPC in blue, POPG in red, and cholesterol in green.

dimers and the interaction of SP-B with the inner core of bilayers at the edge of membrane discs has been suggested in some studies [71,111,189] (see also Fig. 2.7). Our results suggest that such a topology for the SP-B dimer is possible, but due to the limited number of lipids in the system, the transmembrane orientation of the Dim_c in lamellar membranes could not be studied. However, majority of experiments imply a more superficial binding mode of SP-B [67,83,87].

In previous MD simulation studies with monomers and dimers of SP-B, two membrane binding orientations of SP-B on the surface of monolayers and bilayers have been demonstrated [102,104,112]. These studies showed that SP-B can adopt either an extended or a bent conformation, depending on which face of the protein is oriented towards the water phase and which one towards the lipids. Our results show a membrane binding orientation of Dim_c, where the conformation of each monomer of Dim_c is similar to the extended conformation of SP-B monomers and dimers previously shown by Baoukina *et al.* [102,104] or the open conformation suggested by Khatami *et al.* [112].

The bent conformation is denoted as the active form of SP-B monomers that triggers fusion and perturbations in adjacent membranes [104], while the extended conformation has been shown to induce local curvature in lipid monolayers at higher surface pressures in MD simulations [102]. The bent conformation of the SP-B monomer is more like the monomers in our models for the Dim_f and the SP-B hexamer of dimers. Thus the specific lipid interactions and lipid binding orientations of Dim_f are of more interest when studying the higher oligomers.

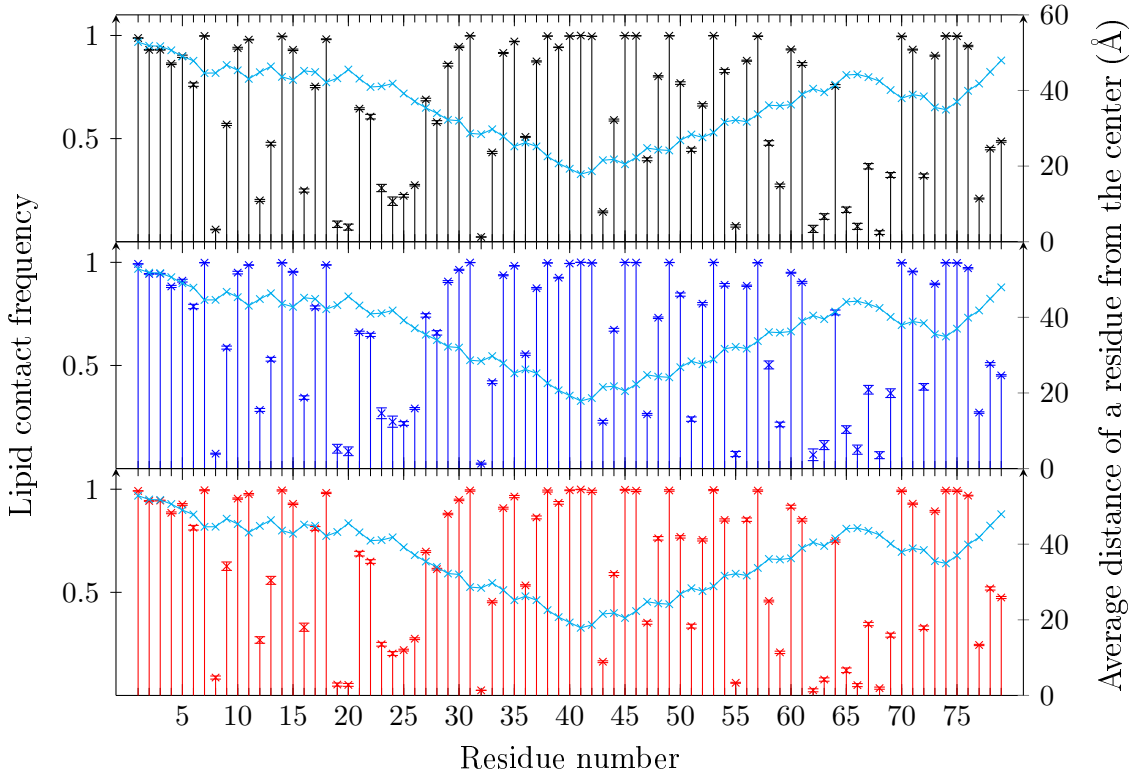


Figure 6.3 The overall lipid contact frequency per residue of Dim_c in lipid self-assembly simulations. DC, EQ, and PHYS lipid compositions, respectively. The average distance of a residue from the center of the 6mer measured from the crystal structure. Results calculated as an average of 24 repetitions each, error as SE.

Functional dimers Dim_f was found to bind lipids in a characteristic orientation with two main lipid interaction sites in the structure (see Fig. 6.4). The membrane binding orientation of Dim_f was independent of the lipid composition, and similar overall lipid contact frequencies were observed with each system (see Fig. 6.5). The four main helical secondary structures in SP-B can be seen in the graph (see *e.g.* Fig. 6.5) for the average distance of a residue from the center of a 6mer as incremental steps at residues from W9 to I22, from A30 to C35, from A42 to G63, and from Q68 to V74. Residues further from the hypothetical center of the 6mer are at the outer face of the monomer, while residues closer to the center of the 6mer are buried within the hydrophobic cavity formed by two adjoined monomers. Thus, the lipid binding orientation of the protein could also be deduced from the difference in trends in lipid binding residues between Dim_c and Dim_f in Figs. 6.3 and 6.5.

Based on our results, the structure of Dim_f has two lipid interaction sites: the *membrane binding* and the *membrane perturbing* site (see Figs. 6.1 and 6.4). Fig. 6.1 shows the structure of Dim_f, where the residues of the membrane binding site and the

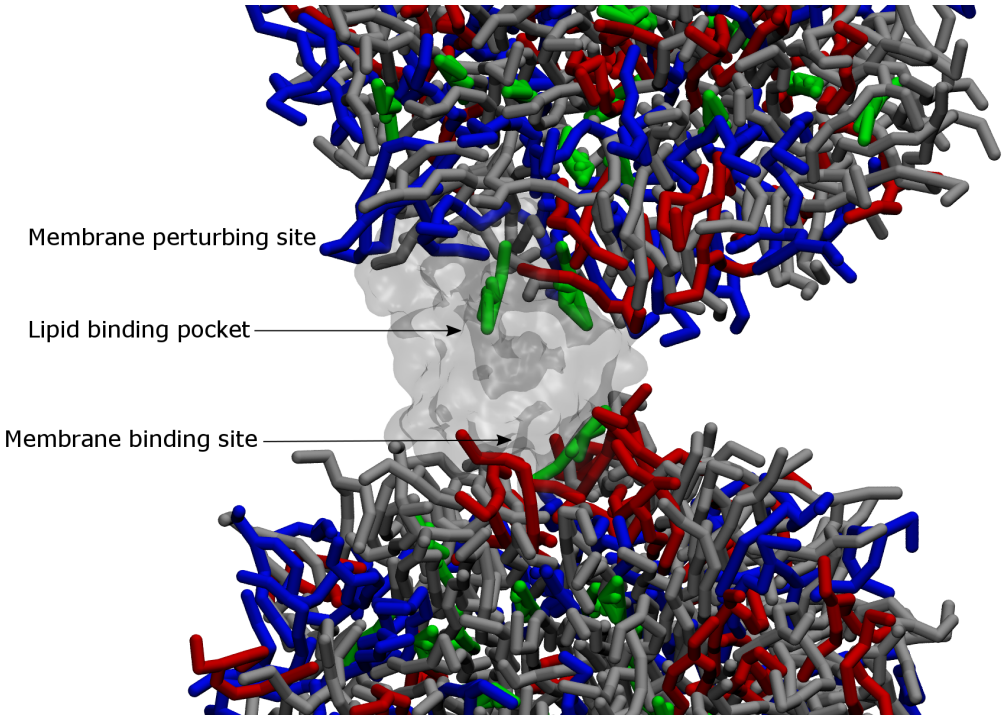


Figure 6.4 Lipid binding configuration of Dim_f shows a possible orientation of binding with lipids for the SP-B dimer based on the SA simulations. Dim_f has two preferential lipid interaction sites on opposite sides of the dimer. A cholesterol molecule (green) is within the lipid binding pocket, which can be seen as a shadow in the center of the protein.

membrane perturbing site are highlighted. The membrane binding sites of individual Dim_f reside at the outer rim of the SP-B hexamer of dimers, while the membrane perturbing sites are facing the center of the hexamer. The experimentally determined important membrane binding residues are located in the membrane binding site. On the other hand, the membrane perturbing site is unique to our model for Dim_f and the SP-B hexamer of dimers. A characteristic feature of the membrane perturbing site observed in the self-assembly simulations was the insertion of lipids in the lipid binding pocket, as seen in Fig. 6.4. Further studies with the SP-B hexamer of dimers show the membrane perturbing effects of the site, discussed further in section 6.2.2.

The membrane binding site of Dim_f refers to the residues located in the N- and the C-termini of the monomers that have been shown to be significant for the correct membrane binding SP-B [94,96]. Based on the overall lipid contact frequency results with Dim_f with different the lipid compositions (see Fig. 6.5), the trends seen in the lipid binding residues of the N-terminus match the proposed lipid binding orientation of SP-B found in experiments [3, 96]. In addition, the lipid binding of the

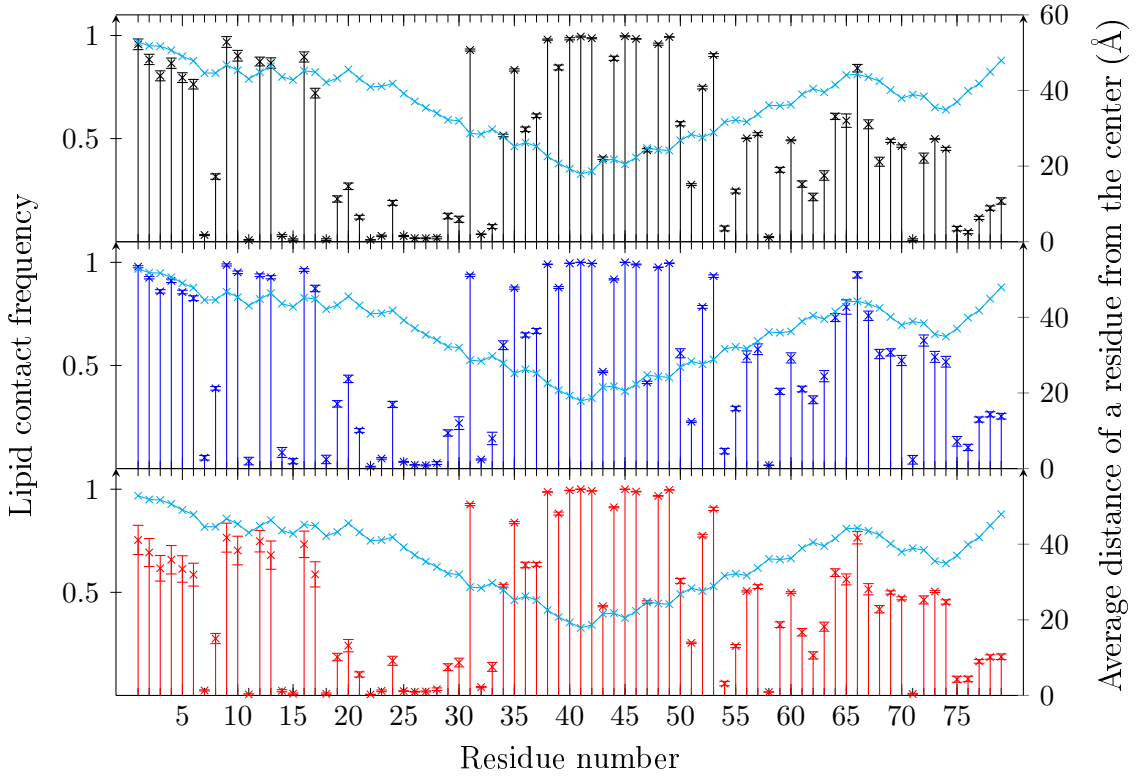


Figure 6.5 The overall lipid contact frequency per residue of Dim_f in lipid self-assembly simulations. DC, EQ, and PHYS lipid compositions, respectively. The average distance of a residue from the center of the 6mer measured from the crystal structure. Results calculated as an average of 24 repetitions each, error as SE.

C-terminus in our Dim_f model is in agreement with experimental and MD simulation studies [88]. The membrane binding site of Dim_f consists of the hydrophobic and positively charged residues in the N- and C-terminal helical structures positioned parallel to the membrane, and the six unstructured hydrophobic residues in the N-terminus. Thus, the suggested orientation of membrane binding of the Dim_f model is in compliance with the superficial membrane interaction topology of SP-B [73].

The lipid-specific interaction contact frequency for each lipid type was analyzed for the Dim_f structure with the EQ lipid composition in the self-assembly simulations for the last 2 µs of the simulation, as shown in Fig. 6.6. The results show specific lipid preference for POPG and cholesterol for certain residues in the previously mentioned interaction sites. POPG shows preferential interaction with the hydrophobic and positively charged membrane binding residues in the N- and C-terminal regions of the protein compared to DPPC or POPC. The positive residues R12, K16, and R17 in the N-terminal region and R72 in the C-terminal region increase the lipid-specific interaction frequency of these residues due to the electrostatic interactions

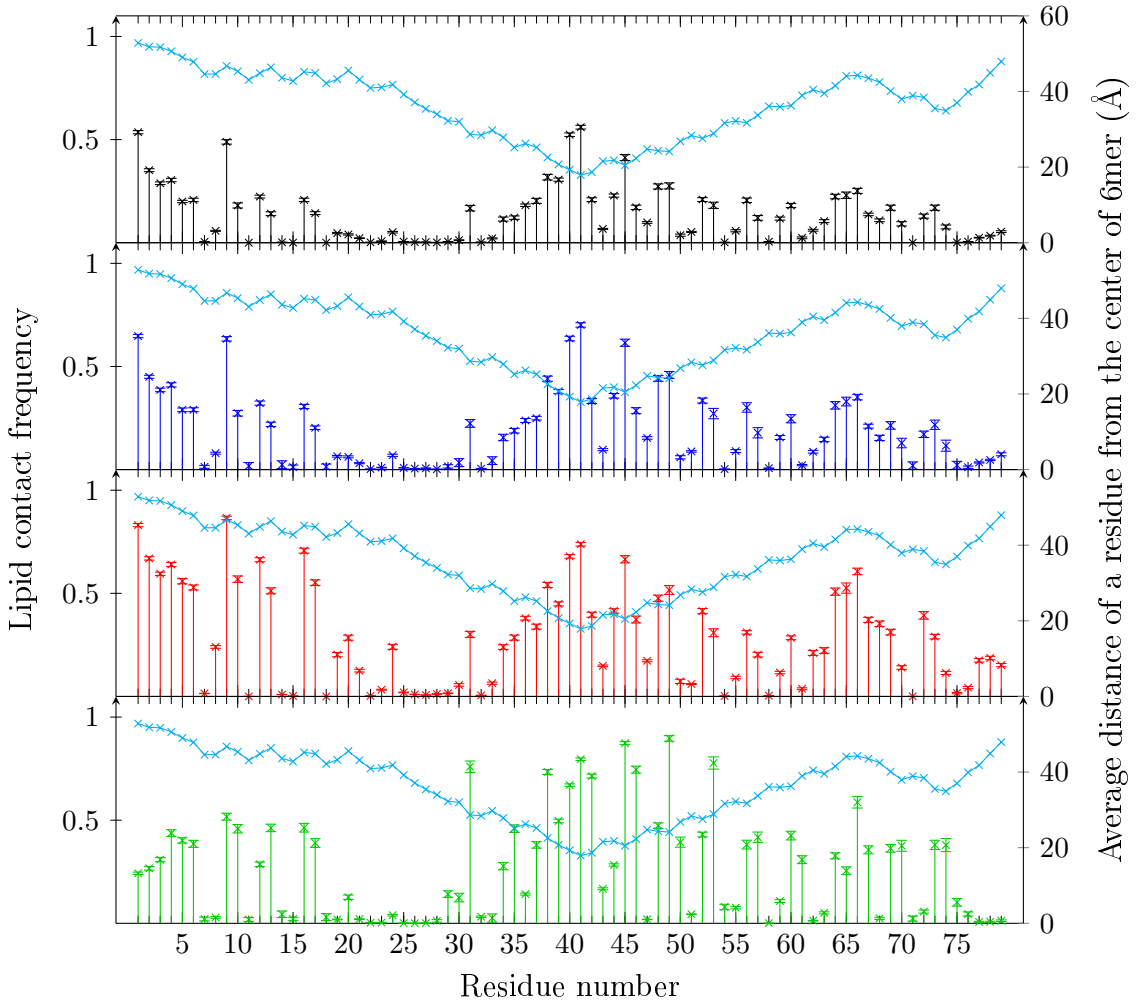


Figure 6.6 Dim_f lipid-specific contact frequency per residue in EQ lipid composition with DPPC, POPC, POPG, and CHOL, respectively. The average distance of a residue from the center of the 6mer measured from the crystal structure. Results calculated as an average of 24 repetitions each, error as SE.

between the anionic phospholipids and the positive residues. An increase in contact frequency with POPG is also seen in the membrane binding sequence from F1 to P6, and with W9. Our results suggest that anionic lipids, especially PG, have an important role in the interaction of the pulmonary surfactant with SP-B, as shown by experiments [72, 190].

The membrane perturbing site of the Dim_f structure is located near the lipid binding pocket of the functional SP-B dimer. To our knowledge, this lipid interaction site in our model for the Dim_f has not been described in previous experimental studies, and its existence is based purely on the homology model with saposin B. The lipid perturbing site of Dim_f is located at the turn in the primary structure,

approximately from V31 to Y53, and includes a high proportion of hydrophobic residues (see Fig. 2.6). The lipid-specific contact frequencies shown in Fig. 6.6 show a preference of interaction for this site with cholesterol. This could be due to a preference of the adjacent lipid binding pocket towards cholesterol. The specific lipid interactions described here for the Dim_f structure indicates a preferential orientation of membrane binding for the SP-B hexamer of dimers and any other possible higher oligomers.

6.1.2 SP-B Hexamer of Dimers Has a Lipid Interaction Site in the Central Pore

Within the 2.5 µs simulation time, the lipid self-assembly simulations with the SP-B hexamer of dimers lead to the formation of both small micellar aggregates, like micelles, bilayer discs and vesicles, and to the formation of lamellar bilayers (see Figs. 6.7 and 6.8). The SP-B hexamer of dimers remained associated with the lipid phase in all simulations, as predicted by the self-assembly simulation results with SP-B dimers. Table 6.1 lists the occurrence of different lipid structures in the simulated systems. Bilayer discs or bicelles were most common with 51 out of 72 simulations. A total of 18 out of these 51 systems with bilayer discs resulted in almost complete bilayers (see Fig 6.7), where SP-B partitioned to the edges of the adjacent membranes, preventing the formation of a complete bilayer. Thus, these systems were calculated as formation of bilayer discs, but are shown also separately as the number of incomplete bilayers out of systems with bilayer discs or bicelles in Table 6.1.

Table 6.1 Formation of different lipid structures in the self-assembly simulations with the SP-B hexamer of dimers. The formation of bilayer discs or bicelles was most common. Incomplete bilayers refer to bilayer discs, which were close to attaining a bilayer structure. Bilayers were observed in a total of 17 simulations out of 72.

Name	**Bilayer discs / bicelles**	**(Incomplete bilayers)**	**Vesicles**	**Bilayers**
DC	22	(8)		2
EQ	13	(4)	2	9
PHYS	16	(6)	2	6
Total (72)	51	(18)	4	17

In the example system in Fig. 6.8, a bilayer assembled within the first 500 ns and maintained an average APL of 63.5 $Å^2$ after assembly for the rest of the 2.5 µs

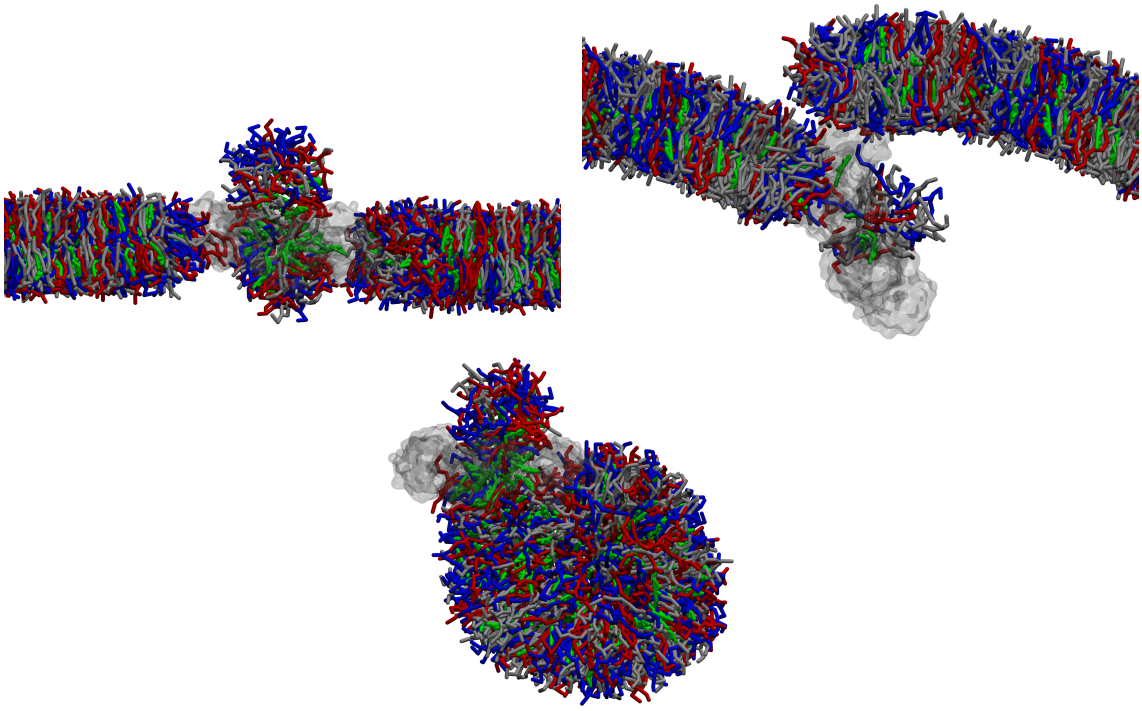


Figure 6.7 Example membrane binding configurations of SP-B hexamer of dimers in lipid self-assembly simulations. Left figure above shows a system where the protein is between two bicelles with a micellar aggregate in the central pore of the ring. Right figure above shows an example of an almost complete bilayer. The figure below shows a vesicle with a bound SP-B hexamer of dimers.

simulation time. The protein remained attached to the membrane while allowing the exchange of a few lipids between the bilayer and the attached micelle through the central pore of the SP-B hexamer of dimers. The formation of a complete bilayer was not observed in the majority of the self-assembly simulations during the simulation time, as shown in Table 6.1.

The self-assembly simulations showed two preferential lipid interaction sites in the SP-B hexamer of dimers, which are equivalent to the sites seen with the Dim_f structure (see Figs. 6.5 and 6.9). The lipid composition of the systems did not significantly affect the overall lipid contact frequency of individual residues. The same lipid interacting residues are involved with membrane binding and lipid interaction in the Dim_f and the 6mer structures, which in the case of the SP-B hexamer of dimers results in a configuration where lipids are within the central pore of the ring (see Figs. 6.1 and 6.8). The lipid-specific contact frequency results (see Appendix B) further show a preference of cholesterol partitioning into the central pore, while POPG interacted preferentially with the membrane binding residues in the N- and

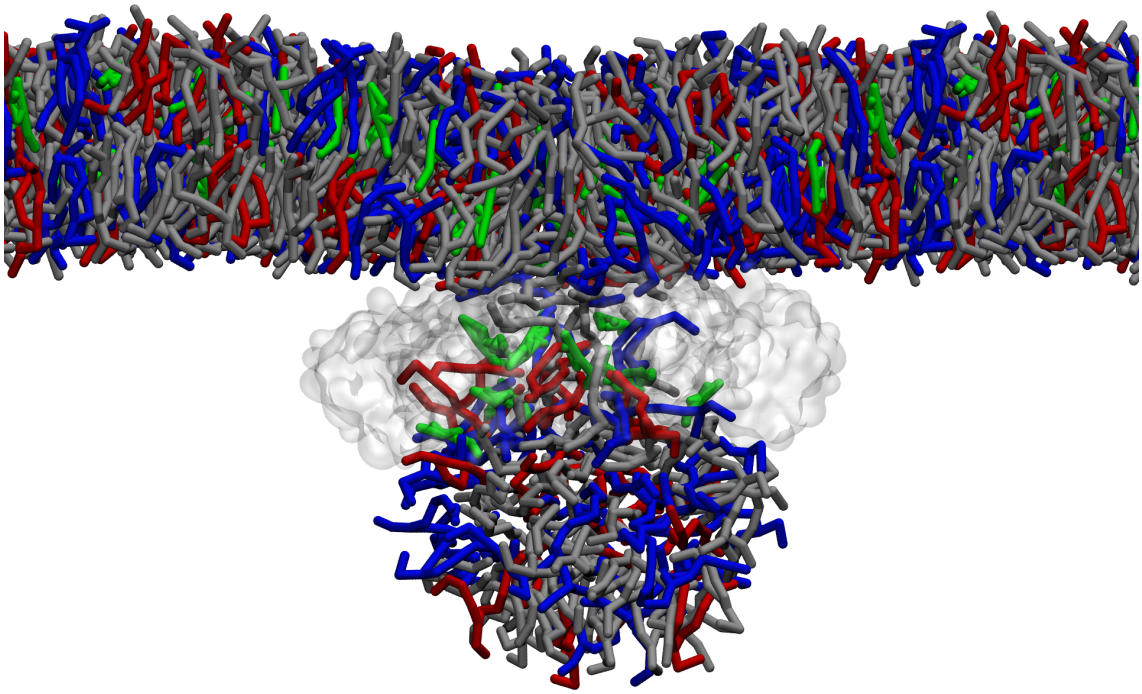


Figure 6.8 SP-B hexamer of dimers connected to a bilayer and a micelle formed in an SA simulation with PHYS lipid composition. Cholesterol (green) partitioned preferentially into the central pore of the hexamer ring in self-assembly simulations (see Appendix B).

C-termini of SP-B. The central pore of the SP-B hexamer of dimers is enclosed by the lipid perturbing sites of the individual dimers. Thus, the lipid binding pockets of the functional dimers are also facing towards the center and may participate in the transfer of lipids across the central pore. Fig. 6.1 recaps these lipid interaction sites in our model for the SP-B hexamer of dimers.

Based on the self-assembly simulation results with our Dim_f and 6mer structures in different lipid compositions, the transfer of lipids between membranes, one of the most important roles of SP-B in the pulmonary surfactant, occurs through the central pore. In the self-assembly simulations with the SP-B hexamer of dimers, we observed an orientation of membrane binding where the axis of the central pore is parallel to the membrane normal (see Fig. 6.8). In this orientation, the active membrane perturbing sites as well as the membrane binding residues on the opposing side of the structure are in contact with their respective membranes. This orientation of membrane binding for the SP-B oligomers is also supported by experimental data [6]. The hydrophobic central pore of the oligomer ring is necessary for the molecular mechanism of function of the SP-B implicated by our results.

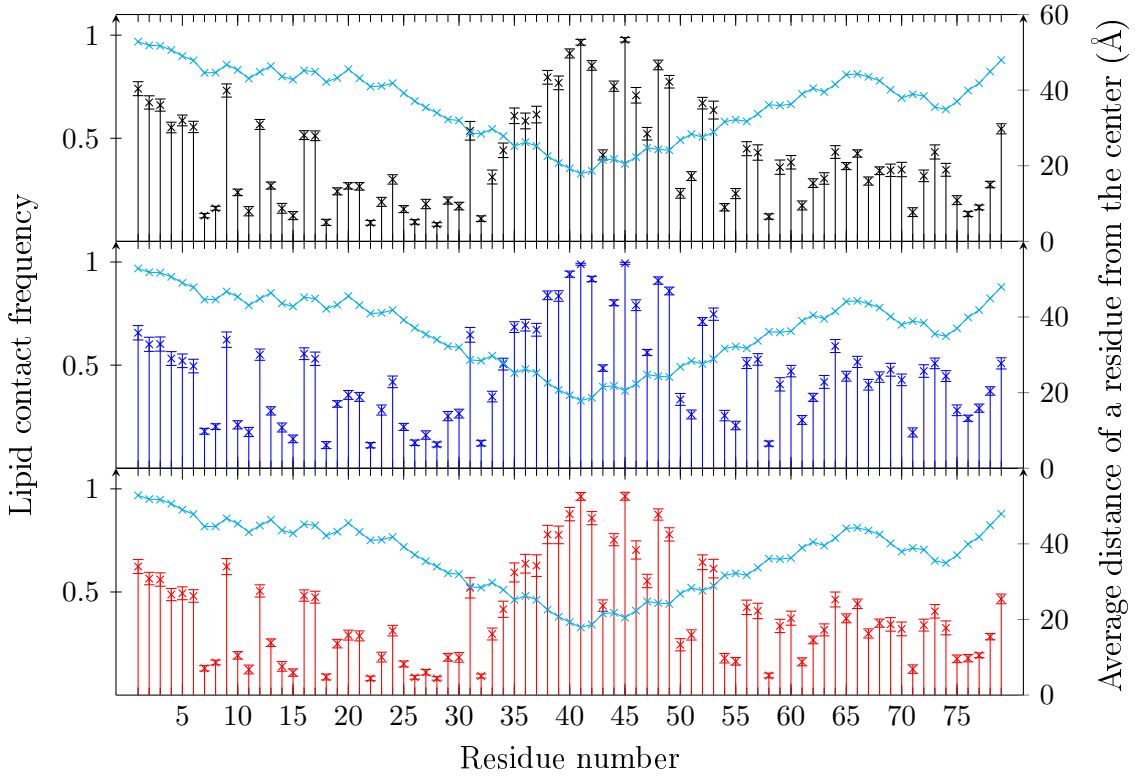


Figure 6.9 SP-B hexamer of dimers overall lipid contact frequency per residue in self-assembly simulations with DC, EQ, and PHYS lipid compositions, respectively. The average distance of a residue from the center of the 6mer measured from the crystal structure. Results calculated as an average of 24 repetitions each, error as SE.

6.2 Monolayer Simulations

The monolayer systems were built to study the effect of SP-B hexamer of dimers on monomolecular films. Pulmonary surfactant forms a monolayer film at the air–liquid interface of the alveoli. We decided that the simplest and computationally least expensive method to study the effect of SP-B hexamer of dimers on monolayers was to build systems, where the monolayers are coupled by the protein in the middle. This experimental simulation setup would allow the study of membrane binding, lipid interactions, and other possible effects of the SP-B hexamer of dimers on monolayers. Different compression states of the pulmonary surfactant were studied by selecting four areas per lipid (APL) *i.e.* surface pressures. All monolayers were identical in the number and type of lipids. Thus, differences seen in the figures are caused by the SP-B hexamer of dimers upon interaction.

6.2.1 SP-B Induces Lateral Lipid Reorganization in Coupled Monolayers

The distance between the selected monolayers ranged from 2.7 nm to 3.8 nm in approximately 0.3 nm steps. In most systems this range was optimal for the interaction of SP-B with both monolayers. The protein remained attached on both sides while a conformational change in the orientation of the dimers turned the membrane binding residues in conjunction towards one monolayer, and the membrane perturbing residues towards the opposite. The area per lipid (APL) did not affect the binding of SP-B with the monolayers.

When forming the monolayers, we used different APLs to investigate how the compressing of the monolayer can effect its interaction with SP-B. The APL of the system affected the integrity of the monolayers. The largest APL of 60 $Å^2$ caused the formation of holes in the monolayers in some simulations. We assume that the holes in the monolayers are mainly caused by the force field's inability to properly account for the energetic penalty of exposing a pure surface, *i.e.* the surface tension ($\sim$30 mN/m in MARTINI [149] vs. the experimental value of 70 mN/m in pure water), as similar behavior was observed in the preparation of the monolayers with APLs of over 60 $Å^2$, even before the addition of SP-B into the system. On the other hand, the formation of proteolipid pores in membranes has been reported in the presence of SP-B [100]. Thus, some of the repetitions of the 60 $Å^2$ APL monolayer simulations might not represent realistic number densities in the monolayer plane and are mostly omitted from the comparison of the results in this section. Based on our results, monolayers with our PHYS lipid composition were not suitable to be simulated with an APL of over 60 $Å^2$ with the MARTINI force field.

The lateral lipid distribution results for two representative systems with 52.5 $Å^2$ and 57.5 $Å^2$ APL are shown in Figs. 6.10 and 6.11, respectively. The results are presented as an average of three repetitions and normalized per lipid type in the PHYS lipid composition. These results show the lateral reorganization of POPG and cholesterol around SP-B in all systems. This lipid specificity was independent of the APL of the monolayers and can be seen in all systems with all four APLs. The lateral lipid distribution results for the rest of the systems are included in Appendix A for reference.

The membrane binding residues caused the enrichment of POPG around the area of the monolayer in contact with the protein, seen as a ring of higher POPG density in the upper monoalyer in Figs. 6.10 and 6.11. This preference for POPG was

also observed in the lipid-specific contact frequency results shown in section 6.1.1. Thus, POPG is likely to have an important role in the correct binding of SP-B to pulmonary surfactant membranes. POPG also binds preferentially to the membrane perturbing site of SP-B, which can be seen as an increased density in and around the central pore on the lower monolayer in Figs. 6.10 and 6.11.

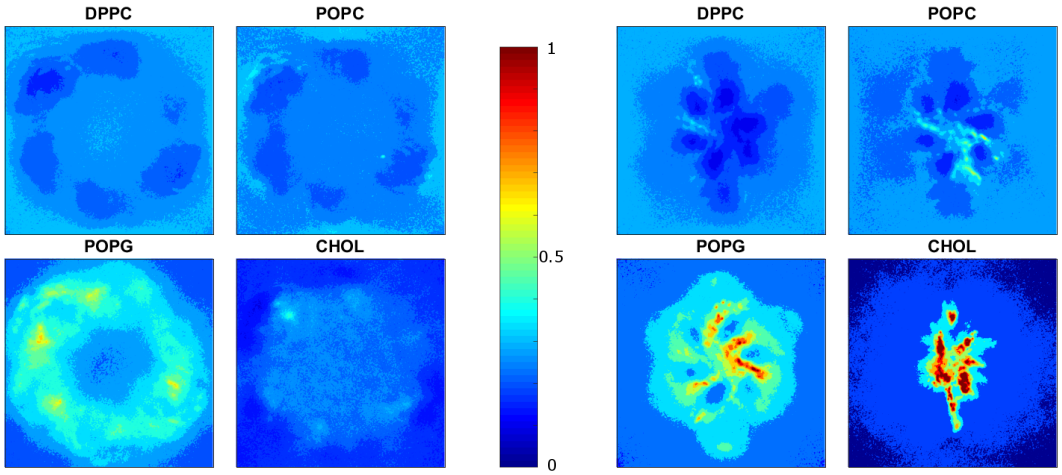


Figure 6.10 Lateral lipid distribution in a 52.5 $Å^2$ APL system, represented by number density maps for the upper (left) and the lower (right) monolayer. The distance of the monolayers was 3.6 nm. Results are normalized per lipid type to show specific lipid interactions with SP-B.

Our model for the SP-B hexamer of dimers induced lateral organization of POPG that complies to experimental results [63, 72]. The enhanced POPG interactions with the membrane binding surface (see Fig. 6.1) of the SP-B oligomer on the upper monolayer in the figures likely plays a role in establishing the correct binding mode with the pulmonary surfactant and directionality of lipid transfer (see Figs. 6.10 and 6.11). POPG interaction is also enhanced on the membrane perturbing surface of the protein (see Fig. 6.1), seen in the figures on the lower monolayer. The membrane binding surface has five or six positively charged residues per monomer (R12, K16, R17, R64, R72, and R76) facing the membrane, while there is one positively charged residue per monomer (R36) in the membrane perturbing surface. The positions of these charged residues in the SP-B hexamer of dimers affected the lateral distribution of POPG in our monolayer simulations. These results are in concordance with the unusually high concentration of anionic phospholipids in the pulmonary surfactant compared to other mammalian membranes [3, 24].

Cholesterol partitioned preferentially into the area of perturbed lipids in the cen-

tral pore of the protein as shown in Figs. 6.10 and 6.11 in the lower monolayers. The 52.5 Å^2–APL system showed an especially high density of CHOL in the central pore in the perturbed area of the monolayer, whereas in the 57.5 Å^2–APL system, an increased density of CHOL was observed also inside the lipid binding pockets of the dimers, which can be seen as local clusters around the perturbed area. In the other systems CHOL clustered into these two sites, but the APL did not affect the preference for either (see appendix A). Based on our resuls, cholesterol seems an important component in the function of SP-B in the surfactant, although it is excluded from some clinical surfactants used in SRT even in the precence of the protein [10].

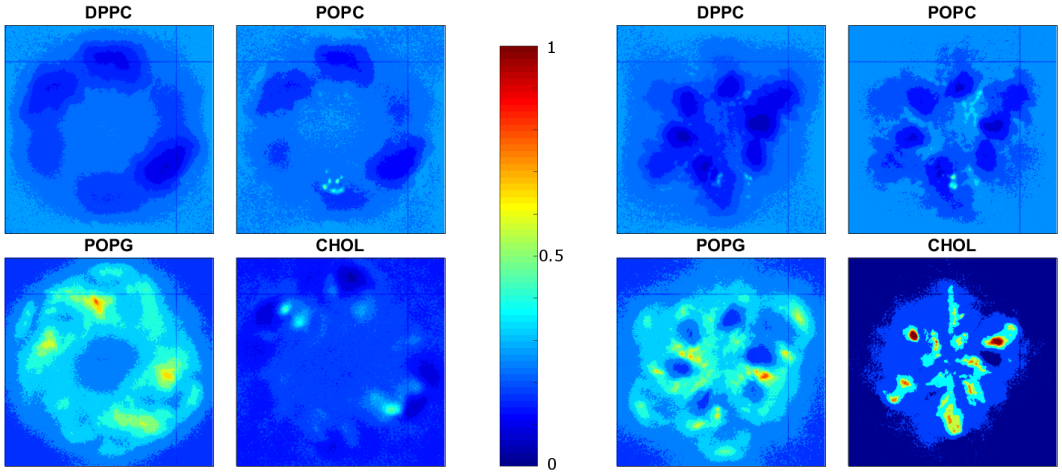


Figure 6.11 Lateral lipid distribution in a 57.5 Å^2 APL system, represented by number density maps for the upper (left) and the lower (right) monolayer. The distance of the monolayers was 3.4 nm. Results are normalized per lipid type to show specific lipid interactions with SP-B.

The compression state of the monolayer affects the packing of the lipids, which leads to a more ordered L_c phase at low molecular areas. In our simulations, the compression state of the monolayers did not substantially affect the lateral distribution of lipids around SP-B. Similar trends in POPG and cholesterol clustering were observed in all systems. Thus, we propose that cholesterol as well as POPG are involved in the molecular mechanism of the function of the protein, and especially POPG is important in the binding of the SP-B oligomer to membranes.

6.2.2 SP-B Mediates Lipid Perturbations in Monolayers

In the monolayer simulations, SP-B bound to the monolayers on both sides of the protein. One monolayer was in contact with the membrane binding surface of the SP-B hexamer of dimers structure, and the other with the membrane perturbing surface of the protein. Upon interaction, the membrane perturbing surface of SP-B caused perturbations and lipid protrusions of variable size to the monolayer in contact with the protein membrane perturbing surface. In some systems, SP-B induced the formation of a larger lipid protrusion or a lipid neck to the monolayer, as shown in Figs. 6.12 and 6.13. This protrusion was dependent on the conformational changes in the protein structure during the simulation, and occurred on the side of the membrane perturbing surface of SP-B in contact with the monolayer. The extent of the lipid neck observed in the simulations varied from small protrusions in the monolayer to an almost complete contact between the lipids of opposite monolayers. In 25 µs each system was simulated for, we did not observe lipid exchange between the monolayers. This is likely due to the lack of a driving force, such as a concentration or stress gradient.

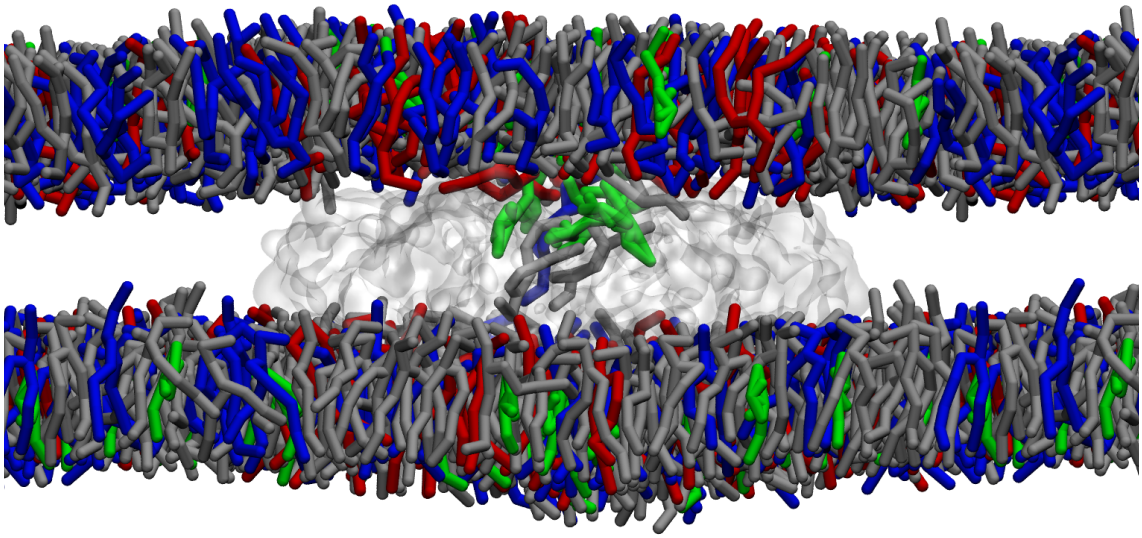


Figure 6.12 Lipid perturbations in monolayers caused by SP-B. Cholesterol (green) partitions preferentially into the hydrophobic central pore of the SP-B hexamer of dimers. Here, the APL is equal to 57.5 $Å^2$.

Formation of the lipid neck was more likely to occur with lower APL, but smaller lipid protrusions were observed with all tested APLs. Thus, the applied surface pressure may cause larger protrusions in the monolayers mediated by SP-B, which would

push lipids through the protein during compression of the monolayer in exhalation. Fig. 6.8 shows a possible intermediate state of lipid flow through SP-B that could form when enough surface pressure is applied to the system. It has been suggested that SP-B promotes connections between surfactant membranes and the monolayer at the air–liquid interface [3, 22]. These SP-B mediated connections (see Fig. 2.8) could allow favorable partitioning of oxygen into the continuous lipid phase, which could permit rapid O_2 diffusion through the hydrophobic core of the membranes, as suggested by Olmeda *et al.* [4]. The formation and the stability of these small lipid reservoirs connected to surfactant monolayers or membranes should be further studied.

As mentioned earlier, higher APL caused the formation of small holes is the monolayers (see Fig. 6.13). There is experimental evidence of SP-B causing pore formation in membranes [100], but in our case we concluded that the holes are more probably caused by the force field's inability to account for the surface tension. On the other hand, the membrane poration activity of SP-B has been suggested to be highly indicative of the oligomeric structure of the protein [100]. Thus some form of pore or hole formation could be characteristic to our model for the SP-B hexamer of dimers.

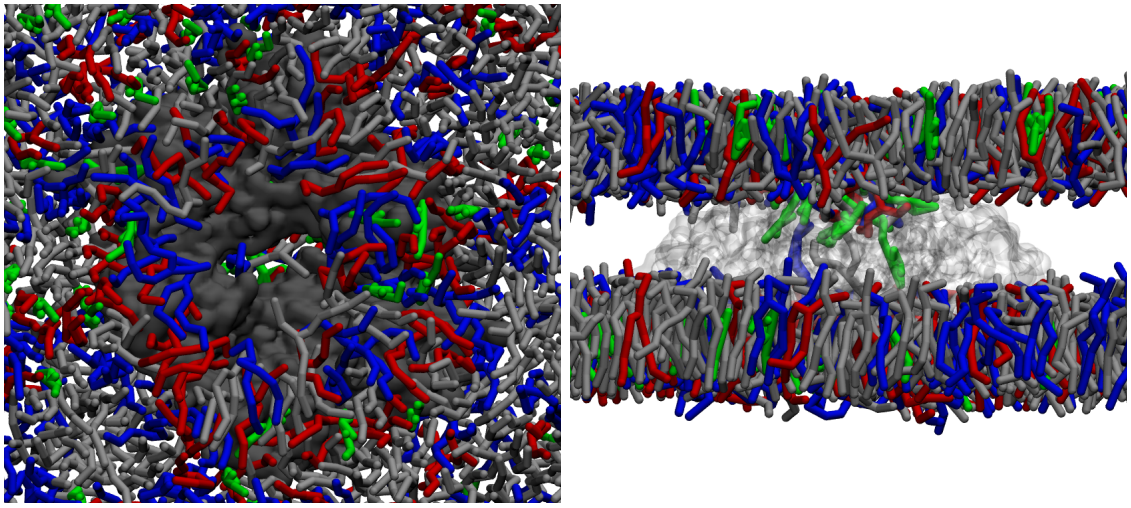


Figure 6.13 High APL caused the formation of holes in monolayers with SP-B. Cholesterol (green) partitioned preferentially in the perturbed lipids in the central pore. Here, the APL is equal to 60 $Å^2$.

7. CONCLUSIONS

The aim of this thesis was to study the specific lipid interactions and th binding of SP-B in pulmonary surfactant membranes. To this end, we used coarse-grained MARTINI MD simulations. Due to the extreme hydrophobicity and difficulty of extraction, only low resolution homology models for monomers and dimers of SP-B exist in the literature. A recent study revealed the presence of a ring-shaped multidimeric SP-B structure in pulmonary surfactant, and provided a novel structure for a SP-B hexamer of dimers [6]. This study concentrates in our refined model for the structure of the SP-B hexamer of dimers that is based on the above-mentioned structure. To our knowledge, this study is pioneering concerning the higher oligomers of SP-B. All previous MD studies have concentrated on monomers and covalent dimers of SP-B [102,104,112], or even smaller portions of the protein [88,191]. These studies have given some insight to the possible molecular mechanism of function of SP-B, but have not provided an unambiguous protein model derived from homology modeling.

The focus of this thesis was on the specific lipid interactions of SP-B, which will work as a basis for further studies concerning the larger oligomeric structures of SP-B. Two types of systems were simulated for this purpose: lipid self-assembly in the presence of the protein, and monolayers coupled by SP-B in a sandwich-like conformation. A total of three protein structures were studied to find out the correct orientation of membrane binding. These structures were the covalently bound dimer, the functional dimer that is the basis for the multidimeric oligomer, and our refined SP-B hexamer of dimers.

The results showed that the dimers have two preferential lipid binding configurations. The covalent dimer prefers a planar orientation on the surface of a membrane and was observed to partition to the edge of bilayer discs. On the other hand, the functional dimer had two membrane binding sites on the opposing sides of the structure. We call these sites the *membrane binding site* and the *membrane perturbing site*. The membrane binding site corresponds to the N- and C-terminal helices in SP-B that have been shown to orient parallel to membranes [88, 191] and contain the residues essential for the correct binding and orientation of SP-B in

surfactant membranes [11, 92]. The membrane perturbing site is on the opposing side of the dimer, adjacent to the lipid binding pocket. The results show that the membrane perturbing site causes protrusions and smaller perturbations in the affected membrane. Residues in the membrane binding site showed preference for POPG, whereas cholesterol was enriched in the membrane perturbing site and the lipid binding pocket.

Self-assembly simulations with the SP-B hexamer of dimers showed a possible planar orientation of membrane binding, which would allow the flow of lipids through the hydrophobic central pore of the oligomer ring, also suggested by Olmeda et al. [6]. These results indicate a possible model for the molecular function for the higher oligomers of SP-B that has not been described earlier. SP-B caused similar perturbations in the coupled monolayers, which lead to the formation of a lipid neck inside the central pore. SP-B induced lateral lipid reorganization in the monolayers and showed preference for POPG and cholesterol. POPG clustered near the membrane binding and the membrane perturbing sites on opposite monoalyers indicating its importance for the binding and function for the protein, while cholesterol was found in increased concentrations in the central pore and the lipid binding pockets of the SP-B hexamer of dimers.

This study can work as a basis for many future MD simulation studies on higher oligomeric structures of SP-B. The implicated molecular model of function for the protein should be further studied, and the systems extended to *e.g.* coupled bilayers, monolayers connected to lipid reservoirs, and other biologically relevant bilayer and monolayer setups. The simulations should also be repeated in atomistic detail to validate the results and to gain better molecular view of the interactions. Free energy profiles of the SP-B membrane binding process would provide more detail of the correct orientation of binding and depth of insertion into membranes. Furthermore, the study of SP-B could provide information about the molecular mechanism of action for all of the SAPLIP proteins.

BIBLIOGRAPHY

[1] **Parra, E. and Pérez-Gil, J.**, Composition, structure and mechanical properties define performance of pulmonary surfactant membranes and films. *Chemistry and Physics of Lipids*, 185: 153–175. 2015.

[2] **Goerke, J.**, Pulmonary surfactant: Functions and molecular composition. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1408(2-3): 79–89. 1998.

[3] **Serrano, A. and Pérez-Gil, J.**, Protein-lipid interactions and surface activity in the pulmonary surfactant system. *Chemistry and Physics of Lipids*, 141(1-2): 105–118. 2006.

[4] **Olmeda, B., Villén, L., Cruz, A., Orellana, G., and Pérez-Gil, J.**, Pulmonary surfactant layers accelerate O2 diffusion through the air-water interface. *Biochimica et Biophysica Acta - Biomembranes*, 1798(6): 1281–1284. 2010.

[5] **Olmeda, B., García-Álvarez, B., and Pérez-Gil, J.**, Structure-function correlations of pulmonary surfactant protein SP-B and the saposin-like family of proteins. *European Biophysics Journal*, 42(2-3): 209–222. 2013.

[6] **Olmeda, B., García-Álvarez, B., Gómez, M., Martínez-Calle, M., Cruz, A., and Pérez-Gil, J.**, A model for the structure and mechanism of action of pulmonary surfactant protein B. *FASEB Journal*, 29(10): 4236–4247. 2015.

[7] **Nogee, L., Garnier, G., Dietz, H., Singer, L., Murphy, A., DeMello, D., and Colten, H.**, A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds. *Journal of Clinical Investigation*, 93(4): 1860–1863. 1994.

[8] **Klein, J.M., Thompson, M.W., Snyder, J.M., George, T.N., Whitsett, J.A., Bell, E.F., Jr., P.B.M., and Nogee, L.M.**, Transient surfactant protein B deficiency in a term infant with severe respiratory failure. *The Journal of Pediatrics*, 132(2): 244 – 248. 1998.

[9] **World Health Organization**, Preterm birth. 2016, available: `http://www.who.int/mediacentre/factsheets/fs363/en/` Cited November 15th 2016.

[10] **Zhang, H., Fan, Q., Wang, Y.E., Neal, C.R., and Zuo, Y.Y.**, Comparative study of clinical pulmonary surfactants using atomic force microscopy. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(7): 1832 – 1842. 2011.

[11] **Sarker, M., Waring, A., Walther, F., Keough, K., and Booth, V.**, Structure of mini-B, a functional fragment of surfactant protein B, in detergent micelles. *Biochemistry*, 46(39): 11047–11056. 2007.

[12] **Marieb, E.N.**, *Essentials of Human Anatomy & Physiology*. Pearson/Benjamin Cummings, San Farncisco, 9th - 10th international 2012. edition. 2009.

[13] **Tortora, G.J. and Grabowski, S.R.**, *Principles of Anatomy and Physiology*. Wiley, New York (NY), 10th edition. 2003.

[14] **Herman, I.P.**, *Physics of the Human Body*. Springer. 2007.

[15] **Haug, E., Sjaastad, Ö.V., Sand, O., Toverud, K.C., and Sillman, K.**, *Ihmisen fysiologia*. Sanoma Pro, Helsinki, 1.-5. p. edition. 2012.

[16] Respiratory system. 2013, available: `https://commons.wikimedia.org/wiki/File:Respiratory_system_complete_no_labels.svg` Cited January 2nd 2017.

[17] **Ochs, M., Nyengaard, J., Jung, A., Knudsen, L., Voigt, M., Wahlers, T., Richter, J., and Gundersen, H.**, The number of alveoli in the human lung. *American Journal of Respiratory and Critical Care Medicine*, 169(1): 120–124. 2004.

[18] **Mason, R.**, Biology of alveolar type II cells. *Respirology*, 11(SUPPL. 1): S12–S15. 2006.

[19] **Van Golde, L., Batenburg, J., and Robertson, B.**, The pulmonary surfactant system: Biochemical aspects and functional significance. *Physiological Reviews*, 68(2): 374–453. 1988.

[20] **Casals, C. and Cañadas, O.**, Role of lipid ordered/disordered phase coexistence in pulmonary surfactant function. *Biochimica et Biophysica Acta - Biomembranes*, 1818(11): 2550–2562. 2012.

[21] **Nelson, D.L., Lehninger, A.L., and Cox, M.M.**, *Lehninger Principles of Biochemistry*. Macmillan. 2008.

[22] **Pérez-Gil, J.**, Structure of pulmonary surfactant membranes and films: The role of proteins and lipid-protein interactions. *Biochimica et Biophysica Acta - Biomembranes*, 1778(7-8): 1676–1695. 2008.

[23] **Daniels, C. and Orgeig, S.**, Pulmonary surfactant: The key to the evolution of air breathing. *News in Physiological Sciences*, 18(4): 151–157. 2003.

[24] **Veldhuizen, R., Nag, K., Orgeig, S., and Possmayer, F.**, The role of lipids in pulmonary surfactant. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1408(2-3): 90–108. 1998.

[25] **Ariki, S., Nishitani, C., and Kuroki, Y.**, Diverse functions of pulmonary collectins in host defense of the lung. *Journal of Biomedicine and Biotechnology*, 2012. 2012.

[26] **Pérez-Gil, J. and Keough, K.**, Interfacial properties of surfactant proteins. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1408(2-3): 203–217. 1998.

[27] **Schmitz, G. and Muller, G.**, Structure and function of lamellar bodies, lipid-protein complexes involved in storage and secretion of cellular lipids. *Journal of Lipid Research*, 32(10): 1539–1570. 1991.

[28] **Yu, S.H. and Possmayer, F.**, Lipid compositional analysis of pulmonary surfactant monolayers and monolayer-associated reservoirs. *Journal of Lipid Research*, 44(3): 621–629. 2003.

[29] **Lopez-Rodriguez, E. and Pérez-Gil, J.**, Structure-function relationships in pulmonary surfactant membranes: From biophysics to therapy. *Biochimica et Biophysica Acta - Biomembranes*, 1838(6): 1568–1585. 2014.

[30] **Vincent, J.**, Raman spectroscopic studies of model human pulmonary surfactant systems: Phospholipid interactions with peptide paradigms for the surfactant protein SP-B. *Biochemistry*, 30(34): 8395–8401. 1991.

[31] **De La Serna, J., Perez-Gil, J., Simonsen, A., and Bagatolli, L.**, Cholesterol rules: Direct observation of the coexistence of two fluid phases in native pulmonary surfactant membranes at physiological temperatures. *Journal of Biological Chemistry*, 279(39): 40715–40722. 2004.

[32] **Wüstneck, R., Perez-Gil, J., Wüstneck, N., Cruz, A., Fainerman, V., and Pison, U.**, Interfacial properties of pulmonary surfactant layers. *Advances in Colloid and Interface Science*, 117(1-3): 33–58. 2005.

[33] **Jacobson, K. and Papahadjopoulos, D.**, Phase transitions and phase separations in phospholipid membranes induced by changes in temperature, pH, and concentration of bivalent cations. *Biochemistry*, 14(1): 152–161. 1975.

[34] **Bernardino De La Serna, J., Vargas, R., Picardi, V., Cruz, A., Arranz, R., Valpuesta, J., Mateu, L., and Pérez-Gil, J.**, Segregated ordered lipid phases and protein-promoted membrane cohesivity are required for pulmonary surfactant films to stabilize and protect the respiratory surface. *Faraday Discussions*, 161: 535–548. 2012.

[35] **Davis, P., Fleming, B., Coolbear, K., and Keough, K.**, Gel to liquid-crystalline transition temperatures of water dispersions of two pairs of positional isomers of unsaturated mixed-acid phosphatidylcholines. *Biochemistry*, 20(12): 3633–3636. 1981.

[36] **Veatch, S.L. and Keller, S.L.**, Seeing spots: Complex phase behavior in simple membranes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1746(3): 172 – 185. 2005.

[37] **Kahya, N., Scherfeld, D., Bacia, K., Poolman, B., and Schwille, P.**, Probing lipid mobility of raft-exhibiting model membranes by fluorescence correlation spectroscopy. *Journal of Biological Chemistry*, 278(30): 28109–28115. 2003.

[38] **Worthman, L.A., Nag, K., Davis, P., and Keough, K.**, Cholesterol in condensed and fluid phosphatidylcholine monolayers studied by epifluorescence microscopy. *Biophysical Journal*, 72(6): 2569–2580. 1997.

[39] **Filippov, A., Orädd, G., and Lindblom, G.**, Lipid lateral diffusion in ordered and disordered phases in raft mixtures. *Biophysical Journal*, 86(2): 891–896. 2004.

[40] **Schief, W., Antia, M., Discher, B., Hall, S., and Vogel, V.**, Liquid-crystalline collapse of pulmonary surfactant monolayers. *Biophysical Journal*, 84(6): 3792–3806. 2003.

[41] **Fleming, B.D. and Keough, K.M.**, Surface respreading after collapse of monolayers containing major lipids of pulmonary surfactant. *Chemistry and Physics of Lipids*, 49(1): 81 – 86. 1988.

[42] **Cruz, A., Worthman, L.A., Serrano, A., Casals, C., Keough, K., and Pérez-Gil, J.**, Microstructure and dynamic surface properties of surfactant protein SP-B/dipalmitoylphosphatidylcholine interfacial films spread from lipid-protein bilayers. *European Biophysics Journal*, 29(3): 204–213. 2000.

[43] **Cruz, A., Casals, C., Keough, K., and Perez-Gil, J.**, Different modes of interaction of pulmonary surfactant protein SP-B in phosphatidylcholine bilayers. *Biochemical Journal*, 327(1): 133–138. 1997.

[44] **Takamoto, D., Lipp, M., Von Nahmen, A., Lee, K., Waring, A., and Zasadzinski, J.**, Interaction of lung surfactant proteins with anionic phospholipids. *Biophysical Journal*, 81(1): 153–169. 2001.

[45] **Schram, V. and Hall, S.**, Thermodynamic effects of the hydrophobic surfactant proteins on the early adsorption of pulmonary surfactant. *Biophysical Journal*, 81(3): 1536–1546. 2001.

[46] **Ross, M., Krol, S., Janshoff, A., and Galla, H.J.**, Kinetics of phospholipid insertion into monolayers containing the lung surfactant proteins SP-B or SP-C. *European Biophysics Journal*, 31(1): 52–61. 2002.

[47] **Haagsman, H. and Diemel, R.**, Surfactant-associated proteins: Functions and structural variation. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 129(1): 91–108. 2001.

[48] **Bräuer, L., Johl, M., Paulsen, F., Börgermann, J., Pleyer, U., and Tsokos, M.**, Detection and localization of the hydrophobic surfactant proteins B and C in human tear fluid and the human lacrimal system. *Current Eye Research*, 32(11): 931–938. 2007.

[49] **Bräuer, L., Kindler, C., Jäger, K., Sel, S., Nölle, B., Pleyer, U., Ochs, M., and Paulsen, F.**, Detection of surfactant proteins A and D in human tear fluid and the human lacrimal system. *Investigative Ophthalmology and Visual Science*, 48(9): 3945–3953. 2007.

[50] **Kingma, P. and Whitsett, J.**, In defense of the lung: surfactant protein A and surfactant protein D. *Current Opinion in Pharmacology*, 6(3): 277–283. 2006.

[51] **Poulain, F., Allen, L., Williams, M., Hamilton, R., and Hawgood, S.**, Effects of surfactant apolipoproteins on liposome structure: Implications for tubular myelin formation. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 262(6 6-6): L730–L739. 1992.

[52] **Korfhagen, T.R., LeVine, A.M., and Whitsett, J.A.**, Surfactant protein A (SP-A) gene targeted mice. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1408(2-3): 296–302. 1998.

[53] **Johansson, J. and Curstedt, T.**, Molecular structures and interactions of pulmonary surfactant components. *European Journal of Biochemistry*, 244(3): 675–693. 1997.

[54] **Wang, Z., Gurel, O., Baatz, J., and Notter, R.**, Differential activity and lack of synergy of lung surfactant proteins SP-B and SP-C in interactions with phospholipids. *Journal of Lipid Research*, 37(8): 1749–1760. 1996.

[55] **Hobi, N., Giolai, M., Olmeda, B., Miklavc, P., Felder, E., Walther, P., Dietl, P., Frick, M., Pérez-Gil, J., and Haller, T.**, A small key unlocks a heavy door: The essential function of the small hydrophobic proteins SP-B and SP-C to trigger adsorption of pulmonary surfactant lamellar bodies. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1863(8): 2124–2134. 2016.

[56] **Wert, S., Whitsett, J., and Nogee, L.**, Genetic disorders of surfactant dysfunction. *Pediatric and Developmental Pathology*, 12(4): 253–274. 2009.

[57] **Clark, J., Wert, S., Bachurski, C., Stahlman, M., Stripp, B., Weaver, T., and Whitsett, J.**, Targeted disruption of the surfactant protein B gene disrupts surfactant homeostasis, causing respiratory failure in newborn mice. *Proceedings of the National Academy of Sciences of the United States of America*, 92(17): 7794–7798. 1995.

[58] **Stahlman, M., Gray, M., Falconieri, M., Whitsett, J., and Weaver, T.**, Lamellar body formation in normal and surfactant protein B-deficient fetal mice. *Laboratory Investigation*, 80(3): 395–403. 2000.

[59] **Melton, K., Nesslein, L., Ikegami, M., Tichelaar, J., Clark, J., Whitsett, J., and Weaver, T.**, SP-B deficiency causes respiratory failure in adult mice. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 285(3 29-3): L543–L549. 2003.

[60] **Pérez-Gil, J. and Weaver, T.**, Pulmonary surfactant pathophysiology: Current models and open questions. *Physiology*, 25(3): 132–141. 2010.

[61] **Vandenbussche, G., Clercx, A., Curstedt, T., Johansson, J., Jörnvall, H., and Ruysschaert, J.M.**, Structure and orientation of the surfactant-associated protein C in a lipid bilayer. *European Journal of Biochemistry*, 203(1-2): 201–209. 1992.

[62] **Johansson, J., Szyperski, T., Curstedt, T., and Wuthrich, K.**, The NMR structure of the pulmonary surfactant-associated polypeptide SP-C in an apolar solvent contains a valyl-rich α-helix. *Biochemistry*, 33(19): 6015–6023. 1994.

[63] **Pérez-Gil, J., Casals, C., and Marsh, D.**, Interactions of hydrophobic lung surfactant proteins SP-B and SP-C with dipalmitoylphosphatidylcholine and dipalmitoylphosphatidylglycerol bilayers studied by electron spin resonance spectroscopy. *Biochemistry*, 34(12): 3964–3971. 1995.

[64] **Bi, X., Flach, C., Pérez-Gil, J., Plasencia, I., Andreu, D., Oliveira, E., and Mendelsohn, R.**, Secondary structure and lipid interactions of the N-terminal segment of pulmonary surfactant SP-C in langmuir films: IR reflection-absorption spectroscopy and surface pressure studies. *Biochemistry*, 41(26): 8385–8395. 2002.

[65] **Creuwels, L., Boer, E., Demel, R., Van Golde, L., and Haagsman, H.**, Neutralization of the positive charges of surfactant protein C: Effects on structure and function. *Journal of Biological Chemistry*, 270(27): 16225–16229. 1995.

[66] **Gonzalez-Horta, A., Andreu, D., Morrow, M., and Perez-Gil, J.**, Effects of palmitoylation on dynamics and phospholipid-bilayer-perturbing properties of the N-terminal segment of pulmonary surfactant protein SP-C as shown by 2H-NMR. *Biophysical Journal*, 95(5): 2308–2317. 2008.

[67] **Oosterlaken-Dijksterhuis, M.**, Characterization of lipid insertion into monomolecular layers mediated by lung surfactant proteins SP-B and SP-C. *Biochemistry*, 30(45): 10965–10971. 1991.

[68] **Bruhn, H.**, A short guided tour through functional and structural features of saposin-like proteins. *Biochemical Journal*, 389(2): 249–257. 2005.

[69] **Johansson, J., Curstedt, T., and Jornvall, H.**, Surfactant protein B: Disulfide bridges, structural properties, and kringle similarities. *Biochemistry*, 30(28): 6917–6921. 1991.

[70] **Wüstneck, N., Wüstneck, R., Pérez-Gil, J., and Pison, U.**, Effects of oligomerization and secondary structure on the surface behavior of pulmonary surfactant proteins SP-B and SP-C. *Biophysical Journal*, 84(3): 1940–1949. 2003.

[71] **Hawgood, S., Derrick, M., and Poulain, F.**, Structure and properties of surfactant protein B. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1408(2-3): 150–160. 1998.

[72] **Baatz, J., Elledge, B., and Whitsett, J.**, Surfactant protein SP-B induces ordering at the surface of model membrane bilayers. *Biochemistry*, 29(28): 6714–6720. 1990.

[73] **Cabré, E., Loura, L., Fedorov, A., Pérez-Gil, J., and Prieto, M.**, Topology and lipid selectivity of pulmonary surfactant protein SP-B in membranes: Answers from fluorescence. *Biochimica et Biophysica Acta - Biomembranes*, 1818(7): 1717–1725. 2012.

[74] **Nag, K., Taneva, S., Perez-Gil, J., Cruz, A., and Keough, K.**, Combinations of fluorescently labeled pulmonary surfactant proteins SP- B and SP-C in phospholipid films. *Biophysical Journal*, 72(6): 2638–2650. 1997.

[75] **Nogee, L.**, Genetics of the hydrophobic surfactant proteins. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1408(2-3): 323–333. 1998.

[76] **Weaver, T.**, Synthesis, processing and secretion of surfactant proteins B and C. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1408(2-3): 173–179. 1998.

[77] **Brasch, F., Ochs, M., Kahne, T., Guttentag, S., Schauer-Vukasinovic, V., Derrick, M., Johnen, G., Kapp, N., Müller, K., Richter, J., Giller, T., Hawgood, S., and Buhling, F.**, Involvement of napsin A in the C- and N-terminal processing of surfactant protein B in type-II pneumocytes of the human lung. *The Journal of Biological Chemistry*, 278(49): 49006–49014. 2003.

[78] **Brasch, F., Johnen, G., Winn-Brasch, A., Guttentag, S., Schmiedl, A., Kapp, N., Suzuki, Y., Müller, K., Richter, J., Hawgood, S., and Ochs, M.**, Surfactant protein B in type II pneumocytes and intra-alveolar surfactant forms of human lungs. *American Journal of Respiratory Cell and Molecular Biology*, 30(4): 449–458. 2004.

[79] **Patthy, L.**, Homology of the precursor of pulmonary surfactant-associated protein SP-B with prosaposin and sulfated glycoprotein 1. *Journal of Biological Chemistry*, 266(10): 6035–6037. 1991.

[80] **Foster, C., Zhang, P., Gonzales, L., and Guttentag, S.**, In vitro surfactant protein B deficiency inhibits lamellar body formation. *American Journal of Respiratory Cell and Molecular Biology*, 29(2): 259–266. 2003.

[81] **Han, S. and Mallampalli, R.**, The role of surfactant in lung disease and host defense against pulmonary infections. *Annals of the American Thoracic Society*, 12(5): 765–774. 2015.

[82] **Larkin, M., Blackshields, G., Brown, N., Chenna, R., Mcgettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Thompson, J., Gibson, T., and Higgins, D.**, Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21): 2947–2948. 2007.

[83] **Morrow, M., Pérez-Gil, J., Simatos, G., Boland, C., Stewart, J., Absolom, D., Sarin, V., and Keough, K.**, Pulmonary surfactant-associated protein SP-B has little effect on acyl chains in dipalmitoylphosphatidylcholine dispersions. *Biochemistry*, 32(16): 4397–4402. 1993.

[84] **Vandenbussche, G., Clercx, A., Clercx, M., Curstedt, T., Johansson, J., Jörnvall, H., and Ruysschaert, J.M.**, Secondary structure and orientation of the surfactant protein SP-B in a lipid environment. A Fourier transform infrared spectroscopy study. *Biochemistry*, 31(38): 9169–9176. 1992.

[85] **Shiffer, K., Hawgood, S., Haagsman, H., Benson, B., Clements, J., and Goerke, J.**, Lung surfactant proteins, SP-B and SP-C, alter the thermodynamic properties of phospholipid membranes: A differential calorimetry study. *Biochemistry*, 32(2): 590–597. 1993.

[86] **Dico, A., Hancock, J., Morrow, M., Stewart, J., Harris, S., and Keough, K.**, Pulmonary surfactant protein SP-B interacts similarly with dipalmitoylphosphatidylglycerol and dipalmitoylphosphatidylcholine in phosphatidylcholine/phosphatidylglycerol mixtures. *Biochemistry*, 36(14): 4172–4177. 1997.

[87] **Cruz, A., Casals, C., Plasencia, I., Marsh, D., and Pérez-Gil, J.**, Depth profiles of pulmonary surfactant protein B in phosphatidylcholine bilayers, studied by fluorescence and electron spin resonance spectroscopy. *Biochemistry*, 37(26): 9488–9496. 1998.

[88] **Bertani, P., Vidovic, V., Yang, T.C., Rendell, J., Gordon, L., Waring, A., Bechinger, B., and Booth, V.**, Orientation and depth of surfactant protein B C-terminal helix in lung surfactant bilayers. *Biochimica et Biophysica Acta - Biomembranes*, 1818(5): 1165–1172. 2012.

[89] **Sharifahmadian, M., Sarker, M., Palleboina, D., Waring, A., Walther, F., Morrow, M., and Booth, V.**, Role of the N-terminal seven residues of surfactant protein B (SP-B). *PloS One*, 8(9). 2013.

[90] **Sylvester, A., MacEachern, L., Booth, V., and Morrow, M.**, Interaction of the C-Terminal Peptide of Pulmonary Surfactant Protein B (SP-B) with a Bicellar Lipid Mixture Containing Anionic Lipid. *PLoS One*, 8(8). 2013.

[91] **Waring, A., Walther, F., Gordon, L., Hernandez-Juviel, J., Hong, T., Sherman, M., Alonso, C., Alig, T., Braun, A., Bacon, D., and Zasadzinski, J.**, The role of charged amphipathic helices in the structure and function of surfactant protein B. *Journal of Peptide Research*, 66(6): 364–374. 2005.

[92] **Palleboina, D., Waring, A., Notter, R., Booth, V., and Morrow, M.**, Effects of the lung surfactant protein B construct Mini-B on lipid bilayer order and topography. *European Biophysics Journal*, 41(9): 755–767. 2012.

[93] **Ryan, M., Qi, X., Serrano, A., Ikegami, M., Perez-Gil, J., Johansson, J., and Weaver, T.**, Mapping and analysis of the lytic and fusogenic domains of surfactant protein B. *Biochemistry*, 44(3): 861–872. 2005.

[94] **Walther, F., Waring, A., Hernandez-Juviel, J., Gordon, L., Wang, Z., Jung, C.L., Ruchala, P., Clark, A., Smith, W., Sharma, S., and Notter, R.**, Critical structural and functional roles for the N-terminal insertion sequence in surfactant protein B analogs. *PLoS One*, 5(1). 2010.

[95] **Manzanares, D., Rodriguez-Capote, K., Liu, S., Haines, T., Ramos, Y., Zhao, L., Doherty-Kirby, A., Lajoie, G., and Possmayer, F.**, Modification

of tryptophan and methionine residues is implicated in the oxidative inactivation of surfactant protein B. *Biochemistry*, 46(18): 5604–5615. 2007.

[96] **Wang, Y., Rao, K., and Demchuk, E.**, Topographical organization of the N-terminal segment of lung pulmonary surfactant protein B (SP-B1-25) in phospholipid bilayers. *Biochemistry*, 42(14): 4015–4027. 2003.

[97] **Serrano, A., Ryan, M., Weaver, T., and Pérez-Gil, J.**, Critical structure-function determinants within the N-terminal region of pulmonary surfactant protein SP-B. *Biophysical Journal*, 90(1): 238–249. 2006.

[98] **Creuwels, L., Van Golde, L., and Haagsman, H.**, Surfactant protein B: Effects on lipid domain formation and intermembrane lipid flow. *Biochimica et Biophysica Acta - Biomembranes*, 1285(1): 1–8. 1996.

[99] **Parra, E., Moleiro, L., López-Montero, I., Cruz, A., Monroy, F., and Pérez-Gil, J.**, A combined action of pulmonary surfactant proteins SP-B and SP-C modulates permeability and dynamics of phospholipid membranes. *Biochemical Journal*, 438(3): 555–564. 2011.

[100] **Parra, E., Alcaraz, A., Cruz, A., Aguilella, V., and Pérez-Gil, J.**, Hydrophobic pulmonary surfactant proteins SP-B and SP-C induce pore formation in planar lipid membranes: Evidence for proteolipid pores. *Biophysical Journal*, 104(1): 146–155. 2013.

[101] **Chang, R., Nir, S., and Poulain, F.**, Analysis of binding and membrane destabilization of phospholipid membranes by surfactant apoprotein B. *Biochimica et Biophysica Acta - Biomembranes*, 1371(2): 254–264. 1998.

[102] **Baoukina, S. and Tieleman, D.**, Lung surfactant protein SP-B promotes formation of bilayer reservoirs from monolayer and lipid transfer between the interface and subphase. *Biophysical Journal*, 100(7): 1678–1687. 2011.

[103] **Poulain, F.R., Nir, S., and Hawgood, S.**, Kinetics of phospholipid membrane fusion induced by surfactant apoproteins A and B. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1278(2): 169 – 175. 1996.

[104] **Baoukina, S. and Tieleman, D.**, Direct simulation of protein-mediated vesicle fusion: Lung surfactant protein B. *Biophysical Journal*, 99(7): 2134–2142. 2010.

[105] **Schürch, D., Ospina, O., Cruz, A., and Pérez-Gil, J.**, Combined and independent action of proteins SP-B and SP-C in the surface behavior and mechanical stability of pulmonary surfactant films. *Biophysical Journal*, 99(10): 3290–3299. 2010.

[106] **Cruz, A., Vázquez, L., Vélez, M., and Pérez-Gil, J.**, Effect of pulmonary surfactant protein SP-B on the micro- and nanostructure of phospholipid Films. *Biophysical Journal*, 86(1): 308 – 320. 2004.

[107] **Fischkoff, S. and Vanderkooi, J.**, Oxygen diffusion in biological and artificial membranes determined by the fluorochrome pyrene. *The Journal of General Physiology*, 65(5): 663–676. 1975.

[108] **Ligeza, A., Tikhonov, A.N., Hyde, J.S., and Subczynski, W.K.**, Oxygen permeability of thylakoid membranes: Electron paramagnetic resonance spin labeling study. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1365(3): 453 – 463. 1998.

[109] **Ivanov, I., Fedorov, G., Gus'kova, R., Ivanov, K., and Rubin, A.**, Permeability of lipid membranes to dioxygen. *Biochemical and Biophysical Research Communications*, 322(3): 746–750. 2004.

[110] **Battino, R., Rettich, T., and Tominaga, T.**, The solubility of oxygen and ozone in liquids. *Journal of Physical and Chemical Reference Data*, 12(2): 163–178. 1983.

[111] **Zaltash, S., Palmblad, M., Curstedt, T., Johansson, J., and Persson, B.**, Pulmonary surfactant protein B: A structural model and a functional analogue. *Biochimica et Biophysica Acta - Biomembranes*, 1466(1-2): 179–186. 2000.

[112] **Khatami, M., Saika-Voivod, I., and Booth, V.**, All-atom molecular dynamics simulations of lung surfactant protein B: Structural features of SP-B promote lipid reorganization. *Biochimica et Biophysica Acta - Biomembranes*, 1858(12): 3082–3092. 2016.

[113] **Baoukina, S. and Tieleman, D.**, Computer simulations of lung surfactant. *Biochimica et Biophysica Acta - Biomembranes*, 1858(10): 2431–2440. 2016.

[114] **Diemel, R.V., Bader, D., Walch, M., Hotter, B., van Golde, L.M., Amann, A., Haagsman, H.P., and Putz, G.**, Functional tests for the characterization of surfactant protein B (SP-B) and a fluorescent SP-B analog. *Archives of Biochemistry and Biophysics*, 385(2): 338 – 347. 2001.

[115] **Baatz, J.E., Zou, Y., Cox, J., Wang, Z., and Notter, R.H.**, High-yield purification of lung surfactant proteins SP-B and SP-C and the effects on surface activity. *Protein Expression and Purification*, 23(1): 180 – 190. 2001.

[116] **Bünger, H., Krüger, R.P., Pietschmann, S., Wüstneck, N., Kaufner, L., Tschiersch, R., and Pison, U.**, Two hydrophobic protein fractions of ovine pulmonary surfactant: isolation, characterization, and biophysical Activity. *Protein Expression and Purification*, 23(2): 319 – 327. 2001.

[117] **Zaltash, S., Griffiths, W., Beck, D., Duan, C.X., Weaver, T., and Johansson, J.**, Membrane activity of (Cys48Ser) lung surfactant protein B increases with dimerisation. *Biological Chemistry*, 382(6): 933–939. 2001.

[118] **World Health Organization and others**, Born too soon: The global action report on preterm birth. 2012.

[119] **Yu, V.**, Developmental outcome of extremely preterm infants. *American Journal of Perinatology*, 17(2): 57–62. 2000.

[120] **Whitsett, J., Wert, S., and Weaver, T.**, Diseases of pulmonary surfactant homeostasis. *Annual Review of Pathology: Mechanisms of Disease*, 10: 371–393. 2015.

[121] **Xu, Y., Wang, Y., Besnard, V., Ikegami, M., Wert, S., Heffner, C., Murray, S., Donahue, L., and Whitsett, J.**, Transcriptional programs controlling perinatal lung maturation. *PLoS One*, 7(8). 2012.

[122] **Rushing, S. and Ment, L.**, Preterm birth: A cost benefit analysis. *Seminars in Perinatology*, 28(6): 444–450. 2004.

[123] **Danhaive, O., Chapin, C., Horneman, H., Cogo, P., and Ballard, P.**, Surface film formation in vitro by infant and therapeutic surfactants: Role of surfactant protein B. *Pediatric Research*, 77(2): 340–346. 2015.

[124] **Mingarro, I., Lukovic, D., Vilar, M., and Pérez-Gil, J.**, Synthetic pulmonary surfactant preparations: New developments and future trends. *Current Medicinal Chemistry*, 15(4): 393–403. 2008.

[125] **Dunbar, A.E., Wert, S.E., Ikegami, M., Whitsett, J.A., Hamvas, A., White, F.V., Piedboeuf, B., Jobin, C., Guttentag, S., and Nogee, L.M.**, Prolonged survival in hereditary surfactant protein B (SP-B) deficiency associated with a novel splicing mutation. *Genetics in Medicine*, 7(7): 275–282. 2000.

[126] **van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A., and Berendsen, H.**, GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16): 1701–1718. 2005.

[127] **Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M., Smith, J., Kasson, P., Van Der Spoel, D., Hess, B., and Lindahl, E.**, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7): 845–854. 2013.

[128] **Abraham, M., Murtola, T., Schulz, R., Páll, S., Smith, J., Hess, B., and Lindah, E.**, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2: 19–25. 2015.

[129] **Abraham, M., Van Der Spoel, D., Lindahl, E., Hess, B., and the GROMACS development team**, *GROMACS Reference Manual Version 5.1.* www.gromacs.org. 2015.

[130] **Leach, A.R.**, *Molecular Modelling: Principles and Application.* Pearson Education Limited, 2nd edition edition. 2001.

[131] **Sansom, M.S. and Biggin, P.C.**, *Molecular Simulations and Biomembranes: From Biophysics to Function.* 20, Royal Society of Chemistry. 2010.

[132] **Schlick, T.**, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, volume Vol. 21. Springer Science & Business Media. 2010.

[133] **Ponder, J. and Case, D.**, Force fields for protein simulations. *Advances in Protein Chemistry*, 66: 27–85. 2003.

[134] **Van Duin, A., Dasgupta, S., Lorant, F., and Goddard III, W.**, ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A*, 105(41): 9396–9409. 2001.

[135] **Chenoweth, K., Van Duin, A., and Goddard III, W.**, ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation. *Journal of Physical Chemistry A*, 112(5): 1040–1053. 2008.

[136] **Jorgensen, W., Maxwell, D., and Tirado-Rives, J.**, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45): 11225–11236. 1996.

[137] **Wang, J., Wolf, R., Caldwell, J., Kollman, P., and Case, D.**, Development and testing of a general Amber force field. *Journal of Computational Chemistry*, 25(9): 1157–1174. 2004.

[138] **Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C.**, Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function and Genetics*, 65(3): 712–725. 2006.

[139] **MacKerell Jr., A., Bashford, D., Bellott, M., Dunbrack Jr., R., Evanseck, J., Field, M., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy,**

D., Kuchnir, L., Kuczera, K., Lau, F., Mattos, C., Michnick, S., Ngo, T., Nguyen, D., Prodhom, B., Reiher III, W., Roux, B., Schlenkrich, M., Smith, J., Stote, R., Straub, J., Watanabe, M., Wiòrkiewicz-Kuczera, J., Yin, D., and Karplus, M., All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 102(18): 3586–3616. 1998.

[140] **Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and Mackerell Jr., A.**, CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, 31(4): 671–690. 2010.

[141] **Scott, W., Hünenberger, P., Tironi, I., Mark, A., Billeter, S., Fennen, J., Torda, A., Huber, T., Krüger, P., and Van Gunsteren, W.**, The GROMOS biomolecular simulation program package. *Journal of Physical Chemistry A*, 103(19): 3596–3607. 1999.

[142] **Oostenbrink, C., Villa, A., Mark, A., and Van Gunsteren, W.**, A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25(13): 1656–1676. 2004.

[143] **Mackerell Jr., A.**, Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, 25(13): 1584–1604. 2004.

[144] **Periole, X. and Marrink, S.**, The Martini coarse-grained force field. *Methods in Molecular Biology (Clifton, N.J.)*, 924: 533–565. 2013.

[145] **De Jong, D., Singh, G., Bennett, W., Arnarez, C., Wassenaar, T., Schäfer, L., Periole, X., Tieleman, D., and Marrink, S.**, Improved parameters for the martini coarse-grained protein force field. *Journal of Chemical Theory and Computation*, 9(1): 687–697. 2013.

[146] **Wassenaar, T.A., Ingólfsson, H.I., Böckmann, R.A., Tieleman, D.P., and Marrink, S.J.**, Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. *Journal of Chemical Theory and Computation*, 11(5): 2144–2155. 2015.

[147] **Wassenaar, T.A., Pluhackova, K., Böckmann, R.A., Marrink, S.J., and Tieleman, D.P.**, Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models. *Journal of Chemical Theory and Computation*, 10(2): 676–690. 2014.

[148] **Marrink, S.J., De Vries, A.H., and Mark, A.E.**, Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B*, 108(2): 750–760. 2004.

[149] **Marrink, S., Risselada, H., Yefimov, S., Tieleman, D., and De Vries, A.**, The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B*, 111(27): 7812–7824. 2007.

[150] **Monticelli, L., Kandasamy, S., Periole, X., Larson, R., Tieleman, D., and Marrink, S.J.**, The MARTINI coarse-grained force field: Extension to proteins. *Journal of Chemical Theory and Computation*, 4(5): 819–834. 2008.

[151] **López, C., Rzepiela, A., de Vries, A., Dijkhuizen, L., Hünenberger, P., and Marrink, S.**, Martini coarse-grained force field: Extension to carbohydrates. *Journal of Chemical Theory and Computation*, 5(12): 3195–3210. 2009.

[152] **López, C., Sovova, Z., Van Eerden, F., De Vries, A., and Marrink, S.**, Martini force field parameters for glycolipids. *Journal of Chemical Theory and Computation*, 9(3): 1694–1708. 2013.

[153] **Uusitalo, J., Ingólfsson, H., Akhshi, P., Tieleman, D., and Marrink, S.**, Martini coarse-grained force field: Extension to DNA. *Journal of Chemical Theory and Computation*, 11(8): 3932–3945. 2015.

[154] **Marrink, S. and Tieleman, D.**, Perspective on the Martini model. *Chemical Society Reviews*, 42(16): 6801–6822. 2013.

[155] **Darden, T., York, D., and Pedersen, L.**, Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12): 10089–10092. 1993.

[156] **Periole, X., Cavalli, M., Marrink, S.J., and Ceruso, M.**, Combining an elastic network with a coarse-grained molecular force field: Structure, dynamics, and intermolecular recognition. *Journal of Chemical Theory and Computation*, 5(9): 2531–2543. 2009.

[157] **Berendsen, H. and Van Gunsteren, W.**, Practical algorithms for dynamic simulations. *Molecular-dynamics simulation of statistical-mechanical systems*, pp. 43–65. 1986.

[158] **Berendsen, H., Postma, J., Van Gunsteren, W., Dinola, A., and Haak, J.**, Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8): 3684–3690. 1984.

[159] **Nosé, S.**, A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2): 255–268. 1984.

[160] **Hoover, W.**, Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3): 1695–1697. 1985.

[161] **Bussi, G., Donadio, D., and Parrinello, M.**, Canonical sampling through velocity rescaling. *Journal of Chemical Physics*, 126(1). 2007.

[162] **Parrinello, M. and Rahman, A.**, Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12): 7182–7190. 1981.

[163] **Nosé, S. and Klein, M.**, Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5): 1055–1076. 1983.

[164] **Stark, A., Andrews, C., and Elcock, A.**, Toward optimized potential functions for protein-protein interactions in aqueous solutions: Osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *Journal of Chemical Theory and Computation*, 9(9): 4176–4185. 2013.

[165] **Campbell, M.K. and Farrell, S.O.**, *Biochemistry*. Brooks/Cole, 7th edition edition. 2011.

[166] **Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P.**, *Essential Cell Biology*. Garland Science. 2013.

[167] **Cuff, J.A. and Barton, G.J.**, Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins: Structure, Function, and Bioinformatics*, 40(3): 502–511. 2000.

[168] **Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E.**, The protein data bank. *Nucleic acids research*, 28(1): 235–242. 2000.

[169] **Kaczanowski, S. and Zielenkiewicz, P.**, Why similar protein sequences encode similar three-dimensional structures? *Theoretical Chemistry Accounts*, 125(3): 643–650. 2010.

[170] **Martí-Renom, M., Stuart, A., Fiser, A., Sánchez, R., Melo, F., and §ali, A.**, Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, 29: 291–325. 2000.

[171] **Pearson, W.R.**, Empirical statistical estimates for sequence similarity searches. *Journal of Molecular Biology*, 276(1): 71–84. 1998.

[172] **Altschul, S., Gish, W., Miller, W., Meyers, E., and Lipman, D.**, Basic local alignment search tool. *Journal of Molecular Biology*, 215(3): 403–410. 1990.

[173] **Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J.**, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17): 3389. 1997.

[174] **Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J.**, Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences*, 23(10): 403 – 405. 1998.

[175] **Fiser, A. and Šali, A.**, Modeller: generation and refinement of homology-based protein structure models. *Methods in Enzymology*, 374: 461–491. 2003.

[176] **Arnold, K., Bordoli, L., Kopp, J., and Schwede, T.**, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2): 195–201. 2006.

[177] **Simons, K.T., Bonneau, R., Ruczinski, I., and Baker, D.**, Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins: Structure, Function, and Bioinformatics*, 37(S3): 171–176. 1999.

[178] **Touw, W.G., Joosten, R.P., and Vriend, G.**, New biological insights from better structure models. *Journal of Molecular Biology*, 428(6): 1375–1393. 2016.

[179] **Andreeva, A., Howorth, D., Chothia, C., Kulesha, E., and Murzin, A.G.**, SCOP2 prototype: a new approach to protein structure mining. *Nucleic Acids Research*, 42(D1): D310. 2014.

[180] **Sillitoe, I., Lewis, T.E., Cuff, A., Das, S., Ashford, P., Dawson, N.L., Furnham, N., Laskowski, R.A., Lee, D., Lees, J.G., Lehtinen, S., Studer, R.A., Thornton, J., and Orengo, C.A.**, CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Research*, 43(D1): D376. 2015.

[181] **Freddolino, P.L., Harrison, C.B., Liu, Y., and Schulten, K.**, Challenges in protein-folding simulations. *Nature Physics*, 6(10): 751–758. 2010.

[182] **Ahn, V.E., Faull, K.F., Whitelegge, J.P., Fluharty, A.L., and Privé, G.G.**, Crystal structure of saposin B reveals a dimeric shell for lipid binding. *Proceedings of the National Academy of Sciences*, 100(1): 38–43. 2003.

[183] **Gray, J.J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C.A., and Baker, D.**, Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *Journal of Molecular Biology*, 331(1): 281–299. 2003.

[184] **Chaudhury, S. and Gray, J.J.**, Conformer selection and induced fit in flexible backbone protein–protein docking using computational and NMR rnsembles. *Journal of Molecular Biology*, 381(4): 1068 – 1087. 2008.

[185] **Zhang, Y., Lervik, A., Seddon, J., and Bresme, F.**, A coarse-grained molecular dynamics investigation of the phase behavior of DPPC/cholesterol mixtures. *Chemistry and Physics of Lipids*, 185: 88–98. 2015.

[186] **Hess, B., Bekker, H., Berendsen, H.J., Fraaije, J.G., et al.**, LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12): 1463–1472. 1997.

[187] **Humphrey, W., Dalke, A., and Schulten, K.**, VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1): 33–38. 1996.

[188] **Koynova, R. and Tenchov, B.**, Transitions between lamellar and non-lamellar phases in membrane lipids and their physiological roles. *OA Biochemistry*, 1(1): 1–9. 2013.

[189] **Krol, S., Ross, M., Sieber, M., Kunneke, S., Galla, H.J., and Janshoff, A.**, Formation of three-dimensional protein-lipid aggregates in monolayer films induced by surfactant protein B. *Biophysical Journal*, 79(2): 904–918. 2000.

[190] **Yu, S.H. and Possmayer, F.**, Effect of pulmonary surfactant protein B (SP-B) and calcium on phospholipid adsorption and squeeze-out of phosphatidylglycerol from binary phospholipid monolayers containing dipalmitoylphosphatidylcholine. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, 1126(1): 26 – 34. 1992.

[191] **Duncan, S.L., Dalal, I.S., and Larson, R.G.**, Molecular dynamics simulation of phase transitions in model lung surfactant monolayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(10): 2450 – 2465. 2011.

APPENDIX A. LATERAL LIPID REORGANIZATION FIGURES

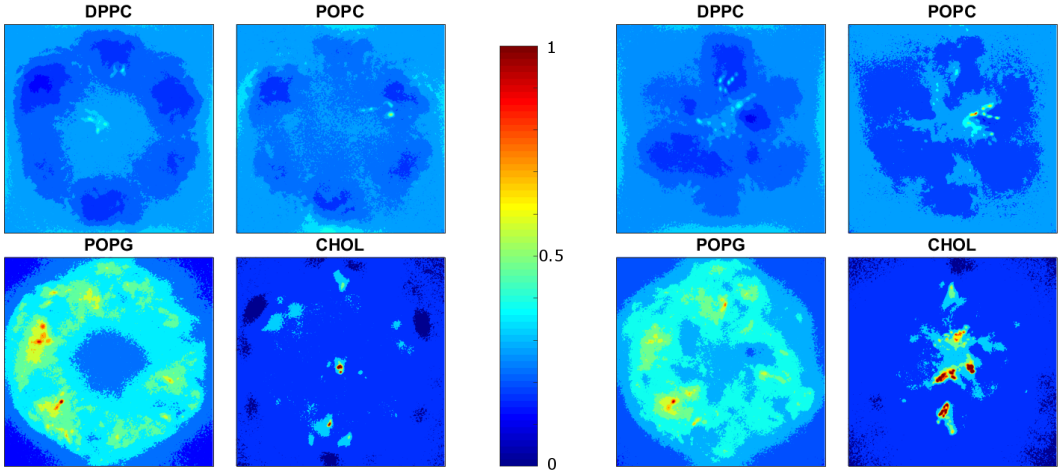


Figure 1 APL 52.5 Å^2, average inter-monolayer distance 3.2 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

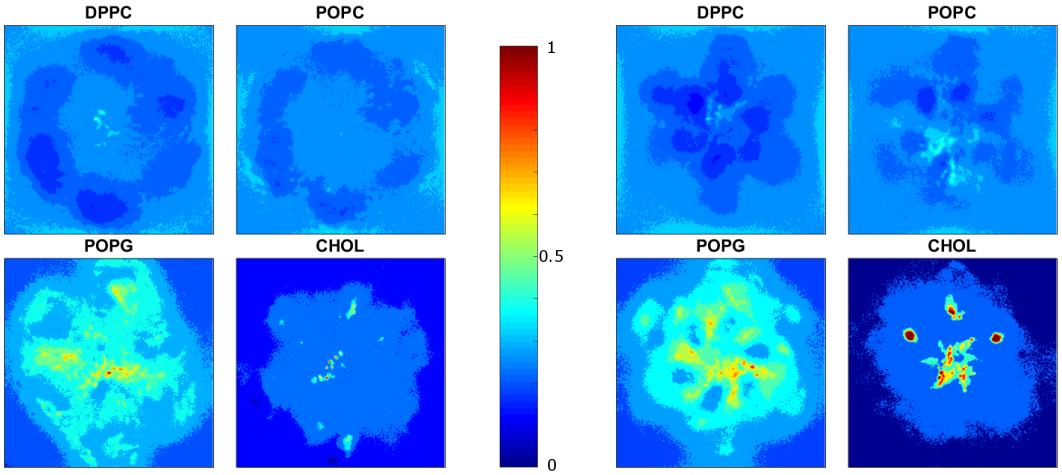


Figure 2 APL 52.5 Å^2, average inter-monolayer distance 3.5 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

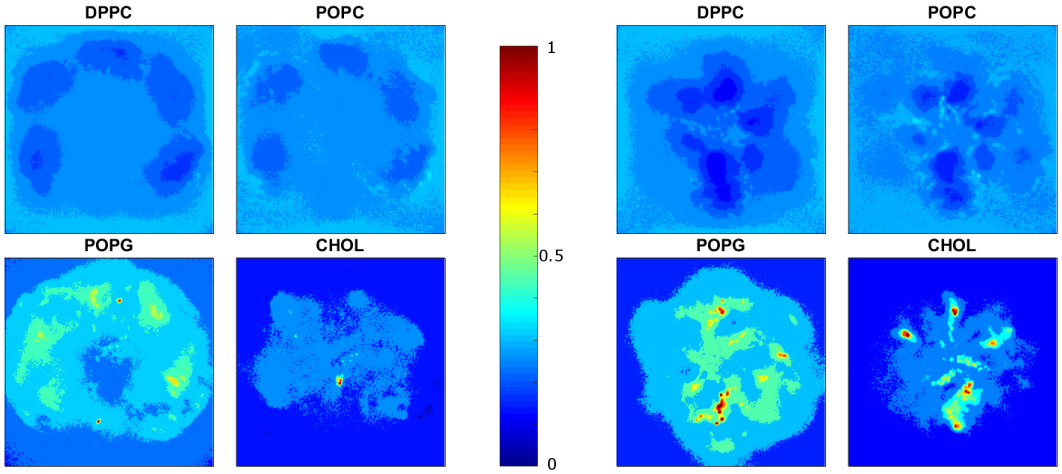


Figure 3 APL 52.5 $Å^2$, average inter-monolayer distance 3.5 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

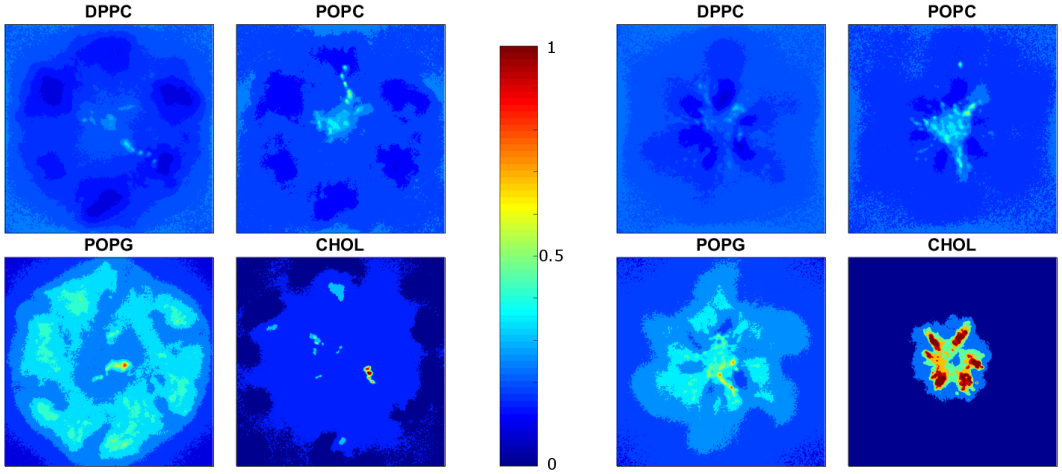


Figure 4 APL 55 $Å^2$, average inter-monolayer distance 2.7 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

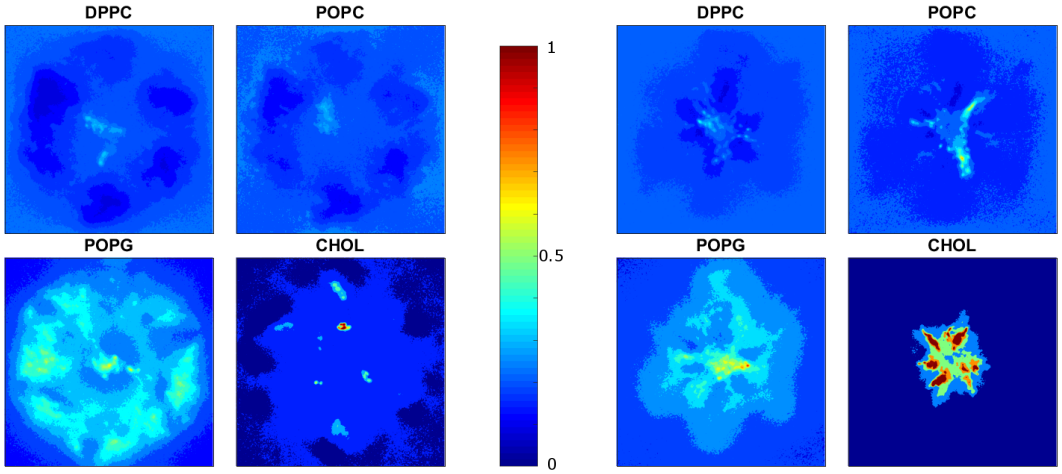


Figure 5 APL 55 Å^2, average inter-monolayer distance 3.1 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

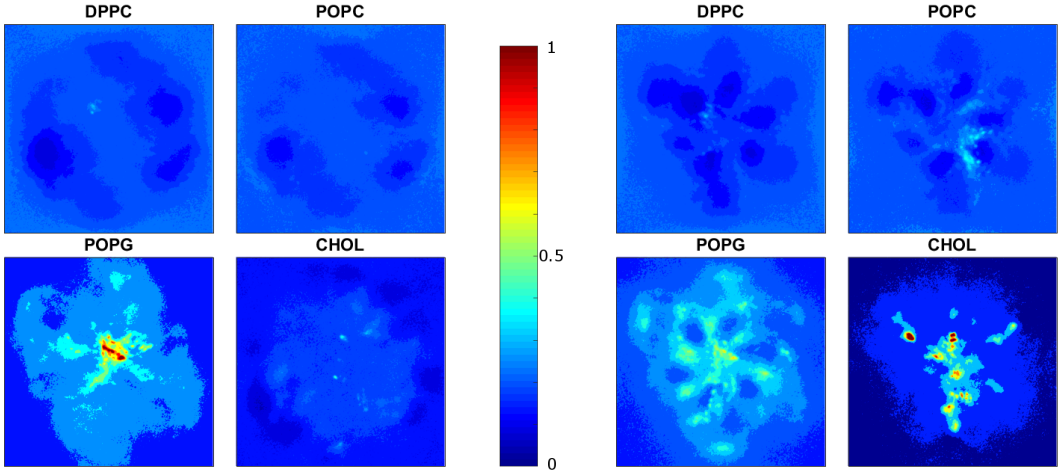


Figure 6 APL 55 Å^2, average inter-monolayer distance 3.3 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

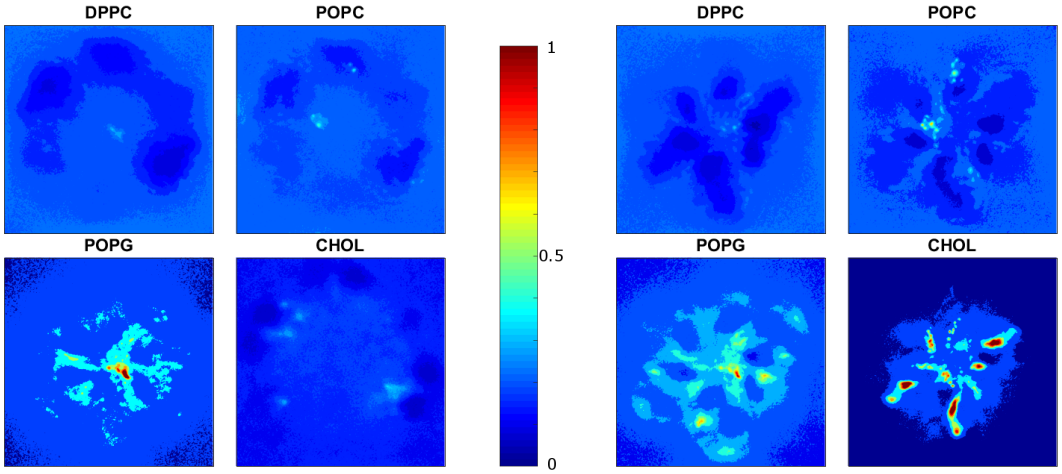


Figure 7 APL 55 Å^2, average inter-monolayer distance 3.6 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

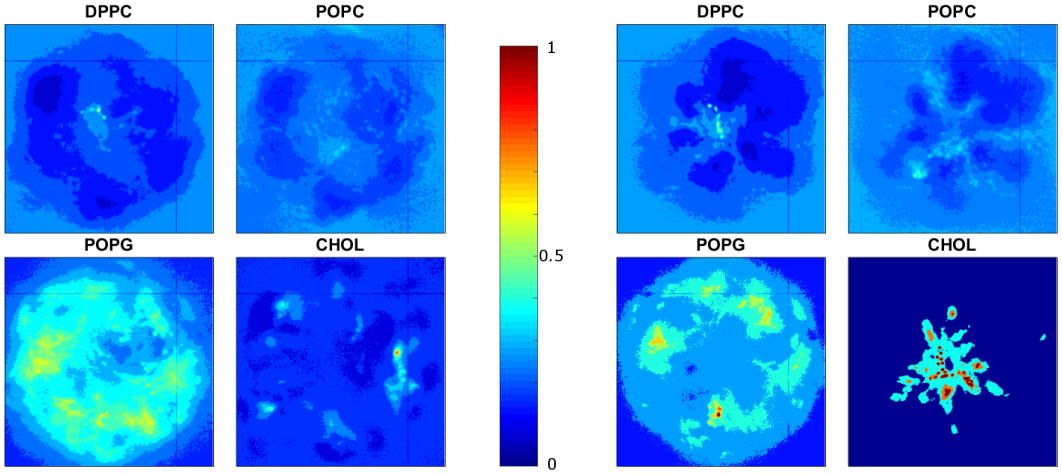


Figure 8 APL 57.5 Å^2, average inter-monolayer distance 3.1 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

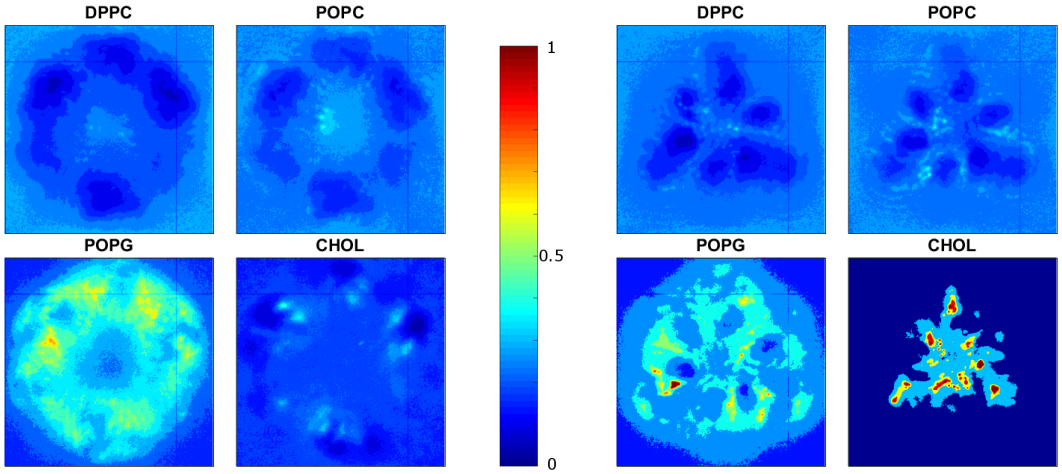


Figure 9 APL 57.5 $Å^2$, average inter-monolayer distance 3.4 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

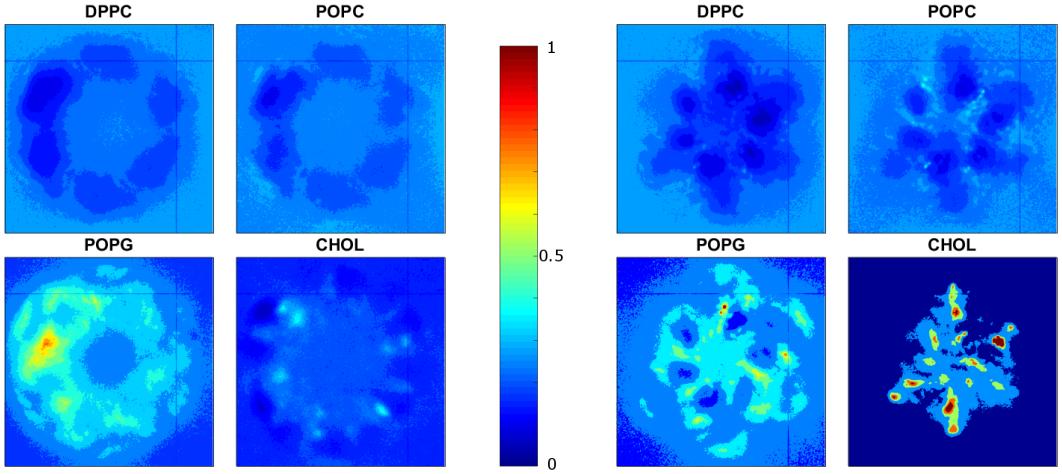


Figure 10 APL 57.5 $Å^2$, average inter-monolayer distance 3.5 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

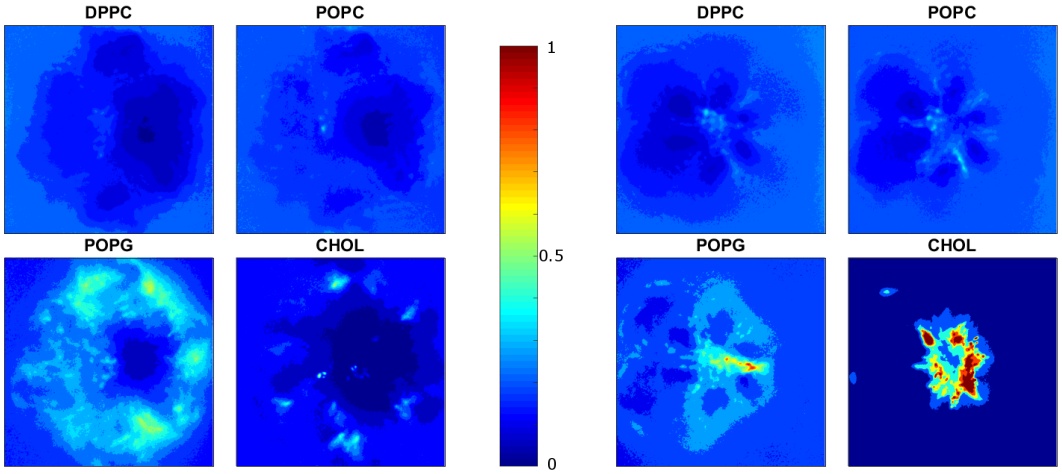


Figure 11 APL 60 Å^2, average inter-monolayer distance 2.7 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

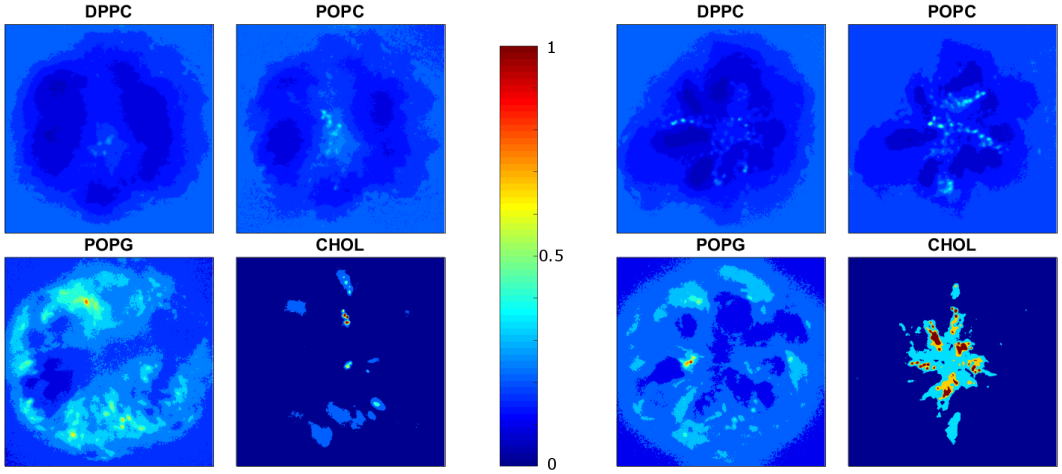


Figure 12 APL 60 Å^2, average inter-monolayer distance 3.0 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

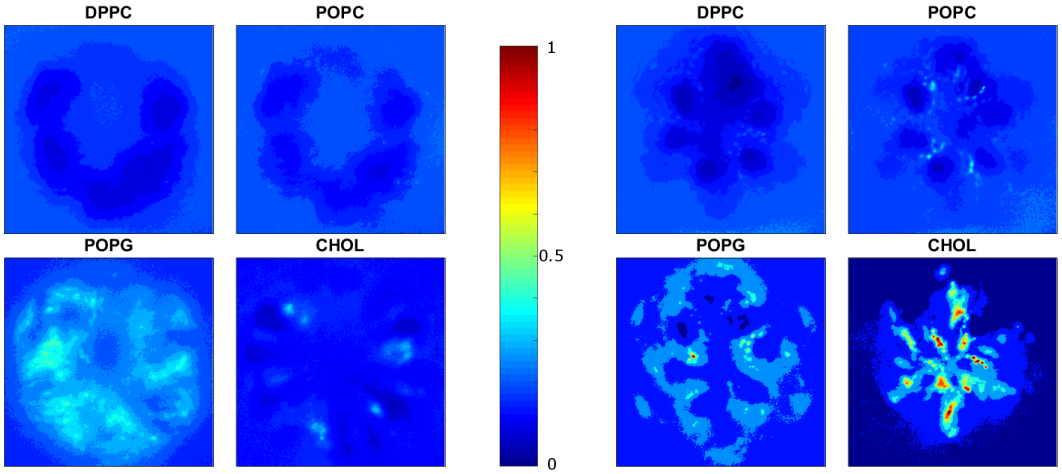


Figure 13 APL 60 Å^2, average inter-monolayer distance 3.3 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

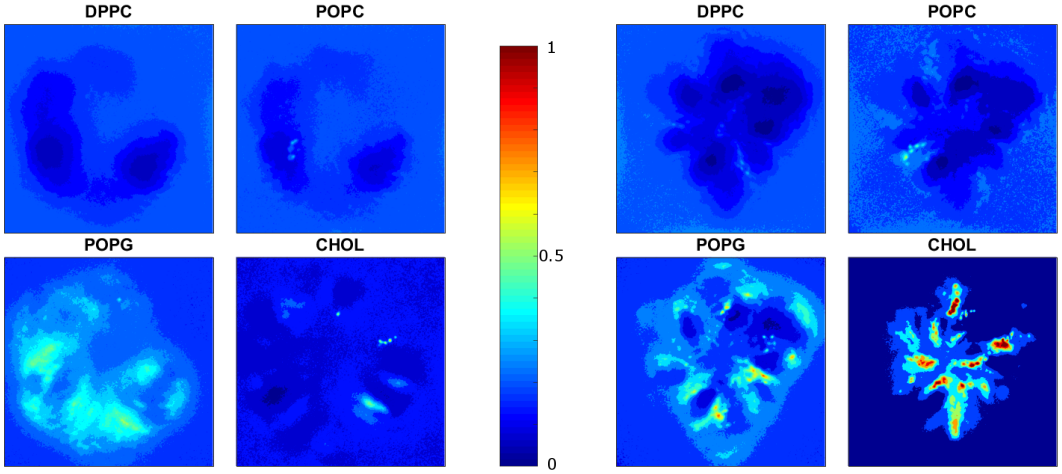


Figure 14 APL 60 Å^2, average inter-monolayer distance 3.6 nm. Left: lipids of the upper monolayer facing the membrane binding surface of the protein, right: lipids of the lower monolayer facing the membrane perturbing surface of the protein. Results are normalized per lipid type to show specific lipid interactions with SP-B.

APPENDIX B. LIDID-SPECIFIC CONTACT FREQUENCY OF SP-B HEXAMER OF DIMERS

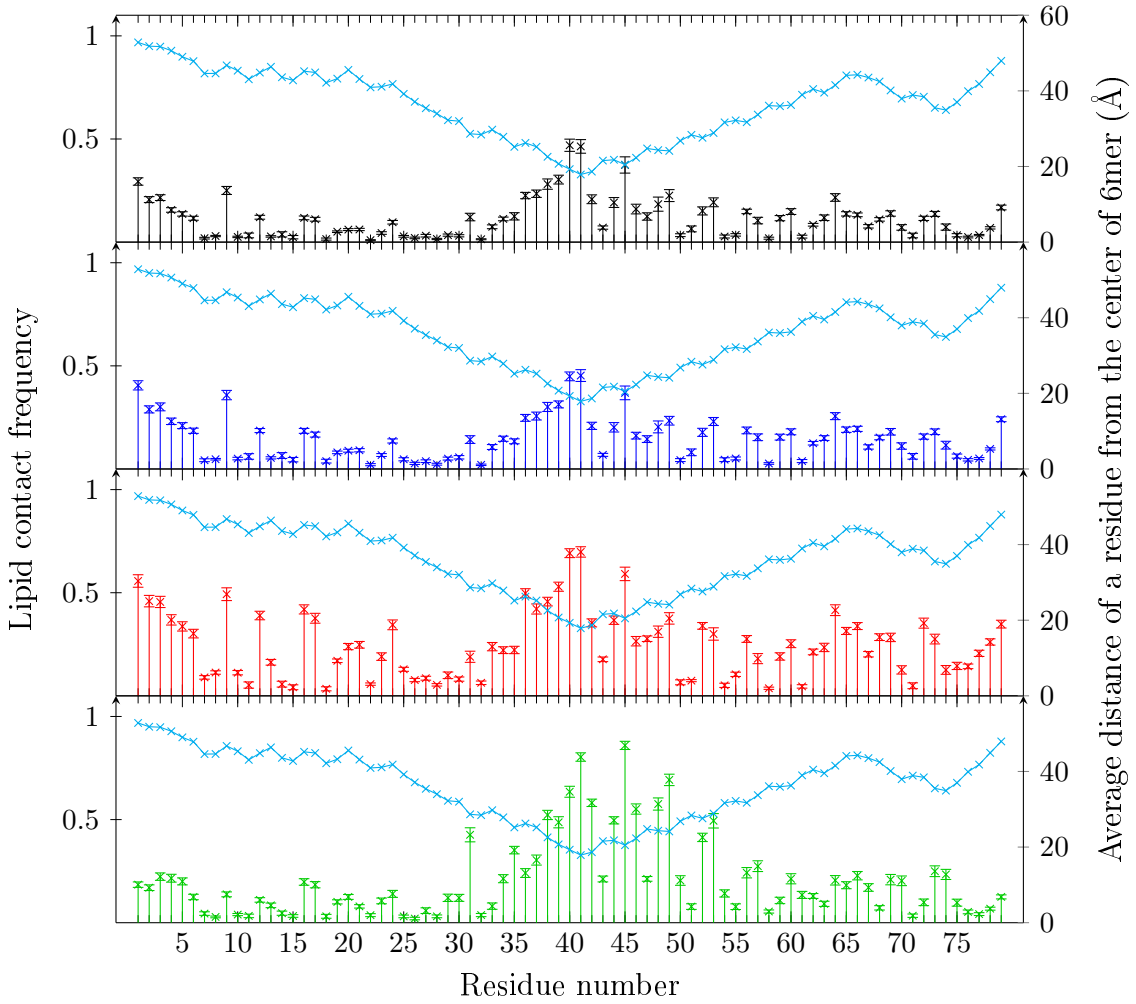


Figure 15 SP-B hexamer of dimers lipid-specific lipid contact frequency per residue in EQ lipid composition with DPPC, POPC, POPG, and CHOL, respectively. Results calculated as an average of 24 repetitions, error as SE.

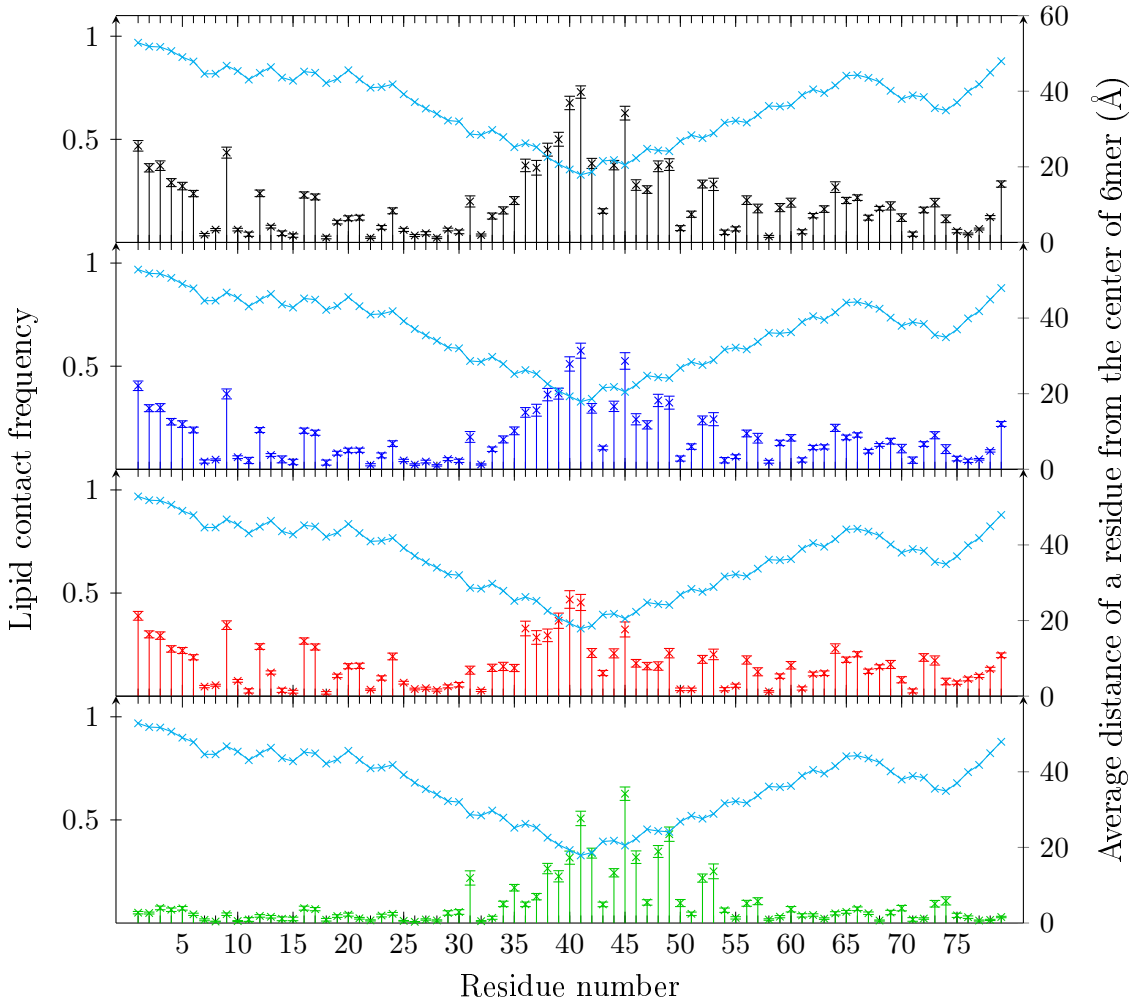


Figure 16 SP-B hexamer of dimers lipid-specific lipid contact frequency per residue in PHYS lipid composition with DPPC, POPC, POPG, and CHOL, respectively. Results calculated as an average of 24 repetitions, error as SE.

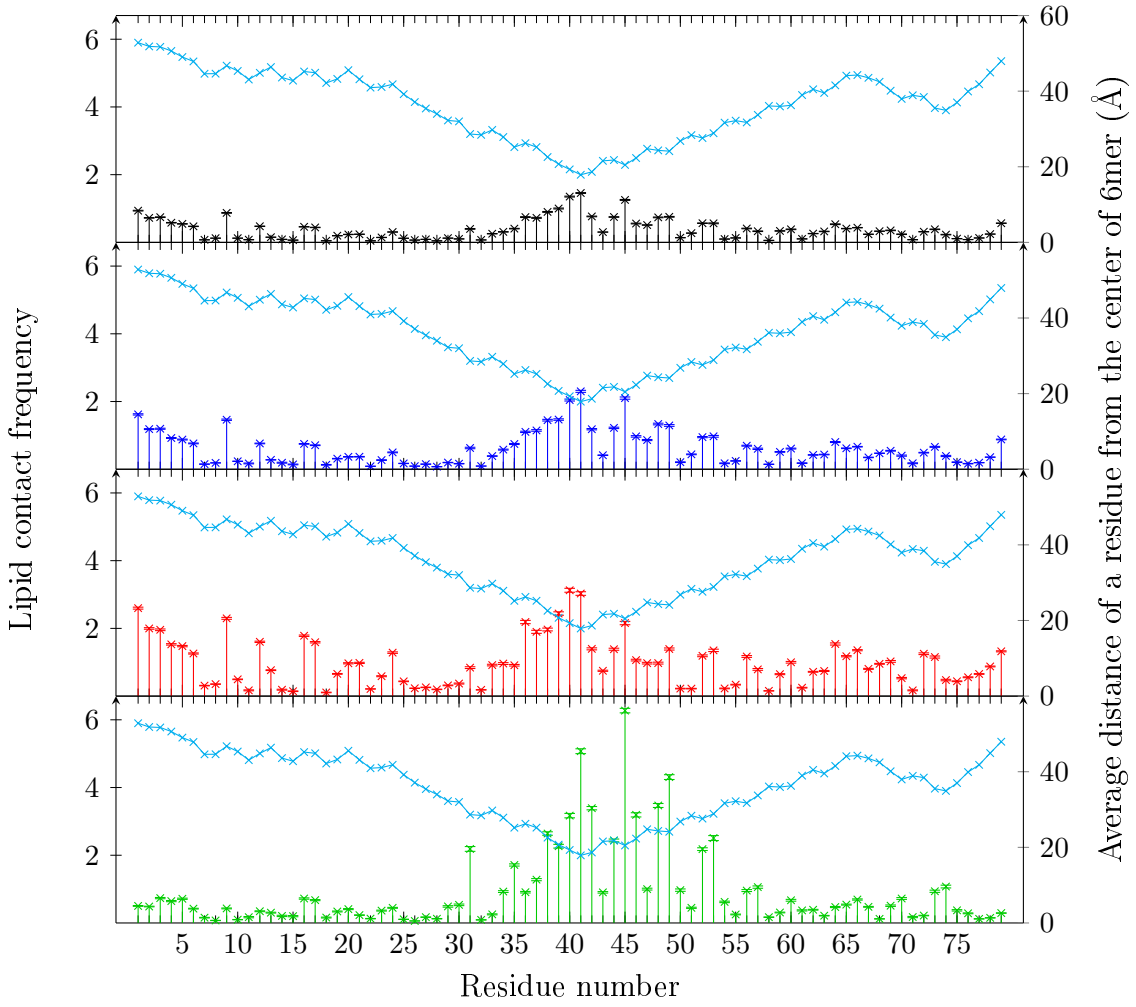


Figure 17 Normalized SP-B hexamer of dimers lipid-specific lipid contact frequency per residue in PHYS lipid composition with DPPC, POPC, POPG, and CHOL, respectively. Results calculated as an average of 24 repetitions and normalized based on the relative concentration of each lipid type, error as SE.