The Expression, Purification and Characterization of Recombinant Human Lung Surfactant Protein-B

By

Tadiwos G. Asrat

A thesis submitted to the School of Graduate Studies in partial fulfilment
of the requirements of the degree of Doctor of Philosophy

Department of Biochemistry
Faculty of Science
Memorial University of Newfoundland
St. John's, Newfoundland and Labrador
2021

Abstract

A way to produce surfactant protein B (SP-B) in the laboratory is highly sought both for basic research on SP-B, as well as for use as a component of surfactant for treatment of patients in respiratory distress. Herein, I hypothesize that SP-B can be expressed in bacteria in its functional form. In the long run, I also anticipate that recombinant SP-B could alleviate the financial burden incurred by the health care system during the treatment of neonatal respiratory distress syndrome patients. Moreover, I envisage, recombinant surfactant protein B variants with an improved capability to resist surfactant dysfunction. In the first part of the study, recombinant DNA technology was used to overexpress the near full-length human SP-B variant, △7NT△M-SP-B-6xHis, in BL21 *Escherichia coli* , strain C43 cells. The results show that recombinant surfactant protein B expresses as inclusion bodies but can be renatured carefully using immobilized metal ion affinity chromatography. Importantly, the thesis work has established a protocol for the production of recombinant surfactant protein B that is amenable to scale up. In the next part of the study, circular dichroism was used to assess the conformation of recombinant SP-B suspensions in different membrane mimetic environments. Likewise, dynamic light scattering was used to characterize sample homogeneity and aggregation propensity of the protein suspension in these conditions. It was found that dodecyl phosphocholine/sodium dodecyl sulfate binary micelles and methanol support native-like secondary structure of recombinant SP-B comparable to animal derived surfactant protein B. In the last portion of the study, the function of recombinant SP-B was tested in lipid environments using the Langmuir-Wilhelmey surface balance. The results indicate that recombinant SP-B possesses the necessary biophysical features to promote the large-scale organization of lipid monolayers that are thought to be critical to SP-B function. In conclusion this work supports the use of recombinant SP-B in a surfactant therapeutic as a potential future alternative to animal derived SP-B.

Acknowledgements

I am always indebted to my supervisor, Dr.Valerie Booth, for her guidance. The assist and mentorship she provided me were of top quality. This work would not have been fruitful without her expert knowledge. My committee members are also greatly acknowledged; their skillful insights and invaluable suggestions made this work better. I also thank the examiners of this thesis for their valuable time and expertise. Staffs and colleagues at the department of Biochemistry are also appreciated for embracing me. Donna Jackmann have been a great help and I thank you. My sincere gratitude goes to my family, their love and support kept me going. Mum you will always be in my heart.

Contents

List of Tables

List of Figures

List of Abbreviations

ABCA1	adenosine triphosphate-binding cassette sub-family A member 1
ABCA3	adenosine triphosphate-binding cassette sub-family A member 3
AC	adenylate cyclase
ACF	auto correlation function
ALI	acute lung injury
Apo A1	apolipoprotein A-I
ARDS	acute respiratory distress syndrome
AT1	alveolar type one
AT2	alveolar type two
ATP	adenosine triphosphate
CaCM-PK	Ca^{2+}-calmodulin dependent protein kinase
CB	composite body
CBS	captive bubble surfactometer
CCT	CTP:phosphocholine cytidylyltransferase
CD	circular dichroism
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CMC	critical micellar concentration
CNBr	cyanogen bromide
CPT	choline phosphotransferase
D_t	translational diffusion coefficient

DAG	diacylglycerol
DCR	derived count rate
DHAP	dihydroacetone phosphate
DLS	dynamic light scattering
DPC	dodecylphosphocholine
DPPC	dipalmytoylphosphatidylcholine
DPPG	dipalmitoylphosphatidylglycerol
E.coli	escherichia coli
ER	endoplasmic reticulum
FAS	fatty acid synthase
FRET	foster resonance energy transfer
FTIR	fourier transform infra red
G3P	glycerol-3 phosphate
GA	Golgi apparatus
GRAVY	grand average hydropathicity index
GWAS	genome wide association study
HDL	high-density lipoproteins
His-tag	histidine tag
HMD	hyaline membrane disease
IB	inclusion body
IMAC	immobilized metal ion affinity chromatography
IMP	integral membrane protein
IP3	inositol triphophate
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kilo dalton

LB	lamellar body
LC	liquid condensed
LDL	low-density lipoprotein
LE	liquid expanded
LPCAT1	lysophosphatidyl choline acetyl transferase
LS	lung surfactant
LUV	large unilamellar vesicle
LWB	Langmuir-Wilhelmy balance
mAb	monoclonal antibody
MMA	mean molecular area
MRE	mean residue molar ellipticity
MSD	mean square displacement
MVB	multivesicular body
MWCO	molecular weight cut-off
NMR	nuclear magnetic resonance
NP	nanoparticle
NRDS	neonatal respiratory distress syndrome
NTA	nanoparticle tracking analysis
OD	optical density
PA	palmitic acid
pAb	polyclonal antibody
PAP	phosphatidic acid phosphatase
PBS	pulsating bubble surfactometer
PC	phosphatidylcholine
PDB	protein data bank

PDI	polydispersity index
PE	phosphatidylethanolamine
PG	phosphatidylglycerol
PI	phosphatidylinositol
PKA	protein kinase A
PKC	protein kinase C
PL	phospholipid
PLC	phospholipase C
PMPC	palmitoyl myristoyl phosphocholine
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
PSD	particle size distribution
PTM	post-translational modification
RDS	respiratory distress syndrome
rSP-B	recombinant surfactant associated protein-B
SAPLIP	saposin-like protein
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SFTPB	surfactant, pulmonary-associated protein B, gene
SFTPC	surfactant, pulmonary-associated protein C, gene
SN	staphylococcus nuclease A
SP	surfactant protein
SP-A	surfactant associated protein-A
SP-B	surfactant associated protein-B
SP-C	surfactant associated protein-C

SP-D	surfactant associated protein-D
T_m	transition temperature
TBS	Tris-buffered saline
TCEP	Tris(2-carboxyethyl)phosphine
TFE	triflouroethanol
TM	tubular myelin
TOF-SIMS	time of flight-secondary ion mass spectrometry
TPA	tetradecanoylphorbol acetate
UTP	uridine triphosphate

1

Introduction

1.1 Overview

This chapter presents detailed literature review on lung surfactant (LS) with particular focus on surfactant associated protein B (SP-B). The chapter starts by highlighting important achievement of researchers that led to the development of lung surfactant science known today. The content of LS, bio-synthesis of LS, secretion and recycling of LS are discussed. The review also addresses the proposed mechanisms how LS achieves a minimum surface tension during breathing. The section, " the structure-function relationships of SP-B" highlight the current gap in knowledge and underlines the importance of SP-B to LS function. The literature also reasons the lack of LS to have dire health consequences and briefly talks about replacement clinical lung surfactant and points to future directions. Finally, the author sets a hypothesis and puts forward some objectives that the current study tries to address. The thesis work is centered around production and characterization of recombinant human SP-B in the laboratory scale. The use of recombinant SP-B at the industrial scale or in replacement lung surfactant is beyond the scope of this study.

1.2 Historical Background

The presence of surfactant in the alveolar air space was proposed by Kurt von Neergard in 1929 following his seminal work which led him to conclude that fluid of no ordinary nature lines the alveolar surface and facilitates the work of breathing [1]. Decades after, in 1954, the Canadian scientist, Charles Macklin asserted the presence of mucopolysaccharide material in the alveolar space likely produced by the granular pnuemocytes [2, 3]. To add to the complexity, the following year, Richard Eric Pattle published a paper postulating that alveoli are covered with surface active insoluble proteins evolved to counteract the high surface tension of the alveolar interface [4]. Later on in 1957, John Clement, in his pioneering work using the Langmuir-Wilhelmy surface balance (LWB), demonstrated that extract from lung tissue is able to reduce interfacial tension *in vitro* and provided evidence that such material could be responsible for preventing alveolar collapse *in vivo* [5]. Such is the insight that lay ground for Maryellen Avery and Jere Mead. Their work in 1959 provided evidence that lack of surfactant could be the cause of neonatal respiratory distress syndrome (NRDS) [6]. Also referred as hyaline membrane disease (HMD) at that point in time: it was the leading cause of preterm infant mortality [7].

These findings, more importantly, the work by Avery *et al.* [6], led surfactant science to burgeon. Of particular interest were the composition of lung surfactant (LS), its relevance to alveolar tension and its clinical implications related to neonatal infant mortality and morbidity. In the wake of this new interest in surfactant, the year 1961 saw progress when Klause *et al.* isolated beef surfactant and showed that LS is a complex mixture of not just phospholipids and proteins but also cholesterol, triglyceride and fatty acids among traces of other molecules [8]. Further, using the LWB, he and his peers showed that i) the phospholipid content of LS has surface activity comparable to that of full spectrum whole lung extract in contrast to lung surfactant that are deprived of the phospholipids which otherwise exhibits poor activity, ii) the surface activity of the phospholipids are compromised in the presence of other lipids, and iii) surfactant activity gradually deteriorates as the surfactant film is exposed to air for more than two hours. All lines of evidence thus suggest that the biophysical activity of lung surfactant is prone to reactive oxygen species and perhaps inhibition of the phospholipids in LS could be the cause of lung related health conditions.

Attempts to treat NRDS using phospholipids were largely unsuccessful in the 1960s. For example, treatment using nebulized dipalmytoylphosphatidylcholine (DPPC), the main component of LS, was attempted meticulously but to no avail [9, 10]. Patrick Kennedy, son of the then American president John F.Kennedy died of NRDS in Boston children hospital in 1963 after unsuccesful treatment using protein-free surfactants [11]. This was a turning point as it led to a huge increase in public awareness and underlined the need for an in-depth investigation into the cause and pathophysiology leading to NRDS. Dozens of trials to treat NRDS thereafter proved to be important. In the early 1970s, for instance, Goran Enhorning and Bent Robertson used an immature diseased rabbit model to show that natural surfactant (including the apoproteins) is effective in ameliorating NRDS [12]. By the year 1972, Richard King succeeded in isolating surfactant apoproteins from dog lung and quantified their size using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [13–15]. Equally important is the finding by Joan Gil and Oscar Reiss [16], the following year. After comparing electron micrographs of rat lung homogenate of different fractions, they proposed the lamellar bodies as the sole intracellular source of intra-alveolar lipids and asserted that the endoplasmic reticulum of type two alveolar (AT2) cells as the source of the apoproteins. Later on that decade, in 1977, the Finnish scientist Mikko Hallman, identified phosphatidylglycerol (PG) as a vital component crucial for surfactant spreading and also reported that PG is scarce in NRDS patients [17].

An important breakthrough came in 1980, in Japan, when Tetsuro Fujiwara published in *The Lancet* that ten preterm infants with severe hyaline membrane disease showed significant improvement after being adminstered with bolus of modified bovine lipid extract endotracheally, eight of which survived, six were birth weight less than 1500 g; two died of unrelated causes [18]. Hence, the mid and late 1980s was the era of many randomized clinical trials. Consequently, new natural and synthetic surfactant formulations were devised that demonstrated a reduction in preterm infant mortality and morbidity. In 1990, for example, poractant alfa (Curosurf), a purified minced porcine extract, was formulated [19] and later approved by FDA for neonates suffering from NRDS in 1997.

1.3 Composition of Lung Surfactant

Lung surfactant is composed of lipids and proteins that vary in content across species [20] as well as from new born to adult life stages [21]. Hence, maturation of lung surfactant is marked by changes in molecular composition both during antenatal and postnatal development periods [22]. Herein, I will discuss the lipid portion first and present the proteins following.

About 90% by mass of lung surfactant are lipids [23]. Glycerophospholipids predominate constituting roughly 80% of the lipid portion [22] among which, phosphatidylcholine(PC), a zwitterionic phospholipid, contributes the most, 60-70% by mass [23]. Four types of PC species are enriched in mammalian lung surfactant; dipalmitoyl (16:0/16:0)-PC, palmitoyl-palmitoleoyl (16:0/16:1)-PC, palmitoyl-myristoyl (16:0/14:0)-PC and myristoyl-palmitoyl (14:0/16:0)-PC [20–22]. Dipalmitoyl PC (DPPC) is the most abundant of all, about 50% of the total PC content. The high content of DPPC offers LS unique behavior and distinct properties compared to that of cell membranes [21]. In particular, the saturated nature of the palmitoyl moieties [24] confer DPPC more rigidity. Therefore, films enriched with DPPC are strong enough to withstand high surface tension at end expiration [25]. In accordance to physiologic needs, content of the three PCs can increase at the expense of DPPC to alter the physical texture of pulmonary surfactant [20–22]. In another words, the higher the DPPC content of LS, the higher the stiffness and hence the lower the compressibility. The opposite is true for PC 16:0/14:0 and PC 16:0/16:1, which confer fluidic property to LS, and therefore more spreadability.

Phosphatidylglycerol (PG) and phosphatidylinositol (PI) are anionic and are the second most abundant phospholipids after PC, about 8-15% of lung surfactant [23]. In humans, the main components are three monounsaturated species, PG/PI 16:0-18:1, PG/PI 18:1-18:1 and PG/PI

18:0-18:1, with small amounts of PG 16:0-16:0 as well. Conversely, many animals, e.g. rat, rabbit and guinea pig, exhibit an increased content of PG 16:0-16:0 and 16:0-18:2. This therfore depicts an interspecies variation of PG and PI across mammals [25]. Compared to PC, PG has less saturated chain content (≈22% palmitoyl chain) and more mono-unsaturated moeity (52% oleoyl chain) [24]. This suggests the PG component of LS may not be as structurally rigid as PC and consequently is less able to withstand high surface tension. Besides the two anionic lipids (PG and PI), minor components of phosphatidylethanolamine(PE) and phosphatidylserine (PS) and even small amounts of other lipids make up the rest of the glycerophospholipids portion [22].

Cholesterol is the most abundant neutral lipid in lung surfactant (≈10 wt% in mammals) followed by trace amounts of free fatty acids and triglycerides. Cholesterol content is thought to be evolutionarily adapted to fine-tune LS function in mammals [26]. One observation supporting this idea is the rapid alteration in cholesterol content. Thus, cholesterol may play important role in adjusting the biophysical properties of LS during seasonal changes, for example, in hibernating animals [22]. Nonetheless, the clinical implications of either elevated or diminished cholesterol level in lung surfactant homoeostasis in humans is not yet clearly established.

The remaining 5-10% of lung surfactant is the protein portion. Four apoproteins named according to order of their discovery as surfactant associated protein (SP)-A/B/C/D are found in lung surfactant. SP-A and SP-D are large water soluble proteins that belong to the collectin family and are important to innate immunity in the lungs. On the other hand, SP-B and SP-C are small hydrophobic proteins involved in facilitating surface activity of lung surfactant lipids. Details of these, specially of SP-B, the focus of this work, will be further discussed in depth later in this chapter. Recently, with the advent of genome wide association study (GWAS) and improved bioinformatics sequencing tools, existence of other SPs referred as SP-G (*SFTA2*: surfactant associated protein 2, UniProtKB: Q6UW10) and SP-H (*SFTA3*: surfactant associated protein 3, UniProtKB: P0C7M3) are also reported. Despite bearing no resemblance in sequence nor structure to the classical SPs, SP-G and SP-H are proposed to have similar functional properties [27, 28]. Although, the exact functions of these two proteins remain unclear as of yet, some preliminary data suggests roles for SP-G/H in regulating inflammation and host defence more like SP-A and SP-D, the two lectins that are involved in immune surveillance in pulmonary surfactant. SP-G and SP-H are highly expressed in the lung and to lesser extent by other human cells and tissues including, but not limited to, kidney and testis. The two proteins, SP-G/H are possibly secreted via the classical secretory pathway instead of the non-classical route by which some of the other SPs are synthesized [29, 30].

In general, lung surfactant is complex mixture of lipids and proteins where environmental factors, lung development and aging, insult to the lung of some sort, as well as nutrition and lifestyle all have implications on the content of LS and quality of life that persist into adult life [31]. Further, interspecies variation and different lipid distribution profiles in the lungs could also play important role in the outcome of cystic fibrosis, lung cancer and asthma [32].

1.4 Biogenesis of Lung Surfactant

Alveolar type two cells synthesize, secrete and recycle all components of lung surfactant [33]. Mammalian lungs consist of more than 40 different cell types [34] and hundreds of millions of alveoli [35]. An alveolus contains two types of epithelial cells, cuboidal alveolar type two (AT2) cells and the squamous, single cell thick, type one cells (AT1) [36]. AT1 and AT2 cells are in direct contact to one another and lie in close apposition to the capillary lumen [34, 37]. While AT1 cells primarily facilitate gas exchange across the air-blood barrier, AT2 cells have multiple functions. Firstly, AT2 cells are stem cells of the alveolar epithelium as they are able to profilerate continously and differentiate into AT1 cells. Differentiation occurs both during regular cell turnover or in response to insult to the lung of some sort to restore normal tissue structure and function of the alveolar epithelium [38]. Secondly, AT2 cells are the source of lung surfactant, i.e. AT2 cells are the only cells that can produce, secrete, and recycle all contents of lung surfactant [34]. Thirdly, AT2 cells are also an integral component of the alveolus. Their ability to interact with resident cells, as well as with neighbouring mobile cells, provides the alveolus with integrity. A proposed model and good example of the latter case is the AT2-AT1 inter-cellular communication during the time of alveolar expansion. This causes AT1 cells to stretch and induce Ca^{2+} oscillation that is transduced to the adjoining AT2 cell via gap junctions and modulates surfactant secretion [37]. Recently, AT1 cells have been found to influence surfactant secretion through $P2X_7R$ mediated paracrine signalling (Figure 1.1) [39]. Next, we turn to the production and processing of lung surfactant by AT2 cells.

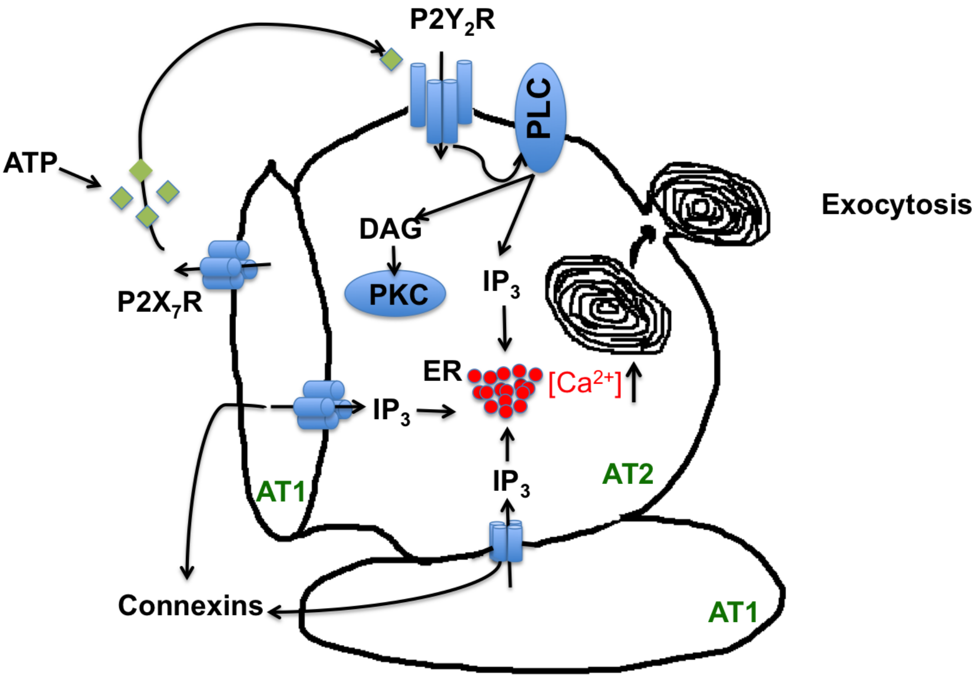


Figure 1.1: Intercellular communication between AT1 and AT2 cells via gap junction and paracrine signalling modulates secretion of LS by AT2 cells. Connexins are believed to transduce IP_3 from AT1 to AT2 cells during mechanical stretching induced by alveolar expansion. As such AT1 cells can be considered as mechanotransducers. The IP_3 released in the AT2 cells then stimulates its receptor (IP_3R) on the ER and promotes calcium release into the cytosol. The other communication is via the ATP released from the $P2X_7R$ of the AT1 cells. The ATP released from AT1 cells stimulates the $P2Y_2R$ on the AT2 cells in paracrine manner. The stimulated $P2Y_2R$ in turn activates PLC which produces the two secondary messengers, DAG and IP_3 that modulate downstream singnalling cascades. More discussion on the pathways that leads to LS secretion is provided in section 1.4. Model concept was adapted from Mishra *et al.* [39] and Koval [40].

Lung surfactant genesis and maturation operate through the traditional secretory pathway. Studies that used isolated AT2 cells in culture have shown that LS production adheres to the classical route, i.e. the endoplasmic reticulum (ER), Golgi apparatus (GA), multivesicular bodies (MVB), the composite bodies (CB) and the lamellar bodies (LB) where LS is ultimately stored (Figure 1.2) [41]. The lamellar bodies are highly organized concentric lipid-protein storage and secretory organelles that are enclosed by an outer (limiting) membrane and contain multilamellar lipid membranes in the core. Furthermore, the LB, which may contain lytic enzymes and apolipoproteins, also exhibits an acidic pH essential for the organelle's lysosomal activity [42]. Unlike the proteins, the phospholipids bypass the Golgi apparatus [43], and hence passage to the lamellar bodies of the proteins and the lipids must take separate route. The ATP-binding transporter cassette A3 (ABCA3), an integral membrane protein highly expressed in the lung and predominantly localized at the limiting membrane of the lamellar bodies, is believed to be responsible for translocating phospholipids and cholesterol from the cytosol into the lamellar bodies [44–47]. ABCA3 therefore plays a critical role in LB biogenesis. The hydrophobic LS proteins, SP-B and SP-C, are transported to the lamellar bodies through the classical route and stored with LS lipids. The lamellar granules are then secreted into the alveolar lumen by regulated exocytosis and also through the constitutive pathway. The level of surfactant pool in the alveolar lumen is likely controlled by GPR116, an orphan G-protein coupled receptor expressed in the apical side of the alveolar epithelial cells [48, 49].

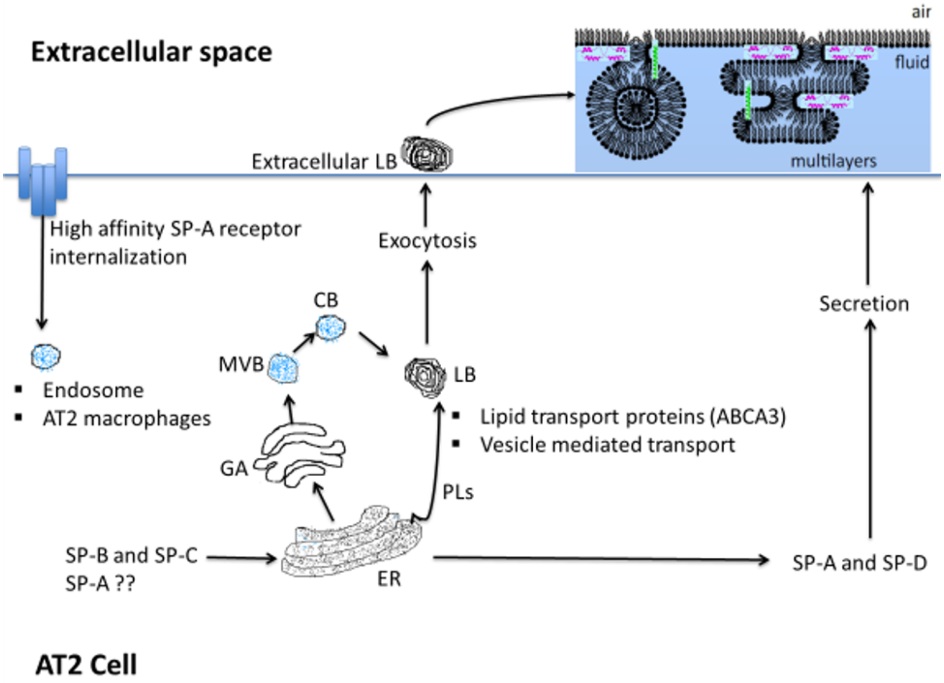


Figure 1.2: Lung surfactant biosynthesis by alveolar type two cells. The synthesis of the two small hydrophobic proteins, SP-B and SP-C, begins in the endoplasmic reticulum (ER) and then they are shuttled through the Golgi apparatus (GA), multivesicular bodies (MVB) and the composite bodies (CB) as they are being processed to their mature form. The two proteins are finally stored in the lamellar bodies (LB). The lamellar bodies are then secreted to the extra-cellular space by regulated exocytosis. Whether SP-A takes this route is a matter of question, however, it is known that SP-A and SP-D are directly secreted from the ER to the outer cellular space. The phospholipids (PLs), on the other hand, bypass the Golgi apparatus and are transported to the LB by lipid transport proteins such as ABCA3, and are also shuttled by vesicle mediated transport. SP-D is exclusively absent from the LB. At the air/water interface, the LB and the hydrophilic proteins are assembled into a complex lipid-protein monolayer along with multilaminated structure beneath as shown to the top right. Depleted surfactant are probably internalized by high affinity SP-A receptor and are acted upon by alveolar macrophages and also degraded in the endosome.

The extracellular LB may be assembled into tubular myelin (TM), an organized meshwork structure of small and large aggregates that comprise most of the surface active material in the hypophase, the surface beneath the interfacial layer [50, 51]. Finally, the TM in the alveolar lumen can unravel and adsorb to the air-water interface forming a monomolecular lipid monolayer lining the air-water interface along with associated multilayer lipid structures underneath [52].

However, whether the secreted vesicles directly adsorb or transit to TM prior to adsorption has been questioned more recently [53, 54]. Figure 1.3 depicts the different models of adsorption of LS after secretion of the LB. Depleted surfactant is cleared from the alveolar surface by three possible mechanisms: i) uptake by AT2 cells for reuse and degradation [55], ii) catabolism by alveolar macrophages [56], and to some extent, iii) transport from the alveolus up the bronchial tree [57, 58].

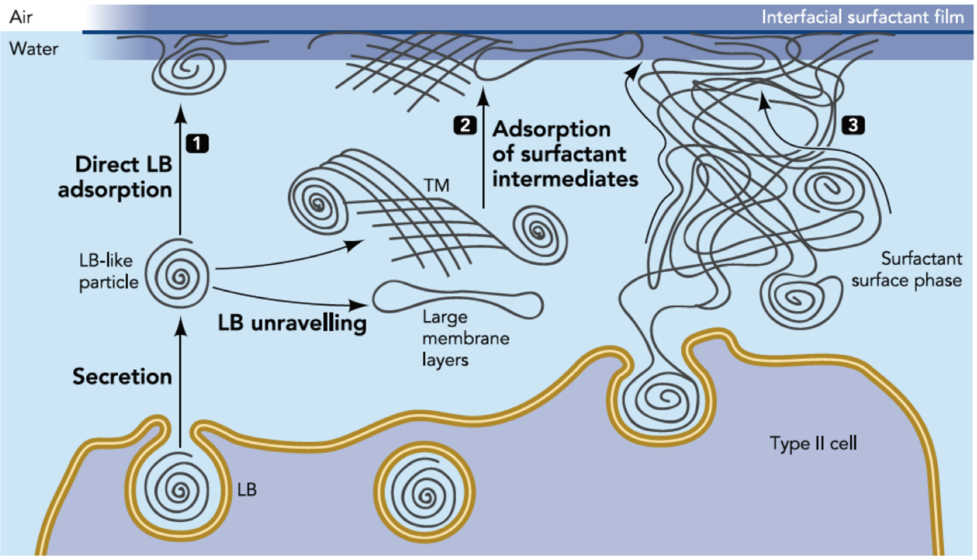


Figure 1.3: Proposed models of lung surfactant secretion and adsorption into the air-water interface. In the first case surfactant secreted as lamellar bodies (LB) by AT2 cells are directly adsorbed to the interface, whereas in the second case the lamellar bodies are assembled into the intermediary tubular myelin prior to adsorption. The last proposition is that the lamellar bodies are unravelled en route from the AT2 cells to the alveolar lumen without forming the intermediate meshwork and simply adsorp to the air-water interface. Adopted from Perez-Gil & Weaver [54] with permission from the publisher.

1.4.1 Protein Synthesis

SP-B is encoded by the *SFTPB* gene on chromosome 2 as a large precursor preproprotein of 381 amino acids (40 kDa). Processing of preproSP-B to mature SP-B (79 residues, 8.7 kDa) involves post-translational modification (PTMs) and extensive proteolysis [59], tightly regulated in optimum acidic pH environment along the exocytotic pathway of surfactant synthesis, Figure 1.4, [60, 61]. Although club (Clara) cells, also known as non-ciliated bronchiolar epithelial cells, and many other cells express the precursor protein, only AT2 cells are able to produce fully matured bioactive SP-B [62, 63]. Many *ex vivo* and animal model studies have implicated three proteinases in SP-B processing, napsin A [64], pepsinogen C [65] and cathepsin H [66]. Napsin A, an aspartic protease, initially known to be highly expressed in renal tubular cells of human kidneys, is also produced within the lung in a tightly controlled manner [67]. Likewise, pepsinogen C is an aspartic protease and is primarily expressed by gastric chief cells, and to a lesser extent by the lung. In the lungs, pepsinogen C is strictly AT2 cell specific and developmentally regulated [65, 68]. Napsin A and pepsinogen C have endopeptidase activity and cut within the amino and carboxy flanking propeptides to leave intermediate 25 kDa and 9 kDa products, respectively [68–70]. However, whether napsin A acts on the carboxyl-terminus of proSP-B is a matter of debate [64, 70]. In contrast, cathepsin H, a cysteine protease abundant in type II pneumocytes, is accepted to be responsible for final maturation of SP-B [66].

Effective silencing of napsin A [70] and pepsinogen C [68] using RNA interference (RNAi), as well as inhibition of cathepsin H with E-64, a common and potent inhibitor of cysteine proteases [66], resulted in decreased levels of mature SP-B in well known models of isolated type II pneumocytes maintained in culture. Conversely, overexpression of pepsinogen C in both explant, as well as fetal and postnatal lungs, increased production of SP-B [68], which supports a role of this enzyme *in vivo*. In addition, the fact that Clara cells, which do not process proSP-B beyond the 25 kDa intermediate, lack pepsinogen C further corroborates the involvement of pepsinogen C in conversion of the 25 kDa precursor to the 9 kDa intermediate product. That said, the exact intracellular compartment(s) in which processing occurs is still ambiguous. Immunohistochemical studies implicated a subcellular localization of napsin A and cathepsin H in the multivesicular and lamellar bodies in cultured cells [70]. What is more, AT2 cells treated with brefeldin A, an inhibitor that blocks trafficking of macromolecules beyond the *cis*-Golgi, do not process proSP-B to smaller fragments. This observation suggests the onset of proteolysis in the Golgi is beyond the cis compartment. Cells treated with mononesin, another inhibitor that hampers protein transport out of the Golgi, exhibit the 25 kDa and 9 kDa intermediate products, but not the mature SP-B, also provides compelling evidence that the N and C-flanking propeptides are truncated within the Golgi, whereas final maturation of SP-B is likely a post Golgi, pre-lamellar event [41, 71].

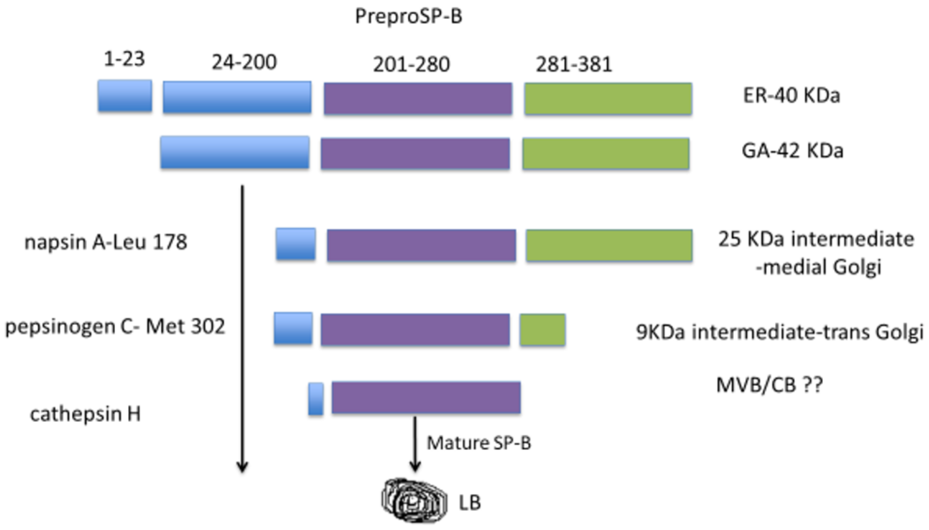


Figure 1.4: Processing of PreproSP-B to the mature form of SP-B by alveolar type two cells. PreproSP-B (40 kDa) is glycosylated in the endoplasmic reticulum and undergoes signal peptide cleavage to produce proSP-B (42 kDa). ProSP-B is then trafficked to the Golgi apparatus intact. In the medial Golgi, the enzyme napsin A cut at leucin-178 and result in a 25 kDa intermediate product. Similarly, in the trans-Golgi a second enzyme, pepcinogen C, cleaves at methionine-302 and leaves another 9 kDa intermediate product. Finally, cathepsin H, probably in the multivesicular bodies and/or composite bodies, trims the N and C flanking pro-peptides and result with the mature protein while still is joined with 13 amino acids at the N-terminus. It is believed that this vestigial peptide assist the mature protein translocate to the lamellar bodies, where mature SP-B is stored.

As it stands now, the widely accepted turns of events taken when preproSP-B is processed to mature SP-B *in vivo* are : i) inside the ER, preproSP-B (40 kDa) undergoes glycosylation of the carboxy terminus and signal peptide (1-23) cleavage producing 42 kDa proSP-B (24-381) [69], which then is shuttled to the Golgi intact, ii) once in the medial Golgi, napsin A acts on the N-terminal flanking propeptide (24-200) of proSP-B cutting after Leu178 to leave a 25 kDa intermediate product coupled with 21 residues at the amino terminal, iii) also in the medial Golgi, further modification of the glycosylated product by manosidase II occurs [41] and renders the 25 kDa product resistant to endo galactosidase activity while it transits from the medial to the trans Golgi, iv) subsequently, in the trans Golgi, pepsinogen C acts on the C-flanking propeptide (280-381), cleaving after Met302 to release the second intermediate product (9 kDa) containing the mature protein, v) finally, in the post Golgi environment, possibly in the multivesicular bodies, [41, 69] the exopeptidase cathepsin H trims the N and C peptides off the 9 kDa precursor protein and the mature SP-B, still coupled with ≈13 residues at the amino terminal side is released. However, the enzyme responsible for removing the vestigial peptide is as yet not known [70]. If an exopeptidase of another sort is present in the pre-lamellar environment, most likely in the composite bodies, is a matter that needs further investigation. Kormilli *et al.* [41] and others suggested a role for the vestigial peptide in trafficking the mature protein to the lamellar bodies, similar to the signal peptide (1-23) that is believed to assist preproSP-B addedto translocate into the ER [59].

Although the exact function of the N and C flanking propeptides are not thoroughly understood, growing bodies of evidence demonstrate that these peptides chaperone the mature protein until it is processed for assembly into surfactant complexes [59, 69, 72]. The N-terminal domain specifically has proved to be crucial in escorting SP-B along the secretory pathway of AT2 cells. Convincingly, transgenic mice lacking the NH_2-terminal propeptide ($\text{SP-B}_{\triangle N(24-200)}$) do not convey SP-B beyond the ER *in vivo*. In contrast, C-flanking knockout mice ($\text{SP-B}_{\triangle C(280-381)}$) secrete fully processed SP-B that supports normal lung function [73, 74], which suggests a profound role of the flanking amino domain in trafficking and correct processing of SP-B. This is also in agreement with the observation that this domain harbours the signal peptide necessary to dispatch SP-B as discussed above. Further evidence to this was provided by Brasch *et al.* [75] who used immunoelectron microscopy and localized mature SP-B and the truncated N-terminus propeptides in the core of lamellar bodies. Hence, mature SP-B and the amino propeptide are shuttled to the lamellar bodies together and thereby may play an important role in the assembly of the lamellar bodies and tubular myelin through the formation of the core particles of these structures. A convincing argument for this is the observation that SP-B knockout mice produce aberrant lamellar bodies, similar to those seen in a congenital genetic disorder in mice [76]. Moreover, the truncated N-terminus has also been detected in bronchoalveolar lavage fluids where microbicidal activity of the peptide is also reported [77]. Despite the C-terminal propeptide's lack of involvement in structure and function of alveolar surfactant, there is an intracellular role of this peptide in the final post-translational processing of proSP-C, as well as maintenance of lamellar body pool size [66, 78, 79].

Correct processing of SP-C relies on the presence of SP-B in the MVB [80]. SP-C is synthesized as a 197 amino acid (21 kDa) proprotein precursor that undergoes endoproteolytic cleavage and palmitoylation of the cysteine residues while en route from the ER to the lamellar bodies, where the final cleavage occurs so as to release a 3.7 kDa (35 residue) mature SP-C [33, 59]. It has been shown that complete lack or inhibition of SP-B leads to the accumulation of MVB, which subsequently hampers proper lamellar body formation. Because the final maturation of SP-C occurs in the LB, improper expression or deficiency of SP-B therefore indirectly leads to the accumulation of the proSP-C precursor intermediate product in the MVB [66, 76, 81]. However, the detailed mechanism of this is not yet described. In contrast, no dependence on SP-B has been documented for the synthesis and processing of the two large lectins, SP-A and SP-D [82], except for in the case of extracellular protein-protein cooperation between SP-B and SP-A [83, 84]. The expression of these lectins, including most of the post-translational modification and oligomerization, happens in the endoplasmic reticulum of type II alveolar epithelial cells [85]. Unlike SP-B and SP-C, the two large hydrophilic proteins largely bypass the lamellar bodies and are constitutively secreted. SP-A and SP-D are also not activated by the same secretagogue that stimulates production of the lipid and the hydrophobic proteins, the details of which will be further discussed in section 1.4.

In conclusion, synthesis of SP-B is highly coupled with lung surfactant biogenesis as a whole and it is also highly developmentally regulated. During AT2 cell differentiation for instance, increased SP-B RNA expression, SP-B processing and maturation, lamellar body formation as well as completion of SP-C processing occurs in chronological order [65, 68, 86]. This is substantiated by the observation that, regardless of high levels SP-C mRNA, AT2 cells of SP-B knockout mice lack mature SP-C and do not show fully formed lamellar bodies [87]. More importantly, SP-B knockout mice are unable to sustain life after birth [82] while SP-C deficient ones survive into adulthood, albeit they develop lung related complications later on in their life [88, 89]. Further, SP-B may also contribute to the assembly of tubular myelin in the alveolar lumen [86]. Recently, extracellular SP-B has also been shown to promote secretion of lamellar bodies by AT2 cells [90]. On the other hand, mice deprived of SP-C exhibit unaltered lamellar bodies and tubular myelin, have the normal pool of the other SPs (SP-A, SP-B and SP-D), and also demonstrate proper morphogenesis of the distal airways [88]. All of this evidence therefore shows that SP-B is an indispensable component of LS, which can compensate for SP-C deficiency. In contrast, SP-C can not compensate for SP-B deficiency, probably because final maturation of SP-C itself is dependent on the expression of SP-B [79]. Clearly, SP-B plays an essential role in lung surfactant biogenesis, which is consistent with the high degree regulation of SP-B processing [63]. Taken together, SP-B processing is a complex post-translationally coordinated sequence of events, perfectly timed to result in a functioning surfactant pool at the right place and in the right time [71].

1.4.2 Lipid Synthesis

The diverse lipid content of lung surfactant ensures optimal function in varied physiological circumstances [59, 91]. The blood supplies AT2 cells with most of the key ingredients for lipid production. The circulation is the major source of the glucose used as the backbone of surfactant glycerolipids, free fatty acids/and triacylglycerol within lipoproteins, as well as the choline and myo-inositol head groups in phospholipids [91, 92]. Nonetheless, AT2 cells can also generate free fatty acids either by hydrolysis of recycled lipids or *de novo* using intracellular glycogen in prenatals or lactate in adults [51]. Figure 1.5 represent major pathways of glycerophospholipid synthesis in mammalian cell. *De novo* lipogenesis involves acetyl coenzyme A (acetyl-CoA) carboxylase, fatty acid synthase (FAS) and citrate lyase. Another important precursor is phosphatidic acid (PA), from which diacylglycerol (DAG), an important substrate to PC and PE, is derived. PA is generated *de novo* within the ER and mitochondria of AT2 cells through the glycerol-3 phosphate (G3P) pathway and dihydroacetone phosphate (DHAP) pathway utilizing glucose or glycogen as a precursor. In addition to this, PA could also be generated by a number of other mechanisms [51]. Phosphatidic acid phosphatase (PAP), the enzyme that converts PA to DAG, plays a key role in regulating phospholipid genesis [93]. Cholesterol, the most abundant neutral lipid in lung surfactant, is primarily supplied by the circulating low- and high-density lipoproteins (LDL and HDL); neverthless AT2 cells can also make cholesterol in peroxisomes [59].

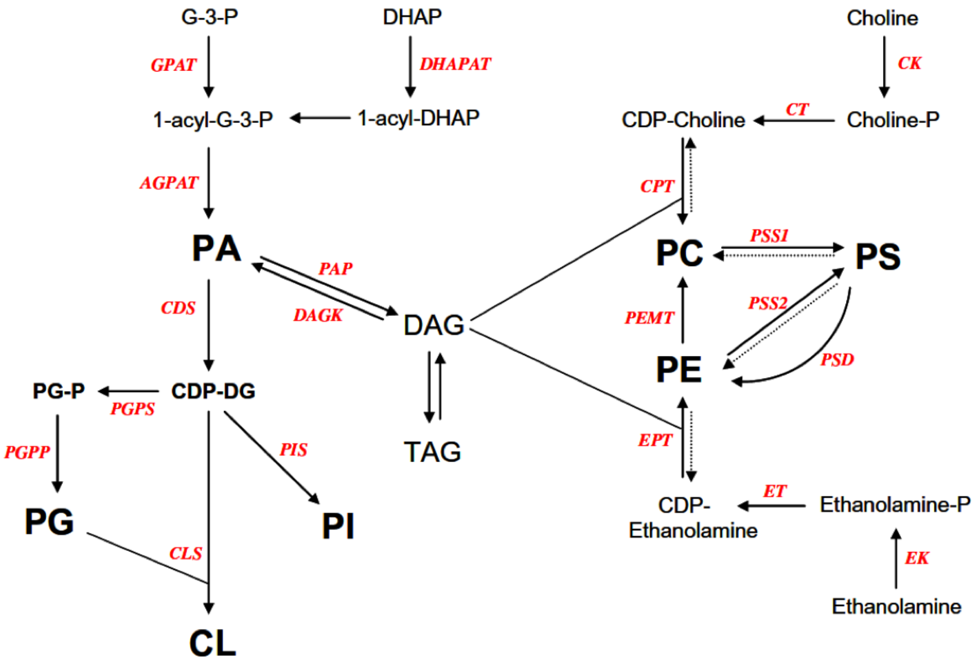


Figure 1.5: Pathways of lipid biosynthesis in mammalian cells. The enzymes catalyzing each reaction are shown in red and the metabolites in black. PG, PI and cardiolipin (CL) are produced from a common precursor, phosphatidic acid (PA) that is dervived from the G3P and DHAP pathways by utilizing glucose and glycogen as a substrate. PA is also useful to generate the DAG that is used in the synthesis of PC and PE. PC is primarily produced in a CDP-choline pathway by using choline as a substrate. Likewise, PE is synthesised in a CDP-ethanolamine pathway using ethanolamine as a substrate. Both PC and PE can be interconverted to phosphatidyl serine (PS). PC can also be generated from PE. Adopted fromHermansson *et al.* [94] with permission from the publisher.

DPPC, the major lipid component of LS ($\approx$40%), is biosynthesized *de novo* via the Kennedy (cytidine diphosphocholine) pathway in the ER [95] or remodeled in the lamellar bodies by the Lands cycle [59, 96]. The *de novo* pathway is a multistep process: initially, choline is phosphorylated by choline kinase (CK) to cholinephosphate, which in turn gets converted to CDP-choline by CTP:phosphocholine cytidylyltransferase (CCT). Finally, in the terminal stage, the reaction of DAG, which is derived from PA, and CDP choline is catalyzed by choline phosphotransferase (CPT) to produce PC. On the other hand, AT2 cells could also take up and remodel lipids in a series of deacylation and reacylation steps. Phospholipase A_2 deacylates unsaturated PC at the *sn-2* position and results in lysophosphatidylcholine (LPC) followed by reacylation of LPC with saturated fatty acid (16:0) by LPCAT1 (lysophosphatidyl choline acetyl transferase) to yield DPPC [97, 98]. About 45% of the DPPC in LS is synthesized *de novo* whereas 55-75% is remodeled for reuse [91, 99, 100]. The enzyme, LPCAT1, which is predominantly expressed in alveolar type II pneumocytes, is known to provide communication between the *de novo* and remodeling pathways such that when LS production is in more demand LCPAT1 expression is upregulated. Over-expression of LPCAT1 desensitizes the *de novo* route by targeting CPT for ubiquitination and degradation, which besides saturated PC, is also responsible for production of other non-surfactant phospholipids [101]. Phospatidylethanolimine (PE), a minor component of LS, is synthesised via a different route of the Kennedy pathway [51].

Phosphatidylglycerol (PG), the most abundant anionic phospholipid in LS, is *de novo* synthesised in the mitochondria and ER from a CDP-diacyglycerol precursor, obtained from the reaction of PA and cytidine triphosphate (CTP). Glycerophosphate phosphatidyltransferase catalyzes the conversion of CDP-diacylglycerol to phosphatidylglycerolphosphate, that subsequently undergoes rapid dephosphorylation to PG [51, 91, 92]. In a like manner to PC remodeling, PG is also remodeled by deacylation and reacylation steps that involve LPCAT1. However, unlike PC remodeling, that also provides less abundant PCs such as palmitoylmyristoyl PC, PG remodeling only leads to the accumulation of dipalmitoyl PG (DPPG) [51], which accounts for $\approx$ 17-38% total PG content [102]. Phophatidylinositol (PI), the next abundant acidic phospholipid after PG, also uses CDP-diacylglycerol as a precursor for initial synthesis; however, there are some studies that suggest PI could be remodeled as well [51]. There is a myriad of information in the literature regarding synthesis of other minor components of lung surfactant, including phosphatidylserine, cardiolipin, and sphigomyelin for reference.

Although there is data to suggest that LS phospholipids are sorted and targeted to the lamellar bodies based on their acyl chain length rather than saturation [103], the details of the trafficking of the phospholipids from the ER to the LB are not yet clear cut. Despite some initial studies indicating vesicle mediated transport through the Golgi and then to the LB, follow-up studies that disrupted the Golgi reported the normal pool of phospholipids in the LB and alveolar lumen [43, 104]. More recently, lipid transfer proteins have been implicated in trafficking the lipids from the ER to the LB [97], or it could be combination of both [105]. While the ABCA3 imports phospholipids and cholesterol to the LB from the cytosol, export from the AT2 cells proceeds via the ABCA1 transporter, which is expressed by and localized at the basolateral membrane of AT2 cells.

1.5 Regulation of Lung Surfactant Secretion

Proteins reach their extracellular target by two secretory pathways, the common constitutive way, or the regulated pathway where stored proteins are exocytosed in response to some sort of stimulus [74]. LS secretion into the alveolar lumen is highly regulated [106, 107] and to some extent, may be constitutive as well [85]. Two types of secretagogues affect secretion. One type are agonists that activate different cell surface receptors. The other are agonists that penetrate the cell and elicit downstream signaling cascades [107]. Figure 1.6 displays the regulation pathways of lung surfactant secretion in AT2 cells. To date, three distinct G-protein dependent signalling pathways are known to stimulate LS secretion in AT2 cells.

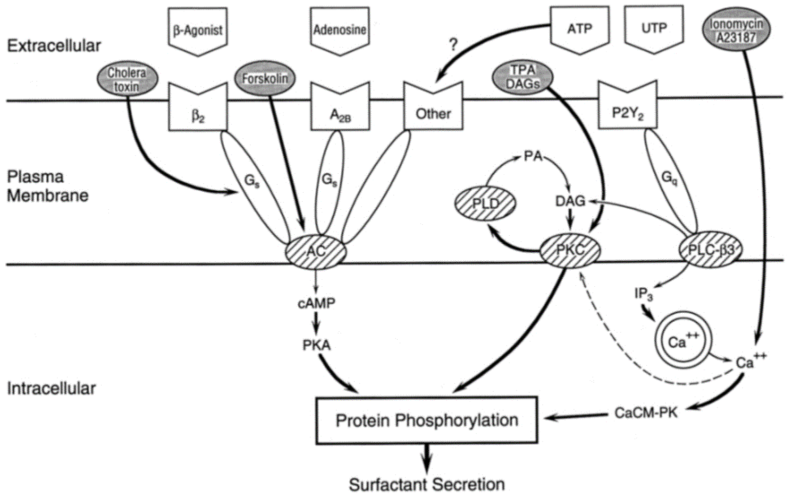


Figure 1.6: Signal transduction pathways of lung surfactant secretion in AT2 cells. The white boxes in the plasma membrane are G-protein coupled receptors that are known to stimulate LS secretion when activated by their extracellular ligands. Three of such receptors are identified so far: the β_2-adrenergic receptor and the adenosine A_{2B} receptors act on adenylyl cyclase (AC) through the heterotrimeric Gs that stimulate the production of the intracellular second messenger, cyclic adenosine monophsphate (cAMP), and activate protein kinase A (PKA) activity. The third receptor, the purinergic $P2Y_2$ receptor, act on phospholipase C (PLC-β3) via Gq to produce the second messenger inositol phosphate 3 (IP_3) and sequester calcium flux into the cytosol. PLC-β3 also produces intracellular DAG to stimulate protein kinase C (PKC). The increase in [Ca^{2+}] in the cytosol thus initiates the calcium-calmodulin protein kinase (CaCM-PK) and PKC activity and increase the secretion of LS. PKC is also directly stimulated by TPA and DAGs that bypass the plasma membrane. Ionophores such as ionomycin and A23187 also leads to the accumulation of cytosolic calcium. Cholera toxin and forskolin escape the plasma membrane and activates the AC pathway of LS secretion. Adopted from Rooney [107] with permission from the publisher.

The first pathway involves activation of adenylate cyclase (AC), which in turn initiates cAMP-dependent protein kinase A (PKA) signalling. This pathway is activated by AC coupled to β-adrenergic and A_{2B} adenosine receptors via Gs, a heterotrimeric G protein subunit that activates AC. In addition to the endogenous ligands that stimulate these receptors i.e. β agonists and adenosine respectively, the AC pathway is also stimulated by other molecules that bypass the plasma membrane. Cholera toxin, which permanently activates Gs, and forskolin, that directly stimulates AC, are examples of such molecules.

The second and major pathway is via stimulation of protein kinase C (PKC). The purinergic receptor, $P2Y_2$ coupled to phospholipase C (PLC)-β3 via Gq, a heterotrimeric G protein subunit that activates PLC, upregulates two secondary messengers. One is intracellular DAG (diacyl glycerol), which directly activates PKC and the second is IP_3 (inositol triphophate) that increases cytosolic calcium and stimulates PKC indirectly. Nucleotides, such as ATP (adenosine triphosphate) and UTP (uridine triphosphate), are equally potent stimulators of the P2Y$_2$ receptor. In addition, cell permeable TPA (12-O-tetradecanoylphorbol-13-acetate) and DAGs are also known to directly act on PKC.

The last mechanism of LS secretion is by activation of Ca^{2+}-calmodulin dependent protein kinase (CaCM-PK). This pathway is triggered by accumulation of intracellular Ca^{2+}. Ionophores such as ionomycin and A23187, which promote extracellular calcium influx into the cell as well as PLC-β3 regulated IP$_3$ dependent accumulation of cytosolic calcium leads to the stimulation of CaCM-PK pathway [52, 107, 108].

ATP activates all three kinases [107, 109], i.e. PKA, PKC and CaCM-PK. Each pathway ultimately leads to protein phosphorylation, and when activated all at once, results in up to 12 to 15 fold surfactant secretion over the basal level [52]. Comparatively, PKC signalling is more potent, inducing approximately five fold secretion of LS, while PKA and CaCM-PK only induce, at the utmost an increase of 2-3 fold each [34, 110, 111]. Besides the agonists, ventilation and labour in the mother are two important physiological factors that trigger LS secretion in the whole lung. However, neither an agonist nor an antagonist is known [108, 112], despite some preliminary data that suggests labor induced secretion, is at least in part, mediated by prostaglandins and β-adrenergic agents [113]. During strenuous excercise surfactant secretion peaks [114]. One deep breath is good enough to stimulate secretion both in adults and newborns [115, 116]. This is also known to be the case *in vitro*, as stretching AT2 cells results in secretion [117, 118]. Nonetheless, the interplay between calcium mediated stretch induced signalling and the pathways leading to kinase activation are not yet clearly established. Further investigation is needed as it is paramount in filling the gap in our understanding of the events leading to LS secretion by AT2 cells [52, 119]. Conversely, SP-A and phosphatidylcholine could inhibit secretion by a sort of negative feedback mechanism; however, the biochemical pathway and regulation by which this process proceeds is not yet delineated [120–124].

Lung surfactant phospholipids and the small hydrophobic proteins are activated by the same agents. In laboratory settings, SP-B and SP-C are stimulated by the same secretagogues as the phospholipids (PLs). Likewise, inhibitors that block PL stimulation also seem to block SP-B and SP-C secretion. Nonetheless, both the extent of stimulation and the time lapse of secretion differs between the phospholipids and the hydrophobic proteins [106]. On the contrary, SP-A and SP-D are not affected by the same agonist that activates PLs or the hydrophobic SPs in primary rat type II cells [125]. Thus, these large lectins may have alternative routing that bypasses the lamellar bodies and perhaps could be secreted constitutively [104]. For instance, SP-D is absent in lamellar bodies and is not involved in the structural organization of LS [108, 126, 127]. Although there seems to be a small amount of SP-A associated with the LB, its source is not pinned down, as SP-A is also internalized by AT2 cells through clathrin mediated uptake from the extracellular space and targeted to the lamellar bodies for resecretion [128, 129]. The question that arises is therefore if newly synthesised SP-A is secreted constitutively or via a regulated route that involves the lamellar bodies as well like the hydrophobic SPs? In this regard, Osani *et al.* [104] reported constitutive secretion and then reuptake to the LB whereas, Fisher *et al.* [130], reported regulated secretion of SP-A that involves the secretory granules.

1.6 Recycling and Degradation of Lung Surfactant

The turnover rate of lung surfactant is rather fast. The time course between secretion and recycling typically ranges from 5 to 11 h [131]. The endocytosis of surfactant complexes by AT2 cells is mediated by a high affinity SP-A receptor through clathrin-coated vesicle pathways and via a non-clathrin, actin dependent pathway [132–134]. The internalized lipid-protein complex dissociates in early endosomes. While most of the SP-A (≈90%) is resecreted quickly; the lipid portion is transported to the MVB where it converges with fresh synthesis to be routed to the lamellar bodies for storage [128]. Furthermore, LS is also cleared by phagocytic alveolar macrophages for degradation. An *in vivo* study in rabbits that used labelled DPPC to estimate the amounts of LS recyling and degradation, reported that about 20% of spent surfactant is cleared by macrophages, while AT2 cells recycle the most (≈65%) [135], and some is also lost up the bronchotracheal airways [57]. However, the mechanisms that sorts lipids and SP-A for internalization by AT2 cells for recycling or catabolism, as opposed to uptake by macrophages for degradation, is unknown [52]. ABCA1 is also believed to play a crucial role in disposing excess surfactant into the circulation [136]. A growing line of evidence suggests that such homeostasis cycle is likely controlled by GPR116/Ig-Hepta signalling [48, 49], possibly by sensing the level of SP-D in the surfactant pool as it appears now that SP-D is a ligand for the orphan receptor, i.e. GPR116/Ig-Hepta [137]. Ablating this receptor results in pulmonary alveolar proteinosis and an emphysema-like pathology [138]. Similar symptoms are also observed in SP-D null mice [139]. About 50-85% of DPPC is internalized for re-utilization by pneumocytes.

1.7 Mechanism of Surface Tension Reduction by Lung Surfactant

Surface tension is a surface characteristic of any surface or interface. In fluids, it is the cohesive force that holds molecule close together at the interface. It is defined by the force per unit length and is expressed in units of mN/m or dynes/cm [140]. In the lungs, gas exchange from inhaled air to the blood occurs via alveoli. The volume and shape of these alveoli are controlled by many factors, among which, alveolar surface tension play a pivotal role. Surface tension determines whether the alveoli remains open for gas exchange or they contract. The lungs does this by lining the alveolar surface with surfactant, a thin and complex fluid layer, and reduce the work of breathing. Lack of this material in the lung impedes breathing and result in collapse of the lungs.

The two prominent biophysical features of lung surfactant are minimal surface tension at the air-water interface during expiration, and efficient film replenishment from the collapsed phases during repeated cycles of breathing. An effective surfactant should therefore rapidly adsorb to the interface and reach its equilibrium surface tension ($\approx$25 mN/m) within a couple of seconds [141, 142]. At the surface, LS spreads and reduces the high surface tension of the air-liquid interface to a near zero value or in another words maximal surface pressure as the surface tension and surface pressure are inversely related (will be discussed in chapter four in more detail). In order to achieve this, LS should be rigid enough to be capable of reaching low surface tension during expiration, and at the same time the mixture must also be fluid enough to spread effectively upon inspiration.

However, neither the lipid nor the protein component of LS demonstrates the two prominent biophysical features on its own. Individual components are only efficient at either achieving low surface tension or fluidizing the monolayer, but not both. Studies have shown that DPPC, which comprises the majority of lung surfactant, is good at achieving very low surface tension, whereas the unsaturated lipids are far better as fluidizing agents [143, 144]. In laboratory settings, for example, DPPC film is able to reduce the surface tension of pristine water (72 mN/m) to near-zero value, at an ambient temperature [145, 146]. Nonetheless, DPPC has a bilayer transition temperature (T_m) of $\approx$41°C, which is above the physiological body temperature [147, 148]. The lipids in a DPPC bilayer are therefore very static and poor at spreading at both ambient and body temperature. Thus, DPPC adsorbs poorly from the hypophase and respreads slowly upon film expansion [149, 150]. However, above the T_m, a DPPC bilayer melts and the lipids can freely diffuse from one leaflet to the other. Likewise, lipid monolayers, can undergo similar, but not identical, phase transitions as the bilayers under lateral compression that depend on the lipid, the length of the acyl chain and the temperature [151]. Figure 1.7 depicts phase transitions in a DPPC monolayer under lateral compression.

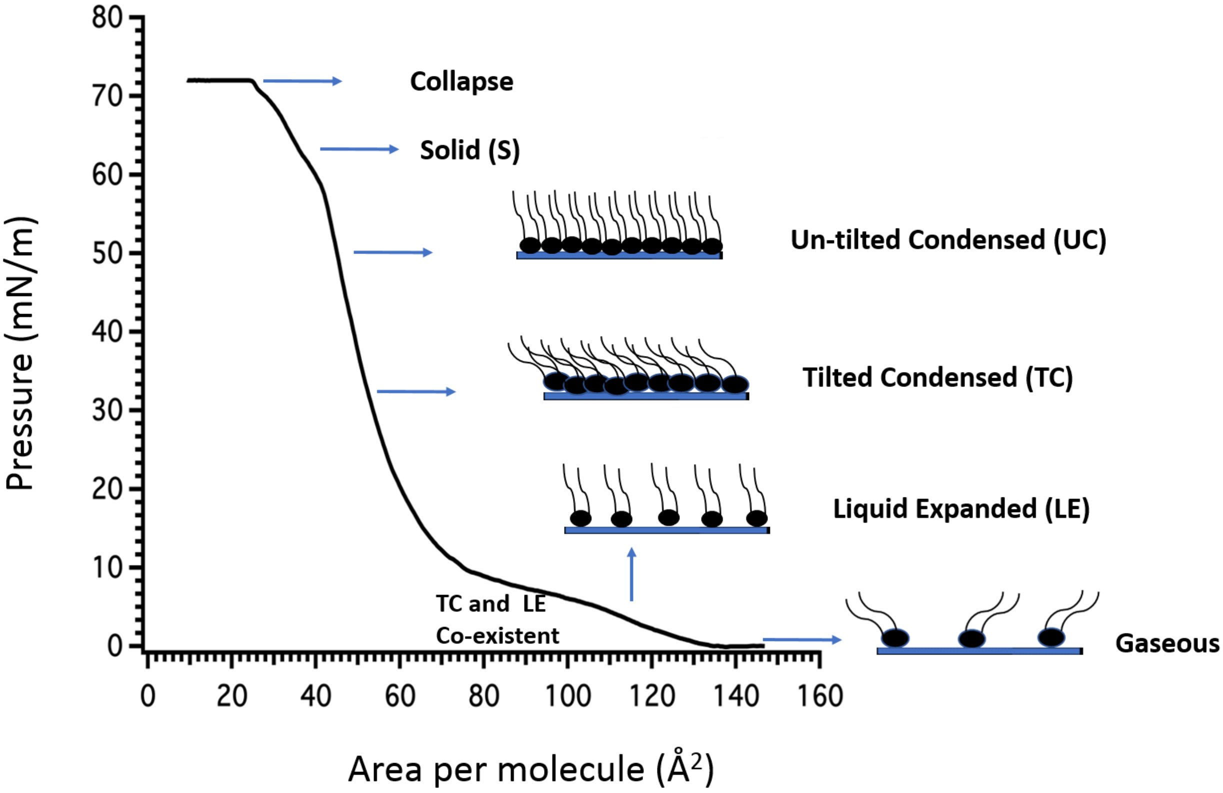


Figure 1.7: Phase transitions of DPPC monolayer at an air-water interface when compressed at room temperature. DPPC is able to reduce the surface tension of pristine water to near zero value. The isotherm was measured by me and the sketch to the right was also drawn by me.

Starting with larger area per lipid molecule, apparent are distinct phase transitions from the random gaseous phase to the liquid expanded (LE) phase, and then to the more tightly packed tilted condensed (TC) and untilted condensed phases (UC), which make DPPC capable of maintaining minimal surface tension at 37°C and sustaining it for a long period of time. Due to the double bond, unsaturated lipids can not tightly pack and hence are unable to generate a TC phase that is rigid enough to reach near-zero surface tension on compression. For an example, palmitoyloleoylphosphocholine (POPC), with a T_m of -6.5°C, is only fluid at room and also physiological temperature and hence can not generate a TC phase capable of reaching low surface tension [152]. The presence of other unsaturated phospholipids in surfactant membrane has been shown to reduce the melting temperature of the membrane to below that observed for pure DPPC [153–155]. In addition, unsaturated lipids add fluidity to LS mixture and increase DPPC's ability to readsorb and respread [148, 150].

Lung surfactant constituents work in a concerted manner, i.e. cooperation between the protein and lipid components is mandatory [145, 156]. The ability of other parts of LS, especially the two small hydrophobic proteins, in facilitating the adsorption and spreading of DPPC back into the interfacial film from the collapsed phases is crucial for proper lung function. [145]. The collapsed phases are multilayer and bilayer aggregates of squeezed-out films from the surface layer when alveoli get flattened at end expiration. The films make lipid-protein protrusions down into the hypophase and mostly stay in close proximity to the air-water interface. In the alveoli, collapse has to be well orchestrated in a way that functional LS material can be readily re-adsorbed to the air-water interface upon inspiration.

Depending on the composition and phase behavior of the lipid(s) that constitute the interfacial layer *in vitro*, monomolecular films can collapse via two folding mechanisms, reversible and irreversible [144]. Monolayers in either liquid condensed (LC) or solid (S) phases collapse irreversibly by fracture. Similarly, films in the more disordered liquid expanded (LE) phase, collapse by solubilization into the hypophase, subsequently forming aggregate materials such as vesicles or liposomes that do not readsorb back into the interfacial layer. On the other hand, films with a continuous LE phase surrounding islands of LC domains or S phase exhibit enough elasticity to buckle and fold as well as the integrity to respread back to the interface upon expansion. Such films therefore are able to collapse reversibly [157]. Most biological monolayers, which are composed of different components of lipids exhibit a bi-phasic nature at high compression, i.e. both fluid and rigid phases co-exist and hence are able to undergo reversible collapse. However, the molecular mechanism of how such folding transitions operate in lung surfactant *in vivo* is not clearly understood. Several different mechanisms have been however suggested. According to the squeeze-out theory, fluid lipids are selectively squeezed out of the interfacial film during compression and at high surface pressure, to ease lateral strain and achieve minimal tension. However, the theory fails to address the possible mechanisms of how LS respreads on subsequent cycling [158]. Schurch *et al.* [141] confirmed experimentally that there is indeed surplus film material that forms surface-associated reservoirs in the third dimension, both in case of *de novo* adsorption or when the film is compressed beyond a certain threshold surface pressure. It was thus evident that these reservoirs either remain associated with, or stay in close proximity to, the interface [159]. Follow-up studies proposed a modified version of the squeeze-out theory with a possible explanation for how these folds, i.e. the multilayers and bilayers stacks that extend down to the hypo or hyperphase of the monolayer are sustained. The theory reasons that the presence of SP-B and SP-C compensates for the squeeze-out effect by promoting reversible folding transitions [144]. These transitions have been referred to differently by many research groups: monolayer (2D) to multilayer (3D) transitions, charge neutralizing docking sites [160] and nanosilos [161]. The protrusions form exclusively from the LE phase, and extend a couple of nanometers down to the hypophase of the alveolar fluid or squeeze onto the upper side of the air-water interface [162]. Such reversible membrane folds are the first such structures seen in nature and hence possibly are unique mechanism by which alveoli retain surfactant near the air-water interface without losing film material irreversibly to the subphase at end expiration of the breathing cycle. In addition, the same folds also appear to facilitate faster readsorption of squeezed-out film materials on subsequent cycles of inspiration [144, 157, 163, 164]. A TOF-SIMS (time of flight-secondary ion mass spectrometry) study provided evidence that the multilayer structures are enriched with unsaturated PLs while the interface remains concentrated with disaturated lipids [158]. Apparently, SP-B has been found to partition exclusively in the LE phase.

Although the molecular basis of how SP-B and SP-C induce and sustain this reversible folding transition is not clearly understood [144], there are some biophysical studies to suggest that SP-B and SP-C proteins alter the packing of the monolayer through electrostatic and hydrophobic interaction with the lipids [33, 165, 166]. Figure 1.8 displays proposed mechanisms of lung surfactant secretion, adsorption and stabilization of the air-water interface.

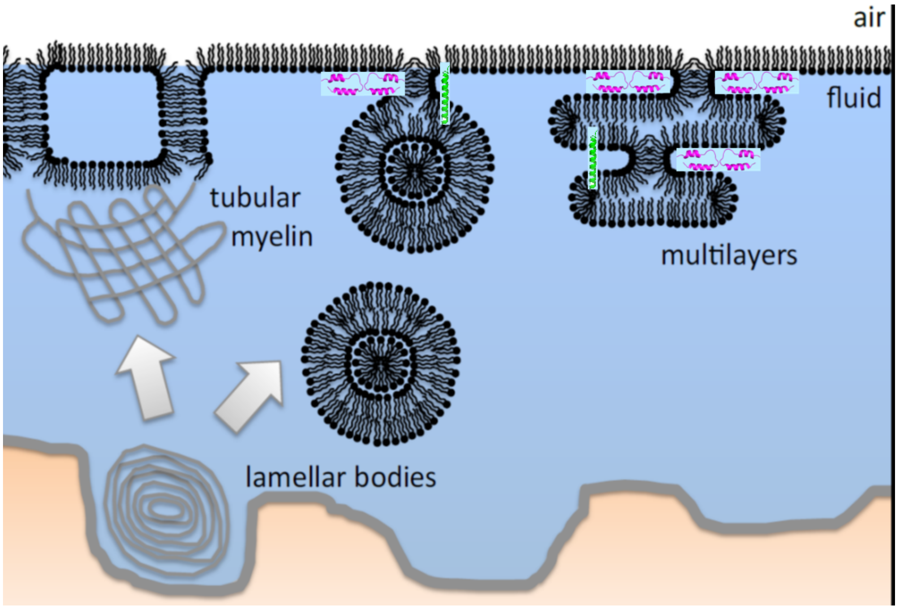


Figure 1.8: Proposed mechanism of lung surfactant adsorption and stabilization during dynamic cycles of breathing. Formation of multilayer stacks and readsorption of the squeezed-out films during compression is shown to the right. *De novo* adsoprtion of newly secreted surfactant material is also depicted to the left. SP-B (magenta) and SP-C (green) are believed to undergo interaction with the lipids to stabilize the protrusions as well as facilitate adsorption of newly secreted lipids into the interfacial layer. The graph was generously provided by Valerie Booth.

SP-B, in particular, alters the size and distribution of ordered lipid domains in the monolayer by raising the line tension and dipole density of the monomolecular film. Lipid domains that form in the presence of SP-B are smaller in size and homogenous, and consequently are more tightly packed with less compression than film without SP-B. Such films are therefore fast at achieving minimal surface tension with the fewest numbers of compression and expansion breathing cycles [162, 167]. The protein-lipid interactions that underlie reversible folding likely also relate to *in vitro* characteristics of SP-B. In laboratory setting, for example, SP-B has been found to facilitate lipid mixing between SP-B containing and fluorescent labelled vesicles at physiological temperature in a concentration dependent manner. Lipid mixing has also been shown to increase when anionic lipids, PG and PI, are increased. Conversely, SP-C shows no sign of lipid mixing ability even at elevated concentrations compared to the physiological amount [168]. In summary, SP-B and SP-C facilitate the formation of lipid protrusions as well as the readsorption and spreading of the lipid patchs back into the superficial film. These properties keep the monolayer intact during consecutive cycles of breathing [164].

Due to the complex nature of lung surfactant, there is no simple *in vitro* assay to test LS activity. To add to the complexity, in contrast to the actual pulmonary surfactant that adsorbs from the hypophase *in vivo*, exogenous surfactants are added to the air side of the interface, which likely introduces artifacts. Efficacy of LS can be tested *in vivo*, *in situ* and *in vitro*. Most *in vivo* methods study efficacy in preterm animals deficient in LS or use diseased adult models, including rabbit [169–172] and lamb [173–175], to investigate how LS dysfunctions (to be discussed in the pathophysiology section below). One common approach is the saline lavage method where the animal is deprived of its own surfactant and exogenous surfactant is instilled to check for the lungs ability to inflate and deflate under normal condition. In addition to this blood oxygen level of the animal can also be monitored [147, 150]. These techniques are best suited for testing efficacy and comparing surfactant preparations [176–178].

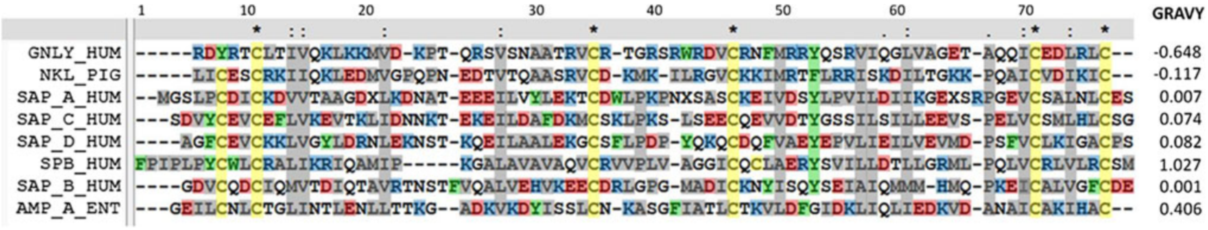
A more simplistic approach to *in vivo* is *in situ*, which measures alveolar surface tension in excised lung tissue. The two *in situ* options are: i) the pressure-volume (P-V) method initially developed by von Neergaard [179] and ii) the more accurate microdroplet technique developed by Schurch *et al.* [180]. Due to convenience of use, ease with which data can be collected and cost of measurement, however, *in vitro* methods are more used than the others.

More importantly, the Langmuir-Wilhelmy balance (LWB), pulsating bubble surfactometer (PBS) and captive bubble surfactometer (CBS) have been widely used and more recently, the constrained sessile drop (CSD) technique has started to gain more attention [147, 150]. All of these *in vitro* methods are suitable to characterize the three prominent biophysical features of LS: i.e. adsorption and spreading kinetics, and interfacial activity. Therefore, it is consideration of the desired features and degree of accuracy required that push researchers to choose one over the other [147].

1.8 Structure-Function Relationships of SP-B

SP-B has been evolutionarily conserved for about 300 million years as it is observed in primitive lungs, in particular in the Australian lungfish, *Neoceratodus forsteri* [181]. The lungfish contains surfactant-like material with SP-A and SP-B-like proteins that bind the same antibodies that bind human SP-A and SP-B. The primary structure of mature human SP-B is highly cationic, with an overall net charge of +7, derived from nine positive and two negativly charged amino acids. A similar charge density has also been observed in SP-B in other animals [182]. Strikingly, around 52% of the residues in mature SP-B are hydrophobic [183].

SP-B shows sequence homology with saposins A-D, lipid binding proteins that activate lysosomal lipid hydrolases [184]. This makes SP-B belong to the saposin-like protein (SAPLIP) superfamily, a diverse group of over 200 proteins that perform varied functions [185]. The SAPLIP superfamily exhibits six strictly conserved cysteines as is shown in Figure 1.9 marked in yellow. The GRAVY index to the right displays that SP-B is exceptionally hydrophobic even amongst the SAPLIP superfamily.



```
          1        10        20        30        40        50        60        70              GRAVY
                    *  ::      :         :       *           *           . :       .  :*  :  *
 GNLY_HUM -----RDYRTCLTIVQKLKKMVD-KPT-QRSVSNAATRVCR-TGRSRWRDVCRNFMRRYQSRVIQGLVAGET-AQQICEDLRLC--  -0.648
  NKL_PIG -----LICESCRKIIQKLEDMVGPQPN-EDTVTQAASRVCD-KMK-ILRGVCKKIMRTFLRRISKDILTGKK-PQAICVDIKIC--  -0.117
SAP_A_HUM --MGSLPCDICKDVVTAAGDXLKDNAT-EEEILVYLEKTCDWLPKPNXSASCKEIVDSYLPVILDIIKGEXSRPGEVCSALNLCES   0.007
SAP_C_HUM ---SDVYCEVCEFLVKEVTKLIDNNKT-EKEILDAFDKMCSKLPKS-LSEECQEVVDTYGSSILSILLEEVS-PELVCSMLHLCSG   0.074
SAP_D_HUM ----AGFCEVCKKLVGYLDRNLEKNST-KQEILAALEKGCSFLPDP-YQKQCDQFVAEYEPVLIEILVEVMD-PSFVCLKIGACPS   0.082
  SPB_HUM FPIPLPYCWLCRALIKRIQAMIP-----KGALAVAVAQVCRVVPLV-AGGICQCLAERYSVILLDTLLGRML-PQLVCRLVLRCSM   1.027
SAP_B_HUM ----GDVCQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPG-MADICKNYISQYSEIAIQMMM-HMQ-PKEICALVGFCDE   0.001
AMP_A_ENT ---GEILCNLCTGLINTLENLLTTKG--ADKVKDYISSLCN-KASGFIATLCTKVLDFGIDKLIQLIEDKVD-ANAICAKIHAC--   0.406
```

Figure 1.9: Sequence alignment of some common SAPLIPs. Human saposins (A-D), NK-lysin from pig, human granulysin and ameobapore A from Entameoba histolytica are shown aligned. Yellow residues depict the conserved cysteines, hydrophobic residues are marked grey. Aromatic residues are shown in green while charged residues are marked blue and red. The GRAVY index to the right is a measure of hydrophobicity. SP-B shows the highest of all. The next hydrophobic protein in the list, ameobapore A, is more than two fold less hydrophobic than SP-B, which indicates that SP-B is extremely hydrophobic. Picture was adopted from Olmeda *et al.* [183] with permission from the publisher.

The saposins are the four proteins (A-D) after which the SAPLIP superfamiliy is named. All four of the saposins are generated from a common precursor protein, prosaposin, by sequential proteolytic cleavage. Saposins (A-D) show the same structural motif, the so called saposin fold also common to other SAPLIPs [186]. The saposin fold is characterized by four or five amphiphatic helices that are connected by the three intrachain disulphide bridges. The disulphides make SAPLIPs extremely resistant to heat, acidic environments and proteolytic degradation [185, 186]. Similar to saposins, all of the SAPLIPs are also able to bind lipids [182, 186, 187]. Apart from SP-B, however, the rest of the SAPLIPs are water soluble and only have one exposed hydrophobic region, the site responsible for binding lipids [188]. Hence, SAPLIPs are only partly lipid associated [189].

SAPLIPs have a variety of lipid binding related functions. Some SAPLIPs, like the amoebapores found in *Entameoba histolytica*, are pore forming polypeptides [190, 191], whereas, NK-lysin and granulysin are lysosomal antimicrobial polypeptides that permeabilize and lyse cells [187, 192, 193]. On the other hand, regions of the enzyme acid sphingomyelinase (ASM) catalyze the breakdown of sphingomyelin to ceramide and phosphorylcholine. Likewise, acyloxyacylhydrolase (AOAH) removes the secondary acyl chains from lipid A of lipopolysaccharides (LPS) [194, 195]. Phytepsin, the plant specific domain, is an aspartic acid proteinase [196–198]. Although there is much variation in SAPLIP function, there is still sequence conservation. For instance, SP-B and NK-lysin share ≈24% sequence identity (49% similarity) and demonstrate the same disulphide bonding pattern and similar alpha-helical content. The rest of the SAPLIPs show 18 to 27% sequence similarity to NK-lysin and SP-B [187]. Neverthless, as of yet, it is not clear how such homologous polypeptides perform quite diverse *in vivo* functions [184, 186, 199]. Notably, the SAPLIPs vary in charge, whilst saposins are overall negative, ameobapore A is neutral, whereas SP-B, NK-lysin and granulysin are all positive [183]. Clearly, the amino acid sequence and charges govern the function [185, 186].

Proteins in the SAPLIP superfamily adopt different overall conformations, a closed globular form or an open configuration. Figure 1.10 shows the closed secondary structure of NK-lysin (Panel A) and the open conformations of saposin C in SDS (Panel B) and saposin A in the presence of the detergent LDAO (Panel C). NK-lysin (PDB code: 1NKL) [192], ameobapore A (PDB code: 1OF9) [200] and the plant specific insert of prophytepsin (PDB code: 1QDM) [198] all demonstrate five amphiphatic alpha helices that fold into a single closed globular domain. Conversely, the crystal structure of saposin B (PDB code: 1N69) exhibits an unusual homodimeric tertiary strucure irrespective of pH or surrounding environment. Each monomer of saposin B folds into four amphiphatic α-helices and adopts a boomerang like open conformation, with the outside convex region composed of hydrophilic amino acids and the inside concave surface with hydrophobic side chains [184]. The presence or absence of the extra helix in SAPLIPs is dependent on the sharpness of the kink in helix 3 at Y54, which is considered to be a breaker of the helix [201]. In the crystal structure of saposin B, for example, the kink at Y54 induces a bend in Chain B. Consequently, the asymmetric dimer of chain A and B encloses a large hydrophobic cavity for lipid binding in saposin B [184]. Thus, based on the type of chains that associate, the dimer of saposin B can be open or closed. On the contrary, unlike saposin B, which takes on the same conformation irrespective of pH and its surroundings, some SAPLIP structures are evidently pH and environment dependent.

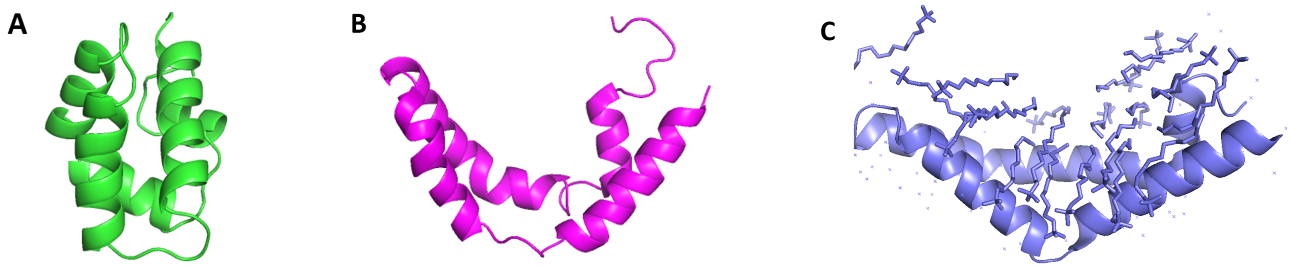


Figure 1.10: A) Solution NMR structure of NK-lysin (1NKL, green) adopting the closed conformation, B) human saposin C in SDS micelles (1SN6, magenta) in the open V-shaped orientation and C) the crystal structure of saposin A (4DDJ, blue) in complex with the detergent lauryldimethylamine-N-oxide (LDAO) in the extended open form. Note: the figures depicted here are all for chain A. The reader is prompted to refer the crystal structure of saposin B (PDB code: 1N69) to look chain B and C. Pictures were drawn using PyMOL.

While the α-helical content of NK-lysin, ameobapore and SP-B does not change in membranes when compared to aqueous solution [187, 191, 192, 202], saposins and some of the SAPLIPs do change their tertiary/quaternary structure according to pH and their environment. Importantly, saposin A and C were found to be monomeric at a pH of 7 in solution, as determined by analytical ultracentrifugation, but exist as a dimer and trimer at a pH of 4.8 in the presence of C_8E_5 detergent [201]. It could be that oligomerization driven by pH and the surrounding environment is mandatory for the *in vivo* function of saposin A and C. Evidently, the monomer solution NMR structure of saposin C in SDS detergent micelles and without detergent have different configurations, an open conformation (Figure 1.10B) in the micelle (PDB code:1SN6) and closed form (PDB code: 1M12) in the detergent free system [188, 203]. In a like manner, saposin A was also observed to take on an extended open form when bound to the detergent LDAO (PDB code: 4DDJ) in acidic pH (Figure 1.10C) [204] as opposed to the closed form (PDB code: 2DOB) without detergent.

Likewise, ameobapore A undergoes pH-dependent dimerization triggered by protonation of the histidine residue (H75) to spark an electrostatic interaction with a neighbouring monomer, particularly residue D63 [200, 205]. In an earlier study, Leippe *et al.* [190] used a chemical cross-linking experiment to verify that ameobapore A self-associates into higher order oligomers in lipid vesicles, a hexamer in this case. The arrangement into hexamer conformation is an important pre-requisite for ameobapore A's pore forming ability and hence, its pathogenicity. The hexameric structure of ameobapore was modelled *in silico* into a ring-like structure that makes a pore, exposing the hydrophobic ring to the outside and hydrophilic residues inside the pore when viewed from above [205]. Nonetheless, pH by itself is not sufficient for saposins to change their quaternary structure, but lipids or detergents are an absolute essential [206].

Self-association of human saposin C and D (PDB code: 2QYP/2Z9A, 2R1Q respectively) also seem to be pH dependent. The kink at Y54 has also been found to play an important role for the conformational flexibility of saposin C. Helix three, which harbours this residue, straightens when binding detergent micelles, which is in contrast to the compact structures reported without detergent [207]. The functional importance of Y54 is supported by the fact that this residue is conserved [183]. On the other hand, activity of the defence proteins NK-lysin and granulysin are pH independent, hence charge distribution and thus surface potential must play crucial role. As for SP-B, Booth *et al.* [208] have observed that the C-terminal SP-B fragment (63-78) in SDS micelles and in organic solvent, hexaflouro-2-propanol (HFIP) (PDB code: 1RG3 and 1RG4 respectively) have subtle structural differences, which could suggest that interactions with lipids alter SP-B's structure. However, due to SP-B's unique hydrophobic nature, the full length SP-B has proven to be very difficult to characterize structurally.

SP-B's overall three-dimensional structure is not yet known. The size and the hydrophobicity of SP-B makes structure determination of the protein very challenging [209]. Hence, the detailed mechanism of SP-B's action is unknown. Whether SP-B makes specific lipid-protein or protein-protein interactions is not yet clearly established. Studies have shown that SP-B likely possesses 4-5 amphiphatic helices similar to the other SAPLIPs. The N-terminus seven residues of SP-B comprise the so-called insertion sequence. The insertion sequence does not share homology with the SAPLIP family and hence it is difficult to predict the secondary structure of the region. According to the sequence alignment by Andersson *et al.* [187], for instance, SP-B conserves residues 8-22, 27-38, 42-50, 67-74 with the SAPLIPs that are predicted to fold into four amphiphatic α-helices per SP-B monomer. On the other hand, a recent model reported by Khatami *et al.* [210] has 5 helices with the insertion sequence (3-7) adopting helical conformation. Figure 1.11 shows the secondary structure taken on by SP-B in the Khatami *et al.* molecular dynamics simulation.

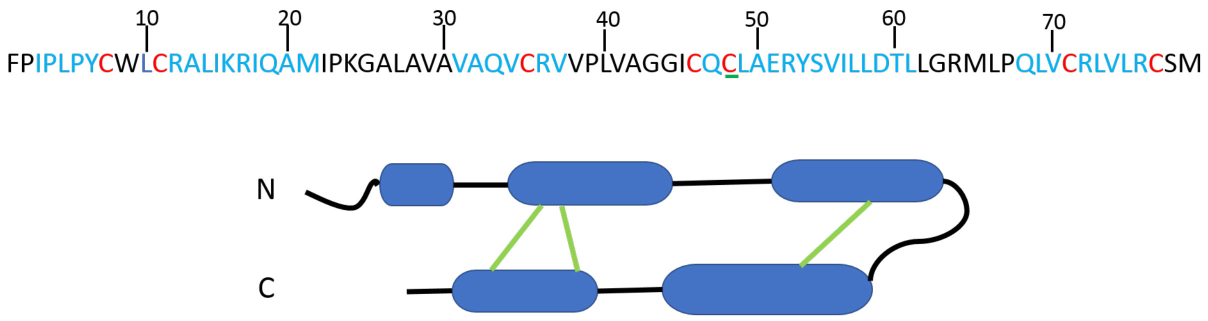


Figure 1.11: The primary and secondary structure of SP-B. The predicted helices are highlighted in blue. The seven cysteines are colored red. The single cysteine residue responsible for interchain disulphide bond is underlined in green. The three intrachain disulphide bonds per SP-B monomer are depicted in green lines. Model concept was adapted from Khatami *et al.* [210]

Like saposins A-D, the secondary structure of SP-B is stabilized by three intrachain disulphide bridges, between Cys8 and Cys77, Cys11 and Cys71, Cys35 and Cys46 [187, 211]. However, contrary to SAPLIPs, SP-B possesses an extra cysteine at Cys48 in the mature SP-B (Cys248 in the proprotein form) [212], which interacts with a neighboring Cys48 on another SP-B molecule at the interface of helix three, to form a covalent homodimer [189, 211]. This makes SP-B the only disulphide linked homodimer in the SAPLIP family [209].

The Khatami *et al.* all-atom molecular dynamic simulation models were built based in part on homology with NK-lysin and in part on Mini-B NMR structure [210]. Helix I (3-7), in contrast to other models from other groups, started as helical at the start of simulation and remained helical at the end of some of the simulations. In the final stable models, five different salt bridges were observed: R64-D59, K24-D59, R52-E51, R36-E51 and R52-D59. The ion pairs could potentially play a role in stabilizing the secondary and tertiary structure of SP-B, as well as possibly aiding in the interaction of SP-B with lipid bilayers. It could also be that the salt bridges stabilize the quaternary structure of SP-B by joining two SP-B dimers together or even one SP-B subunit with another SP-B subunit. Interestingly, the ion pair between E51 and R52 at the surface of helix three was also predicted earlier by Zaltash *et al.* [209]. Supporting evidence of an ion pair between E51 and R52 is the fact that these two residues are all conserved in SP-B's sequence across different species but are absent in all monomeric saposins. Specifically, the acidic Glu51 is strictly conserved. Overall, all of the SP-B models proposed by both groups depict charge density (9 positive and 2 negative) at the opposite poles of SP-B and support a role for SP-B in cross-linking lipid membranes, i.e fusion of two lipid bilayers together.

Mini-B, the largest SP-B fragment for which there is an N-terminal structure, is a 34 residue, N (8-25) and C (63-78) terminal contsruct of SP-B [213]. Figure 1.12 displays the primary and secondary structure of Mini-B.

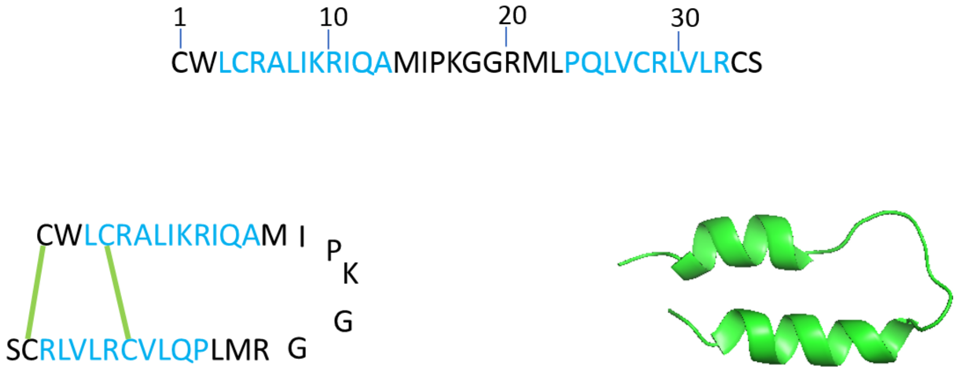


Figure 1.12: The primary and secondary structure of Mini-B. The cartoon representation of Mini-B is shown with two intra-chain disulphide bonds in solid green lines. The experimental structure (PDB code: 2DWF) of Mini-B is shown adapting the expected alpha helical secondary structure [213]. The solution structure was drawn with PyMOL.

Mini-B contains two helical regions interconnected by two disulphide bridges. Like the full length SP-B, Mini-B, also exhibits a net charge of +7 and comprises 41% of the hydrophobic residues [213]. Presumably, for this reason, Mini-B has been found to retain some of the important biophysical activity of full length SP-B, both *in vitro* and *in vivo*. Thus, the N and C-terminal helix structures and charge must play crucial roles in SP-B's function [214]. However, due to lack of the second pair of helices, the structure of Mini-B cannot tell us if SP-B adopts the open or closed conformation observed in other SAPLIP family members [183]. Nonetheless, the secondary structure of SP-B has been found to be stable in a number of environments including fluorinated solvents and detergent micelles.

SP-B shows a consistent predominantly alpha helical conformation in different solution environments. Unlike SP-C, which is highly structurally unstable in the absence of lipids [215], SP-B has more stable secondary structure in different environments [216]. Due to the unusually high hydrophobic residue contents of SP-B and SP-C, the two proteins are water insoluble. Hence, organic solvents or detergents are used to solubilize these proteins but there is no guarantee that the details of the protein structures are the same in solvent or detergent as *in vivo*. Porcine SP-B, for instance, exhibits ≈45% alpha helical content in methanol (MeOH), 60% acetonitrile (ACN) and 70% triflouroethanol (TFE), as observed from CD spectra. Comparable helicity is also deduced when SP-B is reconstituted into lysophosphatidylcholine (LPC) micelles or DPPC vesicles [217]. However, with high amounts of TFE (>70%), SP-B demonstrates as much as 60%

helical content [216]. What is more, a fourier transform infrared spectroscopy (FTIR) study revealed that nearly half of SP-B adopts an α-helix conformation in DPPC/PG model membranes [218]. In contrast, SP-C tends to change conformation in different environments. For example, SP-C takes on an alpha-helical content of nearly 50% in 60% ACN, but is completely β-sheet in 70% TFE [217, 219]. SP-C is perhaps the most hydrophobic protein isolated from mammals [220] and hence is very prone to aggregation and amyloid beta formation. Therefore, in polar solvents, hydrophobic interactions drive more protein-protein interactions. SP-B and SP-C also differ widely from each other in size and charge density, as well as the types of cysteine bonds they exhibit. While intra and inter protein disulfide bonds form in SP-B, SP-C exhibits a thioester palmitoylation of the cysteine residues [211, 221]. Thus solution conditions may affect SP-B's oligomerization and one major unsolved question is the relationship between SP-B oligomerization and function.

SP-B dimerization seems to be important for optimum function. Although it was found that mutating C48 does not have a major effect on SP-B activity, some studies have shown that SP-B dimerization is mandatory for full function [222, 223]. For instance, transgenic mice expressing mutant SP-B, with the putative interchain disulphide forming cysteine replaced by serine (Cys48Ser), are relatively healthy. But, when their SP-B was probed *in vitro* at concentrations lower than 1 µM, the protein was found to exist predominately as a monomer and perform poorly. Whereas, at concentrations above 2 µM, the same SP-B variant was also observed to exist as a non-covalent dimer and demonstrate adequate function as the native form of SP-B. Such observation also supports the prediction made by Zaltash *et al.* [209], i.e., the association of SP-B monomers via the E51 to R52 hydrogen bonds/ion-pairs in a concentration dependent fashion. Even though SP-B null mice rescued with SP-B monomer experience normal longevity, breed successfully and process SP-C, transgenic mice expressing human monomeric SP-B on the other hand exhibit altered lung hysteresis *in vivo* and altered surface activity *in vitro* [224]. Hence, it seems that dimerization of SP-B could be important for normal lung function but may not be so essential for surfactant function. Further, it was also noticed that the dimeric form of the N-terminal 1-25 segment of SP-B supports dynamic cycling, whilst the monomer form of the same N-terminal 1-25 segment of SP-B loses film material during compression in captive bubble experiments [225]. Thus, it could be anticipated that the dimerization of SP-B may be important to support the lipid reservoir near the air-water interface that replenishes the surface active layer upon expansion. In general terms, all of the evidence suggests that from an activity point of view, the loss of C48 covalent bond has no negative implication and that simple hydrogen bonds/ion pairing could be sufficient for SP-B to dimerize. That said, however, the inter-subunit disulphide bonds do further stabilize the dimer structure. Though SP-B is thought to exist as a covalent linked homodimer in most species, higher order oligomers have also been reported [219, 226, 227].

SP-B extracted from lung lavage has a propensity to self-associate into assemblies that are larger than dimers, including trimers and oligomers. Bovine SP-B has been found as a trimer as well as a dimer when isolated by liquid chromatography using a C8 column. In rare cases, SP-B has also been found as higher order oligomers and monomer too. Studies have indicated that such oligomerization behaviour of SP-B is likely triggered by the type of isolation, handling and storage of the protein [226, 227]. In a like manner, ovine SP-B has been detected in varied sizes with SDS-PAGE depending on the method of isolation and type of organic solvent used to extract the protein [219]. Thus, it seems that the purification protocol employed in isolating SP-B, i.e. changing the protein-lipid ratio and solvent polarity, has an effect on the oligomerization propensity of SP-B. Furthermore, porcine SP-B solubilised with the zwitterionic detergent CHAPS in the absence of organic solvents, has been reported to exhibit a supradimeric arrangement, as inferred from blue-native gel electrophoresis, analytical ultracentrifugation and electron microscopy [228]. The group observed a 10 nm diameter ring structure and proposed it to be a higher assembly of SP-B, probably multimers of dimers. A pentameric and hexameric model based on the structure of saposin B fits well to the electron microscopy density of the ring structure. Such structure may be responsible for forming a hydrophobic ring on the surface of phospholipid bilayers and monolayers and potentially transferring lipid molecules during *de novo* secretion or between the monolayer and the bilayer as well as the multilayer structures underneath. This method of preparing SP-B is contrary to the classical harsh organic solvent extraction method, which tends to disrupt weak interactions and hence any supramolecular arrangement. To sum up, oligomerization seem to be a common feature of saposins and the SAPLIP superfamily [207]. Such observations opens up the possibility that particular quaternary structures of SP-B are responsible for particular aspects of SP-B function.

One possibility is that each SP-B subunit provides a surface for membrane binding, and oligomerization of SP-B facilitates membrane apposition, adhesion and subsequent fusion of phospholipid bilayers. In laboratory conditions for example, mature SP-B has been shown to cause aggregation, lysis and fusion of negatively charged liposomes. The cationic region of amphipathic SP-B is thought to be closely associated with phospholipid head groups at the interface thereby providing hydrophobic and hydrophilic helical regions important to anchor SP-B strongly to lipid surfaces [166]. Foster resonance energy transfer (FRET) experiments show that SP-B prefers to associate with anionic lipids over zwitterionic phospholipids [229]. Likewise, FTIR studies done on SP-B/DPPG complex also demonstrate electrostatic interactions between charged SP-B residues and the PO_2^- head group [218, 230]. For an example, SP-B was shown to bind DPPC/PG large unilamellar vesicles (LUVs) containing at least 30% of PG, the minimum amount needed to sustain appropriate morphological and functional reconstitution of surfactant. The affinity between SP-B and the vesicles seemed to increase as the PG content of the LUVs was raised [212]. An earlier study reported that the association of SP-B with the anionic PG is concentration depenedent, as determined by fluorescence anisotropy technique [227]. This is consistent with the unusually high content of PG in the lung (≈3 to 5 mg/ml), the highest when compared with any other organ or tissue in mammals [231]. According to Chang *et al.* [232], there are an estimated ≈600± 300 SP-B binding sites per liposome and as few as 50 bound SP-B molecules were enough to perturb and cause significant leakage of vesicle contents. Ultimately, fusion of vesicles occurred as the protein-lipid ratio was raised.

While SP-B has been shown to promote fusion of bilayer stacks to monolayers, surfactant phospholipid adsorption into the superficial layer, lipid reorganization into specialised structures such as tubular myelin or multilayer structures [153], or to induce membrane curvature [231], it is not known how oligomerization might alter these functional lipid-SP-B interactions. Finding the answers to these questions has been greatly hampered by the inability to make recombinant SP-B and thus the inability to make mutations that affect oligomerization and determine how such changes affect lipid interactions and function.

1.9 Pathophysiology Associated with Lack of Lung Surfactant

Premature neonates can have a partial or total lack of lung surfactant. Lack of functional surfactant in the alveolar space results in lack of proper lubrication and poor lung compliance. Usually, if the condition is not treated adequately and quickly, it leads to atelectasis i.e. partial or full collapse of the alveoli and inability to inflate back to normal lung volume, which ultimately results in respiratory failure [233, 234]. Neonatal respiratory distress syndrome (NRDS), a condition observed in preterm infants, is a consequence of either prematurity or genetic defect [45, 235]. NRDS is the leading cause of perinatal death ≈10% in the US alone, claiming about 49.5% of prematurly delivered babies within the first 14 days of life [32]. The data show that the incidence of NRDS increases with decreased duration of gestation. Less than 5% of preterm neonates delivered after 34 weeks of gestation develop NRDS, however, this number gets to a staggering 60% for those delivered on or before 28 weeks of pregnancy. Studies have estimated that premature babies only have a storage pool of 4-5 mg/kg of body weight of lung surfactant at birth, as opposed to healthy term infants with a pool 100 mg/kg of body weight [233]. Moreover, preterm infants with NRDS cannot precisely metabolize LS, i.e., the half-life of lung surfactant in these infants is about 3 days. Apart from lung immaturity, genetic irregularity in the *SFTPB* (OMIM, 178640), *SFTPC* (OMIM, 178620) or *ABCA3* gene (OMIM, 601615) are also associated with NRDS and interstitial lung disease in both term and pre-term infants [45, 235]. While mutation in the *SFTPB* gene is associated with fatal NRDS, mutation in the SP-C gene, *SFTPC*, is more common and manifests in interstitial disease, lipoproteinosis, NRDS and pulmonary fibrosis in older infants, childrens and adults. On the other hand, genetic defects in the *ABCA3* gene can manifest in both fatal and non-fatal NRDS [45, 236].

Although up until 1973 NRDS was the leading cause of preterm infant mortality [7], the advent of improved therapies, antenatal glucocorticoids and exogenous surfactants, led to a decline in mortality of NRDS patient from nearly 100% to less than 10% in recent years. NRDS treatment can begin before birth by adminstering maternal steroid to women at risk of delivering prematurely, between 24 and 34 weeks of gestation. Transcription factors and glucosteroids enhance lung maturation and regulate surfactant synthesis [237, 238]. Treatment with transcription factors and hormone are believed to stimulate maturation of lung surfactant and may act in part by suppressing the turnover of phopholipids and reducing export of phospholipids from the lung [22]. Preventive treatment with exogenous surfactant is recommended for preterms at high risk of developing NRDS. Neverthless, determining which of the high risk preterms actually develop NRDS is still a challenge [233, 234]. Administering exogenous surfactant raises the pool in preterm infants until enough endogenous surfactant is produced.

Acute respiratory distress syndrome (ARDS) and acute lung injury (ALI) can be experienced by older infants, children and adults. ARDS/ALI result from many forms of insult to the lung and much like asthma, exhibit heterogeneity of syndromes before onset, which complicates treatment. ARDS is marked by a high mortality rate, ≈40-60% of the patients, and is less prevalant than NRDS. The highest occurences are reported in Australia and the United States [239]. Due to insult to the lung, leakage in the distal airways allows potential surfactant inhibitors from plasma to pass into the alveoli and inactivate lung surfactant [234]. Elevated levels of cholesterol, reactive oxygen species and serum lipoproteins are found in lungs of patients with ARDS/ALI [240]. There are data to suggest that lung surfactant from ARDS patients exhibits a shift to unsaturated fatty acyl chains of the phospholipids, which might increase susceptibility to oxidative related dysfunction [241–243]. Though it is known that the causes of ALI/ARDS are multifactorial [239], the underlying mechanism and progonosis leading to ALI/ARDS are poorly understood [244]. Pneumonia and sepsis are two primary risk factors. However, the majority of these patients do not necessarily develop ARDS, which suggests other factors, such as genetic predisposition, play a role in the pathogenesis of ARDS. In fact, more than 40 candidate genes have been identified that directly or indirectly correlate with developing or exacerbating the outcome for ARDS patients. One of these is the *SFTPB* gene [239, 245]. Alteration of macrophage function can lead to the accumulation of surfactant lipid and protein in the alveolar space and manifests as ARDS or pulmonary alveolar proteinosis [246]. Patients who survived ARDS/ALI are at high risk of developing cognitive decline, depression, post traumatic stress disorder and persistent skeletal muscle disorder [239]. Exogeneous surfactant treatment is less efficacious towards ARDS and ALI than NRDS [247]. As of yet, the only effective strategy against ARDS is protective lung ventilation [239, 245, 248, 249].

Altered cholesterol levels have been suggested to contribute to ARDS/ALI. Leonenko *et al.* [240] noticed elevated amount of cholesterol inhibits monolayer to bilayer conversion during dynamic cycling. The detailed process by which such inhibition proceeds is not well described. In essence, the mechanism by which surfactant is inactivated in diseased lungs needs an in-depth investigation as it opens new prospects in ARDS/ALI treatment.

1.10 Clinical Surfactants

Surfactant replacement therapy, the method used to treat NRDS, reduces death, air leak syndromes and intraventricular hemorrhage in preterm infants [11]. Differences in composition between clinical surfactants are a reason why surfactant therapies vary in efficacy [247]. At the moment, animal derived surfactants are the standard care for NRDS [215]. All but one of the animal derived surfactants are bovine in origin and can be prepared by either minced or lung lavage methods. Infasurf (Forest Labs, New York), is produced by lavaging calf lungs and extracting the LS with solvent. Beractant alfa (Survanta, Abbot Laboratories, Columbus, Ohio) is a bovine lipid extract to which colfosceril palmitate, palmitic acid and tripalmitin are added. Here in Canada, the standard therapy is Bovine Lipid Extract Surfactant (BLES, Bles

Biochemicals, London, Ontario, Canada) [11, 250], and recently BLES has also started to be marketed in India. Poractant alfa (Curosurf, Chiesi Pharmaceuticals, Italy) is a porcine lung extract that undergoes additional purification leaving a high content of polar lipid (80 mg/ml), SP-B and plasmalogen. Hence, Curosurf has the lowest volume of adminstration compared to the rest of surfactants available today [251]. The method of preparation and additives used affect performance as well as the adminstration dose. Although it is effective in treating NRDS, animal derived surfactant preparations are expensive to produce [215] and there are concerns over infections and immune rsponse with animal derived products [252]. These concerns therefore provide incentive for non-animal derived alternatives.

By the year 1987, the first generation protein-free synthethic surfactant preparations were devised. This came after a successful two stage randomized clinical trial of 328 very premature infants (delivered between 25 and 29 weeks of gestation) who were treated with Pumactant (also known as ALEC: artificial lung expanding compound, UK). Pumactant only contains DPPC and PG and was found to show promising result in terms of reducing mortality and the need for prolonged respiratory support with no serious side effects [253]. Nonetheless, it was also found that Pumactant was associated with more death discharge than neonates treated with Curosurf [254], henceforth, Pumactant had to be pulled out of the market. Likewise, another first generation surfactant devised was colfosceril palmitate (Exosurf, London, UK), the first FDA approved protein free synthetic surfactant preparation of its kind. Exosurf uses tyloxapol and hexadecanol as alternate additives to SP-B and SP-C, after these two compounds were shown to enhance adsorption and spreading kinetics of LS lipids *in vitro* [255, 256]. Similar to ALEC, however, a meta analysis by Soll and Blanco in 2001 [257] also provided evidence that treatment with Exosurf was correlated with increased mortality, pneumothorax and the need for longer duration for ventillation support compared to treatment with animal derived surfactants. Afterwards, Exosurf was also discontinued post marketing. These findings imply that SP-B and SP-C are superior to artificial additives used in Exosurf and also underlines that protein containing surfactant provides faster extubation and decreased mortality than protein free surfactant preparations [12, 257]. The relatively poor outcome of the first generation synthetic surfactants have led to the next generation surfactant development.

Compounds that try to mimic the structure and function of SP-B and SP-C were designed. Two formulations that contained such compounds entered clinical trials, Lusupultide (Venticute, Takeda Pharmaceuticals, Switzerland) and Lucinactant (Surfaxin, Discovery Laboratories, Pennsylvania). Venticute uses non-acylated recombinant SP-C (2% rSP-C), DPPC/POPG (7:3) and palmitic acid (5%). However, was pulled out of phase three trials after no improved survival rate of patients with ARDS was reported [258]. On the contrary, Surfaxin, a formula that comprises DPPC/POPG in 3:1 ratio, 13.5% palmitic acid and a 21 amino acid hydrophobic and synthetic peptide, sinapultide (KL4, 2.7%) [12], was approved in 2012 for preterm neonates suffering from NRDS [259]. Although, KL4 is deemed to be an SP-B mimic functionally, structure wise, on the other hand, the synthetic peptide more resembles SP-C, adopting a trans-membrane alpha helical orientation as SP-C does [215, 260]. This sparks question in terms of functionality, if KL4 truly simulates SP-B or SP-C or perhaps neither.

When considering exogeonous surfactant for ARDS/ALI treatment, it is also important to think about how resistant a surfactant is to inactivation. For instance, at identical inhibitor concentrations, Surfaxin is more resistant to inactivation by serum proteins and reactive oxygen species than Survanta despite both formulations contained similar phopholipid concentration. However, it was also found that natural surfactant was the least inhibited of all, which makes both Surfaxin and Survanta unattractive for ARDS/ALI treatment. Evidently, surfactant preparations containing only phospholipids were found to be highly inhibited [261]. In 2015, Surfaxin was withdrawn from the market. At the moment, none of the first nor the second generation synthetic surfactants are available to treat NRDS [259]. Hence, the search for better acting and efficacious synthetic surfactant still goes on. Right now, third generation synthetic surfactant containing peptide analogues of both SP-B and SP-C are being investigated. Formulations containing analogs of both SP-B and SP-C are found to perform better in terms of activity and resisting inactivation by serum in acute lung injury animal models when compared to those surfactants prepared from only one mimic of SP-B/SP-C as well as animal derived surfactants. To this end, for instance, it was found that the synthetic surfactant CHF5633, which contains both SP-B and SP-C analogues, demonstrates improved resistance to inactivation compared to poractant alfa in preterm lambs [262, 263].

There is still no general agreement on how endogenous surfactant is inactivated and how to circumvent it. There are different immunogenic/antigenic and infectious complications associated with administering animal derived surfactant prepartions [252]. To improve treatment outcome, exogenous surfactant with better pharmacological profile must be devised. Sufficient dosage, proper timing and targeted delivery as well as inhibition resistant formula will likely prove to be important in combating adult ALI/ARDS in the future [247, 249]. To this end, recombinant production of SP-B along with site directed mutagenesis will potentially play a pivotal role in producing surfactant with improved function, which also opens up the possibility to produce surfactant at a competitive price for patients to receive high doses and multiple doses. Moreover, non-invasive surfactant delivery has yet to be acheived. In this regard, succeseful aerosolization or atomization of surfactant could potentially play vital role. Future directions should explore how surfactant therapy could be used as a vehicle for delivering anti-inflammatory molecules and focus on finding better ways of administering surfactant in less-invasive or non-invasive ways [259]. It is also forecasted that, in the foreseeable future, surfactant could be used to deliver maternal steroids to women at risk of premature delivery. The key therefore lies in establishing the structure-function relationships for both SP-B and SP-C. Nonetheless, owing to the high hydrophobicity of SP-B, so far, it has been difficult to fully characterize it. Crystallizing SP-B has been notoriously difficult and all attempts in the past have failed. To elucidate SP-B's structure using solution NMR, the protein needs to be produced and purified in ample quantities and also retain its native-like functional state. Besides, the sample should also be labelled with NMR active nuclei. Moreover, due to broadening of the line signal, solution NMR is sensitive to protein size. Therefore, the protein complex should as well be small enough to tumble in a time scale that makes it possible to provide meaningful structural information. Although, an SP-B monomer (8.7 kDa) has the right size for solution NMR, no structure has yet been described for the full length SP-B monomer.

Understanding the mechanism of how lung surfactant operates at the atomic level has potential outcomes. Firstly, it may help to devise a better acting surfactant formulation with improved pharmaco-kinetic/dynamic profiles and resistance to inactivation/dysfunction. Secondly, it could alleviate or replace the need for animal derived surfactants that have immunologic and antigenic complication and are expensive to produce. Thirdly, it could help devise a method to use surfactant as a vehicle for targeted drug delivery components. In conjunction with site-directed mutagenesis, a high resolution and biologically active SP-B model could also prove to be key in many other aspects: 1) to identify important residues responsible for specific protein-lipid and protein-protein interactions, 2) to answer how the the tertiary/quaternary structure assembles in lipid bilayers (monolayers), 3) to answer if SP-B takes on an open or closed conformation or a combination of both and whether this configuration is pH dependent or not, as are some of the SAPLIPs or saposins discussed, and 4) to identify dimerization interface(s) and hence the propensity to oligomerize. Before all this however, there has to be a well established technique of producing SP-B. Recombinant DNA technology is a well suited method to introduce specific mutations to study the structure-activity relationships of macromolecules. What is more, the technique is also a viable option for producing full-length SP-B, which is rather difficult to do with protein chemical synthesis. The purity and quantity of the protein needed for structural studies could also be optimized by selecting the right expression host and purification protocols.

1.11 Hypothesis and Objectives

Herein, I hypothesise that SP-B can be expressed in bacteria in its functional form. Specifically, the study aims to: i) devise a method of SP-B expression, ii) characterize the structural aspect of rSP-B and iii) test the function of rSP-B in lipid/detergent environment.

Accomplishing the objectives of this thesis work could be significant to both health care and basic science. A way to produce SP-B in laboratory has a numbers of benefit. Foremost is the therapeutic importance of recombinant SP-B given the great expense of animal-derived SP-B. Second to that, being able to produce SP-B recombinantly, will allow researchers to introduce alterations to SP-B to improve its pharmacological properties, in particular to reduce its inactivation in ARDS conditions. Recombinant SP-B will also enable scientists to produce SP-B for structural and functional analysis that have not been possible with native SP-B, for example the analysis of point mutants or chimeric proteins. To my knowledge, there is no report of recombinant SP-B in the surfactant literature. Such knowledge, I think, will add to the fundamentals of surfactant science immensely.

1.12 References

1. Clements, J. A. Lung surfactant: A personal perspective. *Annu. Rev. Physiol* **59,** 1–21 (1997).
2. Parmigiani, S. & Solari, E. The era of pulmonary surfactant from Laplace to nowadays. *Acta bio-medica : Atenei Parmensis* **74,** 69–75 (2003).
3. Macklin, C. The pulmonary alveolar mucoid film and the pneumocytes. *The Lancet* **263,** 1099–1104 (1954).
4. Pattle, R. E. Properties, function and origin of the alveolar lining layer. *Nature* **175,** 1125–1126 (1955).
5. Clements, J. A. Surface tension of lung extracts. *Proceedings of the Society for Experimental Biology and Medicine* **95,** 170–2 (1957).
6. Avery, M. E. & Mead, J. Surface properties in relation to atelectasis and hyaline membrane disease. *A.M.A. Journal of Diseases of Children* **97,** 517–523 (1959).
7. Farrell, P. M. & Wood, R. E. Epidemiology of hyaline membrane disease in the United States: Analysis of national mortality statistics. *PEDIATRICS* **58,** 167–176 (1976).
8. Klaus, M. H., Clements, J. A. & Havel, R. J. Composition of the surface active material isolated from beef lung. *Proceedings of the National Academy of Sciences of the United States of America* **47,** 1858–1859 (1961).
9. Robillard, E., Alarie, Y., Dagenais-Perusse, P., Baril, E. & Guilbeault, A. Microaerosol administration of synthetic β-γ-dipalmitoyl-L-α-lecithin in the respiratory distress syndrome: A preliminary report. *Canadian Medical Association journal* **90,** 55–57 (1964).
10. Chu, J., Clements, J. A., Cotton, E. K., Klaus, M. H., Sweet, A. Y. & Tooley, W. H. Neonatal pulmonary Ischemia part I: clinical and physiological studies. *Pediatrics* **40,** 709–782 (1967).
11. Stevens, T. P. & Sinkin, R. A. Surfactant replacement therapy. *Chest* **131,** 1577–1582 (2007).
12. Halliday, H. L. Surfactants: past, present and future. *Journal of perinatology* **28,** S47–S56 (2008).
13. King, R. J., Klass, D. J., Gikas, E. G. & Clements, J. A. Isolation of apoproteins from canine surface active material. *American Journal of Physiology* **224,** 788–795 (1973).
14. King, R. J. & Clements, J. A. Surface active materials from dog lung. I. Method of isolation. *American Journal of Physiology* **223,** 707–714 (1972).
15. King, R. J. & Clements, J. A. Surface active materials from dog lung. II. Composition and physiological correlations. *American Journal of Physiology* **223,** 715–726 (1972).
16. Gil, J. & Reiss, O. K. Isolation and characterization of lamellar bodies and tubular myelin prom rat lung homogenates. *Journal of Cell Biology* **58,** 152–171 (1973).
17. Hallman, M., Feldman, B. H., Kirkpatrick, E. & Gluck, L. Absence of phosphatidylglycerol (PG) in respiratory distress syndrome in the newborn. Study of the minor surfactant phospholipids in newborns. *Pediatric research* **11,** 714–720 (1977).

18. Fujiwara, T., Chida, S., Watabe, Y., Maeta, H., Morita, T. & Abe, T. Artificial surfactant therapy in hyaline-membrane disease. *The Lancet* **315,** 55–59 (1980).

19. Robertson, B., Curstedt, T., Johansson, J., Jornvall, H. & Kobayashi, T. Structural and functional characterization of porcine surfactant isolated by liquid-gel chromatography. *Progress in Respiration Research* **25,** 237–246 (1990).

20. Rau, G. A., Vieten, G., Haitsma, J. J., Freihorst, J., Poets, C., Ure, B. M. & Bernhard, W. Surfactant in newborn compared with adolescent pigs: adaptation to neonatal respiration. *American Journal of Respiratory Cell and Molecular Biology* **30,** 694–701 (2004).

21. Bernhard, W., Hoffmann, S., Dombrowsky, H., Rau, G. A., Kamlage, A., Kappler, M., Haitsma, J. J., Freihorst, J., Von der Hardt, H. & Poets, C. F. Phosphatidylcholine molecular species in lung surfactant composition in relation to respiratory rate and lung development. *American Journal of Respiratory Cell and Molecular Biology* **25,** 725–731 (2001).

22. Bernhard, W. Lung surfactant: Function and composition in the context of development and respiratory physiology. *Annals of Anatomy* **208,** 146–150 (2016).

23. Parra, E. & Perez-Gil, J. Composition, structure and mechanical properties define performance of pulmonary surfactant membranes and films. *Chemistry and Physics of Lipids* **185,** 153–175 (2015).

24. Shelley, S. A., Balis, J. U., Paciga, J. E., Espinoza, C. G. & Richman, A. V. Biochemical composition of adult human lung surfactant. *Lung* **160,** 195–206 (1982).

25. Postle, A. D., Heeley, E. L. & Wilton, D. C. A comparison of the molecular species compositions of mammalian lung surfactant phospholipids. *Comparative Biochemistry and Physiology Part A* **129,** 65–73 (2001).

26. Goerke, J. Pulmonary surfactant: functions and molecular composition. *Biochimica et Biophysica Acta* **1408,** 79–89 (1998).

27. Rausch, F., Schicht, M., Bräuer, L., Paulsen, F. & Brandt, W. Protein modeling and molecular dynamics simulation of the two novel surfactant proteins SP-G and SP-H. *Journal of molecular modeling* **20:2053,** 1–12 (2014).

28. Schröder, H., Sollfrank, L., Paulsen, F., Bräuer, L. & Schicht, M. Recombinant expression of surfactant protein H (SFTA3) in Escherichia coli. *Annals of Anatomy* **208,** 129–134 (2016).

29. Mittal, R. A., Hammel, M., Schwarz, J., Heschl, K. M., Bretschneider, N., Flemmer, A. W., Herber-Jonat, S., Königshoff, M., Eickelberg, O. & Holzinger, A. SFTA2-a novel secretory peptide highly expressed in the lung-is modulated by lipopolysaccharide but not hyperoxia. *PLoS ONE* **7,** 1–9 (2012).

30. Schicht, M., Rausch, F., Finotto, S., Mathews, M., Mattil, A., Schubert, M., Koch, B., Traxdorf, M., Bohr, C., Worlitzsch, D., Brandt, W., Garreis, F., Sel, S., Paulsen, F. & Bräuer, L. SFTA3, a novel protein of the lung: three-dimensional structure, characterisation and immune activation. *European Respiratory Journal* **44,** 447–456 (2014).

31. Orgeig, S., Morrison, J. L. & Daniels, C. B. *Effect of environment and aging on the pulmonary surfactant system* in *The Lung: Development, Aging and the Environment* (eds Harding, R. & Pinkerton, K. E.) 2nd ed., 447–469 (Elsevier Inc., Boston, 2015).

32. Dautel, S. E., Kyle, J. E., Clair, G., Sontag, R. L., Weitz, K. K., Shukla, A. K., Nguyen, S. N., Kim, Y. M., Zink, E. M., Luders, T., Frevert, C. W., Gharib, S. A., Laskin, J., Carson, J. P., Metz, T. O., Corley, R. A. & Ansong, C. Lipidomics reveals dramatic lipid compositional changes in the maturing postnatal lung. *Scientific Reports* **7,** 1–12 (2017).

33. Mulugeta, S., Nureki, S.-I. & Beers, M. F. Lost after translation: Insights from pulmonary surfactant for understanding the role of alveolar epithelial dysfunction and cellular quality control in fibrotic lung disease. *American Journal of Physiology - Lung Cellular and Molecular Physiology* **309,** L507–L525 (2015).

34. Fehrenbach, H. Alveolar epithelial type II cell: defender of the alveolus revisited. *Respiratory research* **2,** 33–46 (2001).

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

36. Barkauskas, C. E., Cronce, M. J., Rackley, C. R., Bowie, E. J., Keene, D. R., Stripp, B. R., Randell, S. H., Noble, P. W. & Hogan, B. L. M. Type 2 alveolar cells are stem cells in adult lung. *The Journal of Clinical Investigation* **123,** 3025–3036 (2013).

37. Ashino, Y., Ying, X., Dobbs, L. G. & Bhattacharya, J. [Ca(2+)](i) oscillations regulate type II cell exocytosis in the pulmonary alveolus. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **279,** L5–L13 (2000).

38. Uhal, B. D. Cell cycle kinetics in the alveolar epithelium. *The American journal of physiology* **272,** L1031–L1045 (1997).

39. Mishra, A., Chintagari, N. R., Guo, Y., Weng, T., Su, L. & Liu, L. Purinergic P2X7 receptor regulates lung surfactant secretion in a paracrine manner. *Journal of Cell Science* **124,** 657–668 (2011).

40. Koval, M. Sharing signals: connecting lung epithelial cells with gap junction channels. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **283,** L875–L893 (2002).

41. Korimilli, A., Gonzales, L. W. & Guttentag, S. H. Intracellular localization of processing events in human surfactant protein B biosynthesis. *Journal of Biological Chemistry* **275,** 8672–8679 (2000).

42. Schmitz, G. & Müller, G. Structure and function of lamellar bodies, lipid-protein complexes involved in storage and secretion of cellular lipids. *Journal of lipid research* **32,** 1539–1570 (1991).

43. Osanai, K., Mason, R. J. & Voelker, D. R. Pulmonary surfactant phosphatidylcholine transport bypasses the brefeldin A sensitive compartment of alveolar type II cells. *Biochimica et Biophysica Acta* **1531,** 222–229 (2001).

44. Somaschini, M., Presi, S., Ferrari, M., Vergani, B. & Carrera, P. Surfactant proteins gene variants in premature newborn infants with severe respiratory distress syndrome. *Journal of Perinatology* **38,** 337–344 (2017).

45. Wert, S. E., Whitsett, J. A. & Nogee, L. M. Genetic disorders of surfactant dysfunction. *Pediatric and Developmental Pathology* **12,** 253–274 (2009).

46. Mulugeta, S., Gray, J. M., Notarfrancesco, K. L., Gonzales, L. W., Koval, M., Feinstein, S. I., Ballard, P. L., Fisher, A. B. & Shuman, H. Identification of LBM180, a lamellar

body limiting membrane protein of alveolar type II cells, as the ABC transporter protein ABCA3. *Journal of Biological Chemistry* **277,** 22147–22155 (2002).

47. Yamano, G., Funahashi, H., Kawanami, O., Zhao, L. X., Ban, N., Uchida, Y., Morohoshi, T., Ogawa, J., Shioda, S. & Inagaki, N. ABCA3 is a lamellar body membrane protein in human lung alveolar type II cells. *FEBS letters* **508,** 221–225 (2001).

48. Yang, M. Y., Hilton, M. B., Seaman, S., Haines, D. C., Nagashima, K., Burks, C. M., Tessarollo, L., Ivanova, P. T., Brown, H. A., Umstead, T. M., Floros, J., Chroneos, Z. C. & St.Croix, B. Essential regulation of lung surfactant homeostasis by the orphan G protein-coupled receptor GPR116. *Cell Reports* **3,** 1457–1464 (2013).

49. Bridges, J. P., Ludwig, M. G., Mueller, M., Kinzel, B., Sato, A., Xu, Y., Whitsett, J. A. & Ikegami, M. Orphan G protein-coupled receptor GPR116 regulates pulmonary surfactant pool size. *American Journal of Respiratory Cell and Molecular Biology* **49,** 348–457 (2013).

50. Serrano, A. G. & Pérez-Gil, J. Protein-lipid interactions and surface activity in the pulmonary surfactant system. *Chemistry and physics of lipids* **141,** 105–118 (2006).

51. Agassandian, M. & Mallampalli, R. K. Surfactant phospholipid metabolism. *Biochimica et Biophysica Acta* **1831,** 612–625 (2013).

52. Andreeva, A. V., Kutuzov, M. A. & Voyno-Yasenetskaya, T. A. Regulation of surfactant secretion in alveolar type II cells. *AJP: Lung Cellular and Molecular Physiology* **293,** L259–L271 (2007).

53. Haller, T., Dietl, P., Stockner, H., Frick, M., Mair, N., Tinhofer, I., Ritsch, A., Enhorning, G. & Putz, G. Tracing surfactant transformation from cellular release to insertion into an air-liquid interface. *American journal of physiology-Lung cellular and molecular physiology* **286,** L1009–L1015 (2004).

54. Perez-Gil, J. & Weaver, T. E. Pulmonary surfactant pathophysiology: current models and open questions. *Physiology* **25,** 132–141 (2010).

55. Young, S. L., Fram, E. K., Larson, E. & Wright, J. R. Recycling of surfactant lipid and apoprotein-A studied by electron microscopic autoradiography. *American Journal of Physiology* **265,** L19–L26 (1993).

56. Ikegami, M. Surfactant catabolism. *Respirology* **11,** S24–S27 (2006).

57. Pettenazzo, A., Jobe, A., Humme, J., Seidner, S. & Ikegami, M. Clearance of surfactant phosphatidylcholine via the upper airways in rabbits. *Journal of Applied Physiology* **65,** 2151–2155 (1988).

58. Wright, J. R. Clearance and recycling of pulmonary surfactant. en. *The American journal of physiology* **259,** L1–L12 (1990).

59. Olmeda, B., Martínez-Calle, M. & Pérez-Gil, J. Pulmonary surfactant metabolism in the alveolar airspace: Biogenesis, extracellular conversions, recycling. *Annals of Anatomy* **209,** 78–92 (2017).

60. Voorhout, W. F., Veenendaal, T., Haagsman, H. P., Weaver, T. E., Whitsett, J. A., van Golde, L. M. & Geuze, H. J. Intracellular processing of pulmonary surfactant protein B in an endosomal/lysosomal compartment. *American journal of physiology* **263,** L479–L486 (1992).

61. Bañares-Hidalgo, A., Pérez-Gil, J. & Estrada, P. Acidic pH triggers conformational changes at the NH2-terminal propeptide of the precursor of pulmonary surfactant protein B to form a coiled coil structure. *Biochimica et Biophysica Acta* **1838,** 1738–1751 (2014).

62. Lin, S., Na, C. L., Akinbi, H. T., Apsley, K. S., Whitsett, J. A. & Weaver, T. E. Surfactant protein B (SP-B) -/- mice are rescued by restoration of SP-B expression in alveolar Type II cells but not Clara cells. *Journal of Biological Chemistry* **274,** 19168–19174 (1999).

63. Guttentag, S. H., Beers, M. F., Bieler, B. M. & Ballard, P. L. Surfactant protein B processing in human fetal lung. *American Journal of Physiology - Lung Cellular and Molecular Physiology* **275,** L559–L566 (1998).

64. Brasch, F., Ochs, M., Kahne, T., Guttentag, S., Schauer-Vukasinovic, V., Derrick, M., Johnen, G., Kapp, N., Muller, K. M., Richter, J., Giller, T., Hawgood, S. & 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,** 49006–49014 (2003).

65. Foster, C., Aktar, A., Kopf, D., Zhang, P. & Guttentag, S. Pepsinogen C: a type 2 cell-specific protease. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **286,** L382–L387 (2004).

66. Guttentag, S., Robinson, L., Zhang, P., Brasch, F., Bühling, F. & Beers, M. Cysteine protease activity is required for surfactant protein B processing and lamellar body genesis. *American Journal of Respiratory Cell and Molecular Biology* **28,** 69–79 (2003).

67. Schauer-Vukasinovic, V., Bur, D., Kling, D., Grüninger, F. & Giller, T. Human napsin A: Expression, immunochemical detection, and tissue localization. *FEBS Letters* **462,** 135–139 (1999).

68. Gerson, K. D., Foster, C. D., Zhang, P., Zhang, Z., Rosenblatt, M. M. & Guttentag, S. H. Pepsinogen C proteolytic processing of surfactant protein B. *Journal of Biological Chemistry* **283,** 10330–10338 (2008).

69. Banfi, C. & Agostoni, P. Surfactant protein B: From biochemistry to its potential role as diagnostic and prognostic marker in heart failure. *International Journal of Cardiology* **221,** 456–462 (2016).

70. Ueno, T., Linder, S., Na, C. L., Rice, W. R., Johansson, J. & Weaver, T. E. Processing of pulmonary surfactant protein B by napsin and cathepsin H. *Journal of Biological Chemistry* **279,** 16178–16184 (2004).

71. Guttentag, S. Posttranslational regulation of surfactant protein B expression. *Seminars in Perinatology* **32,** 367–370 (2008).

72. Serrano, A. G., Cabré, E. J., Oviedo, J. M., Cruz, A., González, B., Palacios, A., Estrada, P. & Pérez-Gil, J. Production in Escherichia coli of a recombinant C-terminal truncated precursor of surfactant protein B (rproSP-BΔc). Structure and interaction with lipid interfaces. *Biochimica et Biophysica Acta* **1758,** 1621–1632 (2006).

73. Lin, S., Akinbi, H. T., Breslin, J. S. & Weaver, T. E. Structural requirements for targeting of surfactant protein B (SP-B) to secretory granules in vitro and in vivo. *Journal of Biological Chemistry* **271,** 19689–19695 (1996).

74. Lin, S., Phillips, K. S., Wilder, M. R. & Weaver, T. E. Structural requirements for intracellular transport of pulmonary surfactant protein B (SP-B). *Biochimica et Biophysica Acta* **1312,** 177–185 (1996).

75. Brasch, F., Johnen, G., Winn-Brasch, A., Guttentag, S. H., Schmiedl, A., Kapp, N., Suzuki, Y., Müller, K. M., Richter, J., Hawgood, S. & 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,** 449–458 (2004).

76. Stahlman, M. T., Gray, M. P., Falconieri, M. W., Whitsett, J. A. & Weaver, T. E. Lamellar body formation in normal and surfactant protein B-deficient fetal mice. *Laboratory Investigation* **80,** 395–403 (2000).

77. Yang, L., Johansson, J., Ridsdale, R., Willander, H., Fitzen, M., Akinbi, H. T. & Weaver, T. E. Surfactant protein B propeptide contains a saposin-like protein domain with antimicrobial activity at low pH. *The Journal of Immunology* **184,** 975–983 (2010).

78. Akinbi, H. T., Breslin, J. S., Ikegami, M., Iwamoto, H. S., Clark, J. C., Whitsett, J. A., Jobe, A. H. & Weaver, T. E. Rescue of SP-B knockout mice with a truncated SP-B proprotein : function of the C-terminal propeptide. *Biological Chemistry* **272,** 9640–9647 (1997).

79. Weaver, T. E. & Johnson Conkright, J. FUNCTIONS OF SURFACTANT PROTEINS B AND C. *Annu. Rev. Physiol* **63,** 555–78 (2001).

80. Vorbroker, D. K., Profitt, S. A., Nogee, L. M. & Whitsett, J. A. Aberrant processing of surfactant protein C in hereditary SP-B deficiency. *The American journal of physiology* **268,** L647–L656 (1995).

81. Beers, M. F., Hamvas, A., Moxley, M. A., Gonzales, L. W., Guttentag, S. H., Solarin, K. O., Longmore, W. J., Nogee, L. M. & Ballard, P. L. Pulmonary surfactant metabolism in infants lacking surfactant protein B. *American Journal of Respiratory Cell and Molecular Biology* **22,** 380–391 (2000).

82. Tokieda, K., Whitsett, J. A., Clark, J. C., Weaver, T. E., Ikeda, K., McConnell, K. B., Jobe, A. H., Ikegami, M. & Iwamoto, H. S. Pulmonary dysfunction in neonatal SP-B-deficient mice. *The American journal of physiology* **273,** L875–L882 (1997).

83. Coya, J. M., Akinbi, H. T., Sáenz, A., Yang, L., Weaver, T. E. & Casals, C. Natural anti-infective pulmonary proteins: In vivo cooperative action of surfactant protein SP-A and the lung antimicrobial peptide SP-B N. *The Journal of Immunology* **195,** 1628–1636 (2015).

84. Sarker, M., Jackman, D. & Booth, V. Lung surfactant protein A (SP-A) interactions with model lung surfactant lipids and an SP-B fragment. *Biochemistry* **50,** 4867–4876 (2011).

85. McCormack, F. X. Structure, processing and properties of surfactant protein A. *Biochimica et Biophysica Acta* **1408,** 109–131 (1998).

86. Foster, C. D., Zhang, P. X., Gonzales, L. W. & Guttentag, S. H. In vitro surfactant protein B deficiency inhibits lamellar body formation. *American Journal of Respiratory Cell and Molecular Biology* **29,** 259–266 (2003).

87. Clark, J. C., Wert, S. E., Bachurski, C. J., Stahlman, M. T., Stripp, B. R., Weaver, T. E. & Whitsett, J. A. Targeted disruption of the surfactant protein B gene disrupts surfactant

homeostasis, causing respiratory failure in newborn mice. *Proceedings of the National Academy of Sciences* **92,** 7794–7798 (1995).

88. Glasser, S. W., Burhans, M. S., Korfhagen, T. R., Na, C.-L., Sly, P. D., Ross, G. F., Ikegami, M. & Whitsett, J. A. Altered stability of pulmonary surfactant in SP-C-deficient mice. *Proceedings of the National Academy of Sciences* **98,** 6366–6371 (2001).

89. Glasser, S. W., Maxfield, M. D., Ruetschilling, T. L., Akinbi, H. T., Baatz, J. E., Kitzmiller, J. A., Page, K., Xu, Y., Bao, E. L. & Korfhagen, T. R. Persistence of LPS-induced lung inflammation in surfactant protein-C-deficient mice. *American Journal of Respiratory Cell and Molecular Biology* **49,** 845–854 (2013).

90. Martínez-Calle, M., Olmeda, B., Dietl, P., Frick, M. & Pérez-Gil, J. Pulmonary surfactant protein SP-B promotes exocytosis of lamellar bodies in alveolar type II cells. *The FASEB Journal* **32,** 1–12 (2018).

91. Batenburg, J. J. Surfactant phospholipids: synthesis and storage. *The American journal of physiology* **262,** L367–L385 (1992).

92. Haagsman, H. P. & van Golde, L. M. Synthesis and assembly of lung surfactant. *Annual review of physiology* **53,** 441–464 (1991).

93. Carman, G. M. & Han, G. S. Phosphatidic acid phosphatase, a key enzyme in the regulation of lipid synthesis. *Journal of Biological Chemistry* **284,** 2593–2597 (2009).

94. Hermansson, M., Hokynar, K. & Somerharju, P. Mechanisms of glycerophospholipid homeostasis in mammalian cells. *Progress in Lipid Research* **50,** 240–257 (2011).

95. Kennedy, E. P. Biosynthesis of complex lipids. *Federation Proceedings* **20,** 934–40 (1961).

96. Lands, W. E. M. Metabolism of glycerophospholipides: A comparison of lecithin and triglyceride sythesis. *Journal of Biological Chemistry* **231,** 883–888 (1958).

97. Lin, S., Ikegami, M., Moon, C., Naren, A. P. & Shannon, J. M. Lysophosphatidylcholine acyltransferase 1 (LPCAT1) specifically interacts with phospholipid transfer protein StarD10 to facilitate surfactant phospholipid trafficking in alveolar type II Cells. *Journal of Biological Chemistry* **290,** 18559–18574 (2015).

98. Nakanishi, H., Shindou, H., Hishikawa, D., Harayama, T., Ogasawara, R., Suwabe, A., Taguchi, R. & Shimizu, T. Cloning and characterization of mouse lung-type acyl-CoA: lysophosphatidylcholine acyltransferase 1 (LPCAT1): Expression in alveolar type II cells and possible involvement in surfactant production. *Journal of Biological Chemistry* **281,** 20140–20147 (2006).

99. Den Breejen, J. N., Batenburg, J. J. & van Golde, L. M. The species of acyl-CoA in subcellular fractions of type II cells isolated from adult rat lung and their incorporation into phosphatidic acid. *Biochimica et Biophysica Acta* **1002,** 277–282 (1989).

100. Mason, R. J. & Nellenbogen, J. Synthesis of saturated phosphatidylcholine and phosphatidylglycerol by freshly isolated rat alveolar type II cells. *Biochimica et Biophysica Acta* **794,** 392–402 (1984).

101. Butler, P. L. & Mallampalli, R. K. Cross-talk between remodeling and de novo pathways maintains phospholipid balance through ubiquitination. *Journal of Biological Chemistry* **285,** 6246–6258 (2010).

102. Funkhouser, J. D., Batenburg, J. J. & Van Golde, L. M. Acylation of 1-palmitoyl lysophosphatidylglycerol in alveolar type II cells from rat lung. *Biochimica et Biophysica Acta* **666,** 1–6 (1981).

103. Longmuir, K. & Haynes, S. Evidence that fatty acid chain length is a type II cell lipid-sorting signal. *American Journal of Physiology* **260,** L44–L51 (1991).

104. Osanai, K., Tsuchihara, C., Hatta, R., Oikawa, T., Tsuchihara, K., Iguchi, M., Seki, T., Takahashi, M., Huang, J. & Toga, H. Pulmonary surfactant transport in alveolar type II cells. *Respirology* **11,** S70–S73 (2006).

105. Goss, V., Hunt, A. N. & Postle, A. D. Regulation of lung surfactant phospholipid synthesis and metabolism. *Biochimica et Biophysica Acta* **1831,** 448–458 (2013).

106. Gobran, L. I. & Rooney, S. A. Regulation of SP-B and SP-C secretion in rat type II cells in primary culture. *American Journal of Physiology-Lung Cell and Molecular Physiology* **281,** L1413–L1419 (2001).

107. Rooney, S. A. Regulation of surfactant secretion. *Comparative Biochemistry and Physiology Part A* **129,** 233–243 (2001).

108. Mason, R. J. & Voelker, D. R. Regulatory mechanisms of surfactant secretion. *Biochimica et Biophysica Acta* **1408,** 226–240 (1998).

109. Griese, M., Gobran, L. I. & Rooney, S. A. Signal-transduction mechanisms of ATP-stimulated phosphatidylcholine secretion in rat type II pneumocytes: Interactions between ATP and other surfactant secretagogues. *Biochimica et Biophysica Acta* **1167,** 85–93 (1993).

110. Sanos, K., Voelker, D. R. & Mason, R. J. Involvement of protein kinase C in pulmonary surfactant secretion from alveolar type II Cells. *The journal of biological chemistry* **260,** 12725–12729 (1985).

111. Orgeig, S., Morrison, J. L., Sullivan, L. C. & Daniels, C. B. *Development of the pulmonary surfactant system* in *The Lung: Development, Aging and The Environment* (eds Harding, R. & Pinkerton, K. E.) Second Edi, 183–209 (Elsevier Inc., 2014).

112. Rooney, S. A., Gobran, L. I. & Wai Lee, T. S. Stimulation of surfactant production by oxytocin induced labor in the rabbit. *Journal of Clinical Investigation* **60,** 754–759 (1977).

113. Marino, P. A. & Rooney, S. A. The effect of labor on surfactant secretion in newborn rabbit lung slices. *Biochimica et Biophysica Acta* **664,** 389–396 (1981).

114. Nicholas, T. E., Power, J. H. & Barr, H. A. Surfactant homeostasis in the rat lung during swimming exercise. *Journal of Applied Physiology* **53,** 1521–1528 (1982).

115. Nicholas, T. E., Power, J. H. T. & Barr, H. A. The pulmonary consequences of a deep breath. *Respiration Physiology* **49,** 315–324 (1982).

116. Arold, S. P., Suki, B., Alencar, A. M., Lutchen, K. R. & Ingenito, E. P. Variable ventilation induces endogenous surfactant release in normal guinea pigs. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **285,** L370–L375 (2003).

117. Wirtz, H. & Dobbs, L. Calcium mobilization and exocytosis after one mechanical stretch of lung epithelial cells. *Science* **250,** 1266–1269 (1990).

118. Frick, M., Bertocchi, C., Jennings, P., Haller, T., Norbert Mair, Singer, W., Pfaller, W., Ritsch-Marte, M. & Dietl, P. Ca2+ entry is essential for cell strain-induced lamellar body fusion in isolated rat type II pneumocytes. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **286,** L210–L220 (2004).

119. Copland, I. B. & Post, M. Stretch-activated signaling pathways responsible for early response gene expression in fetal lung epithelial cells. *Journal of Cellular Physiology* **210,** 133–143 (2007).

120. Dobbs, L. G., Wright, J. R., Hawgood, S., Gonzalez, R., Venstrom, K. & Nellenbogen, J. Pulmonary surfactant and its components inhibit secretion of phosphatidylcholine from cultured rat alveolar type II cells. *Proceedings of the National Academy of Sciences of the United States of America* **84,** 1010–1014 (1987).

121. Rice, W. R., Ross, G. F., Singleton, F. M., Dingle, S. & Whitsett, J. A. Surfactant-associated protein inhibits phospholipid secretion from type II cells. *J .Appl. Physiol* **63,** 692–698 (1987).

122. Kuroki, Y., Mason, R. J. & Voelker, D. R. Chemical modification of surfactant protein A alters high affinity binding to rat alveolar type II cells and regulation of phospholipid secretion. *Journal of Biological Chemistry* **263,** 17596–17602 (1988).

123. Strayer, D. S., Pinder, R. & Chander, A. Receptor-mediated regulation of pulmonary surfactant secretion. *Experimental Cell Research* **226,** 90–97 (1996).

124. Krishnaiah, S. Y., Dodia, C., Sorokina, E. M., Li, H., Feinstein, S. I. & Fisher, A. B. Binding sites for interaction of peroxiredoxin 6 with surfactant protein A. *Biochimica et Biophysica Acta* **1864,** 419–425 (2016).

125. Rooney, S. A., Gobran, L. I., Umstead, T. M. & Phelps, D. S. Secretion of surfactant protein A from rat type II pneumocytes. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **265,** L586–L590 (1993).

126. Voorhout, W. F., Veenendaal, T., Kuroki, Y., Ogasawara, Y., Van Golde, L. M. & Geuze, H. J. Immunocytochemical localization of surfactant protein D (SP-D) in type II cells, Clara cells, and alveolar macrophages of rat lung. *Journal of Histochemistry and Cytochemistry* **40,** 1589–1597 (1992).

127. Crouch, E., Persson, A., Chang, D. & Parghi, D. Surfactant protein D. Increased accumulation in silica-induced pulmonary lipoproteinosis. *The American Journal of Pathology* **139,** 765–776 (1991).

128. Wissel, H., Lehfeldt, A., Klein, P., Müller, T. & Stevens, P. A. Endocytosed SP-A and surfactant lipids are sorted to different organelles in rat type II pneumocytes. en. *American Journal of Physiology: Lung Cellular and Molecular Physiology* **281,** L345–L360 (2001).

129. Osanai, K., Mason, R. J. & Voelker, D. R. Trafficking of newly synthesized surfactant protein A in isolated rat alveolar type II cells. *American Journal of Respiratory Cell and Molecular Biology* **19,** 929–935 (1998).

130. Fisher, A. B., Dodia, C., Ruckert, P., Tao, J.-Q. & Bates, S. R. Pathway to lamellar bodies for surfactant protein A. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **299,** L51–L58 (2010).

131. Wright, J. R. & Clements, J. A. Metabolism and turnover of lung surfactant. *American Review of Respiratory Disease* **136,** 426–444 (1987).

132. Bates, S. R. P63 (CKAP4) as an SP-A receptor: Implications for surfactant turnover. *Cellular Physiology and Biochemistry* **25,** 41–54 (2010).

133. Stevens, P. A., Wissel, H., Zastrow, S., Sieger, D. & Zimmer, K. P. Surfactant protein A and lipid are internalized via the coated-pit pathway by type II pneumocytes. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **280,** L141–L151 (2001).

134. Gupta, N., Manevich, Y., Kazi, A. S., Tao, J.-Q., Fisher, A. B. & Bates, S. R. Identification and characterization of p63 (CKAP4/ERGIC-63/CLIMP-63), a surfactant protein A binding protein, on type II pneumocytes. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **291,** L436–L446 (2006).

135. Evelyn, D., Ikegami, M. & Alan, H. Localization of alveolar surfactant clearance in rabbit lung cells. *American Journal of Physiology* **263,** 201–209 (1992).

136. Zhou, J., You, Y., Ryan, A. J. & Mallampalli, R. K. Upregulation of surfactant synthesis triggers ABCA1-mediated basolateral phospholipid efflux. *Journal of Lipid Research* **45,** 1758–1767 (2004).

137. Fukuzawa, T., Ishida, J., Kato, A., Ichinose, T., Ariestanti, D. M., Takahashi, T., Ito, K., Abe, J., Suzuki, T., Wakana, S., Fukamizu, A., Nakamura, N. & Hirose, S. Lung surfactant levels are regulated by Ig-Hepta/GPR116 by monitoring surfactant protein D. *PLoS ONE* **8,** 1–12 (2013).

138. Ariestanti, D. M., Ando, H., Hirose, S. & Nakamura, N. Targeted disruption of Ig-Hepta/Gpr116 causes emphysema-like symptoms that are associated with alveolar macrophage activation. *Journal of Biological Chemistry* **290,** 11032–11040 (2015).

139. Ikegami, M., Na, C.-L., Korfhagen, T. R. & Whitsett, J. A. Surfactant protein D influences surfactant ultrastructure and uptake by alveolar type II cells. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **288,** L552–L561 (2005).

140. Gan, N. N., Gan, M. A., Athirah, N., Taib, M., Krishnan, G. & Dasuki, K. A. A Review on Surface Tension Measurements by Optical Method for Medical Application. *Malysian Journal of Science, Health&Technology* **4,** 2601–2604 (2019).

141. Schurch, S., Qanbar, R., Bachofen, H., Possmayer, F., Schürch, S., Qanbar, R., Bachofen, H. & Possmayer, F. The surface-associated surfactant reservoir in the alveolar lining. *Biol Neonate* **67,** 61–76 (1995).

142. Hermans, E., Saad Bhamla, M., Kao, P., Fuller, G. G. & Vermant, J. Lung surfactants and different contributions to thin film stability. *Soft Matter* **11,** 8048–8057 (2015).

143. Ding, J., Takamoto, D. Y., Von Nahmen, A., Lipp, M. M., Lee, K. Y. C., Waring, A. J. & Zasadzinski, J. A. Effects of lung surfactant proteins, SP-B and SP-C, and palmitic acid on monolayer stability. *Biophysical Journal* **80,** 2262–2272 (2001).

144. Chakraborty, A., Hui, E., Waring, A. J. & Dhar, P. Combined effect of synthetic protein, Mini-B, and cholesterol on a model lung surfactant mixture at the air-water interface. *Biochimica et Biophysica Acta* **1858,** 904–912 (2016).

145. Walters, R. W., Jenq, R. R. & Hall, S. B. Distinct steps in the adsorption of pulmonary surfactant to an air-liquid interface. *Biophysical journal* **78,** 257–266 (2000).

146. Zhang, H., Fan, Q., Wang, Y. E., Neal, C. R. & Zuo, Y. Y. Comparative study of clinical pulmonary surfactants using atomic force microscopy. *Biochimica et Biophysica Acta* **1808,** 1832–1842 (2011).

147. Zuo, Y. Y., Veldhuizen, R. A. W., Neumann, A. W., Petersen, N. O. & Possmayer, F. Current perspectives in pulmonary surfactant - Inhibition, enhancement and evaluation. *Biochimica et Biophysica Acta* **1778,** 1947–1977 (2008).

148. Johansson, J. & Curstedt, T. Synthetic surfactants with SP-B and SP-C analogues to enable worldwide treatment of neonatal respiratory distress syndrome and other lung diseases. *Journal of Internal Medicine* **285,** 165–186 (2018).

149. Poulain, F. R. & Clements, J. A. Pulmonary surfactant therapy. *The Western Journal of Medicine* **162,** 43–50 (1995).

150. Zuo, Y. Y. & Neumann, A. W. Pulmonary surfactant and its in vitro assessment using axis symmetric drop shape analysis. *Tenside Surf. Det.* **42,** 1–21 (2005).

151. Possmayer, F., Hall, S. B., Haller, T., Petersen, N. O., Zuo, Y. Y., Bernardino de la Serna, J., Postle, A. D., Veldhuizen, R. A. W. & Orgeig, S. Recent advances in alveolar biology: Some new looks at the alveolar interface. *Respiratory Physiology and Neurobiology* **173,** S55–S64 (2010).

152. Zuo, Y. Y. & Possmayer, F. How does pulmonary surfactant reduce surface tension to very low values ? *Journal of Applied Physiology* **102,** 1733–1734 (2007).

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

154. Heimburg, T. & Jackson, A. D. On soliton propagation in biomembranes and nerves. *Proceedings of the National Academy of Sciences* **102,** 9790–9795 (2005).

155. De la Serna, J. B., Perez-Gil, J., Simonsen, A. C. & Bagatolli, L. A. Cholesterol rules: Direct observation of the coexistence of two fluid phases in native pulmonary surfactant membranes at physiological temperatures. *Journal of Biological Chemistry* **279,** 40715–40722 (2004).

156. Wang, Z., Hall, S. B. & Notter, R. H. Roles of different hydrophobic constituents in the adsorption of pulmonary surfactant. *Journal of Lipid Research* **37,** 790–798 (1996).

157. Lipp, M. M., Lee, K. Y. C., Takamoto, D. Y., Zasadzinski, J. A. & Waring, A. J. Coexistence of buckled and flat monolayers. *Physical Review Letters* **81,** 1650–1653 (1998).

158. Keating, E., Zuo, Y. Y., Tadayyon, S. M., Petersen, N. O., Possmayer, F. & Veldhuizen, R. A. W. A modified squeeze-out mechanism for generating high surface pressures with pulmonary surfactant. *Biochimica et Biophysica Acta* **1818,** 1225–1234 (2012).

159. Schürch, S., Green, F. H. & Bachofen, H. Formation and structure of surface films: Captive bubble surfactometry. *Biochimica et Biophysica Acta* **1408,** 180–202 (1998).

160. Alig, T. F., Warriner, H. E., Lee, L. & Zasadzinski, J. A. Electrostatic barrier to recovery of dipalmitoylphosphatidylglycerol monolayers after collapse. *Biophys Journal* **86,** 897–904 (2004).

161. Holten-Andersen, N., Michael Henderson, J., Walther, F. J., Waring, A. J., Ruchala, P., Notter, R. H. & Lee, K. Y. KL4 peptide induces reversible collapse structures on multiple length scales in model lung surfactant. *Biophys Journal* **101,** 2957–2965 (2011).

162. Dhar, P., Eck, E., Israelachvili, J. N., Lee, D. W., Min, Y., Ramachandran, A., Waring, A. J. & Zasadzinski, J. A. Lipid-protein interactions alter line tensions and domain size distributions in lung surfactant monolayers. *Biophys Journal* **102,** 56–65 (2012).

163. Lipp, M. M., Lee, K. Y. C., Zasadzinski, J. A. & Waring, A. J. Phase and morphology changes in lipid monolayers induced by SP-B protein and its amino-terminal peptide. *Science* **273,** 1196–1199 (1996).

164. Takamoto, D. Y., Lipp, M. M., von Nahmen, A., Lee, K. Y., Waring, A. J. & Zasadzinski, J. A. Interaction of lung surfactant proteins with anionic phospholipids. *Biophysical journal* **81,** 153–69 (2001).

165. Taneva, S. & Keough, K. M. Pulmonary surfactant proteins SP-B and SP-C in spread monolayers at the air-water interface: II. Monolayers of pulmonary surfactant protein SP-C and phospholipids. *Biophysical Journal* **66,** 1149–1157 (1994).

166. Whitsett, J. A. & Weaver, T. E. Hydrophobic surfactant proteins in lung function and disease. *The New England Journal of Medicine* **347,** 2141–2148 (2002).

167. Lee, K. Y. C., Lipp, M. M., Zasadzinski, J. A. & Waring, A. J. Effects of lung surfactant specific protein SP-B and model SP-B peptide on lipid monolayers at the air-water interface. *Colloids Surfaces A: Physicochem. Eng. Aspects* **128,** 225–242 (1997).

168. Oosterlaken-Dijksterhuis, M. A., van Eijk, M., van Golde, L. M. & Haagsman, H. P. Lipid mixing is mediated by the hydrophobic surfactant protein SP-B but not by SP-C. *Biochimica et Biophysica Acta* **1110,** 45–50 (1992).

169. Kobayashi, T., Curstedt, T., Grossmann, G. & Robertson, B. Inhibition of exogenous surfactant in ventilated immature newborn rabbits. *Respiration Physiology* **76,** 1–12 (1989).

170. Yamada, T., Ikegami, M. & Jobe, A. H. Effects of surfactant subfractions on preterm rabbit lung function. *Pediatric Research* **27,** 592–598 (1990).

171. Rider, E. D., Ikegami, M., Whitsett, J. A., Hull, W., Absolom, D. & Jobe, A. H. Treatment responses to surfactants containing natural surfactant proteins in preterm rabbits. *American Review of Respiratory Disease* **147,** 669–676 (1993).

172. Robertson, B., Berry, D., Curstedt, T., Grossmann, G., Ikegami, M., Jacobs, H., Jobe, A., Jones, S. & Robertson, B. Leakage of protein in the immature rabbit lung; effect of surfactant replacement. *Respiration Physiology* **61,** 265–276 (1985).

173. Ueda, T., Ikegami, M., Rider, E. D. & Jobe, A. H. Distribution of surfactant and ventilation in surfactant-treated preterm lambs. *J. Appl. Physiol.* **76,** 45–55 (1994).

174. Niemarkt, H. J., Kuypers, E., Jellema, R., Ophelders, D., Hütten, M., Nikiforou, M., Kribs, A. & Kramer, B. W. Effects of less-invasive surfactant administration on oxygenation, pulmonary surfactant distribution, and lung compliance in spontaneously breathing preterm lambs. *Pediatric Research* **76,** 166–170 (2014).

175. Hillman, N. H., Kemp, M. W., Miura, Y., Kallapur, S. G. & Jobe, A. H. Sustained inflation at birth did not alter lung injury from mechanical ventilation in surfactant-treated fetal lambs. *PLOS ONE* **9,** 1–10 (2014).

176. Bailey, T. C., Da Silva, K. A., Lewis, J. F., Rodriguez-Capote, K., Possmayer, F. & Veldhuizen, R. A. W. Physiological and inflammatory response to instillation of an oxidized surfactant in a rat model of surfactant deficiency. *Journal of Applied Physiology* **96,** 1674–1680 (2004).

177. Markart, P., Ruppert, C., Wygrecka, M., Colaris, T., Dahal, B., Walmrath, D., Harbach, H., Wilhelm, J., Seeger, W., Schmidt, R. & Guenther, A. Patients with ARDS show

improvement but not normalisation of alveolar surface activity with surfactant treatment: putative role of neutral lipids. *Thorax* **62,** 588–594 (2007).

178. Brackenbury, A. M., Malloy, J. L., McCaig, L. A., Yao, L. J., Veldhuizen, R. A. W. & Lewis, J. F. Evaluation of alveolar surfactant aggregates in vitro and in vivo. *European Respiratory Journal* **19,** 41–46 (2002).

179. V. Neergaard, K. Neue auffassungen über einen grundbegriff der atemmechanik. *Zeitschrift für die gesamte experimentelle Medizin* **66,** 373–394 (1929).

180. Schürch, S., Goerke, J. & Clements, J. A. Direct determination of surface tension in the lung. *Proceedings of the National Academy of Sciences of the United States of America* **73,** 4698–4702 (1976).

181. Power, J. H. T., Doyle, I. R., Davidson, K. & Nicholas, T. E. Ultrastructural and protein analysis of surfactant in the australian lungfish neoceratodus forsteri: evidence for conservation of composition for 300 million years. *The Journal of Experimental Biology* **202,** 2543–2550 (1999).

182. Haagsman, H. P. & Diemel, R. V. Surfactant-associated proteins: Functions and structural variation. *Comparative Biochemistry and Physiology Part A* **129,** 91–108 (2001).

183. Olmeda, B., Garca-Alvarez, B. & Perez-Gil, J. Structure-function correlations of pulmonary surfactant protein SP-B and the saposin-like family of proteins. *European Biophysics Journal* **42,** 209–222 (2013).

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

185. Munford, R. S., Sheppard, P. O. & O'Hara, P. J. Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure. *Journal of Lipid Research* **36,** 1653–1663 (1995).

186. Vaccaro, A. M., Salvioli, R., Tatti, M. & Ciaffoni, F. Saposins and their interaction with lipids. *Neurochemical Research* **24,** 307–314 (1999).

187. Andersson, M., Curstedt, T., Jörnvall, H. & Johansson, J. An amphipathic helical motif common to tumourolytic polypeptide NK-lysin and pulmonary surfactant polypeptide SP-B. *FEBS Letters* **362,** 328–332 (1995).

188. Hawkins, C. A., De Alba, E. & Tjandra, N. Solution structure of human saposin C in a detergent environment. *Journal of Molecular Biology* **346,** 1381–1392 (2005).

189. Johansson, J. & Curstedt, T. Molecular structures and interactions of pulmonary surfactant components. *Eur. J. Biochem* **244,** 675–693 (1997).

190. Leippe, M., Ebel, S., Schoenberger, O. L., Horstmann, R. D. & Müller-Eberhard, H. J. Pore-forming peptide of pathogenic Entamoeba histolytica. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 7659–7663 (1991).

191. Leippe, M., Tannich, E., Nickel, R., van der Goot, G., Pattus, F., Horstmann, R. D. & Müller-Eberhard, H. J. Primary and secondary structure of the pore-forming peptide of pathogenic Entamoeba histolytica. *The EMBO Journal* **11,** 3501–3506 (1992).

192. Liepinsh, E., Andersson, M., Ruysschaert, J. M. & Otting, G. Saposin fold revealed by the NMR structure of NK-lysin. *Nature Structural Biology* **4,** 793–795 (1997).

193. Anderson, D. H., Sawaya, M. R., Cascio, D., Ernst, W., Modlin, R., Krensky, A. & Eisenberg, D. Granulysin crystal structure and a structure-derived lytic mechanism. *J. Mol. Biol.* **325,** 355–365 (2003).

194. Ponting, C. P. Acid sphingomyelinase possesses a domain homologous to its activator proteins: Saposins B and D. *Protein Science* **3,** 359–361 (1994).

195. Munford, R. S. & Hall, C. L. Detoxification of bacterial lipopolysaccharides (endotoxins) by a human neutrophil enzyme. *Science* **234,** 203–205 (1986).

196. Morales, C. R., El-Alfy, M., Zhao, Q. & Igdoura, S. Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in Sertoli cells. *Histology and Histopathology* **10,** 1023–1034 (1995).

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

198. Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A. & Zdanov, A. Crystal structure of plant aspartic proteinase prophytepsin: Inactivation and vacuolar targeting. *EMBO Journal* **18,** 3947–3955 (1999).

199. Vaccaro, A. M., Ciaffoni, F., Tatti, M., Salvioli, R., Barca, A., Tognozzi, D. & Scerch, C. pH-dependent conformational properties of saposins and their interactions with phospholipid membranes. *Journal of Biological Chemistry* **270,** 30576–30580 (1995).

200. Hecht, O., Van Nuland, N. A., Schleinkofer, K., Dingley, A. J., Bruhn, H., Leippe, M. & Grötzinger, J. Solution structure of the pore-forming protein of Entamoeba histolytica. *Journal of Biological Chemistry* **279,** 17834–17841 (2004).

201. Ahn, V. E., Leyko, P., Alattia, J.-R., Chen, L. & Privé, G. G. Crystal structures of saposins A and C. *Protein Science* **15,** 1849–1857 (2006).

202. Morrow, M. R., Pérez-Gil, J., Simatos, G., Boland, C., Stewart, J. & Keough, K. M. Pulmonary surfactant-associated protein SP-B has little effect on acyl chains in dipalmitoylphosphatidylcholine dispersions. *Biochemistry* **32,** 4397–4402 (1993).

203. De Alba, E., Weiler, S. & Tjandra, N. Solution structure of human saposin C: pH-dependent interaction with phospholipid vesicles. *Biochemistry* **42,** 14729–14740 (2003).

204. Popovic, K., Holyoake, J., Pomes, R. & Prive, G. G. Structure of saposin A lipoprotein discs. *Proceedings of the National Academy of Sciences* **109,** 2908–2912 (2012).

205. Leippe, M., Bruhn, H., Hecht, O. & Grötzinger, J. Ancient weapons: The three-dimensional structure of amoebapore A. *Trends in Parasitology* **21,** 5–7 (2005).

206. Hill, C. H., Read, R. J. & Deane, J. E. Structure of human saposin A at lysosomal pH. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **71,** 895–900 (2015).

207. Rossmann, M., Schultz-Heienbrok, R., Behlke, J., Remmel, N., Alings, C., Sandhoff, K., Saenger, W. & Maier, T. Crystal structures of human saposins C and D: Implications for lipid recognition and membrane interactions. *Structure* **16,** 809–817 (2008).

208. Booth, V., Waring, A. J., Walther, F. J. & Keough, K. M. NMR structures of the C-terminal segment of surfactant protein B in detergent micelles and hexafluoro-2-propanol. *Biochemistry* **43,** 15187–15194 (2004).

209. Zaltash, S., Palmblad, M., Curstedt, T., Johansson, J. & Persson, B. Pulmonary surfactant protein B: a structural model and a functional analogue. *Biochimica et Biophysica Acta* **1466,** 179–186 (2000).

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

211. Johansson, J., Jörnvall, H. & Curstedt, T. Surfactant protein B: Disulfide bridges, structural properties, and kringle similarities. *Biochemistry* **30,** 6917–6921 (1991).

212. Hawgood, S., Derrick, M. & Poulain, F. Structure and properties of surfactant protein B. *Biochimica et Biophysica Acta* **1408,** 150–160 (1998).

213. Sarker, M., Waring, A. J., Walther, F. J., Keough, K. M. W. & Booth, V. Structure of mini-B, a functional fragment of surfactant protein B, in detergent micelles. *Biochemistry* **46,** 11047–11056 (2007).

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

215. Curstedt, T. & Johansson, J. New synthetic surfactant - how and when? *Biology of the Neonate* **89,** 336–339 (2006).

216. Cruz, A., Casals, C. & Perez-Gil, J. Conformational flexibility of pulmonary surfactant proteins SP-B and SP-C, studied in aqueous organic solvents. *Biochimica et Biophysica Acta* **1255,** 68–76 (1995).

217. Perez Gil, J., Cruz, A. & Casals, C. Solubility of hydrophobic surfactant proteins in organic solvent/water mixtures. Structural studies on SP-B and SP-C in aqueous organic solvents and lipids. *Biochimica et Biophysica Acta* **1168,** 261–270 (1993).

218. Vandenbussche, G., Clercx, A., Clercx, M., Curstedt, T., Johansson, J., Jornvall, H. & 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,** 9169–9176 (1992).

219. Wustneck, N., Wustneck, R., Perez-Gil, J. & Pison, U. Effects of oligomerization and secondary structure on the surface behavior of pulmonary surfactant proteins SP-B and SP-C. *Biophysical Journal* **84,** 1940–1949 (2003).

220. Kronqvist, N., Sarr, M., Lindqvist, A., Nordling, K., Otikovs, M., Venturi, L., Pioselli, B., Purhonen, P., Landreh, M., Biverstål, H., Toleikis, Z., Sjöberg, L., Robinson, C. V., Pelizzi, N., Jörnvall, H., Hebert, H., Jaudzems, K., Curstedt, T., Rising, A. & Johansson, J. Efficient protein production inspired by how spiders make silk. *Nature Communications* **8,** 1–15 (2017).

221. Curstedt, T., Johansson, J., Persson, P., Eklund, A., Robertson, B., Lowenadler, B. & Jornvall, H. Hydrophobic surfactant-associated polypeptides: SP-C is a lipopeptide with two palmitoylated cysteine residues, whereas SP-B lacks covalently linked fatty acyl groups. *Proceedings of the National Academy of Sciences* **87,** 2985–2989 (1990).

222. Zaltash, S., Griffiths, W. J., Beck, D., Duan, C.-X., Weaver, T. E. & Johansson, J. Membrane activity of (Cys48Ser) lung surfactant protein B increases with dimerisation. *Biological Chemistry* **382,** 933–939 (2001).

223. Ikegami, M., Takabatake, N. & Weaver, T. E. Intersubunit disulfide bridge is not required for the protective role of SP-B against lung inflammation. *Journal of Applied Physiology* **93,** 505–511 (2002).

224. Beck, D. C., Ikegami, M., Na, C. L., Zaltash, S., Johansson, J., Whitsett, J. A. & Weaver, T. E. The role of homodimers in surfactant protein B function in vivo. *Journal of Biological Chemistry* **275,** 3365–3370 (2000).

225. Veldhuizen, E. J., Waring, A. J., Walther, F. J., Batenburg, J. J., Van Golde, L. M. & Haagsman, H. P. Dimeric N-terminal segment of human surfactant protein B (dSP-B1-25) has enhanced surface properties compared to monomeric SP-B1-25. *Biophysical Journal* **79,** 377–384 (2000).

226. Baatz, J. E., Zou, Y., Cox, J. T., Wang, Z. & 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,** 180–190 (2001).

227. Baatz, J. E., Elledge, B., Whitsett, J. A., John E. Baatz, Barry Elledge & Whitsett*, J. A. Surfactant protein SP-B induces ordering at the surface of model membrane bilayers. *Biochemistry* **29,** 6714–6720 (1990).

228. Olmeda, B., Garcia-Alvarez, B., Gomez, M. J., Martinez-Calle, M., Cruz, A. & Perez-Gil, J. A model for the structure and mechanism of action of pulmonary surfactant protein B. *The FASEB Journal* **29,** 1–12 (2015).

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

230. Dluhy, R., Cameron, D., Mantsch, H. H. & Mendelsohn, R. Fourier transform infrared spectroscopic studies of the effect of calcium ions on phosphatidylserine. *Biochemistry* **22,** 6318–6325 (1983).

231. Chavarha, M., Loney, R. W. W., Rananavare, S. B. B. & Hall, S. B. B. Hydrophobic surfactant proteins strongly induce negative curvature. *Biophysical Journal* **109,** 95–105 (2015).

232. Chang, R., Nir, S. & Poulain, F. R. Analysis of binding and membrane destabilization of phospholipid membranes by surfactant apoprotein B. *Biochimica et Biophysica Acta* **1371,** 254–264 (1998).

233. Warren, J. B. & Anderson, J. M. Core concepts: Respiratory distress syndrome. *NeoReviews* **10,** 351–361 (2009).

234. Roberts, D., Brown, J., Medley, N. & Dalziel, S. R. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *The Cochrane database of systematic reviews* **3,** 1–112 (2017).

235. Clark, H. & Clark, L. S. The genetics of neonatal respiratory disease. *Seminars in fetal & neonatal medicine* **10,** 271–282 (2005).

236. Beers, M. F., Hawkins, A., Maguire, J. A., Kotorashvili, A., Zhao, M., Newitt, J. L., Ding, W., Russo, S., Guttentag, S., Gonzales, L. & Mulugeta, S. A nonaggregating surfactant

protein C mutant is misdirected to early endosomes and disrupts phospholipid recycling. *Traffic* **12,** 1196–1210 (2011).

237. Bolt, R. J., van Weissenbruch, M. M., Lafeber, H. N. & Delemarre-van de Waal, H. A. Glucocorticoids and lung development in the fetus and preterm infant. *Pediatric Pulmonology* **32,** 76–91 (2001).

238. Sharma, S., Kho, A. T., Chhabra, D., Qiu, W., Gaedigk, R., Vyhlidal, C. A., Leeder, J. S., Barraza-Villarreal, A., London, S. J., Gilliland, F., Raby, B. A., Weiss, S. T. & Tantisira, K. G. Glucocorticoid genes and the developmental origins of asthma susceptibility and treatment response. *American Journal of Respiratory Cell and Molecular Biology* **52,** 543–553 (2015).

239. Thompson, B. T., Chambers, R. C. & Liu, K. D. Acute respiratory distress syndrome. *New England Journal of Medicine* **377,** 562–572 (2017).

240. Leonenko, Z., Gill, S., Baoukina, S., Monticelli, L., Doehner, J., Gunasekara, L., Felderer, F., Rodenstein, M., Eng, L. M. & Amrein, M. An elevated level of cholesterol impairs self-assembly of pulmonary surfactant into a functional film. *Biophysical Journal* **93,** 674–683 (2007).

241. Al-Saiedy, M., Pratt, R., Lai, P., Kerek, E., Joyce, H., Prenner, E., Green, F., Ling, C.-C. C., Veldhuizen, R., Ghandorah, S. & Amrein, M. Dysfunction of pulmonary surfactant mediated by phospholipid oxidation is cholesterol-dependent. *Biochimica et Biophysica Acta* **1862,** 1040–1049 (2018).

242. Schmidt, R., Meier, U., Yabut-Perez, M., Walmrath, D., Grimminger, F., Seeger, W. & Günther, A. Alteration of fatty acid profiles in different pulmonary surfactant phospholipids in acute respiratory distress syndrome and severe pneumonia. *American Journal of Respiratory and Critical Care Medicine* **163,** 95–100 (2001).

243. Dowling, K. D., Doyle, I. R., Davidson, K. G., Barr, H. A., Nicholas, T. E. & Bersten, A. D. Lung function, permeability, and surfactant composition in oleic acid-induced acute lung injury in rats. *American Journal of Physiology: Lung Cellular and Molecular Physiology* **279,** L1091–L1102 (2000).

244. Nagase, T., Uozumi, N., Ishii, S., Kume, K., Izumi, T., Ouchi, Y. & Shimizu, T. Acute lung injury by sepsis and acid aspiration: a key role for cytosolic phospholipase A2. *Nature immunology* **1,** 42–46 (2000).

245. Meyer, N. J. & Christie, J. D. Genetic heterogeneity and risk of acute respiratory distress syndrome. *Seminars in Respiratory and Critical Care Medicine* **34,** 459–474 (2013).

246. Forbes, A., Pickell, M., Foroughian, M., Yao, L.-J., Lewis, J. & Veldhuizen, R. Alveolar macrophage depletion is associated with increased surfactant pool sizes in adult rats. *Journal of Applied Physiology* **103,** 637–645 (2007).

247. Willson, D. F. & Notter, R. H. The future of exogenous surfactant therapy. *Respiratory Care* **56,** 1369–1386 (2011).

248. Schmidt, R., Markart, P., Ruppert, C., Wygrecka, M., Kuchenbuch, T., Walmrath, D., Seeger, W. & Guenther, A. Time-dependent changes in pulmonary surfactant function and composition in acute respiratory distress syndrome due to pneumonia or aspiration. *Respiratory Research* **8,** 1–11 (2007).

249. Hu, X., Tian, Y., Qu, S., Cao, Y., Li, S., Zhang, W., Zhang, Z., Zhang, N. & Fu, Y. Protective effect of TM6 on lipopolysaccharide-induced acute lung injury in mice. *Scientific Reports* **7,** 1–10 (2017).

250. Singh, N., Halliday, H. L., Stevens, T. P., Suresh, G., Soll, R. & Rojas-Reyes, M. X. Comparison of animal-derived surfactants for the prevention and treatment of respiratory distress syndrome in preterm infants. *The Cochrane database of systematic reviews,* 1–67 (2015).

251. Lemyre, B., Fusch, C., Schmölzer, G. M., Bouali, N. R., Reddy, D., Barrowman, N., Huneault-Purney, N. & Lacaze-Masmonteil, T. Poractant alfa versus bovine lipid extract surfactant for infants 24+0 to 31+6 weeks gestational age: A randomized controlled trial. *PLoS ONE* **12,** 1–13 (2017).

252. Kattwinkel, J. Synthetic surfactants: The search goes on. *Pediatrics* **115,** 1075–1076 (2005).

253. Miller, M. N., Bangham, A. & South, D. M. Ten centre trial of artificial surfactant (artificial lung expanding compound) in very premature babies. *British Medical Journal* **294,** 991–996 (1987).

254. Ainsworth, S. B., Beresford, M. W., Milligan, D. W., Shaw, N. J., Matthews, J. N., Fenton, A. C. & Ward Platt, M. P. Pumactant and poractant alfa for treatment of respiratory distress syndrome in neonates born at 25-29 weeks' gestation: A randomised trial. *The Lancet* **355,** 1387–1392 (2000).

255. Soll, R. F. Synthetic surfactant for respiratory distress syndrome in preterm infants. *Cochrane Database of Systematic Reviews* **2,** CD001149 (1998).

256. Soll, R. & Özek, E. Prophylactic protein free synthetic surfactant for preventing morbidity and mortality in preterm infants. *Cochrane Database of Systematic Reviews,* 1–17 (2010).

257. Soll, R. & Blanco, F. Natural surfactant extract versus synthetic surfactant for neonatal respiratory distress syndrome. *Cochrane Database of Systematic Reviews,* 1–23 (2001).

258. Spragg, R. G., Lewis, J. F., Walmrath, H.-D. D., Johannigman, J., Bellingan, G., Laterre, P.-F. F., Witte, M. C., Richards, G. A., Rippin, G., Rathgeb, F., Hafner, D., Taut, F. J. H., Seeger, W., Häfner, D., Taut, F. J. H. & Seeger, W. Effect of recombinant surfactant protein C-based surfactant on the acute respiratory distress syndrome. *New England Journal of Medicine* **351,** 884–892 (2004).

259. Sardesai, S., Biniwale, M., Wertheimer, F., Garingo, A. & Ramanathan, R. Evolution of surfactant therapy for respiratory distress syndrome: Past, present, and future. *Pediatric Research* **81,** 240–248 (2017).

260. Moya, F. R., Gadzinowski, J., Bancalari, E., Salinas, V., Kopelman, B., Bancalari, A., Kornacka, M. K., Merritt, T. A., Segal, R., Schaber, C. J., Tsai, H., Massaro, J. & D'Agostino, R. A multicenter, randomized, masked, comparison trial of Lucinactant, Colfosceril Palmitate, and Beractant for the prevention of respiratory distress syndrome among very preterm infants. *Pediatric* **115,** 1018–1029 (2005).

261. Manalo, E., Merritt, T. A., Kheiter, A., Amirkhanian, J. & Cochrane, C. Comparative effects of some serum components and proteolytic products of fibrinogen on surface tension-lowering abilities of Beractant and a synthetic peptide containing surfactant KL4. *Pediatric Research* **39,** 947–952 (1996).

262. Seehase, M., Collins, J. J. P., Kuypers, E., Jellema, R. K., Ophelders, D. R. M. G., Ospina, O. L., Perez-Gil, J., Bianco, F., Garzia, R., Razzetti, R. & Kramer, B. W. New surfactant with SP-B and C analogs gives survival benefit after inactivation in preterm lambs. *PLoS ONE* **7,** 1–10 (2012).

263. Walther, F. J., Hernández-Juviel, J. M., Gordon, L. M. & Waring, A. J. Synthetic surfactant containing SP-B and SP-C mimics is superior to single-peptide formulations in rabbits with chemical acute lung injury. *PeerJ* **2,** 1–15 (2014).

2

The Expression and Purification of Recombinant Human Lung Surfactant Protein-B (rSP-B)

2.1 Statement of Authorship

Tadiwos Getachew prepared the samples used for analysis in figures 4, 7 and 8, analyzed and interpreted the data, and also wrote the manuscript. Donna Jackmann prepared the samples used in figures 2, 3, 5 and 6, run SDS-PAGE and blotted the gels and Tadiwos Getachew acquired the pictures and interpreted the results. Tadiwos Getachew trained Selina Freudenberg and used a sample prepared by her for western blotting shown in figure 7.

[1]

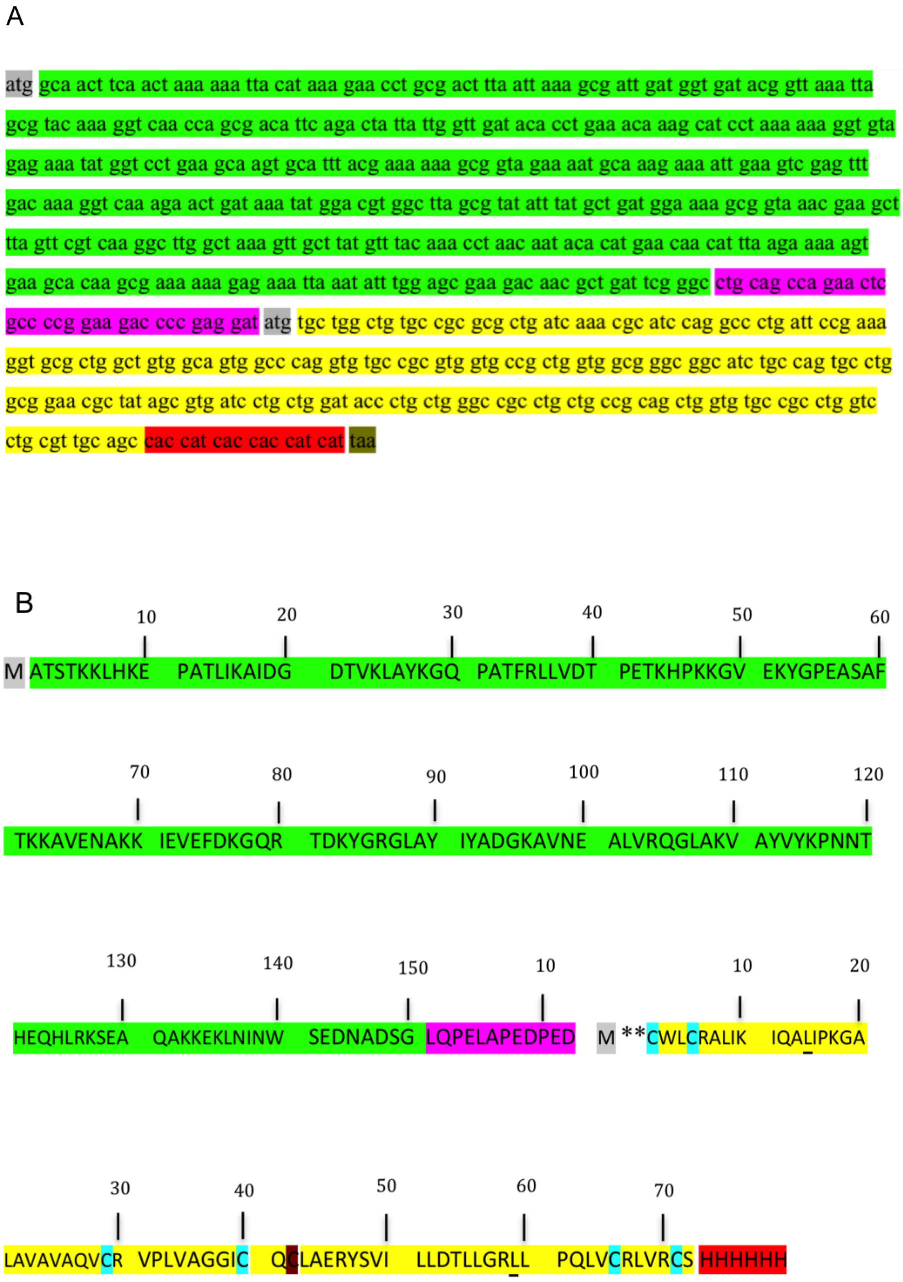
---

[1] In figure captions and legends of this chapter abbreviations can be found. These are used to represent the persons who prepared the protein samples used for analysis. DJ: Donna Jackmann, SF: Selina Freudenberg and TG: Tadiwos Getachew. The numbers following the abbreviations are protein isolation batch numbers.

2.2 Background

Historically, many groups have had challenges producing SP-B, due in large part to its extreme hydrophobicity. Besides us, groups around the world have attempted to express SP-B in bacteria, insect cells and cell-free expression systems, but without success. Also, attempts at chemical synthesis of SP-B have failed. Full length SP-B is big for solid phase peptide synthesis. To work around this problem, one group attempted to synthesized SP-B as fragments of two peptides and ligate the pieces together but this was also unsuccessful ("personal communications from Alan Waring, John Baatz and Jesus Perez-Gil to Valerie Booth, 2004-2017").

Earlier, Jesus Perez Gil (PhD) at Universidad Complutense de Madrid and co-workers managed to express an SP-C variant in bacteria using Staphylococcus nuclease A (SN), a water soluble, single stranded DNA cleaving enzyme of 150 amino acids (≈16.9 kDa), as a fusion tag (applications of fusion tags are discussed in detail below) [1]. Our group adapted the approach and used it to design a plasmid construct to express rSP-B in bacteria and purify it using affinity chromatography [2]. Instead of the thrombin cleavage site that was located between rSN and rSP-C, a method first developed by Hakes and Dixon [3], a cyanogen bromide (CNBr) cleavage site was inserted. The chemical cleavage site allowed rSP-B to be cleaved from rSN in conditions where rSP-B is soluble, but most enzymes are not functional. However, the purification of this construct by LH-60 columns (Sephadex, Pharmacia fine chemicals, Sweden) and other techniques was not successful [2]. The expression construct that is central to the work reported in this thesis was designed to retain the histidine-tag (His-tag) on rSP-B, so that rSP-B could be refolded by using metal ion affinity column after rSN was cleaved off. The vector construct of the clone is shown in Figure 2.1. The construct shown uses rSN (green), as a fusion partner, i.e. just preceding the N-terminus of rSP-B (yellow). A methionine residue (grey) connects the rSN part with rSP-B, and provides the site cleaved by CNBr. The C-terminus of the target protein was fused with a 6xHis tag (Red). Histidines have an imidazole functional group, which has high affinity for bivalent metal ions. As suggested by Sharifahmadian *et al.* [4], the insertion sequence (FPIPLPY), has been removed to improve the solubility of rSP-B. Although we worked with a number of other vector constructs for rSP-B expression, the plasmid rSN-Δ7NTΔM-rSP-B-6xHis is the most thoroughly characterized rSP-B construct in the Booth lab. Unless stated otherwise, herein, I will be referring to this plasmid construct.



atg gca act tca act aaa aaa tta cat aaa gaa cct gcg act tta att aaa gcg att gat ggt gat acg gtt aaa tta
gcg tac aaa ggt caa cca gcg aca ttc aga cta tta ttg gtt gat aca cct gaa aca aag cat cct aaa aaa ggt gta
gag aaa tat ggt cct gaa gca agt gca ttt acg aaa aaa gcg gta gaa aat gca aag aaa att gaa gtc gag ttt
gac aaa ggt caa aga act gat aaa tat gga cgt ggc tta gcg tat att tat gct gat gga aaa gcg gta aac gaa gct
tta gtt cgt caa ggc ttg gct aaa gtt gct tat gtt tac aaa cct aac aat aca cat gaa caa cat tta aga aaa agt
gaa gca caa gcg aaa aaa gag aaa tta aat att tgg agc gaa gac aac gct gat tcg ggc ctg cag cca gaa ctc
gcc ccg gaa gac ccc gag gat atg tgc tgg ctg tgc cgc gcg ctg atc aaa cgc atc cag gcc ctg att ccg aaa
ggt gcg ctg gct gtg gca gtg gcc cag gtg tgc cgc gtg gtg ccg ctg gtg gcg ggc ggc atc tgc cag tgc ctg
gcg gaa cgc tat agc gtg atc ctg ctg gat acc ctg ctg ggc cgc ctg ctg ccg cag ctg gtg tgc cgc ctg gtc
ctg cgt tgc agc cac cat cac cac cat cat taa

M ATSTKKLHKE PATLIKAIDG DTVKLAYKGQ PATFRLLVDT PETKHPKKGV EKYGPEASAF

TKKAVENAKK IEVEFDKGQR TDKYGRGLAY IYADGKAVNE ALVRQGLAKV AYVYKPNNT

HEQHLRKSEA QAKKEKLNINW SEDNADSG LQPELAPEDPED M **CWLCRALIK IQALIPKGA

LAVAVAQVCR VPLVAGGIC QCLAERYSVI LLDTLLGRLL PQLVCRLVRCS HHHHHH

Figure 2.1: Three letter DNA (A) and one amino acid letter code (B) of the expressed construct. Panel A: Methionine residues are highlighted in grey, rSN in green, HSV epitope tag in magenta, rSP-B in yellow, polyhistidine-tag in red and the stop codon in seaweed. Panel B: Cysteines that form intrachain disulphide bridges are marked in cyan while the cysteine that is thought to be involved in interchain covalent homodimerization is highlighted dark purple. The two underlined leucine residues are point mutants (methionine to leucine). Adapted from (with permission) Valerie Booth.

This chapter presents data that illustrate the production of recombinant human lung surfactant protein B (rSP-B) in bacteria. In order to appreciate the challenge with the production of rSP-B, some background on recombinant expression strategies are discussed.

When considering the organism in which to express a protein of interest, each expression system i.e. mammalian, insect, yeast, bacterial or cell-free has its own merits and drawbacks. Eukaryotic cells are the most suitable to express complex proteins. Mammalian cells, for instance, possess the machinery to introduce post-translational modifications (PTMs), but are also the most expensive systems. Eukaryotic systems in general are costly, time consuming, and not straight forward, i.e. strict adherence to experimental protocols is mandatory [5]. As a rule of thumb, large recombinant proteins are often expressed in an eukaryotes and small proteins in prokaryotes [6].

Prokaryotes, comparatively, are less expensive, convenient to use, and can be scaled-up easily. Escherichia coli (*E. coli*), for instance, has fast growing kinetics, with a multiplication time of ≈20 minutes [7, 8]. Also, the microorganism attains high cell density in cultivation media [9, 10]. Transformation of an expression vector into the bacterium is fairly easy. Moreover, the physiology of the enterobacterium is well characterized and documented. For example, *E. coli* has well described transcription and translation mechanisms. Thus, *E. coli*, is often the most convenient and simple system for overproduction of a heterologous gene product at both laboratory and industrial scale. For example, around 30% of recombinant protein pharmaceuticals approved by the FDA as of 2011 [11], and 90% of structures deposited in protein data bank (PDB) as of now [12], are produced in *E. coli.* On the other hand, prokaryotic systems are hampered by the inability to process PTMs and to promote folding of complex mamalian proteins [11]. In general terms, the three major factors to consider when selecting an expression system are i) expression level, ii) solubility, and iii) purification [5, 13].

After selecting the right host for a given gene, the vector type is also paramount to express heterologous protein in ample quantities [7]. An expression vector is characterized by three prominent features, being the promoter, the selection marker, and the origin of replication. The pET expression vector with bacteriophage T7 RNA polymerase is the most common plasmid used in *E. coli.* Most promotors are inducible by chemical stimuli. The T7 RNA polymerase is under the control of the lactose operon (lac-operon) gene, which can be turned on by the chemical inducer, isopropyl β-D-1-thiogalactopyranoside (IPTG). In the presence of IPTG, lacI, the repressor of the *lac-operon* gene, is released, initiating the transcription of the downstream plasmid [14]. Although pET is well tolerated and easily transformed into *E. coli*, the type of protein cloned into the vector may adversely affect the cells, and consequently expression levels because of toxicity. Given the hydrophobic nature and membrane lysing behaviour of SP-B [15], together with the known difficulty of expressing it in standard expression cells, SP-B is likely to be toxic to *E. coli*.

Some recombinant proteins are toxic to host cell profileration and routine functions and hence can only be expressed in very low levels. *E. coli* cells transfected with a plasmid that codes for a toxic protein tend to show slow growth rate, low cell density, and possibly death after induction [16, 17]. Nonetheless, many toxic proteins can be expressed in specialized *E. coli* mutant strains, such as C41(DE3) and C43(DE3). While C41 is a single mutant of BL21, C43 is a double mutant of C41 and demonstrates a better success rate of expressing membrane proteins than its predcessor, C41 [18]. The lacUV5 promotor that controls transcription of the T7 RNA polymerase contains mutations that lead to decreased expression of T7 RNA polymerase in C41 and C43 cells [19]. These strains also solve the issue of plasmid instability during expression of toxic proteins [20] and also possess high transformation efficiency. However, in general, even if a cell line can be found that is capable of expressing a toxic protein, the expressed protein will often locate to inclusion bodies (IBs).

IBs are electron-dense, amorphous aggregates that mostly contain the protein of interest either in a native-like, partially folded state, or in a misfolded state [21, 22]. More often than not, IBs accumulate in the cytoplasm, but occasionally the particles escape into the periplasmic space as well [23]. Studies have shown that proteins with more disulphide bonds and hydrophobic residues are more prone to IB formation than other proteins [11, 24]. Hence, IBs could also be the result of incorrectly formed disulphide bonds [16]. Studies have shown that IBs have both advantages and disadvantages for protein preparation.

IBs offer a number of important benefits. Owing to their large size, inclusion bodies can easily be isolated from other host proteins via centrifugation. Consequently, IBs significantly aid in the purification of the protein of interest [25]. According to Rinas *et al.* [26] IBs have little contaminating host proteins, ribosomal proteins, or nucleic acids, which reduces the amount of purification work needed. Additionally, although they are targets of the proteasome, due to their insoluble aggregated nature, IBs are may be more resistant to proteolytic degradation [27–29].

On the reverse side, IBs also have some undesirable apects. Literature data indicate that IBs also can contain cross-beta structure which imparts a amyloid-like fibril character to them [30]. Amyloid fibrils are associated with different human pathologies including Alzheimer's disease, Parkinson's disease, and type II diabetes [30–32]. Hence, proper care should be taken in the downstream purification and processing of proteins that are prone to aggregation. To disrupt and solubilize the aggregates, IBs need strong chaotropic agents (6-8 M), detergents, and reducing agents. Subsequently, their refolding is cumbersome [33], and requires screening and optimization of a number of renaturation conditions [7, 34]. The data indicates that more than 40% of IBs can be recovered into their bio-active state [35]. The best resolution, if at all possible, however is to produce soluble protein and avoid the challenging renaturation step.

E. coli encodes chaperones of its own that can sometimes help with the folding of expressed recombinant protein. These include the non-ATP mediated Trigger Factor (TF) nascent protein binding chaperone and the ATP-dependent nascent protein binding Dnak chaperone system, as well as the downstream GroEL/GroES co-chaperone system. As soon as nascent protein emerges out of the ribosome, it binds the TF chaperone to assist the protein to fold into its native soluble form [36, 37]. However, such systems may not be compatible to eukaryotic proteins or, even worse, hydrophobic protein may be toxic to the host. Moreover, co-expression of the target and the chaperone protein on the same plasmid is a burden for the host strain [38]. To counteract the solubility and toxicity of some recombinant proteins expressed in *E. coli*, other new methods have been devised. One of the most commonly used are solubility fusion proteins.

Fusion partners are either proteins or peptides that are fused at either of the termini of the target protein to facilitate folding, increase solubility, reduce toxicity, and decrease proteolysis of the target [13, 39]. Different proteins react differently to solubility tags and there are many commericially available fusion proteins to choose from. Among these, the *E. coli* maltose binding protein (MBT), gluthathione S-transferase (GST) and thioredoxin (TRX) have been popular partners for membrane/membrane-associated protein production [40, 41]. Staphylococcus nuclease A (SN) has been used as a chaperone protein for *E. coli* expression of SP-C, one of the most hydrophobic mammalian proteins [1]. Some fusion proteins, however, might adversely affect the structure and function of their partner. Hence, it is customary to devise a method of removing the solubility protein eventually [5]. Fusion proteins can either be removed enzymatically or chemically. Subsequently, following cleavage, the fusion protein must also be separated from the target protein. This is often accomplished with the use of an affinity tag, e.g. a polyhistidine tag on one protein or the other. However, before this can happen, the full expressed protein construct, (i.e. the protein of interest plus the fusion protein) must be purified from the rest of the bacterial cell components.

Affinity chromatography works via reversible interaction with the target protein and a specific ligand on the resin. The interaction could be bio-specific (antibody interaction) or non-bio-specific through immobilized metal ions [13]. Immobilized metal ion chromatography (IMAC) is a robust purification technique with low cost and maximal recovery yield [42]. The metal ions are immobilized on a resin and an affinity tag on the target/fusion protein usually has high affinity for the immobilized divalent metal ions. Because histidine is fairly inert, proteins with polyhistidine tag (His-tag) are suitable for downstream purification and structural studies [43, 44]. Hence, it is not always necessary to remove the His-tag [45]. The fusion protein, on the other hand, generally needs to be separated from the protein of interest. The protein that contains the affinity tag remains bound to the resin while the other without the tag flows out.

Like most integral membrane proteins, SP-B is also highly apolar and is partly membrane associated. The high apolar nature of these proteins presents an inherent problem for characterizing them. To this end, the data indicate that membrane or membrane-associated proteins only represent a minute amount of the structures presented in PDB despite constituting about 20-30% of the proteome of many species. Consequently, the structure-to-function relationships of most of these proteins are poorly understood [25, 46]. Part of the problem is that highly apolar proteins are sparingly soluble in water.

Apolar proteins aggregate in water. There is a driving force to cluster the non-polar residues together and expose the hydrophilic side chains to the aqueous surface, thereby creating non-native hydrophobic contacts. To prevent aggregation and keep their native state, recombinant membrane proteins must be kept in a membrane mimetic environment. However, as each protein is unique, there is no universal solvent. Rather, different lipids and detergents are screened for each hydrophobic protein of interest to identify the conditions that best solubilize and maintain the native function of the protein. As well as being quite a cumbersome procedure, the lipid or detergents used to solubilize and stabilise the protein could also complicate purification of the target [47]. Sodium dodecyl sulphate (SDS), for instance, is one popular detergent used as a membrane mimetic environment [48]. Detergents are amphiphiles and form micelles at the critical micellar concentration (CMC) i.e. the minimum concentration above which micelles start to form [49]. Micelles have both lipophilic and hydrophilic entities and thus resemble membrane surroundings. Expression of a membrane protein in *E. coli* could proceed in two possible ways. Either the protein inserts into the bacterial membrane in functional form or the protein expresses as an inclusion body [25]. The latter, however, is often preferable, as purifying a membrane inserted hydrophobic protein can be more challenging. Although a cell line might be found that expresses a toxic protein, the expression is often hampered by low yield and usually requires a lot of optimization and careful choice of expression conditions to maximize yield and solubility of apolar proteins.

Concentration of induction agents, incubation temperature, types of cultivation media, and host strains can all be optimized [7, 13, 50, 51]. In bacteria, lower temperature promotes a reduced rate of protein synthesis, which is known to promote protein folding and solublity [52, 53]. High temperature, on the other hand, may promote stress that leads to protein aggregation and hence IB formation. Also, important considerations come from studies that have reported that expression at physiological temeprature is more susceptible to proteolytic degradation than expression at lower temperature [54, 55]. By the same token, incomplete or partial induction of promoters also leads to a slow rate of protein synthesis at low temperature [56]. Hence, decreased concentration of inducers may increase the solubility of heterologous genes [57]. The type of media also plays a vital role as many bacteria thrive differently in various media [53]. Thus, the level of expression of a given gene varies from media to media [58]. Increasing glucose concentration in incubation media for an example, has been reported to enhance expression levels via decreasing promoter suppression [7, 59–61]. Lastly, some strains, such as BL21, which is the most common *E. coli* strain for recombinant expression of heterologous genes, is known to improve plasmid stability [62]. There are more strains with different traits, which improve upon the parent BL21 cell. For instance, strains that improve correct cytoplasmic disulphide bond formation and strains capable of regulating expression level are often helpful. Importantly, for SP-B work, C41(DE3) and C43(DE3) have demonstrated considerable success in over-production of hydrophobic and toxic proteins [63]. As C43 cells form an extra membrane during expression of heterologous DNA, it could be that such a membrane protect C43 cells from the toxicity of hydrophobic proteins. In addition, a portion of the target protein may also associate with the extra membrane.

Expression of SP-B has been very challenging. Mature SP-B is highly hydrophobic, membranolytic, and fusogenic. SP-B has one of the highest hydropathy (GRAVY: grand average of hydropathicity) index [64]. The higher is the hydropathy index, the greater is the hydrophobicity and vice versa is true for hydrophilic proteins [65]. In other words, a positive GRAVY index indicates that the protein is a hydrophobic and negative index indicates a hydrophilic protein. Most SAPLIPs have negative GRAVY indices. This indicates that SP-B is unique. If not the most hydrophobic, SP-B, undoubtedly, is one among the most hydrophobic proteins. Thus, while many other integral membrane proteins proved to be very difficult to overexpress and characterize, SP-B is even more so. Here, we report the recombinant expression and purification of surfactant protein B (rSP-B) in *E. coli*, strain C43 using SN as a fusion partner for the first time.

2.3 Materials and Methods

2.3.1 Transformation

To transform the cells with the plasmid, 50 µl of *E. coli* strain C43 (>1×10^6 cfu/µg DNA) chemically competent cells (VWR International, Edmonton, Canada) was mixed with 1 µl of the vector DNA (pET11a-SN-SP-B). The mixture was placed on ice for 30 minutes and then heat shocked at 42°C for 1 minute. One millilitre of the expression recovery medium (lactose minus, Lucigen) was added and incubated for an hour at 37°C and 175 RPM. Next, the mixture, was plated onto 2XYT (16 g/L tryptone, 10 g/L yeast extract and 5 g/L sodium chloride) plus ampicillin (1 µg/ml, Fisher BioReagents) agar medium and was incubated overnight at 37°C. An isolated colony was inoculated into 5 ml of 2XYT ampicillin selective media and cultured overnight at 37°C at 150 RPM. Aliquotes of 1:1 (v/v) glycerol:culture stock solution were kept at -80°C.

2.3.2 Expression

To grow the culture for expression, 75 ml of 2XYT starter culture was inoculated with 100 µl of the glycerol stock cells and was incubated overnight at 30°C while shaking at 150 RPM. Next, 60 ml of the overnight suspension culture was used to inoculate 6 litres of pre-autoclaved (liquid cycle at 121°C, 15 lb PSI for 20 minutes) expression media (2XYT). Prior to inoculation, filter sterile ampicillin, a final concentration of 0.05 mg/ml, was added to the sterile media after it cooled to room temperature. The incubation was continued until the optical density (OD 600 nm) reached 0.6 i.e. mid-log phase. Since the plasmid was under the control of the lac-operon gene, IPTG (isopropyl-β-D-thiogalactopyranoside, Gold Biotechnology), was added to a final concentration of 0.4 mM (400 µl per litre) to induce the expression of rSN-SP-B. After three hours incubation at 37°C and 175 RPM, the cells were harvested by centrifugation at 4690 g for 15 minutes at an ambient temperature. The supernatant was discarded and the pellet was suspended in 20 ml of 1XTBS (50 mM TRIS, 150 mM NaCl, pH 7.4) with 5 mM PMSF (phenylmethylsulfonylfluoride).

2.3.3 Cell Lysing

Cells were passed through a French press (hydraulic pressure liquid shearing force, 10,000 psi) at 4°C in three consecutive cycles. Alternatively, cells were disrupted by sonication (four cycles of 20 seconds each). Samples were kept on ice water at all times during sonication. Since rSP-B expresses mainly as inclusion bodies, urea and 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS, AG Scientific) were added to the lysed cells to a final concentration of 6 M urea and 0.2% (w/v) CHAPS. The lysate was agitated for 2 hours at room temperature on an orbital shaker. To remove most of the DNA from the sample, 10 ml of Ion-Sep DE52, pre-swollen DEAE (diethyl amino ethyl) cellulose (Biophoretics, Sparks, NV) was added to the lysate. The mixture was left on a rotary shaker at room temperature for a minimum of an hour and then centrifuged at 7580 g at 25°C for 15 minutes. After discarding the resin pellet, the supernatant was retained to enrich the target protein.

2.3.4 Purification

An immobilized metal ion affinity chromatography (IMAC) column was used in two stages of the protein enrichment (Figure 2.2): first to purify the fusion protein from the whole bacterial cell mass and second, to separate the target protein (rSP-B) from the fusion protein (rSN) after cleavage. Five ml of IMAC Sepharose 6 Fast Flow resin (GE healthcare, UK) was applied to a PD-10 column (GE healthcare,UK). The resin was loaded with 3 ml of 0.1 M nickel sulfate solution. To remove any unbound nickel metal ions from the column, the resin was washed with 5 column volumes of distilled water and equilibrated with a minimum of 5 column volumes (25 ml) of the loading buffer (5 mM imidazole, 6 M urea, 0.2% (w/v) CHAPS, 1XTBS, pH 7.5).

In the first stage of purification, separation of the fusion protein (rSN-SP-B) from the rest of bacterial cell mass was achieved. The sample, after the cell lysis and ion exchange steps, was applied to a 5 ml pre-equilibrated IMAC resin and was poured into a 50 ml centrifuge tube. The mixture was gently agitated overnight at room temperature. Following that, the resin was taken into a PD-10 column. To wash off the non-bound bacterial cell mass, the column was washed with five column volumes (25 ml) of the loading buffer. Next, rSN-SP-B, was eluted from the column using 10 ml aliquots of an imidazole concentration increment (5, 10, 20, 50, 100, 200 and 300 mM imidazole in the loading buffer). Fractions (≈1.5 ml/tube) were collected. Absorbance at UV 280 nm of each fraction was taken using a clear quartz cuvette after blanking the spectrophotometer with the elution buffer and at any time the urea concentration was changed in the sample. Finally, absorbance against elution volume was plotted.

To confirm the expression level and successful enrichment of rSN-SP-B, a 1:1 (v/v) mixture of sample and sample buffer (1.3% (w/v) SDS, 65 mM TRIS, 13% (w/v) glycerol, 0.1% bromophenol blue, 0.02% sodium azide (w/v)) was boiled for 2-3 minutes after which, 30 µl samples were loaded onto a precast 12% non-reducing Tris-Glycine SDS-PAGE (Bio-Rad) and run for 1 hour at 180 V. The gel was stained using Coomassie Brilliant Blue R-250 for one hour and de-stained until the desired bands were observed. Fractions that showed the expected bands on the gel were pooled together.

Protein was quantified with a Bradford assay by using a Bio-Rad protein assay dye concentrate. To read-off the unknowns, bovine serum albumin (1 mg/ml) was used to make a standard curve in the same buffer as in which the protein was suspended. The concentrate stock was diluted (1:10) with deionised water. A protein suspended in the buffer of choice (i.e. the buffer plus the detergent used to stabilise the protein) was taken and 1:10, 1:20, 1:50, and 1:100 aliquots were made in the dilute dye. An absorbance reading of the aliquots was taken at 595 nm using UV-Visible spectrophotometer. Each time prior to reading sample absorbance, the spectrophotometer was zeroed with a mixture of the dilute dye plus the buffer. The dilution factor was taken into consideration during calculation of the concentration of the unknown. The accuracy was monitored by comparing how well the concentrations of each aliquot agreed with each other.

To remove the CHAPS, salt, imidazole and urea, the pooled sample was dialysed using Spectra POR-4 12,000 to 14,000 molecular weight cut-off (MWCO) dialysis tubing (VWR International, Edmonton Canada) against 3 litres of distilled water overnight in a cold room (+4°C) while stirring. The dialysed sample was flash frozen using liquid nitrogen and lyophilized overnight

(Labconco freeze drier, Kansas City, MO, USA). In order to prepare rSN-SP-B for cleavage, the lyophilized sample was resuspended in a minimal volume (2 ml) of 70% formic acid. A few small crystals of cyanogen bromide (CNBr 97%, Acros Organics) were added to it. Cyanogen bromide is a volatile toxic compound and should be handled in the fume hood with all the necessary safety precautions to avoid inhalation and skin contact. The vial was wrapped in aluminium foil and was left for 1 to 2 days at room temperature. The reaction was stopped by diluting the sample with distilled water (1:15 dilution). The solution was flash frozen and lyophilized (mass after lyophilization (≈9.3 mg). The resulting pellet was resuspended in 5 ml of the second loading buffer (6 M urea, 0.5% (w/v) CHAPS, 1XTBS, 5 mM imidazole, pH 7.5) to prepare it for the next IMAC column.

In the next purification step, an IMAC was used to separate rSP-B from rSN, to renature rSP-B on-column, to exchange rSP-B into the desired detergent or lipid, and finally to elute rSP-B. The cleaved sample was incubated with 5 ml of IMAC resin previously pre-equilibrated with the loading buffer. This was left to rotate on a rotary shaker overnight at room temperature in the PD-10 column. The next day, the column was allowed to drain and then was washed with 15 ml of the loading buffer. This allowed the rSN to flow through, while the rSP-B, which retains the His-tag, remained bound to the column. Next, the rSP-B was renatured while still bound to the column by decreasing the concentration of urea in the column. This was done by washing the column with 10 ml of washing buffer (6 M urea, 0.5% CHAPS, 1XTBS and 5 mM imidazole) in stepwise 1 M decrements of urea from 6 M to 0 M. To remove the salt and CHAPS, the column was washed with 30 ml of 20 mM Tris-HCl buffer (pH 7.5). Following this, detergent exchange was performed by washing the resin with 3 column volumes (15 ml) of the desired detergent or lipid and the Tris-HCl buffer (pH 7.5). The lipids used underwent 5 freeze thaw cycles either in liquid nitrogen or in a freezer. After each cycle of freezing, the lipids were thawed at 50°C to break the large multilamellar vesicles into smaller vesicles. Finally, rSP-B was eluted from the column by washing it with 5 column volumes of the elution buffer, either with increasing imidazole concentration or with decreased pH (pH 5). Absorbance (UV 280 nm) readings of the fractions were taken and a graph was plotted against the fraction volumes. The fractions with highest absorbance were selected to run on a 16% Tris-Tricine SDS-PAGE (Bio-Rad) for 2 hours at 80 V in duplicate. One gel was retained for western blotting and the second gel was silver stained using a kit from Bio-Rad as per the manufacturer's instructions. The sample was then ready for characterization.

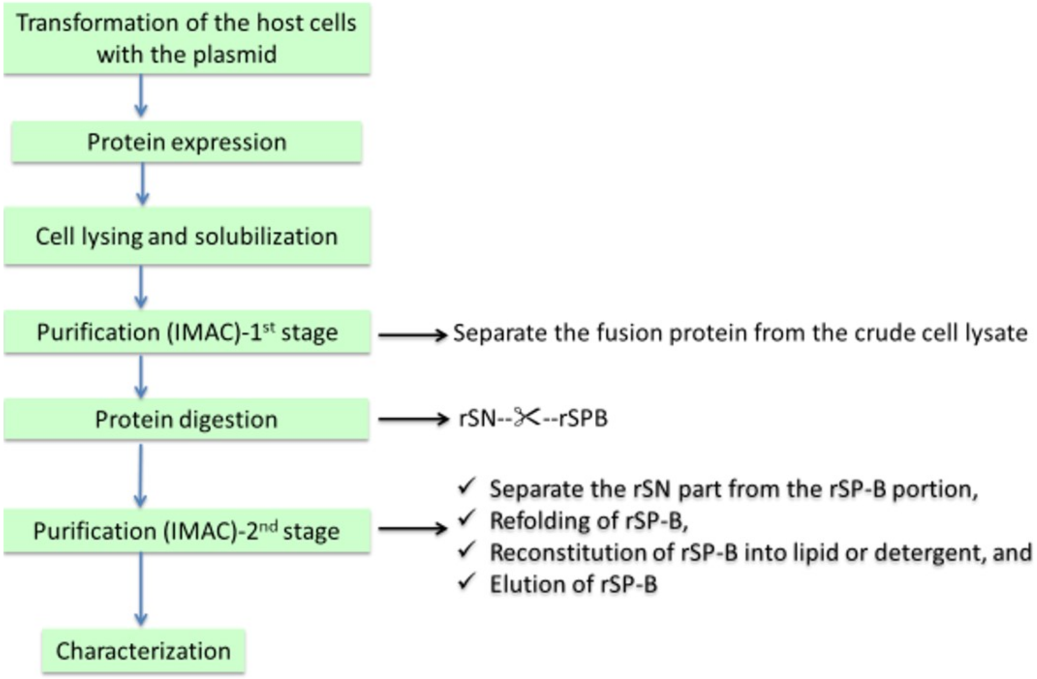


Figure 2.2: Schematic illustration of the production and characterization of SP-B. The purification of recombinant SP-B is carried in two stages. In the first stage of the purification, the fusion protein is separated from the crude cell lysate. Whereas, in the second stage, recombinant SP-B is separated from its fusion partner, refolded while being attached to the column and reconstituted into the desired environment before being eluted from the IMAC column. The characterization stage involves many techniques such as gel electrophoresis and immunoblottings, dynamic light scattering (DLS), circular dichroism (CD) and surface activity characterization of the protein with the Langmuir-Wilhelmey blodget (LWB) balance. These techniques for characterizing SP-B, especially DLS, CD and LWB will be further discussed in the proceeding chapters.

2.3.5 Immunoblotting

PVDF (polyvinyldene fluoride, GE Healthcare) membrane was soaked in methanol for 10 seconds and then was briefly rinsed in de-ionized water twice. The PVDF membrane and the SDS-PAGE gel to be transferred were equilibrated in 10mM of 100 ml CAPS (3-(cyclohexylamino)-1 propane sulfonic acid) transfer buffer for half an hour. Transferring of the protein bands from the gel to the PVDF membrane was performed by electrophoresing the stacking (the gel with the membrane) for 2 hour at 50 V in 100 mM CAPS transfer buffer. The membrane after electrophoresis was left in a blocking solution, 3 gm of casein milk powder (casein + BSA) plus 100 ml of 1XTTBS (50 mM Tris.HCl, pH 7.4, 150 mM NaCl, 0.1% Tween 20), for two hours. The membrane was incubated with primary antibody (6 µl of anti-SN/SP-B/His antibody plus 0.12 g of milk powder in 30 ml of 1xTTBS) overnight. To remove non-specific binding, the membrane was washed twice with 100 ml of 1XTTBS for 10 minute. Following, the membrane was probed with the secondary antibody (6 µl in 30 ml of 1xTTBS) for an hour. Again, the membrane was washed twice with 100 ml of 1XTTBS for 10 minutes and was left to equilibrate in $NaHCO_3$-buffer (0.1 M $NaHCO_3$, 1 mM $MgCl_2$). Color development was done in a solution of 20 ml of $NaHCO_3$, 60 µl BCIP (Fisher Reagents) and 120 µl of NBT (Fisher Reagents) in the dark. The reaction was allowed to go from 20 minutes to 60 minutes. Once the intensity of the band is strong enough, the reaction was stopped with TE buffer (10 mM TRIS, 1 mM EDTA, pH 8) and the blot was photographed. The blotted membrane was stored inside a sealing plastic bag for future reference.

2.4 Results

To confirm expression and to determine the portion of the C43 cell that contained rSP-B, the cells were harvested by centrifugation, suspended in 1XTBS, and lysed. The sample was then pelleted down via centrifugation and the soluble portion was retained. The remaining pellet was resuspended in 0.2% (w/v) CHAPS. Once more, the sample was pelleted by centrifugation and the pellet was resuspended in 0.2% (w/v) CHAPS and 6 M urea. The three samples were then analysed by non-reducing SDS-PAGE. The protein, rSN-SP-B, was not apparent in the Coommassie stained gel (Figure 2.3A). The western blot of the duplicate gel probed with anti-SP-B antibody (Figure 2.3B) indicated most of the protein lies in the 6 M urea, 0.2% (w/v) CHAPS solubilized sample, that is, in inclusion bodies (IBs). The band at ≈30 kDa in the western blot was in the expected mass range for the fusion protein, i.e. 25.6 kDa. The negative control, BSA, does not show on the immunoblot. Next, the focus was turned to determining if rSN-SP-B could be enriched from the cell lysate.

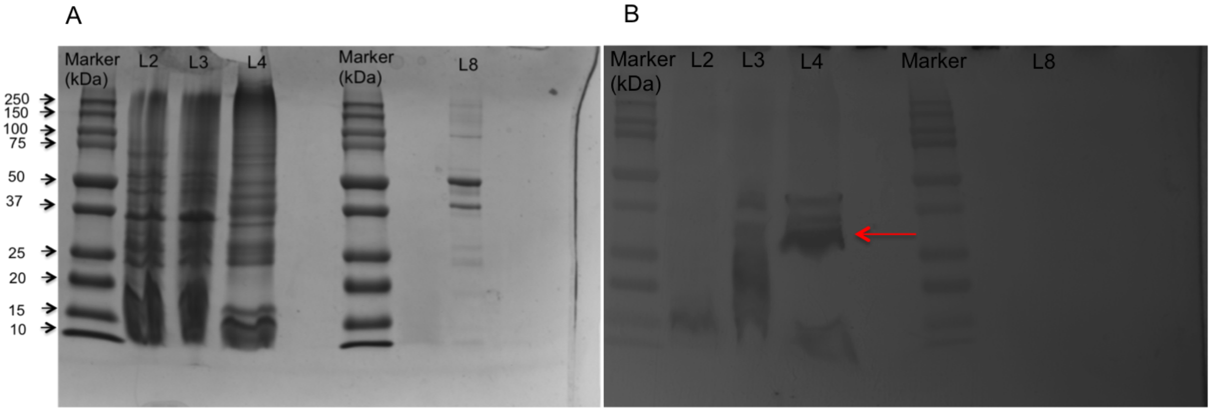


Figure 2.3: rSN-SP-B expresses mainly in inclusion bodies. Panel A: Non reducing 12% Tris-Glycine SDS-PAGE of french-pressed cells before purification (prep DJ-40). L2: Tris-buffered saline soluble portion of the cell lysate. L3: 0.2% (w/v) CHAPS soluble fraction of the pellet from L2. L4: 0.2% (w/v) CHAPS and 6 M urea soluble fraction of the pellet from L3. L8: 40 µl of bovine serum albumin (BSA). The gel was stained with Coomassie Brilliant Blue R-250 (Bio-Rad). Panel B: duplicate gel of A probed with polyclonal anti-SP-B rabbit antibody (Seven Hills Bio-reagents, Cincinnati, OH). L8 shows the BSA used as negative control does not show any reaction. The red arrow in L4 indicates the most intense band close to the expected mass of rSN-SP-B. Protein molecular-weight size standard: Kaleidoscope Precision Plus Protein^{TM} (Bio-Rad).

Elution of rSN-SP-B starts at low concentration of imidazole. The 12% Tris-Glycine SDS-PAGE in Figure 2.4 exhibits protein bands in the expected mass range for rSN-SP-B (25.6 kDa). The bands in L3 and L4 at ≈30 kDa and in L5 and L6 at ≈26 kDa were from fractions eluted with 20 mM and 50 mM imidazole respectively. These bands are in the expected mass range for rSN-SP-B. Unless indicated otherwise, all of the rSP-B preparations were continued via pooling of the 20-100 mM imidazole fractions together. To confirm at least one of the bands that appeared close to the expected molecular mass for rSN-SP-B indeed contained the target, rSN-SP-B was probed for either rSP-B or rSN with specific anti-SP-B/SN antibodies. Probing with antibodies can also confirm whether cyanogen bromide cleavage results in fragments of the expected masses or not.

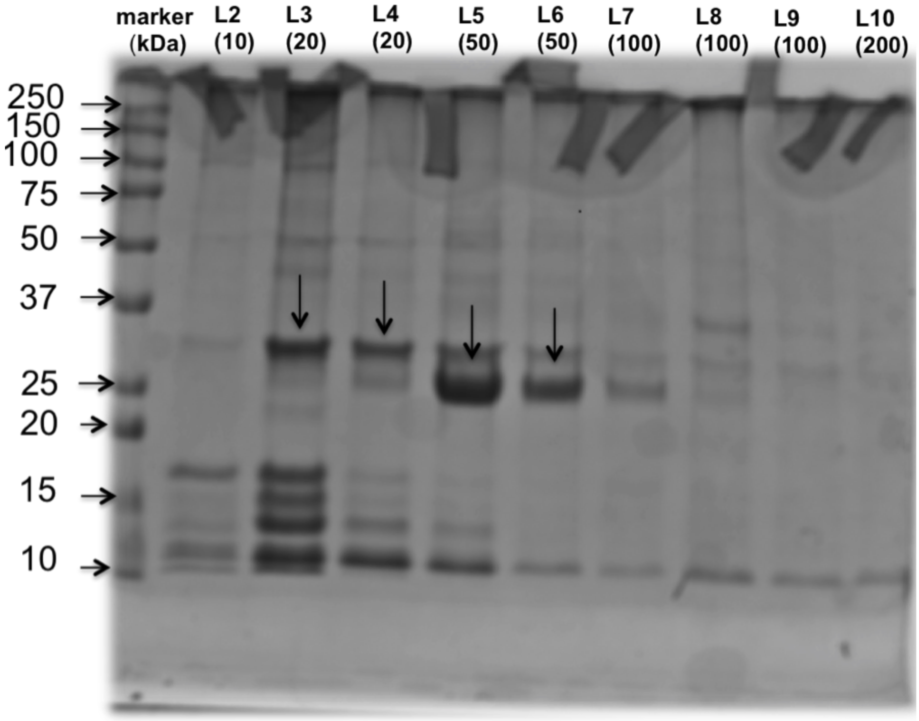


Figure 2.4: Elution of rSN-SP-B from the nickel column starts at low concentrations of imidazole. Non reducing 12% Tris-Glycine SDS-PAGE of purified rSN-SP-B (Prep. DJ-11). Numbers in parentheses are the concentration of imidazole in millimolar. The arrows indicate intense bands close to the expected mass of the rSN-SP-B protein at ≈30 kDa and 26 kDa. Normally all the fractions from 20 to 100 mM imidazole are pooled together for cyanogen bromide digestion. The molecular-weigh size marker used is Precision Plus Protein standard (Bio-Rad) and the gel was stained with Coomassie Brilliant Blue R-250 (Bio-Rad).

The western blot in L2 of Figure 2.5B displays a strong band for the fusion protein after incubating it with an anti-SP-B antibody. This result therefore confirms the expression of rSN-SP-B. It looks as though the antibody recognized the slightly higher molecular mass band of the two in the pooled fraction shown in the duplicate gel to the left. Likewise, the sample also shows a positive band when rSN-SP-B was probed with an anti-SN antibody (data not shown: it is more customary to show data that directly show when rSP-B, the target molecule of this study, is probed instead of proofing the presence of the molecule by probing for rSN). Next, rSN-SP-B was cleaved to separate the rSN chain from the rSP-B chain.

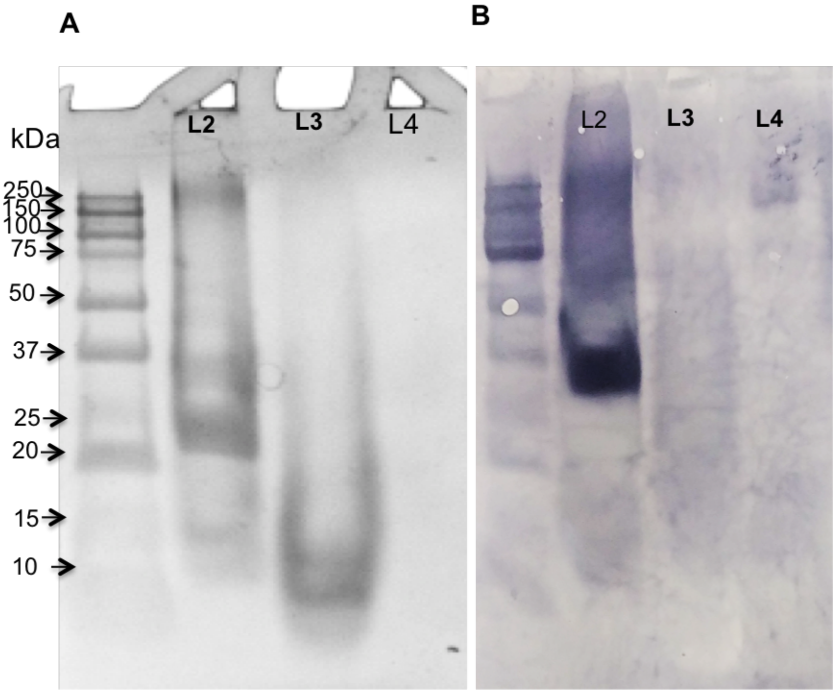


Figure 2.5: rSN-SP-B shows positive immunoblot when probed with polyclonal anti-SP-B antibody. Panel A: non reducing 12% Tris-Glycine SDS-PAGE of rSN-SP-B protein and its digest (prep. TG-13). L2: undigested rSN-SP-B suspended in the loading buffer (1XTBS, 6 M urea and 0.2% CHAPS, pH 7.5). L3: after digestion and lyophilization of the rSN-SP-B digest, the powder was resuspended in the same loading buffer as in L2. L4 is methanol (100%) extracted digest. Panel B: immunoblot of the duplicate gel of Panel A probed with polyclonal rabbit anti-SP-B antibody (Seven Hills Bioreagents, Cincinnati, OH) and polyclonal goat anti-rabbit (H+L) secondary antibody (Pierce, USA). The molecular-weight size marker used was Kaleidoscope Precision Plus Protein standard (Bio-Rad) and the gel was stained with Coomassie Brilliant Blue R-250 (Bio-Rad)

Cyanogen bromide (CNBr) cleaves at every methionine residue in polypeptides under highly acidic environments [66, 67]. As there was only one methionine in the expressed construct, this was expected to result in two polypeptide chains, one with rSN, and the other one with rSP-B plus the His-tag. The success of the cleavage was confirmed by 16% Tris-Tricine SDS-PAGE. Samples before (L2) and after cleavage (L3, L4) are shown in Figure 2.6A. The gel shows that the cyanogen bromide digestion resulted in the separate masses for rSN (≈14 kDa) and rSP-B at (≈8.5 kDa) as shown in L3 and L4. The reason why the rSN portion over-migrated might be attributed to the SDS and the gel sieve used. Different antibodies recognizing different SP-B epitopes were used to detect the fusion protein and recombinant SP-B after being cleaved from its fusion partner. Monoclonal antibodies, unlike polyclonal antibodies recognize a single SP-B epitope. Thus, for increased specificity, rSP-B were mainly probed with monoclonal antibodies. Also, to monitor cross-reactivity, blots were probed with only secondary antibodies and no reaction were detected. The bands in L1 and L2 of the immunoblot shown in Figure 2.6B indicated that rSP-B reacted positively to anti-SP-B monoclonal antibody (mAb) after cleavage. This result confirmed that rSP-B was present in the eluted fractions. The result also matches positively with the results observed for human SP-B by Lin *et al.* [68] and Guttentag *et al.* [69].

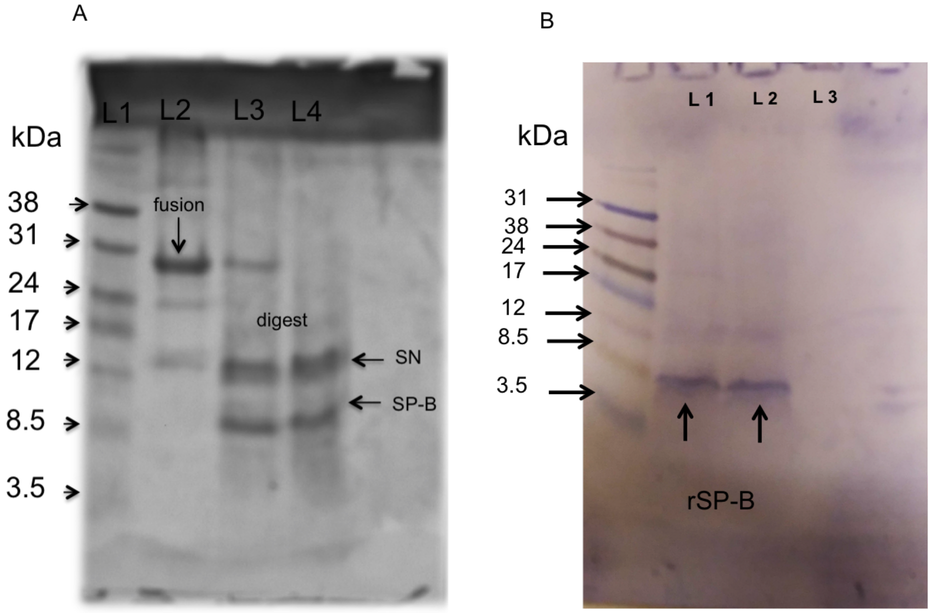


Figure 2.6: Digesting rSN-SP-B separates rSN from rSP-B. Panel A: non reducing 16% Tris-Tricine SDS PAGE of rSN-SP-B before and after cyanogen bromide digestion (Prep. May 24, 2013). L2 is rSN-SP-B before digestion suspended in the loading buffer (1XTBS, 6 M urea, 0.2% (w/v) CHAPS). L3 and L4 are the cyanogen bromide digested protein subjected to a Bligh and Dyer chloroform-methanol extraction with the organic soluble fraction loaded into L3 and the aqueous fraction into L4. The two bands in L3 and L4 are consistent with molecular mass of rSP-B (≈8.7 kDa) and SN (≈17 kDa). The gel was stained with a silver staining kit. Panel B: immunoblot of the rSN-SP-B digest transferred from a 16% Tris-Tricine SDS-PAGE (not a duplicate of panel A). L1 and L2 are lyophilized cyanogen bromide digest of rSN-SP-B re-suspended in Tris-Tricine sample buffer (200 mM Tris-HCl, pH 6.8, 40% glycerol, 2% SDS, 0.04% Coomassie Blue G-250) and applied in duplicate. L3 is the methanol soluble fraction of the lyophilized digest. The blot was probed with monoclonal anti-SP-B mouse antibody (Seven Hills Bioreagents, Cincinnati, OH) and goat anti-mouse secondary antibody (Seven Hills Bio-reagents, Cincinnati, OH). Both L1 and L2 show positive reaction at the expected mass for rSP-B monomer, ≈8.7 kDa. The molecular-weight size marker used is Rainbow Low-Range (GE Healthcare).

Another western blot was run using a sample from the same protein preparation, but this time, the samples were run on a Tris-Glycine gel and part of the sample was subjected to methanol extraction (Figure 2.7). The sample in L3 was the methanol soluble portion and in L4 was the insoluble portion dried under nitrogen gas stream and re-suspended in 6 M urea and 1XTBS. For reference, a Coomassie stained duplicate gel of the same samples applied on the western blot is shown in panel A. Compared to the duplicate gel, which showed two distinct bands for rSN and rSP-B after digestion, the blot in Figure 2.7B, was only reactive with the protein bands at the expected mass of rSP-B. These results demonstrated the successful cleavage of the rSN portion from rSP-B part.

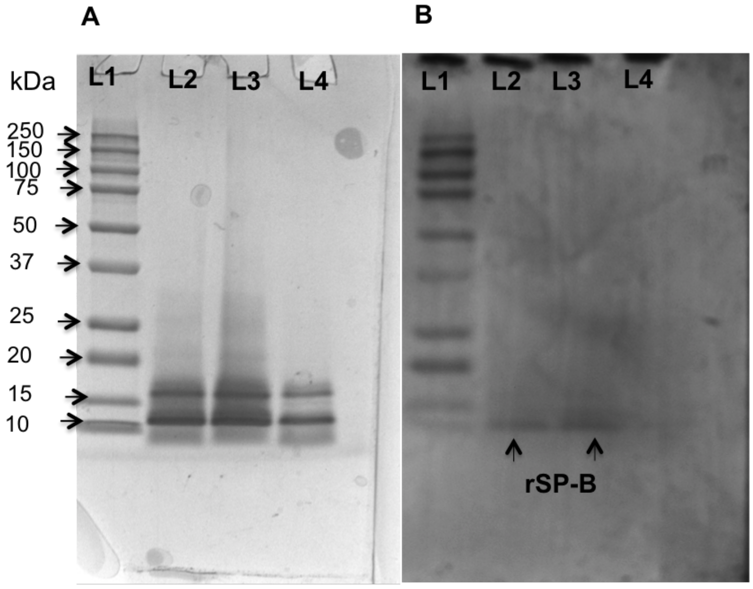


Figure 2.7: rSP-B shows positive immunoblot after digestion. Panel A: 12% Tris-Glycine non reducing SDS-PAGE of the fusion digest. L2: lyophilized digest of rSN-SP-B dissolved in Tris-Glycine sample buffer (31.5 mM Tris-HCl, pH 6.8, 10% glycerol, 1% SDS, 0.005% Bromophenol Blue). L3 is the methanol-soluble fraction of the lyophilized digest powder. L4 is the methanol insoluble fraction of L3 dried and subsequently resuspended in 1XTBS and 6 M urea. The gel was stained with Coomassie Brilliant Blue R-250 (Bio-Rad). Panel B: western blot of the duplicate gel of panel A probed with monoclonal anti-SP-B mouse antibody (Seven Hills Bio-reagents, Cincinnati, OH). The molecular-weight size marker used is Kaleidoscope Precision Plus Protein standard (Bio-Rad).

An important observation here was that rSP-B is dependent on the surrounding environment to react with anti-SP-B antibody after cleaving it from rSN. For instance, it was observed that in the presence of the detergent CHAPS, rSP-B does not react with either anti-SP-B mAb or polyclonal antibody (pAb). Both of the blots in Figure 2.6B and Figure 2.7B were therefore dissolved without CHAPS. On the other hand, it was observed that rSN-SP-B does interact with anti-SP-B pAb in the presence of CHAPS. This is demonstrated in Figure 2.5, which shows rSN-SP-B before and after fusion protein cleavage in the presence (L2 and L3) and absence (L4) of CHAPS. The strong band in L2 of the western blot (panel B) confirms that rSN-SP-B reacts with the anti-SP-B antibody in the presence of CHAPS, whereas L3, the lane that contains digested sample (rSN and rSP-B), did not show any reaction. A possible explanation could be if CHAPS bind the epitope recognized by the antibodies or induce conformational change in rSP-B so that the antigenic site(s) were inaccessible.

While I was successful in showing that rSP-B was present in IMAC fractions, the concentration of imidazole needed to elute most of rSN-SP-B i.e. 20-50 mM, was low compared to many other His-tagged proteins, which typically elute between 150 to 300 mM imidazole. Given how little rSN-SP-B is expressed, it is possible that I was missing some of the rSN-SP-B that may have eluted at higher imidazole fractions. I and other workers (Donna Jackman, Stella Barth and Selina Freudenberg) routinely ran SDS-PAGE gel using samples taken from fractions eluted with 150 to 300 mM imidazole but never detected a protein band corresponding to rSN-SP-B when stained with Coomassie Brilliant blue. Hence, we decided to proceed and run an immunoblot anyway.

Interestingly, it was found that all samples taken from 300 mM elution reacts to pAb specific to SP-B. The blot in Figure 2.8 demonstrates that rSN-SP-B exhibits different band sizes to suggests that the protein exhibits mixture of different masses in the loading buffer.

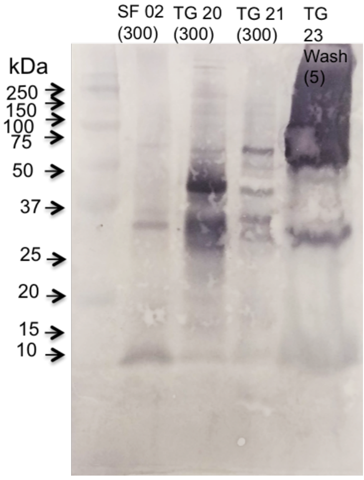


Figure 2.8: rSN-SP-B eluted with 300 mM imidazole does not stain with Coomassie Brilliant Blue R-250 (Bio-Rad) but shows a positive immunoblot when probed with polyclonal anti-SP-B rabbit antibody (Preps. SF-02, TG-20 and TG-21). The antibody used was from Seven Hills Bio-reagents, Cincinnati, OH. The molecular-weight size marker used is Kaleidoscope Precision Plus Protein standard (Bio-Rad).

The wash also exhibits bands, supporting the fact that, even after an overnight incubation with the IMAC resin in batch mode, binding of rSN-SP-B to the IMAC resins is not as efficient.

Furthermore, fractions collected with 20, 50 and 300 mM imidazole were electrophoresed in duplicate. One gel was stained with Coomassie and the other half was probed with anti-SP-B pAb as shown in Figure 2.9. Again, fractions eluted with 300 mM imidazole were not apparent on the Coomassie stained gel, whereas, samples eluted with 20 mM and 50 mM, displayed distinct bands at ≈ 50 kDa and 25 kDa respectively. The western blot, on the other hand was different in that a sample taken from the 300 mM elution show bands at ≈30 kDa, 37 kDa and 50 kDa. The pH of all fraction elutions are found to be within the buffer range ≈7.5 and hence the phenomenon must be driven by the surrounding environment rather than pH. Considering that these samples were exposed to CHAPS earlier in the protein preparation, it could be that the detergent plays a role, as others have found. Still, there is no clear explanation for why the 300 mM imidazole fractions do not react with Coomassie and this is a matter that needs more investigation.

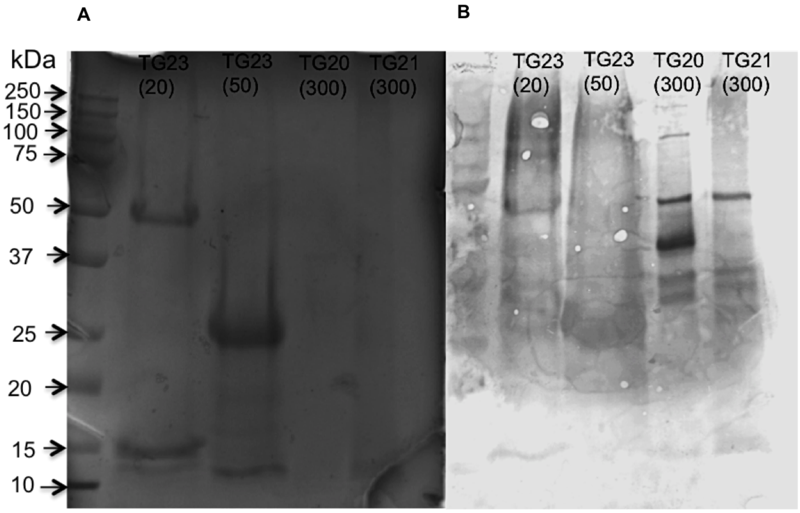


Figure 2.9: rSN-SP-B from three different preparations eluted with different imidazole concentrations (Prep. TG-20, TG-21 and TG-23). Panel A: non reducing 12% Tris-Glycine of samples eluted at 20 mM, 50 mM and 300 mM imidazole stained with Coomassie Brilliant Blue R-250 (Bio-Rad). The gel shows bands for the 20 mM and 50 mM eluted fractions. Panel B: western blot of of the duplicate gel of panel A incubated with anti-SP-B mouse polyclonal antibody (Seven Hills Bio-reagents, Cincinnati, OH). The immunoblot shows strong positive bands for the 300 mM imidazole fraction that are not evident on the coomassie stained gel. The molecular-weight size marker used is Kaleidoscope Precision Plus Protein standard (Bio-Rad).

Finally, the lyophilized dry digest was suspended in the next loading buffer (6 M urea, 0.5% CHAPS, 1XTBS, pH 7.5) and the rSN portion, with no His-tag, was flushed out by washing the column with the loading buffer. The protein was refolded on the column by decreasing the urea concentration and buffer exchange was performed by washing the column with the detergent/lipid buffer of choice. Following, rSP-B was eluted with an increment of imidazole concentration or decreased pH (pH 5). However, it was observed that enriching the protein by pH gradient was not as efficient as eluting the protein by an imidazole gradient. The protein was also observed in the fractions eluted by imidazole concentration after that same protein batch was eluted by lowering the pH. This might be due to the prior Tris-HCl wash, which may have removed all the salt present. Perhaps elution by decreased pH might be maximised by adding more salt. The drawback, however, is that, like imidazole, salt also interferes with the next step in characterization of rSP-B.

2.5 Discussion

In this chapter, I present the first ever successful attempt at expressing and purifying rSP-B in bacteria using a heterologous DNA expression system. Data presented were run three times to check for the reproducibility. This has been demonstrated by data shown in Figure 4,5,6,7 for which, samples were taken from different protein isolates as well as samples taken from aliquots of the same protein preparation that were produced by my self and co-workers Donna Jackman, Stella Barth and Selina Freudenberg (unpublished). Here, some of the key highlights of my findings and insights are discussed.

rSN-SP-B (Figure 2.1) expresses in *E. coli* strain C43 cells. Of the different *E. coli* strains tested, including BL21, origami, and C41, only C43 cells were able to overexpress rSN-SP-B. C43 cells have been found to produce excess intracellular membrane (ICM) rich in cardiolipin within their cytosol during over-production of some membrane proteins [63, 70]. In these studies, the ICM were asscociated with almost all of the overexpressed protein and could be isolated by single centrifugation. In light of SP-B's membranolytic and fusogenic nature [15, 71], formation of such membranes might be the reason for the cell's ability to express rSP-B.

rSN-SP-B seems to consists of multiple populations that vary in their ability to bind to the metal ions. It was observed, after an overnight incubation of rSN-SP-B with the resin, the elution of the protein seems to have two distinct ranges. There is a portion of the protein that elutes at a relatively low imidazole concentrations (20-50 mM) and also the other portion of it that elutes at higher imidazole concentrations (300 mM). This observation suggests a difference in the binding strength of the two rSN-SP-B populations to the metal ions. That is to say that a portion of the protein binds the metal ions weakly, while some of it also binds tightly. One potential reason why this could be is if the histidine-tags are not equally available for the metal ions. This is possible if the protein self-associates either by covalent bond or via weaker non-covalent interactions.

It is known that an SP-B monomer is stabilized by three intrachain disulfide bonds, while the dimer form is stabilized by one interchain disulfide bridge [72, 73]. Thus, even in high denaturing condition, proteins with disulfide linkage such as SP-B could still be intact through covalent bonding. Hence, there is a chance that the affinity tag can be occluded [45, 74]. To overcome this issue, a future direction could be purifying the protein after treating it with a reducing agent.

The other case is that rSN-SP-B might self-associate through non-covalent interactions. For example, studies have proposed the possibility of ion pairing (salt-bridges) between the monomer units of SP-B [73, 75]. Such a scenario could lead to a complex mixture of different oligomers in the binding buffer. Clearly, a dimer has more affinity tag than a monomer unit and likewise for higher order oligomers and thus oligomerization may contribute to increased binding strength. On the contrary, there is also the chance that the affinity tag can be hidden because of oligomerization. Both of these events could therefore lead to the difference in the binding affinity between the different populations. However, non-covalent dimerization is doubtful given the high denaturing media used, which disrupts weak protein-protein interactions between subunits.

Another possible explanation for the weak rSN-SP-B histidine binding is non-specific cleavage. A previous study demonstrated that some membrane proteins might undergo non-specific cleavage during expression, which might lead to the loss of the His-tag. Consequently, proteins without the tag could elute at low imidazole concentration, while the protein with the tag elutes at high imidazole concentration [76]. Given that rSN has four histidines, it is likely, especially in the denaturing condition that some are exposed at the surface for binding. On the other hand, the fact that I was able to successfully separate the cleaved rSN from rSP-B using the rSN-SP-B population eluted at low imidazole concentration during the first purification suggests that most of the affinity tag is intact.

Yet a different potential explanation could be the detergents used to solubilize and stabilize the protein. Often times, detergents used to solubilize membrane proteins also bind the protein and the deteregent micelles may cover part or all of the His-tag [47, 77, 78] and restrict access of the His-tag to the ligand. Thus, at this point, it is not clear why rSP-B elutes at such a broad range of imidazole concentrations. To try circumvent this issue and promote more uniform binding, the His-tag was moved to the N-terminus but to no avail as the protein did not express. Another way of troubleshooting this problem is to optimize the affinity tag itself.

Increasing the length of the affinity tag has been shown to improve the binding affinity and expression level. A study by Lee *et al.* [79] demonstrated that an increase in the length of the affinity tag to 8 or 10 histidine residues increased the binding affinity of the target to the metal ligands. This is especially useful for low expressing proteins as it may allow washing with high salt buffers, which is known to promote an increase in the purity of the target. Interestingly, Lee *et al.* designed a T7 pET expression vector suitable to modify the length of the affinity tag at the C-terminus. More importantly, from an rSP-B point of view, the vector designed by the Lee group was found to successfully express in BL21(DE3) cells. In connection with the amount of protein expression and length of the His-tag, however, the data in the literature are not clearcut. A study reported that 6X-His-tag gave better yield than 10X-His-tag [80], whereas an earlier study reported improved yield for the 10X-His-tag [77]. However, both of these studies were done on a different membrane protein, hence it could be that the length of the His-tag and the amount of expression could be protein specific.

Purifying rSP-B under both denaturing and reducing conditions may improve binding. Though, the purification was carried out under highly denaturing condition (6 M urea), it is also imperative to consider the three intrachain disulphide bridges of rSP-B, which are likely to keep the monomer unit of the protein connected in a non-reducing environment. In addition to this, SP-B is also known to be stabilized by an interchain disulphide bridge. Hence, the His-tag on rSP-B might not be as readily accessible as denatured soluble proteins or membrane proteins without disulphide linkage would be. Perhaps purifying rSP-B under both denaturing and reducing conditions should be considered. Although IMAC is amenable to low concentrations of reducing agents [44, 81], high amounts have been found to adversely affect purification of proteins with IMAC. DTT, for instance, is tolerable up to 10 mM, but the resin has been shown to change color to brown due to the reduction of the nickel ions by the reducing agent. TCEP is odorless, potent, selectively reduces disulphide bonds, and is recommended up to concentrations of 5 mM for Ni-IMAC [47]. The difficulty here is that successfully reducing rSP-B in the presence of the detergent used to solubilize the protein could be challenging. Re-oxidization of the cysteines is also an issue to think about. Thus, careful choice of reducing agent and an optimum concentration is mandatory. Besides, as said, the nature of the detergent used to solubilize rSP-B may also hinder access of affinity tag by the ligands. Therefore, the choice of detergent is also crucial for the success of the purification.

Both rSN-SP-B and rSP-B require CHAPS to bind the nickel ions. Interestingly, binding was not successful in the presence of other detergents or lipids in place of CHAPS or absence of any lipid or detergent. In such cases, rSP-B was found in the flow through. This is interesting in its own right as it depicts some form of interaction between CHAPS and rSP-B, which allows the rSP-B/CHAPS complex to expose the His-tag to the ligands that does not seem to occur when rSP-B is alone or is associated with other detergent/lipids. Studies indicate that CHAPS increases the propensity of native SP-B to self-associate [82]. Hence, the zwitterionic detergent may mediate conformational flexibility of rSP-B in a way to make the His-tag accessible. In a separate study, the C-terminal segment of SP-B was also observed to expand when in SDS detergent micelles [83]. Hence, similar to other SAPLIPs, SP-B may also exhibit a change in conformation when in detergents. To further investigate the role of CHAPS, we also probed rSN-SP-B and rSP-B with both polyclonal and monoclonal anti-SP-B antibodies with and without CHAPS.

Anti-SP-B mAb and pAb do not recognize rSP-B in the presence of CHAPS but react to rSN-SP-B with CHAPS present (Figure 2.7). It could be that either CHAPS binds the epitope recognized by the antibodies or induces an assembly of rSP-B in a way that the antigenic site(s) are hidden from the antibodies. Conversely, the fusion protein rSN-SP-B does bind anti-SP-B pAb in the presence of CHAPS. Here, the reason why the anti-SP-B pAb binds rSN-SP-B is not clear, but two potential reasons are suspected. One could be that, since the pAb recognizes different sites on rSP-B, it stands a better chance of binding than the mAb. Second to that, the presence of rSN may influence the ability of rSP-B to undergo conformational change and hence the epitopes might be better exposed. Another aspect of rSP-B that merits further investigation is the heterogeneity of apparent molecular masses observed in SDS-PAGE.

rSN-rSP-B exhibits different sizes in the binding buffer. The blot in Figure 2.8 shows bands at different sizes of rSN-SP-B. The observation is in keeping with studies of native SP-B extracted from bovine and ovine sources, which have been found to adopt different oligomeric state depending on the method of protein preparation and type of solvents used for the solubilization [84–87]. However, it is also imperative to consider that the binding buffer is highly denaturing and hence affects weak interactions. Whether the different masses observed are due to covalent bond or weak interaction is a matter that needs further study. The ≈50 kDa and 25 kDa bands observed in the SDS-PAGE in Figure 2.9 could be dimers and monomers of rSN-SP-B respectively. The bands that appear in all lanes at ≈10 kDa in Figure 2.4 could be either a histidine rich bacterial protein or rSP-B on its own. It could be that C43 (DE3) may have some endogenous enzymes to proteolytically cleave at the transcript. The author would like to point that due to lack of slaughter houses in the island of Newfoundland and Labrador, Canada, especially in St.John's area during time course of this study, it was difficult for us to get animal lungs to extract and purify native SP-B. Native SP-B would have been a good positive control for immunoblotting detection of recombinant SP-B. Such experiments would add value by providing a point of comparison. Experiments presented in this chapter are repeated a minimum of three times. Results were found to be consistent with samples taken from different isolates as well as aliquots taken from the same sample.

To sum up, herein, I report the successful expression of rSP-B in bacteria for the first time. It was found that denatured rSP-B renatures while on column and some membrane mimetic conditions support rSP-B native conformation as deduced from the CD scans (as will be discussed in chapter 3), which depicts comparable percent helicity to animal derived SP-B. More importantly, *in vitro* measure of rSP-B function demonstrate an activity comparable to native SP-B and clinical surfactant (as will be discussed in chapter 4).

2.6 References

1. Lukovic, D., Plasencia, I., Taberner, F. J., Salgado, J., Calvete, J. J., Pérez-Gil, J. & Mingarro, I. Production and characterisation of recombinant forms of human pulmonary surfactant protein C (SP-C): structure and surface activity. *Biochimica et Biophysica Acta* **1758,** 509–518 (2006).

2. Moulton, J. *Expression and functional studies of a recombinant surfactant protein B.* Masters Thesis (Memorial University of Newfoundland, 2013), 1–121.

3. Hakes, D. J. & Dixon, J. E. New vectors for high level expression of recombinant proteins in bacteria. *Analytical Biochemistry* **202,** 293–298 (1992).

4. Sharifahmadian, M., Sarker, M., Palleboina, D., Waring, A. J., Walther, F. J., Morrow, M. R. & Booth, V. Role of the N-terminal seven residues of surfactant protein B (SP-B). *PloS ONE* **8,** 1–8 (2013).

5. Esposito, D. & Chatterjee, D. K. Enhancement of soluble protein expression through the use of fusion tags. *Current Opinion in Biotechnology* **17,** 353–358 (2006).

6. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances* **27,** 297–306 (2009).

7. Kaur, J., Kumar, A. & Kaur, J. Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements. *International Journal of Biological Macromolecules* **106,** 803–822 (2018).

8. Sezonov, G., Joseleau-Petit, D. & D'Ari, R. Escherichia coli physiology in Luria-Bertani broth. *Journal of Bacteriology* **189,** 8746–8749 (2007).

9. Riesenberg, D. & Guthke, R. High-cell-density cultivation of microorganisms. *Applied Microbiology and Biotechnology* **51,** 422–430 (1999).

10. Choi, J. H., Keum, K. C. & Lee, S. Y. Production of recombinant proteins by high cell density culture of Escherichia coli. *Chemical Engineering Science* **61,** 876–885 (2006).

11. Berlec, A. & trukelj, B. Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells. *Journal of Industrial Microbiology & Biotechnology* **40,** 257–274 (2013).

12. Bird, L. E. High throughput construction and small scale expression screening of multi-tag vectors in Escherichia coli. *Methods* **55,** 29–37 (2011).

13. Costa, S., Almeida, A., Castro, A. & Domingues, L. Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: The novel Fh8 system. *Frontiers in Microbiology* **5,** 1–20 (2014).

14. Overton, T. W. Recombinant protein production in bacterial hosts. *Drug Discovery Today* **19,** 590–601 (2014).

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

16. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology* **5,** 1–17 (2014).

17. Doherty, A. J., Connolly, B. A. & Worrall, A. F. Overproduction of the toxic protein, bovine pancreatic DNaseI, in Escherichia coli using a tightly controlled T7-promoter-based vector. *Gene* **136,** 337–340 (1993).

18. Miroux, B. & Walker, J. E. Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol* **260,** 289–298 (1996).

19. Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Gbom, M. H., van Wijk, K. J., Slotboom, D. J., Persson, J. O. & de Gier, J.-W. Tuning Escherichia coli for membrane protein overexpression. *Proceedings of the National Academy of Sciences* **105,** 14371–14376 (2008).

20. Dumon-Seignovert, L., Cariot, G. & Vuillard, L. The toxicity of recombinant proteins in Escherichia coli: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification* **37,** 203–206 (2004).

21. Ramón, A., Señorale-Pose, M. & Marín, M. Inclusion bodies: Not that bad... *Frontiers in Microbiology* **5,** 2010–2015 (2014).

22. Gonzalez-Montalban, N., Natalello, A., García-Fruitós, E., Villaverde, A. & Doglia, S. M. In situ protein folding and activation in bacterial inclusion bodies. *Biotechnology and Bioengineering* **100,** 797–802 (2008).

23. Miot, M. & Betton, J. M. Protein quality control in the bacterial periplasm. *Microbial Cell Factories* **3,** 1–13 (2004).

24. Murby, M., Samuelsson, E., Nguyen, T. N., Mignard, L., Power, U., Binz, H., Uhlén, M. & Ståhl, S. Hydrophobicity engineering to increase solubility and stability of a recombinant protein from respiratory syncytial virus. *European Journal of Biochemistry* **230,** 38–44 (1995).

25. Cunningham, F. & Deber, C. M. Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods* **41,** 370–380 (2007).

26. Rinas, U. & Bailey, J. E. Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli. *Applied Microbiology and Biotechnology* **37,** 609–614 (1992).

27. Vera, A., Arís, A., Carrió, M., González-Montalbán, N. & Villaverde, A. Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies. *Journal of Biotechnology* **119,** 163–171 (2005).

28. Carrió, M. M., Cubarsi, R. & Villaverde, A. Fine architecture of bacterial inclusion bodies. *FEBS Letters* **471,** 7–11 (2000).

29. Cubarsí, R., Carrió, M. M. & Villaverde, A. In situ proteolytic digestion of inclusion body polypeptides occurs as a cascade process. *Biochemical and Biophysical Research Communications* **282,** 436–441 (2001).

30. Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D. & Riek, R. Bacterial inclusion bodies contain amyloid-like structure. *PloS Biology* **6,** 1791–1801 (2008).

31. Wang, L. Towards revealing the structure of bacterial inclusion bodies. *Prion* **3,** 139–145 (2009).

32. Rambaran, R. N. & Serpell, L. C. Amyloid fibrils. *Prion* **2,** 112–117 (2008).

33. Middelberg, A. P. J. Preparative protein refolding. *Trends in Biotechnology* **20,** 437–443 (2002).

34. Sørensen, H. P. & Mortensen, K. K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial Cell Factories* **4,** 1–8 (2005).

35. Singh, S. M. & Panda, A. K. Review: Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering* **99,** 303–310 (2005).

36. Deuerling, E., Patzelt, H., Vorderwülbecke, S., Rauch, T., Kramer, G., Schaffitzel, E., Mogk, A., Schulze-Specking, A., Langen, H. & Bukau, B. Trigger factor and DnaK possess overlapping substrate pools and binding specificities. *Molecular Microbiology* **47,** 1317–1328 (2003).

37. Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevaux, P., Agostini, F., Tartaglia, G. G., Hayer-Hartl, M. & Hartl, F. U. DnaK functions as a central hub in the E. coli chaperone network. *Cell Reports* **1,** 251–264 (2012).

38. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H. & Yura, T. Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli. *Applied and Environmental Microbiology* **64,** 1694–1699 (1998).

39. Waugh, D. S. Making the most of affinity tags. *Trends in Biotechnology* **23,** 316–320 (2005).

40. Costa, S. J., Almeida, A., Castro, A., Domingues, L. & Besir, H. The novel Fh8 and H fusion partners for soluble protein expression in Escherichia coli: A comparison with the traditional gene fusion technology. *Applied Microbiology and Biotechnology* **97,** 6779–6791 (2013).

41. Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification* **48,** 1–13 (2006).

42. Magnusdottir, A., Johansson, I., Dahlgren, L. G., Nordlund, P. & Berglund, H. Enabling IMAC purification of low abundance recombinant proteins from E. coli lysates. *Nature Methods* **6,** 477–478 (2009).

43. Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. His-tag impact on structure. *Acta Crystallographica Section D: Biological Crystallography* **63,** 295–301 (2007).

44. Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H. & Maertens, B. Purification of His-tagged proteins. *Methods in Enzymology* **559,** 1–15 (2015).

45. Bornhorst, J. A. & Falke, J. J. Purification of proteins using polyhistidine affinity tags. *Methods in Enzymology* **326,** 245–254 (2000).

46. Stevens, T. J. & Arkin, I. T. Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins: Structure, Function, and Genetics* **39,** 417–420 (2000).

47. Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J. & Schäfer, F. Immobilized-metal affinity chromatography (IMAC): A review. *Methods in Enzymology* **463,** 439–473 (2009).

48. Melnyk, R. A., Partridge, A. W. & Deber, C. M. Retention of native-like oligomerization states in transmembrane segment peptides: Application to the Escherichia coli aspartate receptor. *Biochemistry* **40,** 11106–11113 (2001).

49. Hanafy, N., El-Kemary, M. & Leporatti, S. Micelles structure development as a strategy to improve smart cancer therapy. *Cancers* **10,** 238 (2018).

50. Jana, S. & Deb, J. K. Strategies for efficient production of heterologous proteins in Escherichia coli. *Applied Microbiology and Biotechnology* **67,** 289–298 (2005).

51. Gopal, G. J. & Kumar, A. Strategies for the production of recombinant protein in escherichia coli. *Protein Journal* **32,** 419–425 (2013).

52. Schein, C. H. & Noteborn, M. H. M. Formation of soluble recombinant proteins in escherichia coli is favored by lower growth temperature. *Nature Biotechnology* **6,** 291–294 (1988).

53. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry* **307,** 249–264 (2008).

54. Chesshyre, J. A. & Hipkiss, A. R. Low temperatures stabilize interferon α-2 against proteolysis in Methylophilus methylotrophus and Escherichia coli. *Applied Microbiology and Biotechnology* **31,** 158–162 (1989).

55. Spiess, C., Beil, A. & Ehrmann, M. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97,** 339–347 (1999).

56. Baneyx, F. & Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. *Nature Biotechnology* **22,** 1399–1407 (2004).

57. García-Fraga, B., da Silva, A. F., López-Seijas, J. & Sieiro, C. Optimized expression conditions for enhancing production of two recombinant chitinolytic enzymes from different prokaryote domains. *Bioprocess and biosystems engineering* **38,** 2477–2486 (2015).

58. Broedel, S. E., Papciak, S. M. & Jones, W. R. The selection of optimum media formulations for improved expression of recombinant proteins in E . coli. *Technical Bulleti-Athena Enzyme SystemsTM Group* **2,** 8 (2001).

59. Yang, Q.-H., Wu, C.-L., Lin, K. & Li, L. Low concentration of inducer favors production of active form of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase in Escherichia coli. *Protein expression and purification* **10,** 320–324 (1997).

60. Chan, V., Dreolini, L. F., Flintoff, K. a., Lloyd, S. J. & Mattenley, A. a. The effects of glycerol , glucose , galactose , lactose and glucose with galactose on the induction of β-galactosidase in Escherichia coli. *Journal of Experimental Microbiology and Immunology* **2,** 130–137 (2002).

61. Mansey, M. S., Ghareeb, K. A., Moghazy, A. N., Tawfick, M. M., Fouda, M. M., El Marzugi, N. A., Othman, N. Z. & El Enshasy, H. A. Glucose concentration affects recombinant interferon α-2b production in Escherichia coli using thermo-induction system. *Journal of Applied Pharmaceutical Science* **4,** 1–5 (2014).

62. Sørensen, H. P. & Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology* **115,** 113–128 (2005).

63. Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., De Kruijff, B., Runswick, M. J. & Walker, J. E. Characterisation of new intracellular membranes in Escherichia coli accompanying large scale over-production of the b subunit of F1Fo ATP synthase. *FEBS Letters* **482,** 215–219 (2000).

64. Olmeda, B., Garca-Alvarez, B. & Perez-Gil, J. Structure-function correlations of pulmonary surfactant protein SP-B and the saposin-like family of proteins. *European Biophysics Journal* **42,** 209–222 (2013).

65. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157,** 105–132 (1982).

66. Andreev, Y. A., Kozlov, S. A., Vassilevski, A. A. & Grishin, E. V. Cyanogen bromide cleavage of proteins in salt and buffer solutions. *Analytical Biochemistry* **407,** 144–146 (2010).

67. Villa, S., De Fazio, G. & Canosi, U. Cyanogen bromide cleavage at methionine residues of polypeptides containing disulfide bonds. *Analytical Biochemistry* **177,** 161–164 (1989).

68. Lin, S., Akinbi, H. T., Breslin, J. S. & Weaver, T. E. Structural requirements for targeting of surfactant protein B (SP-B) to secretory granules in vitro and in vivo. *Journal of Biological Chemistry* **271,** 19689–19695 (1996).

69. Guttentag, S. H., Beers, M. F., Bieler, B. M. & Ballard, P. L. Surfactant protein B processing in human fetal lung. *American Journal of Physiology - Lung Cellular and Molecular Physiology* **275,** L559–L566 (1998).

70. Carranza, G., Angius, F., Ilioaia, O., Solgadi, A., Miroux, B. & Arechaga, I. Cardiolipin plays an essential role in the formation of intracellular membranes in Escherichia coli. *Biochimica et Biophysica Acta* **1859,** 1124–1132 (2017).

71. Hawgood, S., Derrick, M. & Poulain, F. Structure and properties of surfactant protein B. *Biochimica et Biophysica Acta* **1408,** 150–160 (1998).

72. Johansson, J. & Curstedt, T. Molecular structures and interactions of pulmonary surfactant components. *Eur. J. Biochem* **244,** 675–693 (1997).

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

74. Cheung, R. C. F., Wong, J. H. & Ng, T. B. Immobilized metal ion affinity chromatography: A review on its applications. *Applied Microbiology and Biotechnology* **96,** 1411–1420 (2012).

75. Zaltash, S., Palmblad, M., Curstedt, T., Johansson, J. & Persson, B. Pulmonary surfactant protein B: a structural model and a functional analogue. *Biochimica et Biophysica Acta* **1466,** 179–186 (2000).

76. Jormakka, M. & Byrne, B. *Solubilization and purification of membrane proteins* in *Structural Genomics on Membrane Proteins* (ed Lundstrom, K. H.) 179–198 (CRC Press, 2006).

77. Grisshammer, R. & Tucker, J. Quantitative evaluation of neurotensin receptor purification by immobilized metal affinity chromatography. *Protein Expression and Purification* **11,** 53–60 (1997).

78. Grisshammer, R. Purification of recombinant G-protein-coupled receptors. *Methods in Enzymology* **463,** 631–645 (2009).

79. Lee, J. & Kim, S. H. High-throughput T7 LIC vector for introducing C-terminal polyhistidine tags with variable lengths without extra sequences. *Protein Expression and Purification* **63,** 58–61 (2009).

80. Mohanty, A. K. & Wiener, M. C. Membrane protein expression and production: Effects of polyhistidine tag length and position. *Protein Expression and Purification* **33,** 311–325 (2004).

81. Kimple, M. E., Brill, A. L. & Pasker, R. L. Overview of affinity tags for protein purification. *Current Protocols in Protein Science* **73,** 1–26 (2013).

82. Olmeda, B., Garcia-Alvarez, B., Gomez, M. J., Martinez-Calle, M., Cruz, A. & Perez-Gil, J. A model for the structure and mechanism of action of pulmonary surfactant protein B. *The FASEB Journal* **29,** 1–12 (2015).

83. Booth, V., Waring, A. J., Walther, F. J. & Keough, K. M. NMR structures of the C-terminal segment of surfactant protein B in detergent micelles and hexafluoro-2-propanol. *Biochemistry* **43,** 15187–15194 (2004).

84. Baatz, J. E., Zou, Y., Cox, J. T., Wang, Z. & 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,** 180–190 (2001).

85. Baatz, J. E., Elledge, B., Whitsett, J. A., John E. Baatz, Barry Elledge & Whitsett*, J. A. Surfactant protein SP-B induces ordering at the surface of model membrane bilayers. *Biochemistry* **29,** 6714–6720 (1990).

86. Wustneck, N., Wustneck, R., Perez-Gil, J. & Pison, U. Effects of oligomerization and secondary structure on the surface behavior of pulmonary surfactant proteins SP-B and SP-C. *Biophysical Journal* **84,** 1940–1949 (2003).

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

3

Characterizing rSP-B using Circular Dichroism, Dynamic Light Scattering and Nanoparticle Tracking Analysis

3.1 Statement of Authorship

Tadiwos Getachew prepared the samples used for analysis, measured, analyzed and interpreted the data, and also wrote the manuscript. Donna Jackmann trained Caroline Stange and Tadiwos Getachew used a sample prepared by Caroline Stange and Donna Jackmann for analysis. Tadiwos Getachew trained Stella Barth and used a sample prepared by Stella Barth for analysis.

[1]

---

[1] In figure captions and legends of this chapter abbreviations can be found. These are used to represent the persons who prepared the protein samples used for analysis. DJ: Donna Jackmann, SB: Stella Barth, CS: Caroline Stange and TG: Tadiwos Getachew. The numbers following the abbreviations are protein isolation batch numbers.

3.2 Background

Most *in vitro* studies of membrane proteins involve the use of detergents to solubilize the protein and maintain it in a functional form. A well folded, functional, and homogeneous protein is a prerequisite for most structural studies. This chapter uses circular dichroism (CD) to characterize the secondary structure of rSP-B in different environments. Furthermore, dynamic light scattering (DLS) and nanoparticle tracking analysis (NTA) were used to measure the size of rSP-B-detergent/lipid complexes.

3.2.1 Circular Dichroism (CD)

CD has been an invaluable tool to study the secondary structure of proteins for a long while now. Circulary polarized light contains two orthogonal field components that are 90 degrees out of phase. Optically active asymmetric molecules, also called as chiral molecules, absorb left and right handed circularly polarized light to a different extent. This differential absorption is what CD takes advantage of. Proteins are chiral molecules and exhibit different levels of circular dichroism at different wavelengths (λ). While peptide bonds absorb below 240 nm, i.e. the far-UV (180-250 nm) region, the side chains of aromatic residues (tryptophan, phenylalanine and tyrosine) absorb in the near-UV (250-400 nm) region. Hence, based on the surrounding environment of the aromatic chromophores, the near-UV CD of macromolecules also provides tertiary structural information [1]. Disulphide bonds also absorb close to 260 nm.

The far-UV CD (180-250 nm) spectrum is a fingerprint of the secondary structure of biomolecules. Figure 3.1 depicts characteristic features of the common secondary structures of proteins. The CD spectra of α-helices, anti-parallel β-structure, disordered proteins, as well as the triple-helix of collagen in both native and denatured forms are presented. An alpha helical protein gives a CD spectrum with two chacteristic minima at 222 nm and 208 nm and a maximum at 192 nm. Full α-helical proteins (i.e. proteins that only adopt α-helix structure) should exhibit equal magnitudes of the two negative peaks at 222 and 208 nm. Likewise, β-helix proteins show a negative peak at 218 nm and a maximum at 195 nm. On the other hand, disordered proteins exhibit a negative peak at 195 nm and a slight positive peak at 210 nm [2]. Hence, CD provides insight into the secondary structure of proteins, although, the technique cannot provide atomic detail structural information like NMR and X-ray crystallography. Thus, CD is especially valuable for proteins with unknown structure or proteins that are difficult to determine the structure of using the other techniques. In addition, folding behaviour, changes in stability due to mutations, and binding of proteins can be well characterized by CD [3]. Therefore, CD is particularly suited to characterize the secondary structure of rSP-B.

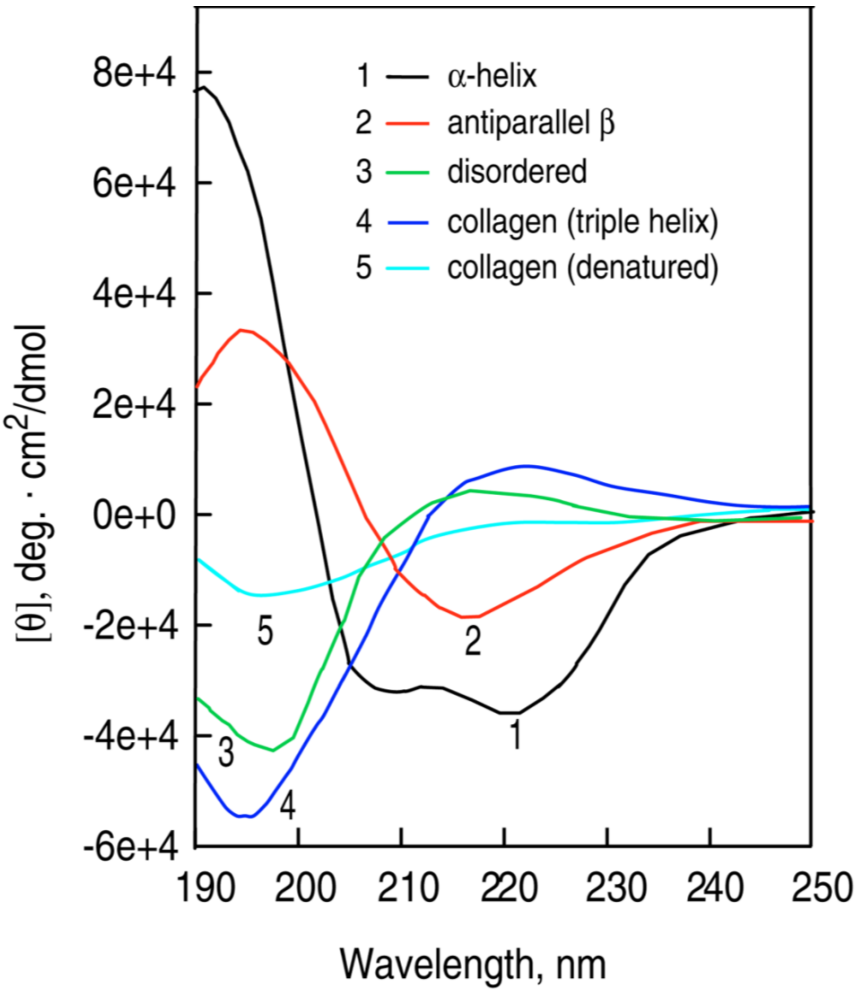


Figure 3.1: Characteristic circular dichroism spectra of an α-helical protein (black), antiparallel β-pleated sheet (red), disordered protein (green) and triple-helix (blue) and denatured (cyan) forms of collagen. Picture was adopted from Greenfield [2] with permission from the publisher.

The challenge with using CD to obtain the fraction of each type of secondary structure is determining the correct protein concentration. The best method of determining accurate protein concentration is quantitative amino acid analysis [4]. Other protein quantification techniques such as colorimetric assays can also give reliable results; however, some inconsistencies have been reported [2]. With cuvettes of different path lengths, protein samples with concentrations ranging from 0.05 to 5 mg/ml can be measured for characterizing secondary structure. Given the accurate concentration, the technique is a rapid way of analysing biomolecules in solution and with a minimal amount of sample. Also advantageously, measurements can be done in a wide range of pH and temperature conditions. However, for best results, samples for CD must be at least 95% pure. Usually, samples must be dialysed and filtered (0.1 to 0.2 µm pore filter) prior to CD measurement. The dialysis removes optically active material that absorbs in the far-UV region and filtering reduces light scattering [2].

Dialysis of a CD sample is important to obtain quality CD spectrum. Imidazole at concentrations above 100 mM absorbs strongly in the far-UV region. In the case of His-tagged proteins eluted using an imidazole gradient, extensive dialysis is thus important to remove the imidazole prior to CD measurement. Ideally, the concentration of imidazole for best CD results should be $\leqslant$1 mM. The same is true for chaotropic agents, such as urea and guanidium chloride, that absorb strongly below 210 nm [3]. What is more, detergents used to solubilize membrane proteins should not absorb strongly in the far-UV region. Careful choice of buffers and detergents are therefore mandatory for CD analysis.

The nature of the detergent(s)/buffers matters. Some alkyl glycoside based detergents such as lauryl maltoside absorb far less in the far-UV region and so are suitable for CD measurenment. However, these detergents are expensive. Most often, such buffers are exchanged into the protein after protein solubilization and purification. Detergents like Triton X-100 could be difficult to remove from proteins and absorb strongly at 280 nm and thus interfere with protein quantification. Sodium dodecyl sulfate (SDS) on the other hand, does not absorb in the far-UV region and is also a good detergent in which to study secondary structure of polypeptides [3]. Besides detergents, some organic compounds were found to increase proteins' propensity to form secondary structure, among which triflouroethanol (TFE), hexaflouroisopropanol (HFIP), ethyleneglycol, and glycerol are commonly used [1]. Of these, TFE is the best characterized [5, 6].

CD is more accurate for measuring the secondary structure of α-helical proteins than β-sheet proteins [2, 7]. The amide chromophores of peptides undergo a $\eta \rightarrow \Pi^*$ transition that is electrically forbidden, but magnetically allowed, as well as $\Pi \rightarrow \Pi^*$ transitions that are electrically permitted but magnetically disallowed [8]. While the $\eta \rightarrow \Pi^*$ transition is primarily responsible for the 222 nm signal in an α- helical protein and the 218 nm in β-pleated sheet structures, the $\Pi \rightarrow \Pi^*$ transition is mostly responsible for the intense positive band around 190 nm of an α- helical protein [3]. These transitions are influenced by the backbone geometry of the protein. The signal harnessed from such transition is therefore indicative of the ψ and ϕ backbone dihedral angles and thus the structure of the molecules. Different algorithms have been developed to deconvolute the various pieces of information contained in the spectrum, which are then combined together as a measure of the overall secondary structure of the protein. Some of the most widely used algorithms are SELCON, K2D2, CONTIN, CDSSTR, DICHROWEB [7]. Most of these methods use datasets from standard CD libraries of structures determined by X-ray crystallography [3, 7]. Whilst most alpha-helical proteins have tightly defined ψ, ϕ values, β-sheet structures on the other hand, can take on either parallel, anti-parallel or a combination of both. Hence, β-structures demonstrate more variable ψ and ϕ values [1], which is major part of the reason why CD is more reliable for estimating the secondary structure of α-helical proteins than β-sheets.

In mathematical terms, the CD signal is expressed as the difference in the absorption between the left-handed (A_L) and right-handed (A_R) circularly polarised light by an asymmetric molecule. Thus, CD is reported as either the change in the absorption (ΔA) of the left and right-handed circularly polarized light as ΔA= A_L - A_R or in degrees of ellipticity (θ) since the resultant vector of the two absorptions results in an ellipsoid [1–3]. The resulting beam is therefore said to be elliptically polarized. The ellipticity (θ) is given as the tangent of the ratio between the minor (b) to the major (a) axis of the ellipse ($\tan(\theta) = \frac{b}{a}$). Figure 3.2 displays the circular dichroism effect of chiral molecules showing an elliptical rotation (dashed line).

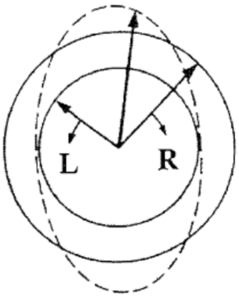


Figure 3.2: Illustration of the unequal absorption of left and right circularly polarized light by chiral molecules. The resultant vector is an ellipse (dashed line). Picture was adopted from Kelly *et al.* [3] with permission from the publisher.

CD follows the Beer-Lambert law [1], A=$\mathcal{E}$ × C × ℓ, where $\mathcal{E}$ is the molar extinction coefficient in $M^{-1}cm^{-1}$, C is the concentration in M and ℓ is the optical path length in cm. Thus, the CD signal can also be expressed as

$$\Delta A = A_L - A_R = \mathcal{E}_L \times C \times \ell - \mathcal{E}_R \times C \times \ell = \Delta\mathcal{E} \times C \times \ell, \tag{3.1}$$

where $\Delta\mathcal{E}$ is the difference in the molar extinction coefficient of the left and right circularly polarized light, also in $M^{-1}\ cm^{-1}$. The molar ellipticity $[\theta]$ and ΔA are related as

$$[\theta] = 3298\Delta A, \tag{3.2}$$

where $[\theta]$ is the CD signal corrected for concentration and is measured in degrees, ΔA is the difference in the absorption of the left and right circularly polarized light.

The molar ellipticity is normalized against the number of residues in the protein and mostly reported as mean residue molar ellipticity (MRE). The MRE is a standardized way of comparing CD spectra of different samples and is given in $deg.cm^2.dmol^{-1}$ [1],

$$MRE = [\theta] = \frac{\theta \times 100 \times M}{C \times \ell \times n}, \tag{3.3}$$

where θ is the ellipticity in degrees, ℓ is the path length of the cell in cm, M is the molecular mass of the protein, C is the concentration of the protein in mg/ml and n is the number of residues in the protein.

The fraction of α-helix, β-strand and disordered region within a secondary structure can be estimated with good accuracy. The fraction of helicity (fH) of a protein or peptide is taken to be proportional to the molar residue ellipticity at 222 nm, $[\theta]_{222}$ nm. According to the equation provided by Rohl *et al.* [9], fH can be calculated as

$$fH = \frac{\theta_{222nm} - \theta_C}{\theta_H - \theta_C}, \tag{3.4}$$

the random coil (θ_C) and whole helix (θ_H) are given by

$$\theta_C = 2220 - 53T \qquad \theta_H = (-44000 + 250T)(1 - \frac{1}{3}N_r), \tag{3.5}$$

where T is the temperature in degrees Celsius and N_r is the number of residues in the polypeptide.

3.2.2 Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS), also known as quasi-elastic light scattering (QELS) or photon correlation spectroscopy (PCS), is a technique used to measure the size of particles ranging from proteins to detergent complexes suspended in a fluid medium [10, 11]. Particle suspensions in solution are undergoing constant Brownian motion, which is a random movement of particles due to the bombardment of the particles by solvent molecules. Smaller particles are kicked harder and thus move faster than the larger ones and hence the speed at which particles diffuse in solution correlates with their size. Due to the Doppler effect, the light scattered from particles in motion exhibits spectral broadening on the order of $\frac{D_t}{\lambda^2}$ [12, 13]. Either by measuring the line broadening of the spectral laser lines or the intensity fluctuation correlation function, the translational diffusion coefficient (D_t) of the partcicles can be determined. The D_t and the hydrodynamic diameter (d_h) are related by the Stokes-Einstein relation (equation 3.6) [14, 15],

$$D_t = \frac{k_B T}{3\pi\eta d_h}, \tag{3.6}$$

where k_B is the Boltzmann constant, T is the absolute temperature, η is the viscosity of the medium, and d_h is the hydrodynamic diameter.

Given the temperature and viscosity of the medium, it is possible to estimate the hydrodynamic diameter, i.e. the diameter of a hard sphere that diffuses at the same speed as the particles in suspension by measuring the D_t of each subpopulation. Thus, DLS can resolve submicron particle sizes according to their translational diffusion. This study uses DLS to estimate the average size and distribution of rSP-B in detergents.

The digital correlator used in DLS compares the intensity fluctuations of singly scattered light over a period of time to give the auto correlation function (ACF) [16], a function that describes how long a given signal remains the same [17]. The ACF (equation 3.7) converts intensity fluctuation of the electric field into a correlogram,

$$G(\tau) = \langle I(t)I(t+\tau)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T I(t)I(t+\tau)dt, \tag{3.7}$$

where $I(t)$ is the intensity of the scattered light at a given time t, and $I(t+\tau)$ is the intensity at a later delay time, t+τ, where τ is the delay time.

Embedded in the ACF is the signal decay rate(s) of each size class in the system. The D_t and the decay rate (Γ) are related as in equation 3.8.

$$\Gamma = D_t q^2, \tag{3.8}$$

where q is the magnitude of the scattering vector and is given by

$$q = (4\pi n/\lambda) sin(\theta/2), \tag{3.9}$$

where n is the refractive index of the dispersant, λ is the wavelength of the incident laser in vacuum and θ is the scattering angle.

Thus, by combining equation 3.6 and 3.9, the decay rate can also be expressed as

$$\Gamma = \frac{k_B T}{3\pi\eta d_h}((4\pi n/\lambda) sin(\theta/2))^2. \tag{3.10}$$

For a large number of monomodal and monodisperse systems, the ACF is a single exponential decay as in equation 3.11:

$$g^{(1)}(\tau) = e^{-\Gamma\tau}, \tag{3.11}$$

and for multimodal and polydisperse systems, the ACF can be considered as collection of monomodal components as in equation 3.12 [18, 19]:

$$g^{(1)}(\tau) = \int_0^\infty G(\Gamma) e^{-\Gamma\tau} d\Gamma, \tag{3.12}$$

where τ is the lag time and Γ is the decay rate.

Thus for a multicomponent system, the ACF is the summation of the intensity weighted decay of each component in the system. A fit to the ACF using different algorithms extracts the decay rate(s). In simple terms, the ACF of monomodal systems can be fit with a simple function such as

$$f(x) = yo + A \times exp(-\Gamma x), \tag{3.13}$$

where yo is the experimental baseline and A is the y-intercept.

Likewise, multicomponent systems can be fit with a sum of decaying exponentials. For instance, taking a three component system, the function takes the form

$$f(x) = yo + A1 \times exp(-\Gamma_1 x) + A2 \times exp(-\Gamma_2 x) + A3 \times exp(-\Gamma_3 x), \tag{3.14}$$

where $A1$, $A2$ and $A3$ represents the Y-intercept for each size class. The summation of these values should add up to one and an initial guess for each is provided by the user. Γ_1, Γ_2 and Γ_3 are the decay rates of each particle class in the system.

Different algorithms have been used to estimate the intensity-weighted mean hydrodynamic diameter (Z-average) and particle size distribution (PSD), among which, the cumulant fit analysis [20], non-negative least square (NNLS) [18] and constrained regularization (CONTIN) [21] methods are the most common. The cumulant analysis is a fit of a single exponential function to the ACF and is used to calculate the Z-average and the polydispersity index (PDI), a dimensionless parameter that describes the broadness of the size distribution. The NNLS and CONTIN, on the other hand, are multi-exponential fits to the ACF that are used to calculate the PSD for broad size distribution and multicomponent systems [11].

The intensity PSD of DLS, can also be converted to volume and number PSD by using the Mie theory. While the intensity PSD is a more accurate estimation of the particle size distribution, the volume and number PSD, on the other hand, provide valuable information on the relative abundance of each particle class in a multicomponent sample. The accuracy of the estimation diminishes in the order intensity $>$ volume $>$ number.

Multiple light scattering (MLS), particle-particle interaction, different modes of diffusion and florescence affect the estimation of the particle size . Some vendors use non-invasive back scatterer technology (NIBS) where the detector is located at 173 degrees away from the incident beam. Because large particles scatter light in the forward direction, the NIBS optic system is especially useful to avoid signal from large artifacts such as dust [22]. The NIBS setup also reduce MLS effect, the tendency of a photon scattered by a particle to be also scattered by another particle, as the diffracted laser travel through the sample the shortest path. MLS can be diminished by diluting highly concentrated solutions. Also, at high concentration, particle-particle interaction might affect the result. Hence, it is customary to determine an optimum working concentration for best DLS results. However, compared to the classical 90 degree detectors, the

NIBS optic do allow to characterize more concentrated samples. What is more, in the case of charged macromolecules/polymers, an electrical double layer, the so-called as the Debye layer, could form that leads to overestimation of the result [10]. To counteract this effect, often times, counterions are added to disperse the double layer or minimize the Debye length (K^{-1}). This is especially important when measuring the surface potential. There is also an inherent error in conventional DLS distribution for samples that are not truly spherical, and for which sizing can be both under or overestimated [14].

In general terms, DLS is non-invasive, fast and provides estimates of the apparent particle size(s) in a suspension. Given that the technique is sensitive to larger particles, DLS is mostly the method of choice to study sample homogeneity or to monitor aggregation and oligomerization propensity or the presence of higher order assemblies of macromolecules in solution. This is especially true for proteins in detergent. That said, samples for DLS analysis must be as clean as possible and should be free from dust and bubbles, which dominate the distribution to give erroneous results if care is not taken. Hence, careful sample preparation is mandatory.

3.2.3 Nanoparticle Tracking Analysis (NTA)

Nanoparticle tracking analysis (NTA) tracks submicron particle suspensions in Brownian motion. Each particle, for example each protein/lipid complex, in the field of view is tracked in parallel and accurate particle by particle sizing in the range of 10 nm to 1000 nm is provided. Since thousands of particles are tracked individually, high resolution number particle size distributions (PSD) are obtained [23, 24]. The technique combines laser light scattering and microscopy with a charge-coupled device (CCD) camera mounted at a right (90°) angle to the irradiation plane. The CCD camera records a short video (20-60 seconds) at 30 frames per seconds to provide real time visualization of nanoparticle particles (NPs) in motion. The NPs in the field of view are tracked until they go off view or overlap with adjacent particles [25]. Frame by frame analysis of the video (Panel A in Figure 3.3) by the software program determines the mean square displacement (MSD) of an individual particle with time interval Δt [24, 26]. Although particles move in three dimensions, NTA tracks the D_t in only two dimensions, $\overline{\langle x, y\rangle}^2$, which is the spatial deviation of the particles over time with respect to an initial reference point. The hydrodynamic diameter is then calculated by using a formula derived from the Stokes-Einstein equation,

$$D_t = \frac{\overline{(x, y)}^2}{4}, \tag{3.15}$$

$$\overline{(x,y)}^2 = \frac{4Tk_Bt}{3\pi\eta d_h}, \tag{3.16}$$

where $\overline{(x,y)}^2$ is the MSD in two dimensions, k_B is the Boltzmann constant, t is the period $(1/framerate)$, η is the viscosity and d_h is the hydrodynamic diameter.

All types of NPs in any solvent including protein/lipid and protein/detergent complexes can be analysed. The minimum detection limit is dependent on the refractive index of the particle suspension. For good scatterers, sizes down to ≈10 nm diameter can be detected, whereas for weak scatterers like most biological molecules, the lower limit is 30-50 nm diameter. The upper limit of diameters measurable with NTA is defined by the point at which diffusion becomes too slow to track [24]. The measured intensity fluctuation of the light scattered by each particle with the corresponding size is plotted in two dimensions (Panel B in Figure 3.3). The analytical software rejects invalid tracks, tracks with short durations, and those above the detection limit, as well as trajectory crossovers [25]. All valid tracks are considered in order to generate a number particle size distribution as shown in Panel D of Figure 3.3. Data can also be plotted in 3D with the third dimension being the number distribution. The particles in Panel A are rSP-B molecules suspended in DPC/SDS detergent micelles.

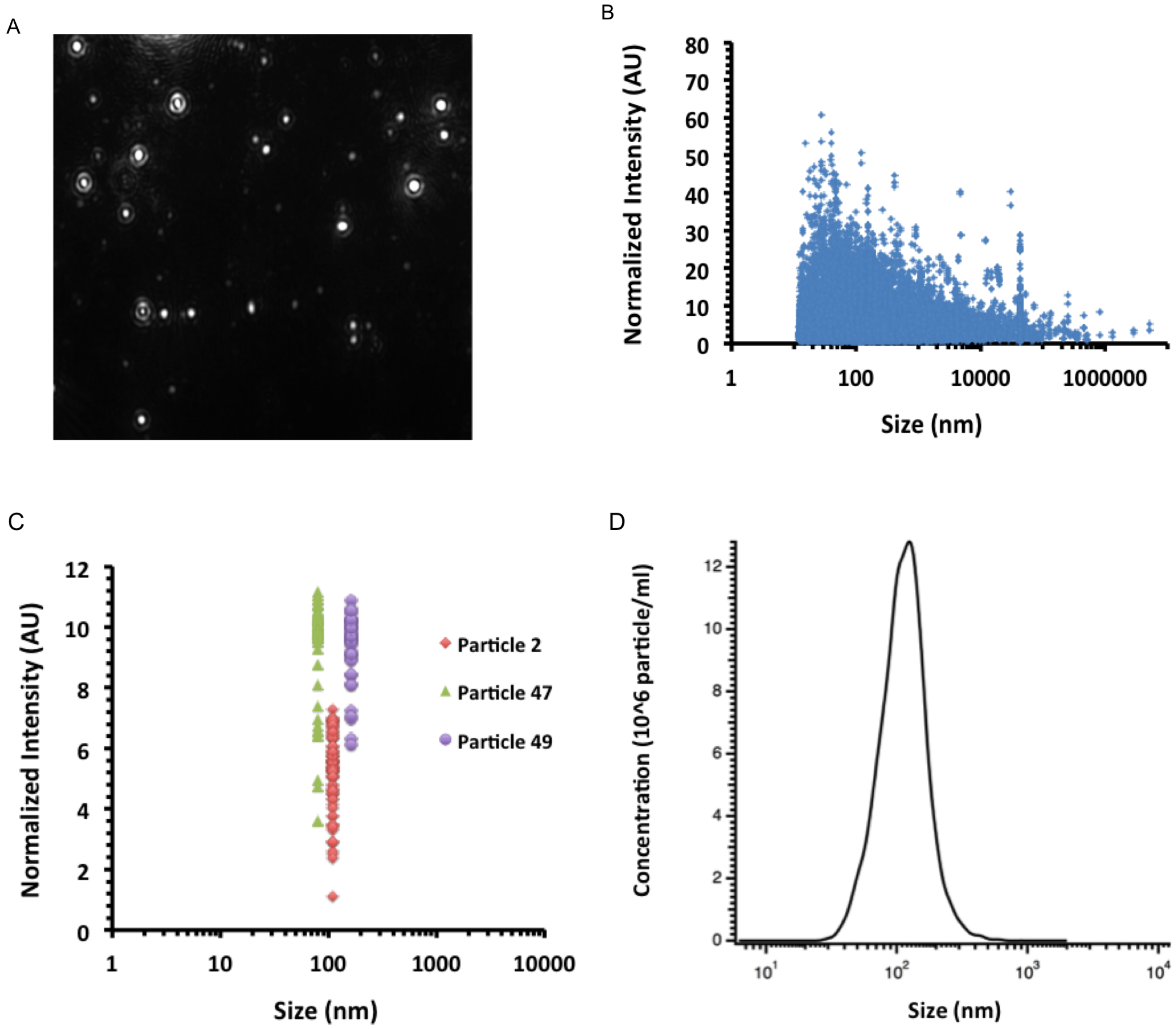


Figure 3.3: An overview of nanoparticle tracking analysis (NTA). NTA of rSP-B suspended in 0.2% DPC/SDS detergent micelles. Panel A: snapshot of the suspended particles undergoing Brownian motion (magnification scale x20), Panel B: the intensity fluctuation of scattered light from all particles detected in the scattering volume, Panel C: particle by particle analysis of three valid tracks. The numbers represent the particles under the scattering volume and is automatically given by the analytical software, Panel D: number particle size distribution of valid tracks. The data displayed were measured by me at the department of Earth Science, Memorial University of Newfoundland.

NTA complements DLS. Although DLS is user friendly and provides consistent fast results, it is also an ensemble technique and cannot well resolve distributions of particle sizes if they are less than a factor of about 3 different in size. For particles that are close in size, DLS gives a broad distribution that is weighted towards the bigger sized NPs [27]. NTA, on the other hand, tracks single particles and therefore provides better size resolution for polydisperse suspensions. Importantly, fluorescent capable NTA is a robust method for visual detection and quantification of a tagged target against a complex background [23]. Hence, thorough analysis of biological samples and protein aggregates can be performed. Conversely, unlike DLS that automatically adjusts the attenuator to measure a wide range of concentration, NTA is limited by concentration.

The recommended concentration is between 10^5 and 10^9 particles per ml [25]. Optimization of the right concentration and parameter settings to capture a high resolution video is paramount for accurate and reproducible results. Although NTA reports concentration as numbers of particles in suspension per cm^3, the measured concentration varies between consecutive measurements of the same sample, which provides an estimate of the uncertainty. As DLS analyzes large numbers of particles, it provides a smooth PSD and hence less deviation between sequential measurements. On the other hand, NTA only accumulates thousands of particles for the distribution and therefore there is more deviation between measurements and the distribution is less smooth [27]. Recently, anomalous diffusion (sub-diffusion) has been suggested to be an issue of NTA, if not considered in the calculation of the D_t from the MSD with a proper method, especially for biological samples [28]. Sample treatment in DLS is less aggressive than NTA; hence, DLS is practically more useful for studying unstable biological samples. That being said however, given a sample that meet the requirements for NTA analysis and with careful measurement, NTA can provide very accurate sizing of both monodisperse and polydisperse samples and provide high resolution number PSD.

3.3 Materials and Methods

3.3.1 Detergent Exchange and Organic Extraction

The rSP-B left bound to the IMAC column after cleavage and renaturation was washed with 30 ml of TRIS-HCl buffer (pH 7.5) to remove the salt and CHAPS. Following this, the protein was exchanged into the desired detergent and buffer of choice (i.e. in most cases is DPC/SDS adjusted to a pH of 7.5 with TRIS-HCl) by washing the rSP-B bound to the resin with three column volumes of the detergent. Finally, rSP-B was eluted either by lowered pH or imidazole gradient using the same detergent and buffer. The sample was then taken for analysis by circular dichroism, dynamic light scattering and nanoparticle tracking. Likewise, in order to put the protein in 100% methanol, the digest after CNBr cleavage was mixed with the detergent SDS, weight by weight (5 times the mass of the protein digest). The mixture was then suspended in high purity water and was left to incubate in a circulating water bath at 42°C for an hour and half. The rSP-B portion was extracted by 2:1 chloroform:methanol (v/v) and the chloroform part was decanted using a separating funnel. To remove all the chloroform from methanol, the rSP-B in methanol was dried by nitrogen gas stream to minimal volume or until the sample becomes turbid (i.e. sign of precipitation was observed). Next, the dried solution was diluted to 15 times the volume of the concentrate with pure methanol and was left to dry again under nitrogen gas stream. To ensure that all of the chloroform is removed, the same procedure was repeated one more time. The final solution was concentrated by drying the resulting solution until the first initial sign of precipitation was noticed. The solution was then ready for analysis.

Reduction of the disulphide bond experiments were done using Tris 2-carboxyethyl phosphine hydrochloride (TCEP, Sigma Aldrich), a potent and selective disulphide bond reducing agent. The TCEP was pre-made in water and was used in the required amount to reduce the disulphide linkage of rSP-B suspended in methanol-water solution.

3.3.2 CD

The Far-UV (260 to 190 nm) CD spectra of rSP-B were recorded using a Jasco J-810 spectropolarimeter (Tokyo, Japan) at a continuous scanning speed of 100 nm/min and standard sensitivity (100 mdeg). The spectrometer was purged for a minimum of 10 minutes with nitrogen gas and the lamp was allowed to warm up for at least 5 minutes. Protein concentration was determined using the Bradford assay. Prior to CD measurement, all samples were dialysed overnight against deionised water with either 1 kDa or 6-8 kDa Spectra/Por dialysis tubing. The cuvette was thoroughly cleaned with ultra pure water and was rinsed with 95% ethanol and subsequently dried using a nitrogen gas stream. Care was taken while dispensing sample into the cell to avoid bubble formation. Unless stated otherwise, all scans were carried out at ambient temperature. Ten acquisitions were taken at a band width of 1.0 nm in mdeg. The spectra were processed with the Jasco Spectra Manager^{TM} software. The spectra were baseline corrected by subtracting a blank, the detergent used to solubilize the protein. Conversion of the raw data (mdeg) to MRE and estimation of the secondary structure content was done in a Microsoft Excel spreadsheet using equations 3.3, 3.4 and 3.5.

3.3.3 DLS

The PSD of rSP-B/detergent complexes was measured with a Zetasizer Nano ZS (Malvern instruments Ltd, Worcestershire, UK). The cuvette was thoroughly cleaned with high purity water and rinsed with ethanol. All dispersants and solvents were filtered with a sintered glass filter (Fisherbrand). One millilitre of sample was slowly dispensed onto the side of the disposable cell. The temperature was allowed to equilibrate for two minutes. Measurements were taken on automatic mode at a scattering angle of 173°. Three measurements were acquired for each sample. Fitting of the ACF was done with IGOR Pro 7 (WaveMetrics, Portland, Oregon) built-in functions, as well as with the Malvern zetasizer software version 7.03. The viscosity and refractive index, which are specific to the dispersing medium used, and the measurement temperature (25°C) were put into the standard operating procedure by the user.

3.3.4 NTA

Nanoparticle tracking analysis was done using the Nanosight NS500 (Malvern, UK) in the lab of Dr. Tao Cheng in the department of Earth Science. The thermoelectric peltier element controls the temperature, which is adjustable by user software. All measurements were taken at room temperature. The flow cell chamber and tubings were flushed with the dispersant/solvent and wash fluids for few minutes prior to loading the sample. Sample was loaded into the flow cell viewing chamber using the built-in fluidics pump controlled by the NTA software. Between measurements, the flow cell was flushed with the diluent. rSP-B sample were measured in DPC/SDS detergent micelles prior to dialysis. Two new protein preparations (rSP-B/DPC/SDS) were used for the analysis.

Remark:throughout this chapter abbreviations can be found in the figure legends. These are meant to represent the persons who prepared the samples. SB: Stella Barth, CS: Caroline Stange, TG: Tadiwos Getachew and DJ: Donna Jackmann. The numbers following the abbreviations are the batch numbers. Please see the authorship statement on page 94.

3.4 Statement of Purpose

Forty years past its discovery there is still a dearth of knowledge on SP-B's quaternary structure. As yet not enough literature data are available to provide information whether SP-B adopts the open or closed conformation or if the protein switches between these two forms based on the surrounding environments as some other SAPLIPs do. Herein, some insight are put forward why studying the protein suspension in solution might be useful. The size of rSP-B in solution especially was of high interest for different reasons: 1) structure determination by solution NMR. As crystallizing SP-B has been impossible, liquid state NMR is the method of choice for determining the structure of SP-B. However, the technique is also sensitive to protein size, which makes it important to measure the size of SP-B in a solution prior NMR measurement. If the protein suspension is beyond some threshold size, typically above 10 nm, then the sample is not suitable for characterization by solution NMR as line broadening of the NMR signal becomes significant and hampers structural study, 2) sample homogeneity is also another important consideration. Normally, a homogeneous protein that is stable for days or best for weeks is required for best results. Even if a protein can be obtained in the lowest energy, a mixture of the same protein in different conformations could also exist in solution. This results in complex NMR spectrum difficult to resolve. Different populations of micelles inside a solution provides different environment for the protein and thus decreases sample homogeneity. Protein-protein or protein-detergent interactions also adds to the heterogeneity of the system. Particle size distribution of a sample is an important parameter to judge homogeneity/heterogeneity of particle suspensions, and 3) the propensity of the protein to oligomerize can also be studied by size measurement. For instance, a protein in an organic solvent and in detergent micelles may demonstrate a similar secondary structure but may exhibit different sizes of the protein. By carefully monitoring the sample before and after treatment, for example after dialysis or filtration, the change within the sample can be probed. For instance, the location of a protein in a polydisperse sample can be carefully analysed. Particle-to-particle interactions may be predicted from the change in the translational diffusion of the particles. Overall, a careful analysis of a complex system may provide important insights not only on the particle suspensions but also on the surrounding solvent. This chapter therefore presents an in-depth investigation of rSP-B suspension in DPC/SDS micelles.

3.5 Results

DLS was used to assess the size of rSP-B/detergent micelle complexes. The particle size distribution of rSP-B suspended in 0.2% (w/v) DPC/SDS (9:1) detergent micelles is polydisperse and shows four peaks. The overlay of the PSD from three different batches prior to dialysis and filtration are shown in Figure 3.4. The y-axis is intensity percentile of light scattered by each size class and the x-axis represents hydrodynamic size of the particles in suspension. The size at 4.1 nm (modal size) is consistent with the expected size for detergent micelles. The next two populations display heterogeneity. There are two peaks that range in modal diameter size from 68 nm to 400 nm in all three samples tested. The sizes are similar enough that we expect the analysis software to have difficulty resolving them. The species may also vary somewhat from sample to sample. Moreover, samples prior to dialysis or ultracentrifugation exhibit a sharp peak at 5560 nm, which is common in DLS data and is normally attributed to dust. To make sure that the sizes coming out of the acquisition system software (from Malvern) are supported by the data, the auto correlation function was checked using a separate software package (IGOR Pro 7).

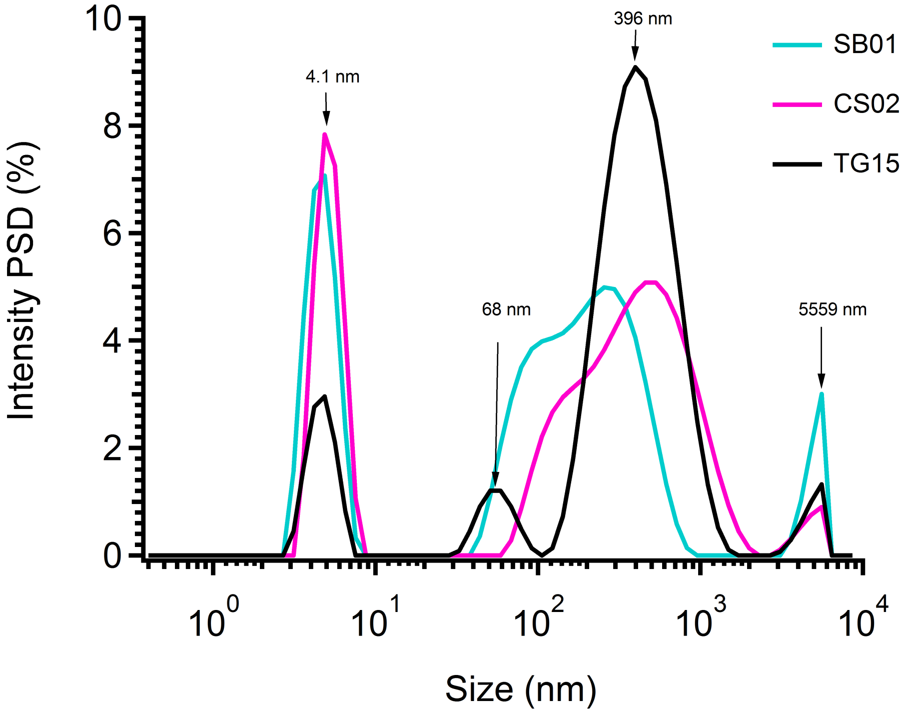


Figure 3.4: DLS of rSP-B in 0.2% (w/v) DPC/SDS (9:1) detergent micelles shows a multimodal size distribution prior to filtering or dialysis. Particle size distribution by intensity is shown for three different protein preparations. The population with the smallest size at ≈4 nm is consistent with the size of detergent micelles. The second and third populations show heterogeneity, with hydrodynamic diameter that range in size from ≈68 nm to ≈400 nm. The unit on the y-axis is percentile, which represents intensity of light scattered by each size class.

The auto correlation functions (ACF) for the same samples display at least three decay rates, consistent with a multicomponent and polydisperse system (Figure 3.5). The arrows mark the lag-times where the correlation starts to decay, depending on the hydrodynamic size of each particle class. The fastest decay rate is consistent with the micelles, whereas the latter two slow decay rates are within the time frame of the 68 to 400 nm particle sizes. By eye, there is no apparent decay corresponding to the micron sized particles (5560 nm), between 10^3 and 10^4 µsec. As the intensity PSD is weighted toward the larger particles that contribute more intensity, we also looked at the data adjusted to show the size distribution by particle volume and particle number.

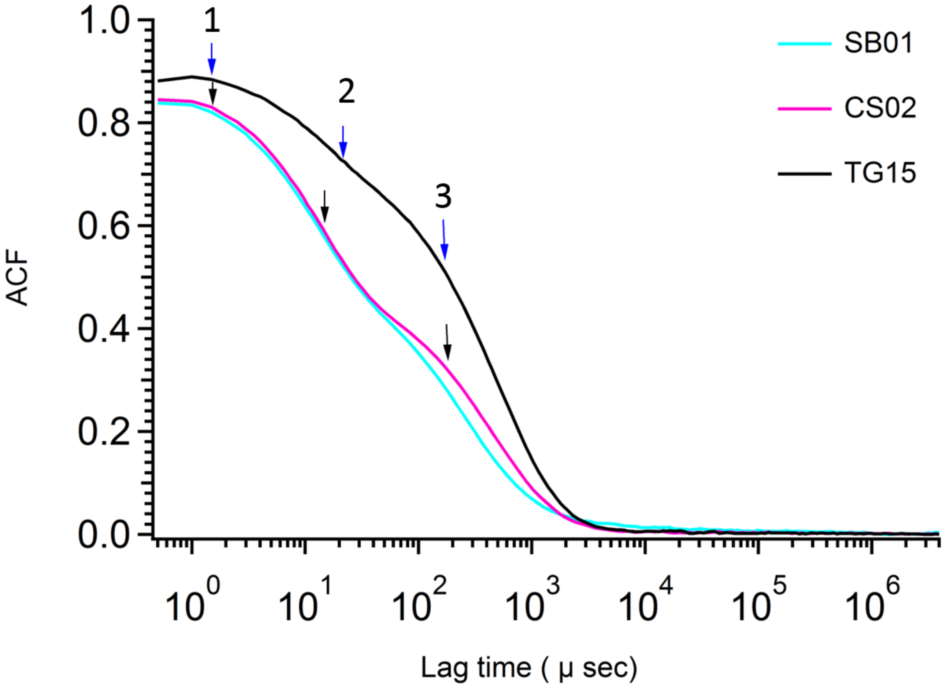


Figure 3.5: The auto correlation function of the particle size distribution shown in Figure 3.4. The three decay rates are consistent with the presence of different particle sizes in all of the samples. The arrows mark the beginning of the intensity decay of scattered light by each size class.

The PSD plotted by volume or number is shown in Figure 3.6. The data indicate that the micelle-sized particles dominate the system. This indicates that the small complexes are much more numerous than the heterogeneous and larger populations. Thus, I wanted to determine if the smaller complexes contain rSP-B or if they are composed of detergent alone.

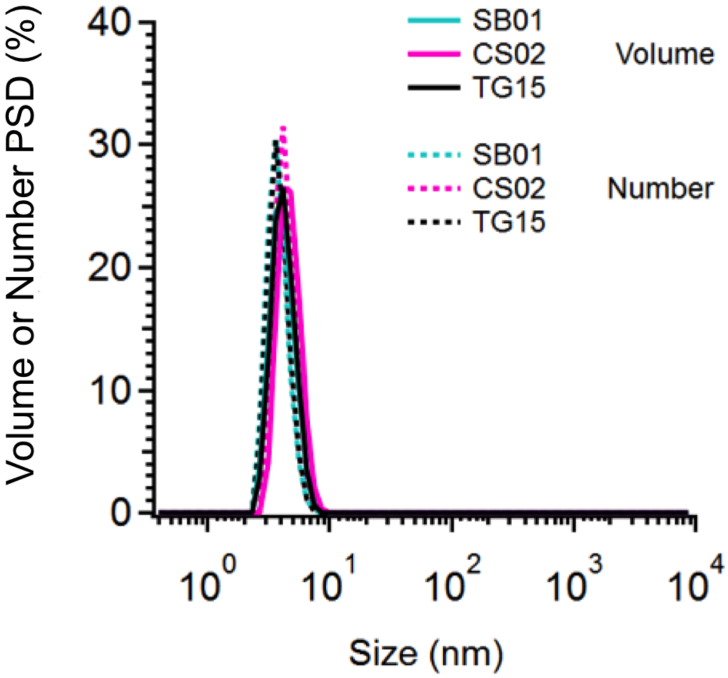


Figure 3.6: Detergent micelle sized particles are the most abundant species in the rSP-B plus 0.2% (w/v) DPC/SDS (9:1) system. Shown are the intensity distribution of the data presented in Figure 3.4 expressed as both volume and number particle size distributions.

To help find out which species contains rSP-B, the samples were dialysed with 1 kDa pore size dialysis tubing against deionised water overnight. The pore sizes are too small for rSP-B or micelles to pass through the membrane, but individual detergent molecules freely diffuse across. Panel A in Figure 3.7 shows an intensity PSD before and after dialysis. The dialysed sample exhibits a monodisperse population (dotted line) and no particles with micellar size. Convincingly, Panel B depicts reproducible size measurements from different batches of rSP-B after dialysis when presented in terms of intensity, volume or number.

This observation indicates that i) the 4 nm species are removed by dialysis and hence do not contain protein tightly associated with them and ii) rSP-B/DPC/SDS complexes are much larger than would be expected for monomeric rSP-B in micelles. The same result was also obtained when samples were dialysed with 6-8 kDa dialysis tubing. Another piece of evidence to consider are the derived count rates (DCRs).

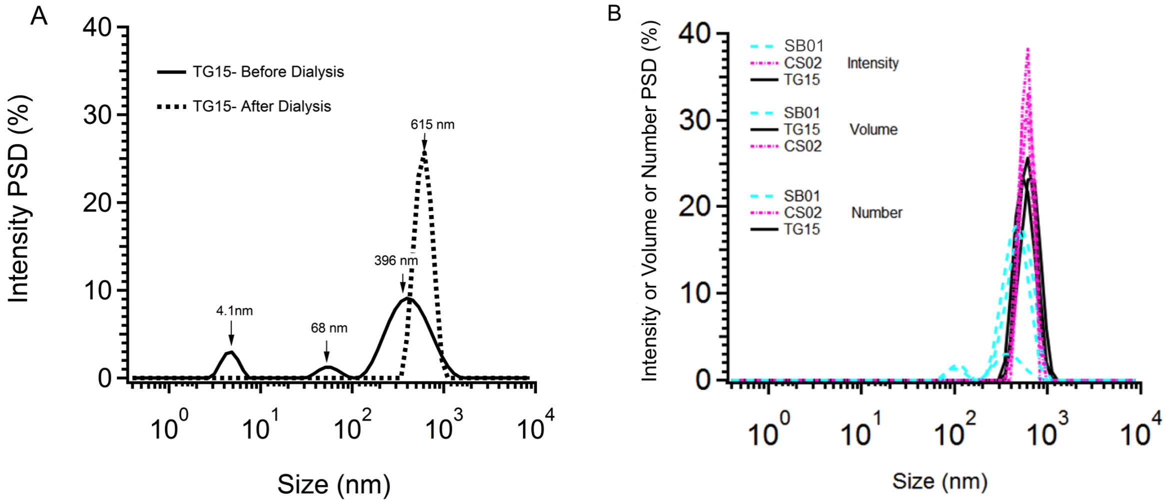


Figure 3.7: Dialysis removes detergent micelle sized species. Panel A: overlay of the intensity PSD before and after dialysis of TG15 (rSP-B suspended in 0.2% (w/v) DPC/SDS (9:1)). Panel B: the intensity, volume and number PSD overlay of three different batches shown in Figure 3.4 after dialysis. Sample was dialysed against three litres of deionised water stirring in cold room (4°C) overnight using 1 kDa molecular weight cut-off (MWCO) tubing. Dialysis leads to the disappearance of the 4 nm sized populations shown in Figure 3.4.

The derived count rate, the translational diffusion coefficient (D_t) and average sizes of all dialysed samples increased with dialysis. Table 3.1 compares the parameters before and after dialysis. Here, it is imperative to note that the Z-average size is the weighted intensity average size of the ensemble cumulative of all sizes measured by DLS and is derived by cumulant analysis of the ACF with a single exponential fit and hence only has one mean value. The mean count rate (MCR) represents the number of scattered photons counted at the detector after the incident laser light has been attenuated by some factor set by the machine based on the number of photon counts of a given sample. This prevents the avalanche photon detector (APD) from being damaged and also sets the photon counts to be well within the linear dynamic range of the APD. Most automatic mode machines adjust the counts to be between 200 to 500 kilo counts per seconds (KCPS). The DCR on the other hand is the theoretical count rate of the scattered photons detected at the detector assuming no signal attenuation (100% transmittance) and hence corrects for the attenuation factor. Higher DCR therefore could indicate one of either: i) higher concentration, ii) larger particles or iii) both larger particles and higher concentration. The DCR before dialysis were in the range of 800 to 1500 KCPS whereas, after dialysis, the DCR is higher, as much as 28,000 to 49500 KCPS. The increased DCR after dialysis could indicate that either there are more larger protein complexes than micelles or it could also be that as the bulk of the micelles are removed, space is created for the protein-detergent complexes (PDCs) to undergo attractive interaction. The other argument could be that dialysis concentrates the sample. Assuming water also diffuses into the system, however, the concentration factor could be ambiguous. Besides, the larger D_t and sizes of the dialysed sample support the conclusion that dialysis removes the smaller particles but not the larger ones. Interestingly, filtering the dialysed sample seems to cause all the D_t, DCR and Z-average to diminish. This could suggest that filtering removes the particles that were responsible for the DCR increase, the larger particles. Importantly, the PDI of the dialysed-only samples and of the sample that was dialysed and filtered agree very well with each other compared to the not-dialysed sample, with PDI of almost twice as high. Thus, the samples prior to dialysis are more polydisperse than the dialysed samples.

Table 3.1: Summary of the effects of dialysis and filtration on rSP-B/DPC/SDS DLS size measurements

Before dialysis					
Batch	Mean count rate (MCR, in KCPS)	Derived count rate (DCR, in KCPS)	Translational diffusion coefficient (m^2sec^{-1})	Z-average (nm)	PDI
SB 01	393	1400	8.78×10^{-12}	56	0.80
CS 02	235	836.50	11.20×10^{-12}	43	1.00
TG 15	166	1495.60	2.32×10^{-12}	211.80	0.74
After dialysis					
SB 01	357.60	28350.70	9.70×10^{-11}	507	0.47
CS 02	110	30654.60	3.50×10^{-11}	1405	0.46
TG 15	179	49511.90	5.62×10^{-11}	875.20	0.45
After dialysis and filtration					
CS 02	383	377	5.97×10^{-12}	82.40	0.47
CS 02(2X)	294.50	1048	3.76×10^{-12}	131	0.46

[a] 2X represents sample that was filtered twice

The ACF of all dialysed samples (Figure 3.8, Panel A) also show a single decay rate at a long lag time, which indicates that the samples are now more uniform in size and the particles are large. The lag-time of the particles in the dialysed sample (10^2-10^3 µsec) corresponds with the third arrow in Figure 3.5. The volume distribution (Figure 3.8, Panel B) indicates that filtering the dialysed sample with a 0.2 micron filter pore causes a shift in size indicating loss of particles larger than the pore size.

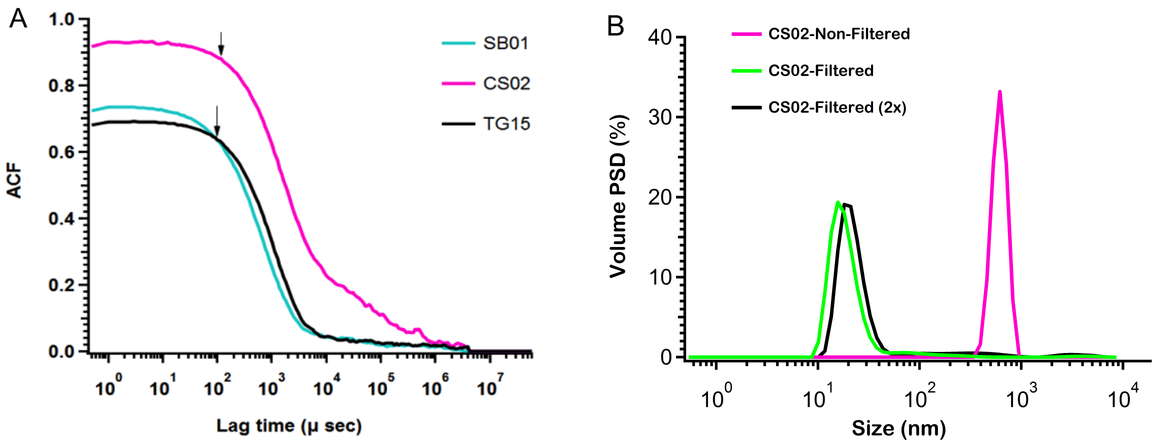


Figure 3.8: The ACF after dialysis depicts one decay rate corresponding to larger size. Panel A: The time frame of the ACF at the beginning of the decay corresponds to the lag-time of the heterogeneously populated particles (≈68 to 400 nm) as shown at the later delay time (10^2-10^3 µsec, the third arrow) in Figure 3.5. Panel B: filtering causes loss of bigger particles in rSP-B/DPC/SDS system. Volume PSD of dialysed sample (CS 02) before and after filtering it with a 0.2 micron (200 nm) pore size syringe filter (Sartorius, Germany). The hydrodynamic diameter of the particles shifts from 615 nm (before filtering) to ≈20-100 nm after filtering. The ACF in Panel A (CS 02, pink) indicates the presence of larger sedimenting particles in between the lag times of 10^4-10^5 µsec.

These observations demonstrate that either the protein-detergent complexes (PDCs) are bigger than 200 nm or the PDCs sticks to the filter membrane. As a rule of thumb, a sample for DLS measurement is not recommended to be filtered. The dispersant on the other hand must be filtered prior to suspending the analyte. Dialysis alone, as well as dialysis in combination with filtration, causes the PDI of the sample to decrease to a more acceptable range for DLS analysis, ≈ 0.45. Since it is challenging to get accurate sizes out of DLS data if the PDI is larger than 0.3, I decided to validate the sizes the Malvern acquisition software derived from its fitting of the data with fits I did myself using IGOR Pro 7 fitting functions.

The ACF of the undialysed, unfiltered sample in Figure 3.5 (black) was fitted with a triple exponential function using the IGOR Pro 7 built-in function (equation 3.14), which fits a sum of decaying exponentials. Panel A in Figure 3.9 is the intensity PSD of the fitted sample and Panel B is the semi log plot of the fit. Panel C is the same data in log-log scale. At a very large delay time the log-log plot indicates the fit does not agree well with the data. Panel D is the same fit shown in Panel C with the data that did not converge well (noise) cut-off, also presented in log-log scale.

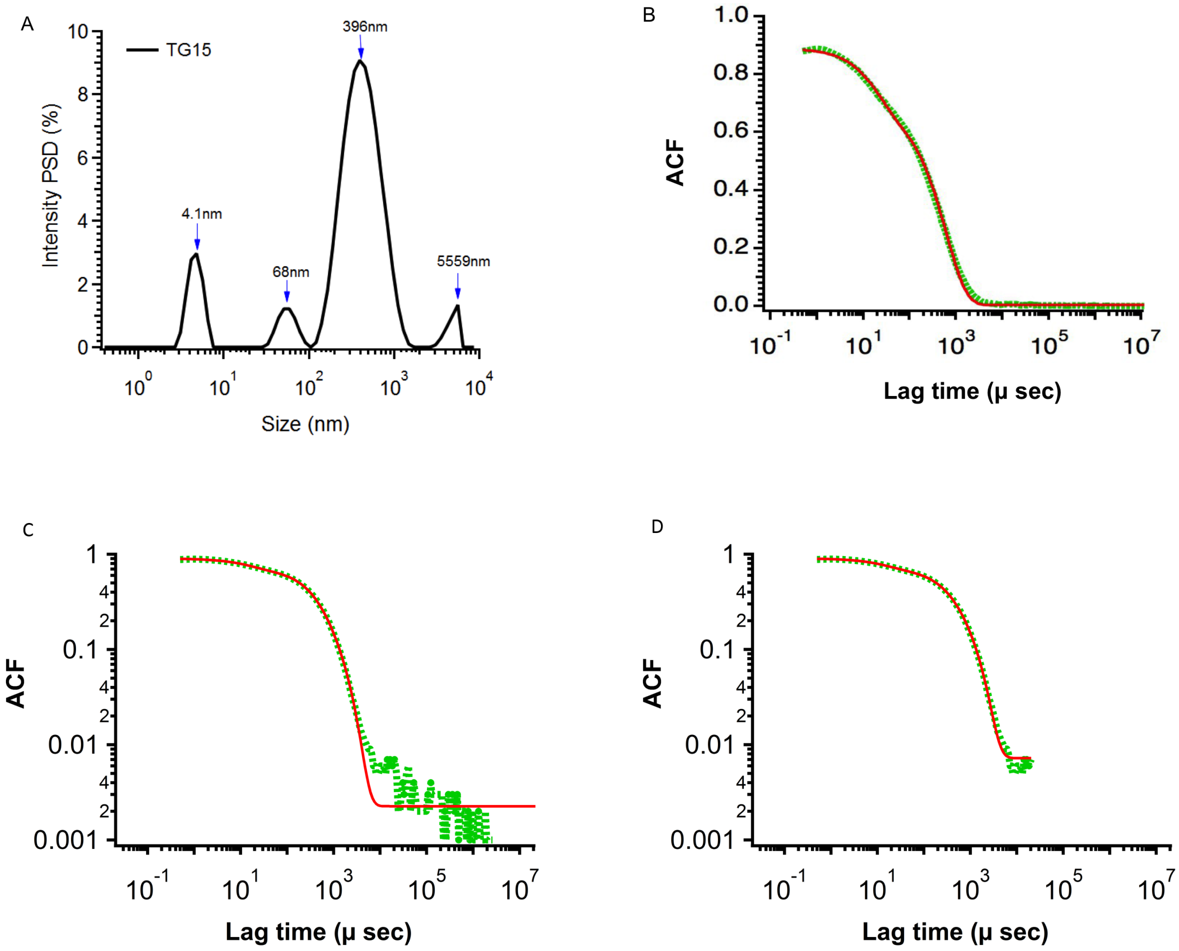


Figure 3.9: Manual curve fit to the auto correlation function of TG15 prior to dialysis and filtering. The apparent sizes were compared from the fitting generated by Malvern software. The fit panels show data in green and triple exponential fit in red. Panel A: is the intensity PSD of TG15 prior dialysis. Panel B: demonstrates comparison of the data and the fit on a semi log plot. Panel C: a log-log plot of the same data presented in panel B reveal that the fit at larger lag-time is not good. Panel D indicates the fit that converged with the noisy data cutoff. Thus, the data for the apparent size at ≈5560 nm reported by the Malvern software at a much larger lag time is not reliable. Curve was fitted using IGOR Pro 7.

A new fit was also attempted on the cut-off data and the fit function was found to converge properly. The decay rates were extracted from the fit function and the hydrodynamic diameters (d_h) were calculated by using equation 3.10. Finally, the D_t were calculated using equation 3.6.

Similarly, Figure 3.10 shows that the dialysed sample fits well with a single exponential function (equation 3.13). As for the non-dialysed sample, the log-log plot in Panel B exhibits a poor fit at the higher delay-time scale. According to the literature, this is mostly due to the correlation function approaching zero and $\log(0) = -\infty$. When an intensity fluctuation is analyzed by computing the correlation function, the baseline value is used to normalize the data. By subtracting the mean of the intensity, I(t), from the data prior to calculating the correlation function, the baseline is set to start at a value of 1 for a lag time, τ=0 and gradually decays to zero at a larger lag time. Because of noise, baseline-subtracted, normalised values can dip below zero, which can not be computed as logarithms of a negative number are not allowed [17]. All in all, the data demonstrates that after dialysis the rSP-B/DPC/SDS system is homogeneous with a Z-average size of ≈615 nm.

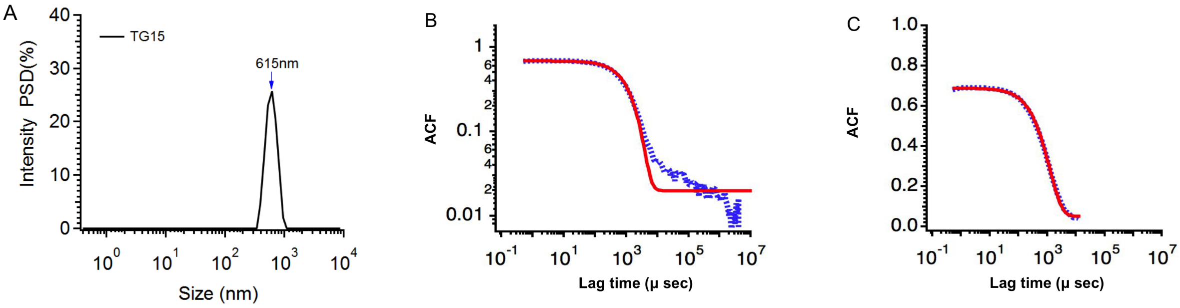


Figure 3.10: Dialysed sample fits well with a single exponential function. The fit to the correlation curve of TG15 shown in Figure 3.8A after dialysis and prior to filtering. Both panels show data in blue and the fit in red. Panel A is the PSD of the fitted sample, Panel B is log-log plot of the fit and Panel C is the semi log plot of the fit in Panel B after excluding the data that did not converge.

The decay rates, D_t, and sizes calculated from the fit before and after dialysis are compiled in Table 3.2. All sizes calculated using the fit function fall within the PSD generated by the NNLS algorithm (Malvern fitting). The sizes calculated from the best fit, however, were all smaller than those calculated from the full range data fit for both non-dialysed and dialysed samples. Importantly, the micron sized (5560 nm) species falls in the time scale of the noisy range with no meaningful decay rate. Hence, the manual fitting supports the interpretation of the data in terms of modal particle diameter sizes of 5 nm, 125 nm, and 386 nm prior dialysis. The data therefore underlies that the ≈5560 nm population should be disregarded from the distribution analysis. Besides, this population was not evident in the dialysed samples. In essence, the calculated sizes are in the range of the size distribution provided by the Malvern software.

Table 3.2: Calculated parameters using the data generated from the curve fittings of TG15 before and after dialysis

Fitting of the auto correlation curve of TG15 before dialysis			
	Decay rates (1/ µsec)	Translational diffusion coefficient ($m^2 sec^{-1}$)	Diameter (nm)
Full range of the ACF data fit (Figure 3.9B)	0.0688 ± 0.00139	8.24×10^{-11}	5.30
	0.0029 ± 0.00009	3.50×10^{-12}	125
	0.0009 ± 0.00003	1.13×10^{-12}	386
ACF data cut-off at a lag-time of 10^4 µsec (Figure 3.9D)	0.0710 ± 0.00130	8.51×10^{-11}	5.13
	0.0035 ± 0.00014	4.20×10^{-12}	104
	0.0011 ± 0.00027	1.37×10^{-12}	319
Malvern fitting			4.10 (2.60-7.50)
			68 (37-122)
			396 (141-1281)
			5559 (3090-6438)
Fitting of the auto correlation curve of TG15 after dialysis			
Full range of the ACF data (Figure 3.10B)	0.0007 ± 0.00103	8.69×10^{-13}	502
ACF data cut-off at a lag-time of 10^4 µsec (Figure 3.10C)	0.0008 ± 0.000006	9.75×10^{-13}	448
Malvern fitting			615 (341-1106)

Next, since the PDI of the non-dialysed or filtered samples was very high, I wanted to complement the DLS data with another technique. As NTA measures diffusion of many individual complexes, the technique is well suited to polydisperse samples. However, the smaller complexes, e.g. sizes as small as individual micelles, are below the detection limit of the NTA to measure.

The top panels in Figure 3.11 display reproducible NTA data of rSP-B/DPC/SDS complexes within one batch (Panel A) and across two batches (Panel B). The bottom panels (Panel C and D) show the 3D intensity profile of scattered light and the surface fit to the 3D scatter plot that gives rise to average size distributions shown in panel A. The surface fit in Panel D depicts at least three highly scattering populations with different concentrations (black arrows). What is more, some of the the surface fits also have different sizes (blue arrow). Overall the plots display different scattering populations in keeping with a multimodal system.

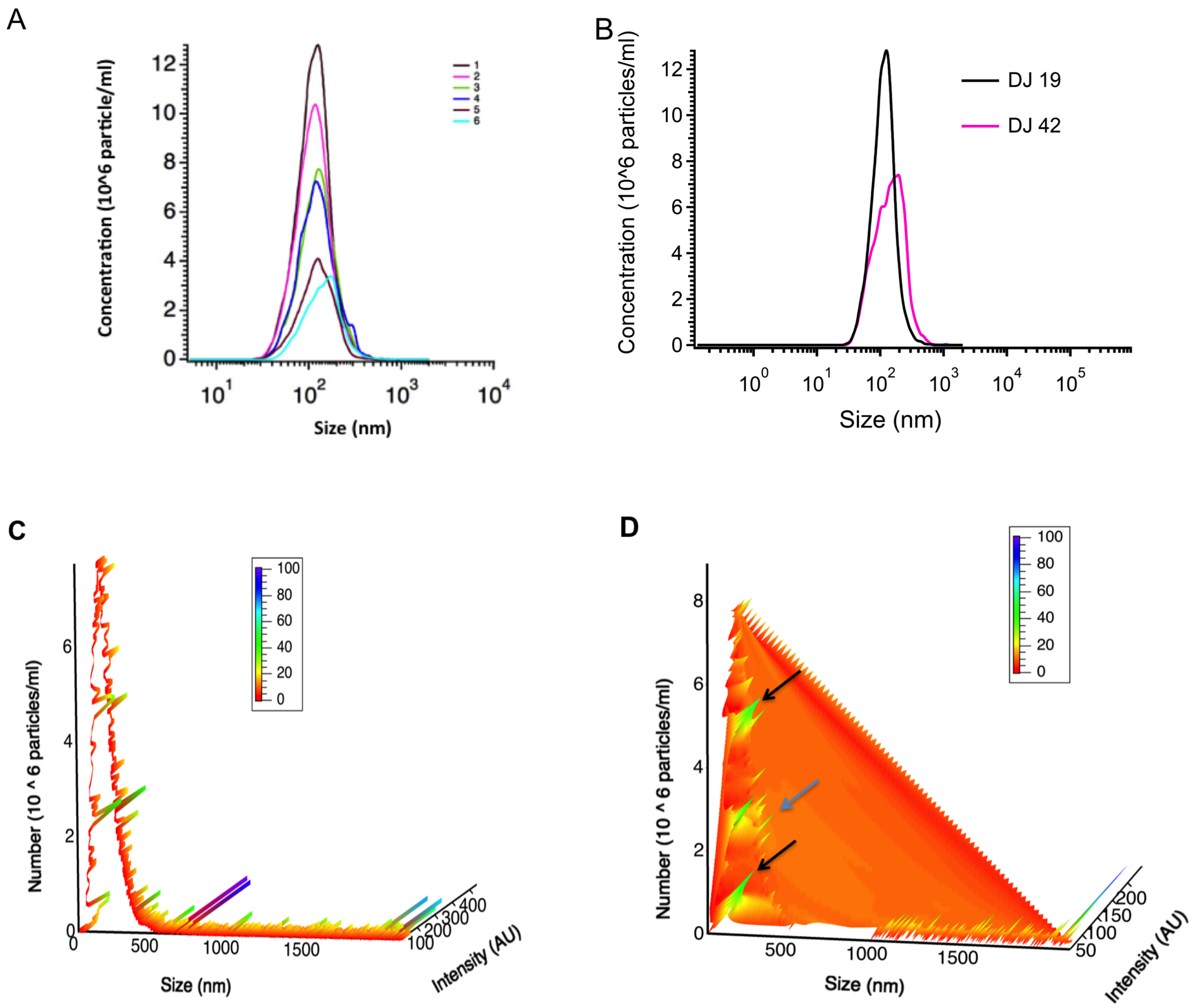


Figure 3.11: Nanoparticle tracking analysis shows repeatable size measure of rSP-B in DPC/SDS detergent micelles. Panel A is the PSD of different measurements of DJ 19 (full length rSP-B). The legend indicate the number of measurement of the same sample. Panel B is the overlay of batch 19 with batch 42. The two batches are different that DJ 42 does not contain the Cys48 considered to be important for covalent homodimerization, whereas DJ 19 is a full-length (including the insertion sequence: the 7N-terminus residues) rSP-B with Cys48 included. Panel C is the 3D intensity path plot of one of the PSD shown in Panel A and Panel D is the surface fit to the 3D scatter plot of the distribution shown in Panel C. The scale bar indicates the intensity of the light scattered by each species.

Table 3.3 represents the diffusion coefficients, particle count (concentrations) and sizes obtained from NTA of rSP-B suspension in DPC/SDS. The mode is ≈120 nm for batch 19 and ≈180 nm for batch 42. The D_t obtained from NTA are comparable in magnitude to that obtained for the 68-400 nm species by DLS measurement before dialysis. The concentration, however, was observed to noticeably vary between repeated measurement of same sample.

Table 3.3: Characterization of each nanoparticle tracking analysis of DJ 19 and DJ 42 shown in Figure 3.11.

DJ 19	Completed tracks	Mode (nm)	Concentration (10^8 particles/ml)	Translational diffusion coefficient ($m^2 sec^{-1}$)
1	3315	128	10.43	3.45×10^{-12}
2	1922	120	10.12	3.34×10^{-12}
3	10779	125	14.70	3.80×10^{-12}
4	3322	117	11.99	3.55×10^{-12}
5	4492	125	5.27	7.41×10^{-12}
DJ 42				
1	7541	193	16.08	1.87×10^{-12}
2	3297	183	7.31	2.70×10^{-12}

Another way of interpreting the NTA data is by looking at the cumulative undersize distribution in Figure 3.12. The cumulative undersize distribution is mostly described by the D_{10}, D_{50} and D_{90} values. The D_{10} value indicates that 10% of the particles are below some threshold value while 90% of the particles are above the threshold. Similarly, D_{50} is the median size of the distribution, where half are below and half are above the given threshold. The D_{90} on the other hand is the opposite of D_{10}, that is, 90% of the particles are below threshold while 10% of the particles are above the threshold. Therefore, the interpretation of the result in Figure 3.12A demonstrates that 99% (D_{99}) of the complexes have sizes less than 430 nm in batch 19 and less than 499 nm in batch 42.

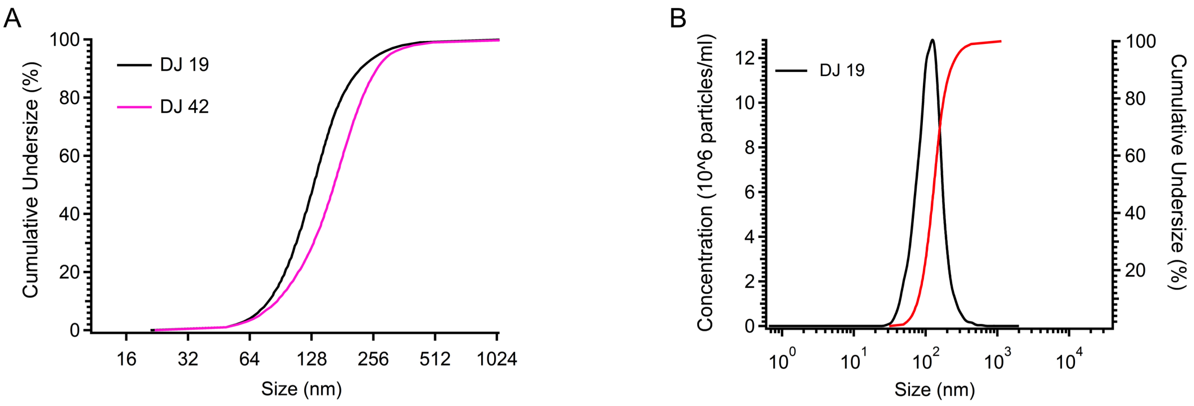


Figure 3.12: Cumulative undersize distribution of rSP-B/DPC/SDS. Panel A: overlay of batch 19 and 42. Panel B: an overplot of the number PSD with the cumulative undersize distribution of batch 19. According to the cumulative undersize distribution, 99% of the particles are below size of 500 nm

3.5.1 CD of rSP-B in Detergents

Next, we questioned whether the rSP-B in complex with the detergent(s) takes on the native-like conformation. To answer this question, besides DPC/SDS, other membrane mimetic lipids/detergents were also exchanged as described in the method section. Although CHAPS was important for rSP-B to bind to the metal ions, it was found that the detergent also results in rSP-B with no secondary structure. Not only that, with CHAPS present, other detergents/lipids were not able to induce the secondary conformation of the protein. Hence, the CHAPS used to facilitate rSP-B binding to the column had to be removed. This was done by washing the column with Tri-HCl buffer (pH 7.5) before the protein was exchanged into any other detergent/lipids. Such an experiment also helps to check whether or not rSP-B adopts the expected secondary structures in different environments. In addition to this, the behaviour of rSP-B in different surroundings could also give important information.

The helicity of rSP-B varies with the type of detergent/lipid used. The protein exhibits alpha helical conformation in 0.5% (w/v) dodecylphosphocholine (DPC), 0.2% (w/v) DPC/SDS (9:1) and 0.2% (w/v) LMPG (1-myristoyl-2-hydroxy-sn-glycero-3-phospho-(1-rac-glycerol) micelles. Panel A in Figure 3.13 is in mdeg and Panel B is in MRE, depicting helicity of 13% in DPC, 25% in LMPG and 37% in DPC/SDS. The helicity appears to be highest in anionic containing detergent/lipid. Though many conditions were screened, most were not successful since no secondary structure was recorded for the protein in different conditions attempted thus far.

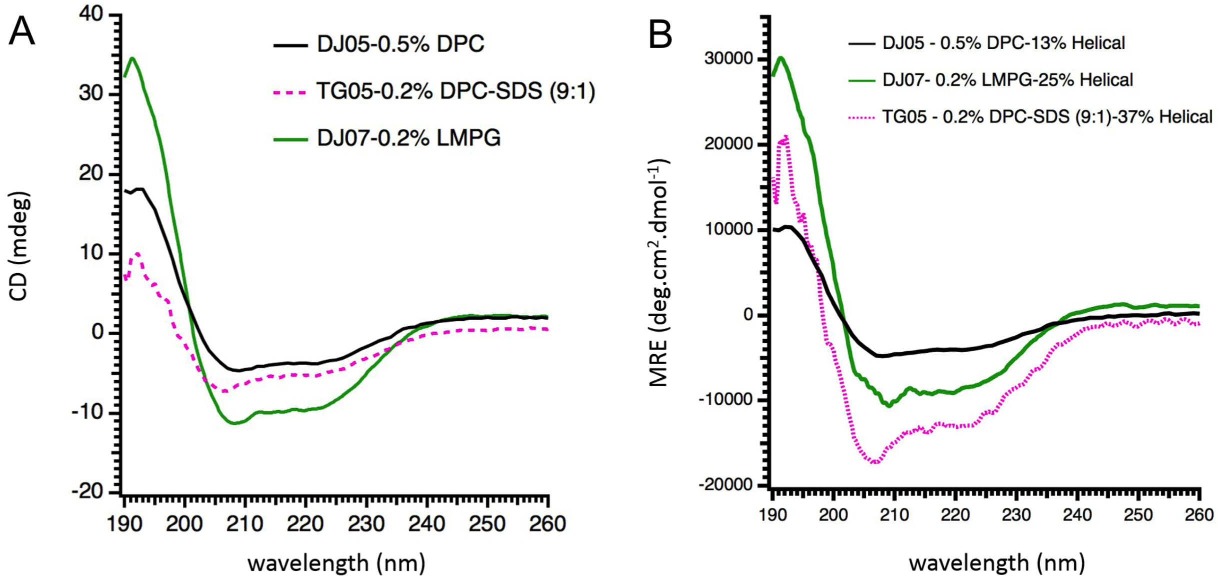


Figure 3.13: rSP-B retains the expected secondary structure in a number of environments. Panel A: CD scans of rSP-B in 0.5% DPC, 0.2% DPC/SDS (9:1) and 0.2% LMPG. Panel B: the estimation of the MRE of the CD scans shown in Panel A.

Since it was observed that dialysis removes the micelle-sized species from the rSP-B/DPC/SDS system and resulted with homogeneous PSD as shown in Figure 3.10, an important question raised was if there is rSP-B in this population and if so, does it have the right secondary structure? To answer this question, the dialysed sample was taken for CD measurement.

Figure 3.14 shows that rSP-B exhibits a characteristic alpha-helical CD spectrum with minima at 208 nm and 222 nm and an estimated helicity of 47% after dialysis. These data therefore confirms not only the presence rSP-B but also a protein-detergent complex bigger than anticipated. In essence, DLS and CD together indicated that rSP-B has native-like secondary structure and exists in detergents in size complexes ranging from 340 nm to 1106 nm. The next question addressed was finding out if organic solvent could break up the big complexes and still maintain the helicity of rSP-B?

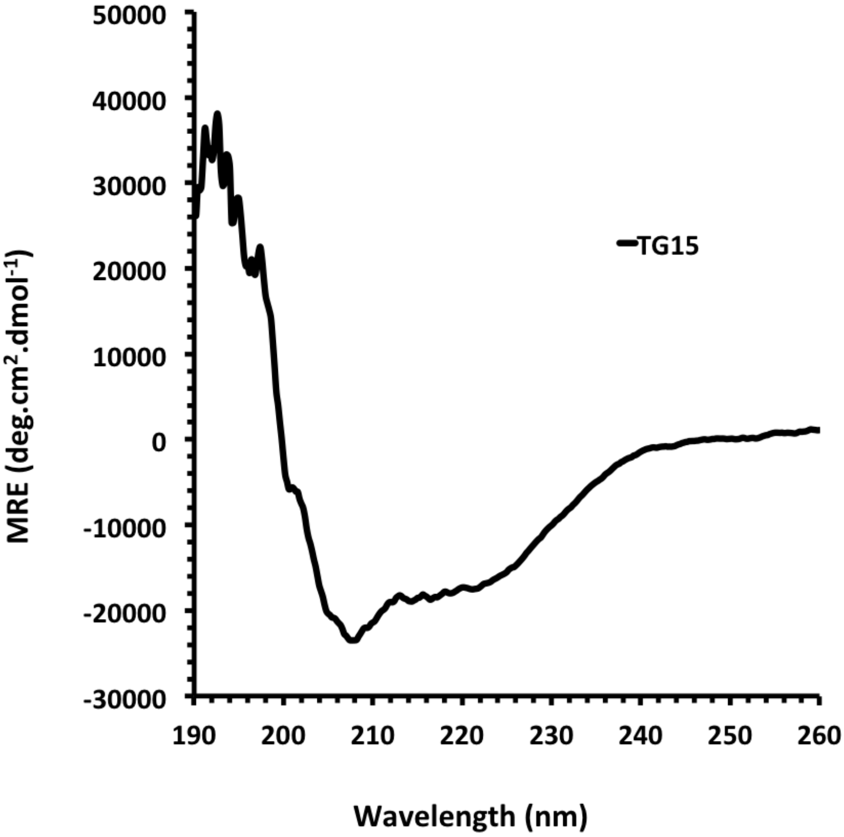


Figure 3.14: rSP-B exhibits the expected alpha-helical conformation in DPC/SDS detergent micelles after dialysis and prior filtering. According to estimation of the secondary structure, rSP-B exhibit ≈47% alpha-helical content. The concentration (0.02 mg/ml) used for estimating was determined by using the Bradford assay. Sample was dialysed with 1K MWCO tubing overnight stirring in cold room prior to analysis.

3.5.2 CD of rSP-B in Methanol

To put the rSP-B into an organic solvent, 2:1 chloroform:methanol (v/v) extraction of the rSN-SP-B cyanogen bromide digest was performed as described in the method section. The chloroform portion was removed by repeatedly drying the extract and resuspending it with methanol. It was found that rSP-B also exhibits an alpha helical secondary structure in pure methanol as shown in Figure 3.15.

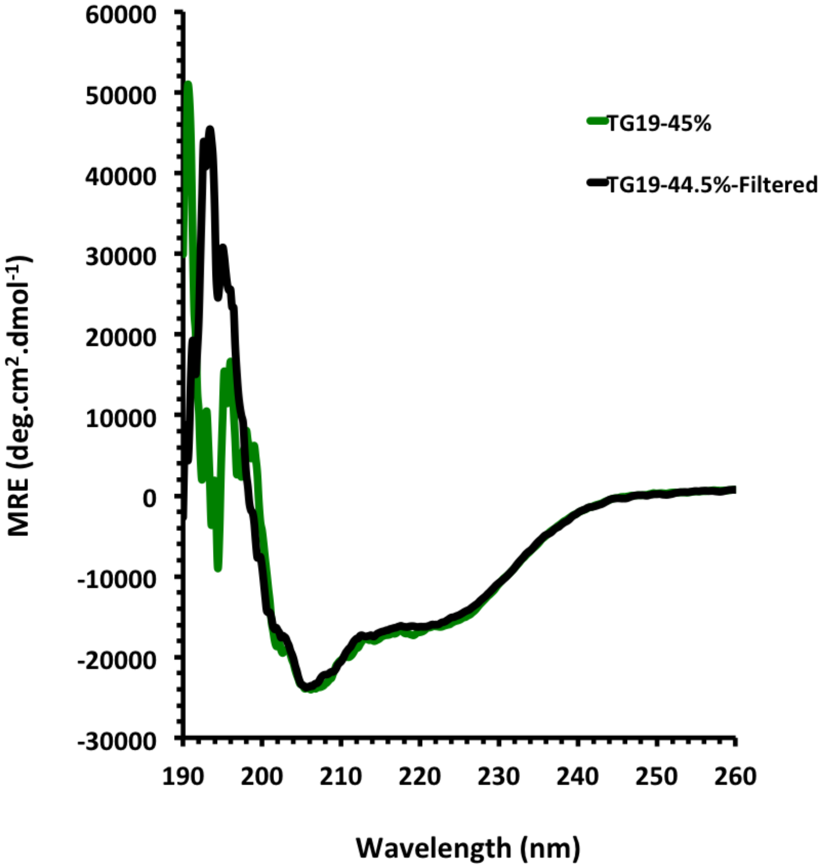


Figure 3.15: rSP-B exhibits the expected alpha-helical conformation in methanol. Far-UV circular dichroism spectra of the TG 19 expressed in mean residue ellipticity. Lyophilized cyanogen bromide digest of the 20 to 100 mM imidazole elution fractions pooled together and extracted using 2:1 chloroform:methanol solution and then exchanged into pure methanol. Green spectrum is prior filtering sample and the black spectrum is a scan taken after filtering same sample with a 0.2 micron (200 nm) pore size syringe filter. According to estimation of the secondary structure content, rSP-B exhibit ≈45% alpha helical content. Concentration used for the calculation (0.22 mg/ml) was determined by Bradford assay.

The rSP-B suspension in methanol that was filtered with 200 nm pore size (black spectrum) retraces the spectrum of the not-filtered sample (green spectrum). Both with a percent helicity estimation of ≈45%. This observation was interesting because it was in contrast to the finding for the rSP-B suspension in DPC/SDS (Figure 3.14). The complex of rSP-B/DPC/SDS behaves differently after being filtered with the same filter pore size. To further investigate this, DLS of rSP-B suspension in methanol was measured and analysed.

Figure 3.16 is the DLS size measure of rSP-B in 100% methanol. Table 3.4 compiles the count rates, diffusion coefficients, and hydrodynamic sizes. The filtrate of rSP-B in methanol exhibits high DCR and a PDI less than 0.3, an acceptable range for DLS analysis. Also, the concentrations determined before and after filtration were similar (0.22 mg/ml). Conversely, rSP-B/DPC/SDS behaves differently as filtration removed much of the CD signal, and decreased the DCR and concentration. In addition to this, unlike the case for rSP-B in methanol, the MRE of rSP-B/DPC/SDS complex after filtration was never the same when measured in the same path length cuvette as that in which the CD signal of the non-filtered sample was acquired, which is an indication of sample loss after filtration. Possibly, this is the case when the rSP-B/DPC/SDS complex is either bigger than the pore size of the filter or perhaps it could be that the complex sticks to the filter itself. Another possible explanation could be that methanol disrupts weak protein-protein interactions between rSP-B subunits thus making the protein smaller in size and while maintaining its secondary structure. Finally, we wanted to confirm whether rSP-B has non-native disulphide bonds or not. To do that, the rSP-B suspension in methanol was chosen for a preliminary test since accessing rSP-B, especially the disulphide linkages, in the presence of detergents is difficult.

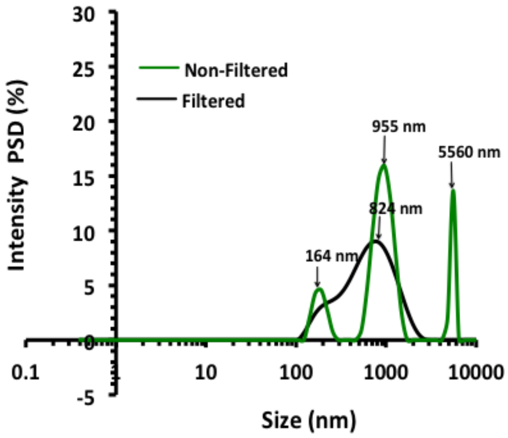


Figure 3.16: rSP-B suspended in methanol has comparable size as in DPC/SDS micelles when judged by the intensity distribution before filtering. The intensity particle size distribution of TG 19 (rSP-B suspended in methanol) before and after filtering sample with 0.2 micron pore size syringe filter.

Table 3.4: The count rate, diffusion coefficients and sizes of rSP-B suspended in methanol before and after filtration with a 200 nm filter.

TG 19	Mean count rate (MCR, in KCPS)	Derived count rate (DCR, in KCPS)	Translational diffusion coefficient ($m^2 sec^{-1}$)	Z-average (nm)	PDI
Non-Filtered	205	16246	3.71×10^{-11}	2148	0.74
Filtered	156	12422	1.61×10^{-12}	495	0.248

High concentrations of Tris 2-carboxyethyl phosphine (TCEP) reduce the disulphide bonds of rSP-B suspension in methanol. As a preliminary test for the disulphide bond linkage, rSP-B in methanol (TG 19) was subjected to a reducing agent TCEP as is shown in Figure 3.17.

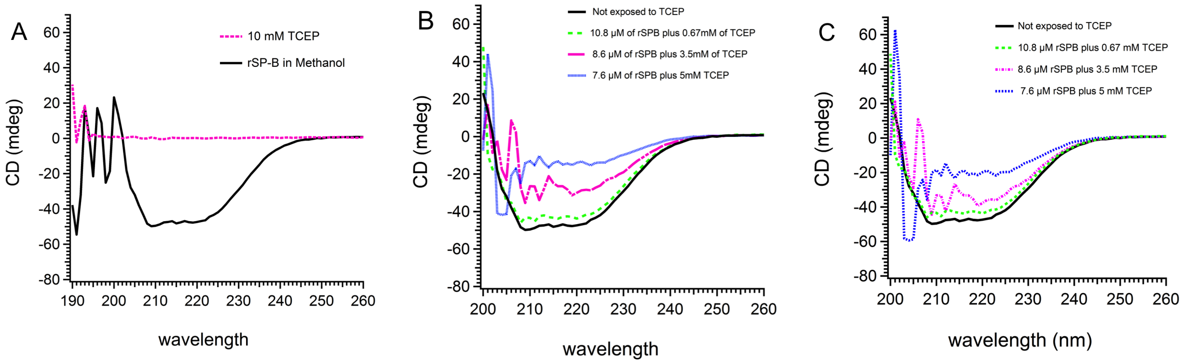


Figure 3.17: The CD signal of rSP-B in methanol diminishes as a function of Tris 2-carboxyethyl phosphine (TCEP) concentration. TG 19 before and after exposure to TCEP. Panel A: CD scan of rSP-B suspended in methanol overlay with that of 10 mM TCEP (blank). Panel B: depicts the effect of TCEP on rSP-B suspended in methanol. Panel C: concentration corrected relative to the 10.8 µM rSP-B. The spectra show same trend as panel B.

TCEP is a more potent and irreversible reducing agent than dithiothreitol (DTT) and β-mercaptoethanol (BME) [29–31]. Panel A to the left shows 10 mM TCEP (pink) and rSP-B in methanol (black) in mdeg. 10 mM TCEP did not give background noise in the far-UV region. The middle spectra demonstrate scans as a function of TCEP concentration. In contrast to the reference (not-treated), the spectra taken in the presence of the reducing agent diminish in intensity as the TCEP concentration increases. To correct for protein concentration, the spectra were normalized with respect to the 10.8 µM rSP-B (Panel C). The sample treated with 5 mM TCEP looks the least helical and started to resemble a random coil.

Finally, a hypothetical model of DPC/SDS micelles is presented for an illustration purpose. It is known that DPC and SDS are synergic in solution when mixed above the CMC value of each detergent. Figure 3.18 shows a model for mixed micelles in dynamic equllibrium with the detergent molecules. As DPC is zwitterionic and SDS anionic, the mixture of the two makes zwitteranionic mixed micelles. The charge distribution of the micelles may be altered by changing the stoichiometry of the constituent molecules. A more thorough analysis of DPC/SDS micelles is presented in the discussion section below.

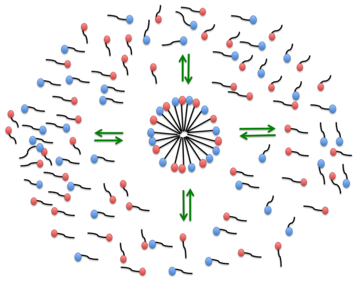


Figure 3.18: A hypothetical model DPC/SDS micelles in dynamic equilibrium with the detergent molecules. DPC/SDS make synergistic micelles when mixed at an appropriate molar ratio. The molecular mass of the micelle depends on the aggregation number.

3.6 Discussion

After successful expression of rSP-B, the next task was to determine if the structure of recombinant SP-B bears similarity with native SP-B. Luckily, the literature has good amount of data on the secondary structure of animal SP-B. But these data were mostly from organic extracts of SP-B from lung lavage [32, 33] and very few had information on detergent(s) suitable to stabilize SP-B into its native form. Although the structure of Mini-B was elucidated by solution NMR in SDS micelles [34], it is not clear if the same detergent is suitable for the near full length SP-B. Hence, the challenge with rSP-B work was while rSP-B was carefully refolded when bound to the IMAC column, the protocol may not lead to native-like conformation, making it important to check the structure. CD is invaluable in determining the secondary structure of proteins in solutions, while DLS characterizes sample homogeneity and aggregation of particle suspension. Thus, to assess the secondary structure and sample homogeneity, CD and DLS of the sample were measured after each trial of refolding rSP-B into a membrane mimetic environment.

The secondary structure of rSP-B is environment dependent. The CD of rSP-B in 0.5% (w/v) DPC, 0.2% (w/v) LMPG and 0.2% (w/v) DPC/SDS (9:1) micelles demonstrated an alpha helical secondary structure (Figure 3.13A). Since the MRE is normalized against the number of residues of the protein and is also corrected for concentration as per equation 3.3, it is a useful parameter to compare the degree of rSP-B helicity a particular detergent induces. Based on the nature of the head group and length of the acyl chains, each of these surroundings presents a unique environment. rSP-B looks to be more helical when in DPC/SDS than in either DPC or LMPG (Figure 3.13B), as judged by the intensity of the two characteristic local minima at 222 nm and 208 nm. The fraction of helicity (fH) of rSP-B in these environments calculated as per equations 3.4 and 3.5, indicates that rSP-B is ≈37% helical in DPC/SDS micelles as compared to ≈25% in LMPG or ≈13% in DPC. Thus, the presence of the anionic detergent, SDS, seems to be making the difference: it could be that rSP-B makes more contact with the DPC/SDS binary micelles.

Studies support that DPC and SDS detergents are synergistic in solution and readily form mixed micelles [35, 36]. Due to electrostatic interactions, there is a mutual association between anionic and zwitterionic surfactants [37, 38]. At room temperature, for instance, DPC/SDS in 9:1 molar ratio exhibits a CMC value (0.94 mM) less than pure DPC (1.33 mM) in phosphate-buffered saline (PBS). Likewise, the CMC of pure SDS in salt, at least in 150 mM NaCl, is about 1-2.5 mM, also above that for DPC/SDS [36]. The lesser CMC value of DPC/SDS as compared to the pure detergents confirms the synergistic nature of the two detergents. Certainly, in TBS(1X) and/or in Tris-HCl, the buffers used in this study, DPC/SDS forms zwitteranionic mixed micelles. Since major phospholipids of lung surfactant are also zwitterionic and anionic by nature, the mixed micelles of DPC/SDS may approximate the natural environment for rSP-B. This could be part of the reason why the fH of rSP-B in DPC/SDS (Figure 3.14) is quite similar to that of native SP-B [32, 33, 39, 40]. Because most of the native SP-B work was in organic solvent in the past, we also extracted rSP-B using 2:1 chlorofom:methanol (v/v) and later exchanged it into pure methanol (Figure 3.15). The spectra obtained both in DPC/SDS and in methanol are in agreement with other studies of SP-B in different environments [40–43]. Since methanol is likely to disrupt the rSP-B interactions that stabilize its 3D structure, DPC/SDS was chosen as the environment in which to characterize rSP-B in depth.

Since an important long term goal is to obtain the 3D structure of SP-B, it is of high interest to consider the rSP-B data in this chapter, in light of the various techniques available to obtain its tertiary/quaternary structure. Due to the extereme apolar nature of SP-B, crystallizing the protein has been impossible. Another option is to use liquid state NMR for which the sample must yields narrow line width. To achieve narrow line width, the size of the protein/detergent complexes (PDCs) must be sufficiently small. Another consideration for solution NMR is isotope labeling. Although native SP-B can be extracted in pure form with organic solvent, there is no way to label the native protein with ^{15}N and ^{13}C isotopes, a requirement for solution NMR structure determination. The ability to produce SP-B recombinantly in bacteria provides, for the first time, an opportunity to incorporate isotope labels into SP-B for structural studies. Yet another important aspect to consider for solution NMR is sample homogeneity. In the simplest form, the rSP-B/DPC/SDS system, consists the monomer units of each detergent, the free micelles and the PDCs. Ideally, for solution NMR, the PDCs are best to be only slightly bigger than free micelles. In reality, however, the complexity of the system can further be increased by protein-protein interactions. For instance, a protein in complex with a micelle may bind another neighbouring protein in complex or the protein might self associate and bind more detergent molecules. In that case, the diffusion of the PDCs would be slowed due to their larger size. Hence, the translational diffusions of the species in many component system can give important clues. I was therefore particularly intrigued to investigate the behaviour of rSP-B in DPC/SDS micelles.

Typically, micelles have sizes that range from 2 to 20 nm [44]. As such, due to their short rotational correlation times, micelles present an ideal system to study biological molecules with conventional liquid state NMR [45]. A study by Manzo *et al.* [35] has reported a DLS diameter size of 4.3 ± 0.9 nm for DPC/SDS micelles. Considering that a biological bilayer in the liquid crystalline state has a bilayer thickness of ≈4 nm [46], DPC/SDS micelles can be a good model for membrane proteins. Hypothetically, a protein-detergent complex of 4-5 nm is small enough to tumble in solution on a time scale such that liquid state NMR can provide structural information to construct a 3D model. Nonetheless, it is noteworthy to stress that SP-B is not a transmembrane protein hence the interaction of SP-B with lipid bilayers or micelles could be significantly different from that of integral membrane proteins. Particularly considering the fusogenic and lytic nature of SP-B [40], it is possible that the shape and size of DPC/SDS micelles may change considerably when binding rSP-B. DLS, despite some of its limitations, is non-invasive and thus particularly suited to monitor protein oligomerization and aggregation.

DLS size measurements of the rSP-B/DPC/SDS system display polydisperse distributions with modal particle sizes of ≈4, 70 and 400 nm (Figure 3.4). While the CD spectra of the same sample shows percent helicity comparable to native SP-B, it is not clear if rSP-B lies in some or all of the ≈4, 70 or 400 nm complexes seen by DLS.

The first population with a diameter size of 4.1 nm is quite similar to the DLS size reported for DPC/SDS micelles (4.3 nm) in phosphate-buffered saline by Manzo *et al.* [35]. A general assumption is that a micelle can only accomodate a single peptide [47]. For small proteins, the PDCs look more like a protein embodied within a micelle, whereas, for bigger proteins, the detergents shields the proteins by surrounding them [46]. Assuming an aggregation number of 50 for DPC [48, 49] and 62 for SDS [50] micelles, at room temperature, the molecular masses of the micelles of each surfactant can be calculated to be ≈17.5 kDa. According to Sikorska *et al.* [36], the addition of SDS only slightly affects the size of DPC micelles. Thus, the molecular mass of DPC/SDS micelle can be roughly estimated to be between ≈18-19 kDa. According to Garcia *et al.* [51] the hydrodynamic diameter of globular proteins can be calculated from their atomic structure. Similary, Langer *et al.* [52] have also shown that the hydrodynamic diameter of proteins can be calculated in angstroms precision. For instance, the hydrodynamic diameter of ubiquitin (8.5 kDa), as calculated by these methods, gives a size of 3.2 ± 0.2 nm. Likewise, TNFα (17.5 kDa), demonstrates a hydrodynamic diameter of 4.4 nm using both methods. Extrapolating these data to SP-B monomer (8.7 kDa) or dimer (17.5 kDa) therefore roughly gives an expected hydrodynamic size range of ≈3-4.5 nm. The fact that DPC/SDS micelles only demonstrate a hydrodynamic diameter of ≈4 nm by DLS, also gives a clue that the integration of an rSP-B monomer or dimer into a DPC/SDS micelle is only expected to increase the DLS size of the complex by a small amount. Of course, such an assumption is only vague and should be taken with caution as it is dependent on a numbers of factors such as the shape of the protein and the complex, the oligomeric state of the protein and hence the numbers of detergent micelles that associate with the protein as well as any form of attractive or repulsive interaction between the complexes. The same data can also be viewed in terms of the translational diffusion constant (D_t). The D_t calculated for the micelle-sized species, $8.51 \times 10^{-11} m^2 sec^{-1}$ (Table 3.2, best fit), is similar to the self-diffusion NMR results for DPC micelles reported by Sikorska *et al.* [36] and Kallick *et al.* [49] to be $10.69 \times 10^{-11} m^2 sec^{-1}$ and $9.17 \times 10^{-11} m^2 sec^{-1}$ respectively. The fact that the calculated D_t for DPC/SDS micelles is quite similar to that of pure DPC micelles, and considering the 4.3 ± 0.9 nm DLS size of pure DPC/SDS micelles in PBS buffer reported by the Manzo group [35], suggests that the 4.1 nm (modal size) or the 5.1 nm (Table 3.2, best fit) particles in the rSP-B/DPC/SDS system may not have protein associated with them. It could be that these species are DPC/SDS micelles that are freely diffusing in solution. But still, such an assumption is inconclusive and needs further work.

The other larger size complexes detected by DLS at ≈70 and 400 nm were better resolved in some samples than others (Figure 3.4), but two populations consistently show in all of the samples in these size ranges. Likewise, the ACF generated by the automated software shows two decay rates consistent with two distinct sub-populations (Figure 3.5, second and third arrows). Not only that but also the manual fit to the ACF using equation 3.14 as shown in Figure 3.9 also demonstrates two particle sizes of 104 nm and 319 nm diameter that diffuse with speeds of 4.2×10^{-12} m^2sec^{-1} and 1.37×10^{-12} m^2sec^{-1} respectively (Table 3.2, best fit). Moreover, NTA size measurements of rSP-B/DPC/SDS complexes also exhibit diffusion coefficients comparable in magnitude to these subpopulations observed by DLS, e.g. 3.8×10^{-12} m^2sec^{-1} with a modal size of 125 nm. For polydisperse sample such as this, NTA provides the advantage of tracking individual particles, allowing for better definitions of the sizes present [27]. While, NTA cannot detect particle sizes as small as micelles (≈4 nm), it did indicate that most of the remainder of the particles were between modal sizes of 120-200 nm (Table 3.3).

To add more, DLS size measurement of rSP-B in methanol exhibits a Z-average diameter of 495 nm and D_t of 1.61×10^{-12} m^2sec^{-1} (Table 3.4) after being filtered with a 200 nm pore filter. Although the viscosity of the solvent, sample concentration, and particle shape affect the diffusion coefficient, the D_t in methanol, has a similar magnitude to the D_t obtained from fitting the bigger bimodal DLS distribution or to that obtained for an individual complex by NTA. Because of the non-micellar nature of the methanol system, one can contemplate that the protein could lie in the bigger bimodal population in the DPC/SDS system as well. Nonetheless, it should be taken into account that the difference in the solvent system could have profound effects on the D_t. Hence, assuming the protein has similar size in different surroundings based on the D_t could be inaccurate. Yet another way of interpreting a multicomponent and polydisperse sample is to look at the volume and number PSD, which provides information on the relative abundance of the species with respect to each other in the suspension.

The bulk of the rSP-B/DPC/SDS system is predominantly micelle-sized (Figure 3.6). The number distribution shows the numbers of particles in the bulk solution that have similar size, whereas the volume distribution is signature of particles that are likely to have similar shape and size. Thus, the good agreement between the number and volume distributions suggests that the micelle-sized particles are monodisperse and likely with similar shape. These data support a well orchestrated formation of micelles that have similar aggregation number. In addition, the same data also indicate that the micelle-sized population are numerous relative to the bigger particles observed in the intensity distribution (Figure 3.4). To test which populations contained the protein, the size of rSP-B/DPC/SDS solution was measured as a function of buffer salt, pH, reducing agents, and temperature. Preliminary data showed that the micelle-sized paricles are unaltered within the physiological condition and at conditions suitable for solution NMR measurements. The broader distribution, with particle sizes of 70 and 400 nm, however, was found to vary based on conditions. It could be that the PDCs in these population are stabilized by networks of weak interactions that are prone to alteration. Nonetheless, as micelles are also affected by pH, salt, and temperature as proteins are, an experiment like this may not precisely answer for the presence or absence of the protein within the micelles. However, it was noticed that at extreme acidic or basic pH, and high salt, the micelle-sized particles are not observed in the DLS size measurements. Still, the possibility that some rSP-B also lies in the micelled-sized population could not be overruled. If, however, there is no protein in the micelle-sized populations, the freely diffusing micelles are just debris that may interfere with downstream processing [53, 54], and may also interfere with protein quantification [53]. What is more, after stabilizing a protein with a detergent(s), it is recommended to remove the excess detergent to increase sample homogeneity. This is specially important for structural biology [54]. One way of removing the free micelles is by dialysis. Since some detergents or detergent micelles are non-dialysable [55], it was also interesting to investigate whether DPC/SDS micelles are dialysable. If so, the dialysate can then be analysed on its own.

Dialysing the rSP-B/DPC/SDS sample with a 1 kDa MWCO tubing removes the micelles-sized particles from the bulk of the solution (Figure 3.7A). Thus, considering the size of an rSP-B monomer (8.7 kDa), which cannot pass through a 1 kDa membrane pore, we can conclude that these species have no protein associating with them. This observation also suggests the protein must be in the populations with size much larger than a single micelle. A hypothetical model of monomer-micelle equilibrium is shown for DPC/SDS mixed micelles in Figure 3.18. According to the study by Manzo *et al.* [35], DPC and SDS molecules are in fast exchange between the bulk and the mixed micelles in solution. The result is a rather quick relaxation process (τ_1), on the time scale of microseconds [56]. In addition, micelles themselves are spontaneously forming and disintegrating [57], which is a rather long relaxation process (τ_2), on the time scale of milliseconds and it is the rate limiting step of micelle formation and dissolution [56]. Hence, dialysis must be disrupting the equilibria between the monomer unit in the bulk and within the micelles. Extensive dialysis therefore likely leads all the micelles to disintegrate and diffuse. Ideally, only the detergent molecules/micelles tightly bound to the protein remain. The detergent, SDS for instance, tightly binds cationic molecules such as rSP-B and in such case the bound SDS may not be removed by dialysis. Thus, an important question is whether these big complexes retain the expected secondary structure of the protein after dialysis and how dialysis might have affected the size and translational diffusion of the particles. To investigate that the dialysate was further analysed by DLS and CD.

The dialysate demonstrated well formed and homogeneous PDCs that tumble slowly. After dialysis, DLS size measurements showed a homogeneous population (Figure 3.7B) and CD showed native-like protein secondary structure (Figure 3.14). The ACF (Figure 3.10C, best fit) fits very well with a single exponential function (equation 3.13). The PDC were calculated to have sizes of 448 nm and D_t of 9.75×10^{-13} m^2sec^{-1} (Table 3.2, best-fit). These data support big PDCs that diffuse an order of magnitude more slowly than the slowest diffusing complex prior to dialysis. A potential reason for the improved homogeneity is considered to be due to the removal of the monomer units of the surfactants and the micelles. However, the reason why the bimodal trend of the subpopulations changes is not precisely clear. Generally, however, it can be taken to be because of the change in the dynamics of the complexes. The PDCs leverage more space to diffuse freely without being hampered after dialysis. This creates better chance for the PDCs to come to close proximity and associate. Hence, the larger size of the PDCs may implicate the presence of higher assemblies probably driven by protein-protein interactions. The correct secondary conformation could suggest the protein is likely in the functional form. Thus, it can be argued that such complexes are well defined oligomers, rather than non specific aggregates. In addition, the increased homogeneity and uniform D_t, suggests that oligomerization of SP-B may be a consequence of specific protein-protein or protein-detergent interactions rather than random nonspecific interactions. Although some trimeric and oligomeric forms of SP-B have been reported, many studies have found that SP-B mostly arranges as a homodimer [58–60]. Assuming a dimer of rSP-B forms in DPC/SDS, then, higher order oligomeric arrangments could be possible. Notably, a supramolecular arrangement of SP-B has been suggested for native SP-B in the zwitterionic detergent CHAPS [61]. The other possibility is that rSP-B promotes fusion of DPC/SDS micelles, similar to that observed for native SP-B and phospholipid vesicles [40]. In such a case, the shape of the PDCs changes, which also increases the size. Moreover, it is also important to consider that a higher order oligomer of the protein could bind more micelles, which also increases the size of the complex. To test this hypothesis (i.e. if there is supramolecular arrangement of rSP-B complex with detergents), the rSP-B suspension in methanol and the dialysate of DPC/SDS were filtered with same filter pore (200 nm) and analysed.

The PDCs of the rSP-B/DPC/SDS dialysate are larger in size than the rSP-B suspension in methanol. It was found that rSP-B in methanol exhibits similar CD spectra before and after filtering. Also, filtering does not reduce DCR very much relative to DCR for the non-filtered sample (Table 3.4). Conversely, for the rSP-B/DPC/SDS, filtration reduces the concentration of the protein in the filtrate to a greater extent. Likewise, the DCR of the dialysate decreased significantly after filtering (Table 3.1). More importantly, the secondary conformation of rSP-B/DPC/SDS was destabilized by filtration when measured by CD (data not shown). These observations suggest that, while a susbstantial part of the rSP-B suspension in methanol is small enough to pass through the filter pore, the rSP-B/DPC/SDS complex is not. This makes sense, because according to DLS (Figure 3.16), rSP-B in methanol before filtering had a component at 164 nm, which should pass through the 200 nm filter. On the other hand, the volume, number and intensity size measure of the rSP-B/DPC/SDS dialysate is well above 200 nm (Figure 3.7B), ≈450 nm according to the best fit. Thus, despite showing similar secondary structure, rSP-B exhibits different sizes in these two environments. Due to the harsh nature of organic solvents, methanol, although known to stabilize the secondary structure of proteins, may also disrupt weak protein-protein interactions between subunits [61, 62]. Thus, it could be that weak protein-protein interactions are the main reason for the bigger PDCs. These results could therefore suggest that the quaternary structure of SP-B could form better in detergent environments than in organic solvents.

For reasons mentioned in chapter two, it was difficult for us to get native SP-B to use as a positive control. Reconstituting purified animal SP-B into 0.2% DPC/SDS and performing the same experiment would have been good to compare the behavior of recombinant SP-B to the native SP-B in DPC/SDS micelles. Although it should be noted that even for native SP-B, the structure and function have been shown to be highly sensitive to how SP-B is extracted from lung surfactant [61]. Experiments presented in this chapter are repeated a minimum of three times. Results were found to be consistent with samples taken from different isolates as well as aliquots taken from the same samples.

In conclusion, rSP-B exhibits a larger size than what is expected for an rSP-B dimer in 0.2% DPC/SDS (9:1). A protein-detergent complex of such size is expected to lead to spectral broadening with conventional solution NMR and has been limiting the NMR work so far. . Either further optimization of rSP-B/DPC/SDS system is required or other appropriate techniques must be pursued to get sensible structural data. The functional aspect of rSP-B in lipid environment will be discussed in the next chapter.

3.7 References

1. Corrêa, D. H. A. & Ramos, C. H. I. The use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research* **3,** 164–173 (2009).
2. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **1,** 2876–2890 (2006).
3. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta* **1751,** 119–139 (2005).
4. Dunn, B. Quantitative amino acid analysis. *Current Protocols in Protein Science* **76,** 10.1.11–10.1.13 (2006).
5. Buck, M. Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Quarterly Reviews of Biophysics* **31,** 297–355 (1998).
6. Luo, P. & Baldwin, R. L. Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36,** 8413–8421 (1997).
7. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89,** 392–400 (2008).
8. Woody, R. W. & Koslowski, A. Recent developments in the electronic spectroscopy of amides and α-helical polypeptides. *Biophysical Chemistry* **101-102,** 535–551 (2002).
9. Rohl, C. A. & Baldwin, R. L. Comparison of NH exchange and circular dichroism as techniques for measuring the parameters of the helix-coil transition in peptides. *Biochemistry* **36,** 8435–8442 (1997).
10. Bhattacharjee, S. DLS and zeta potential - what they are and what they are not? *Journal of Controlled Release* **235,** 337–351 (2016).
11. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews* **8,** 409–427 (2016).
12. Langevin, D., Raspaud, E., Mariot, S., Knyazev, A., Stocco, A., Salonen, A., Luch, A., Haase, A., Trouiller, B., Relier, C., Lozano, O., Thomas, S., Salvati, A. & Dawson, K. Towards reproducible measurement of nanoparticle size using dynamic light scattering: Important controls and considerations. *NanoImpact* **10,** 161–167 (2018).
13. Kerker, M. *The scattering of light and other electromagnetic radiation: physical chemistry: a series of monographs* Rev. ed. (ed Loebl, E. M.) pp. 688 (Acadamic press, 2013).
14. Brar, S. K. & Verma, M. Measurement of nanoparticles by light-scattering techniques. *Trends in Analytical Chemistry* **30,** 4–17 (2011).
15. Einstein, A. *Investigation on the theory of the Brownian movement* (ed R.Furth) (Dover Publications INC, 1956).
16. Lim, J., Yeap, S. P., Che, H. X. & Low, S. C. Characterization of magnetic nanoparticle by dynamic light scattering. *Nanoscale Research Letters* **8,** 381 (2013).
17. Russo, P. A practical minicourse in dynamic light scattering. *Lsu. edu,* 1–48 (2012).

18. Morrison, I. D. & Grabowski, E. F. Improved techniques for particle size determination by quasi-elastic light scattering. *Langmuir* **1,** 496–501 (1985).

19. Naiim, M., Boualem, a., Ferre, C., Jabloun, M., Jalocha, A. & Ravier, P. Multiangle dynamic light scattering for the improvement of multimodal particle size distribution measurements. *Soft Matter* **11,** 28–32 (2015).

20. Koppel, D. E. Analysis of macromolecular polydispersity in intensity correlation spectroscopy: The method of cumulants. *The Journal of Chemical Physics* **57,** 4814–4820 (1972).

21. Provencher, S. W. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Computer Physics Communications* **27,** 213–227 (1982).

22. Kaszuba, M., Connah, M. T., McNeil-Watson, F. K. & Nobbmann, U. Resolving concentrated particle size mixtures using dynamic light scattering. *Particle and Particle Systems Characterization* **24,** 159–162 (2007).

23. Griffiths, D., Hole, P., Smith, J., Malloy, A., Carr, B., Ave, M. & Mesa, C. Counting and sizing of virus and protein aggregates by nanoparticle tracking analysis (NTA). *NSTI-Nanotech* **3,** 176–179 (2010).

24. Carr, R., Hole, P., Malloy, A., Smith, J., Weld, A. & Warren, J. The real-time, simultaneous analysis of nanoparticle size, zeta potential, count, asymmetry and fluorescence. *NSTI-Nanotech* **1,** 866–870 (2008).

25. Malloy, A. & Carr, B. Nanoparticle tracking analysis -The halo system. *Particle and Particle Systems Characterization* **23,** 197–204 (2006).

26. Finder, C., Wohlgemuth, M. & Mayer, C. Analysis of particle size distribution by particle tracking. *Particle and Particle Systems Characterization* **21,** 372–378 (2004).

27. Filipe, V., Hawe, A. & Jiskoot, W. Critical evaluation of nanoparticle tracking analysis (NTA) by NanoSight for the measurement of nanoparticles and protein aggregates. *Pharmaceutical Research* **27,** 796–810 (2010).

28. Majka, M., Durak-Kozica, M., Kamiska, A., Opaliska, A., Szczch, M. & Stpie, E. The effects of subdiffusion on the NTA size measurements of extracellular vesicles in biological samples. *Biological Physics* **2,** 1–18 (2017).

29. Han, J. C. & Han, G. Y. A procedure for quantitative determination of Tris(2-carboxyethyl)phosphine, an odorless reducing agent more stable and effective than dithiothreitol. *Analytical Bio-Chemistry* **220,** 5–10 (1994).

30. Fischer, W. H., Rivier, J. E. & Craig, A. G. In situ reduction suitable for matrix-assisted laser desorption/ionization and liquid secondary ionization using Tris(2-carboxyethyl)phosphine. *Rapid Communications in Mass Spectrometry* **7,** 225–228 (1993).

31. Burns, J. A., Butler, J. C., Moran, J. & Whitesides, G. M. Selective reduction of disulfides by Tris(2-carboxyethyl)phosphine. *Journal of Organic Chemistry* **56,** 2648–2650 (1991).

32. Cruz, A., Casals, C. & Perez-Gil, J. Conformational flexibility of pulmonary surfactant proteins SP-B and SP-C, studied in aqueous organic solvents. *Biochimica et Biophysica Acta* **1255,** 68–76 (1995).

33. Perez Gil, J., Cruz, A. & Casals, C. Solubility of hydrophobic surfactant proteins in organic solvent/water mixtures. Structural studies on SP-B and SP-C in aqueous organic solvents and lipids. *Biochimica et Biophysica Acta* **1168,** 261–270 (1993).

34. Sarker, M., Waring, A. J., Walther, F. J., Keough, K. M. W. & Booth, V. Structure of mini-B, a functional fragment of surfactant protein B, in detergent micelles. *Biochemistry* **46,** 11047–11056 (2007).

35. Manzo, G., Carboni, M., Rinaldi, A. C., Casu, M. & Scorciapino, M. A. Characterization of sodium dodecylsulphate and dodecylphosphocholine mixed micelles through NMR and dynamic light scattering. *Magnetic Resonance in Chemistry* **51,** 176–183 (2013).

36. Sikorska, E., Wyrzykowski, D., Szutkowski, K., Greber, K., Lubecka, E. A. & Zhukov, I. Thermodynamics, size, and dynamics of zwitterionic dodecylphosphocholine and anionic sodium dodecyl sulfate mixed micelles. *Journal of Thermal Analysis and Calorimetry* **123,** 511–523 (2016).

37. Li, F., Li, G.-Z. & Chen, J.-B. Synergism in mixed zwitterionic-anionic surfactant solutions and the aggregation numbers of the mixed micelles. *Colloids and Surfaces A-Physicochemical and Engineering Aspects* **145,** 167–174 (1998).

38. Ghosh, S., Khatua, D. & Dey, J. Interaction between zwitterionic and anionic surfactants: Spontaneous formation of zwitanionic vesicles. *Langmuir* **27,** 5184–5192 (2011).

39. Oosterlaken-Dijksterhuis, M. A., Haagsman, H. P., van Golde, L. M. & Demel, R. A. Characterization of lipid insertion into monomolecular layers mediated by lung surfactant proteins SP-B and SP-C. *Biochemistry* **30,** 10965–10971 (1991).

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

41. Morrow, M. R., Pérez-Gil, J., Simatos, G., Boland, C., Stewart, J. & Keough, K. M. Pulmonary surfactant-associated protein SP-B has little effect on acyl chains in dipalmitoylphosphatidylcholine dispersions. *Biochemistry* **32,** 4397–4402 (1993).

42. Sharifahmadian, M., Sarker, M., Palleboina, D., Waring, A. J., Walther, F. J., Morrow, M. R. & Booth, V. Role of the N-terminal seven residues of surfactant protein B (SP-B). *PloS ONE* **8,** 1–8 (2013).

43. Bañares-Hidalgo, A., Pérez-Gil, J. & Estrada, P. Acidic pH triggers conformational changes at the NH2-terminal propeptide of the precursor of pulmonary surfactant protein B to form a coiled coil structure. *Biochimica et Biophysica Acta* **1838,** 1738–1751 (2014).

44. Joseph, M., Trinh, H. M. & Mitra, A. K. *Peptide and protein-based therapeutic agents* in *Emerging Nanotechnologies for Diagnostics, Drug Delivery and Medical Devices* 145–167 (Elsevier, 2017).

45. Marassi, F. M. & Opella, S. J. NMR structural studies of membrane proteins. *Current Opinion in Structural Biology* **8,** 640–648 (1998).

46. Warschawski, D. E., Arnold, A. A., Beaugrand, M., Gravel, A., Chartrand, É. & Marcotte, I. Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochimica et Biophysica Acta* **1808,** 1957–1974 (2011).

47. Strandberg, E. & Ulrich, A. S. NMR methods for studying membrane-active antimicrobial peptides. *Concepts in Magnetic Resonance Part A* **23A,** 89–120 (2004).

48. Lazaridis, T., Mallik, B. & Chen, Y. Implicit solvent simulations of DPC micelle formation. *Journal of Physical Chemistry B* **109,** 15098–15106 (2005).

49. Kallick, D. A., Tessmer, M. R., Watts, C. R. & Li, C. Y. The use of dodecylphosphocholine micelles in solution NMR. *Journal of Magnetic Resonance, Series B* **109,** 60–65 (1995).

50. Turro, N. J. & Yekta, A. Luminescent probes for detergent solutions. A simple procedure for determination of the mean aggregation number of micelles. *Journal of the American Chemical Society* **100,** 5951–5952 (1978).

51. García De La Torre, J., Huertas, M. L. & Carrasco, B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophysical Journal* **78,** 719–730 (2000).

52. Langer, A., Kaiser, W., Svejda, M., Schwertler, P. & Rant, U. Molecular dynamics of DNA-protein conjugates on electrified surfaces: Solutions to the drift-diffusion equation. *Journal of Physical Chemistry B* **118,** 597–607 (2014).

53. Linke, D. Detergents: An overview. *Methods in Enzymology* **463,** 603–617 (2009).

54. Liang, L., Nachtergaele, S., Seddon, A., Tereshko, V., Ponomarenko, N. & Ismagilov, R. Simple host-guest chemistry to modulate the process of concentration and crystallization of membrane proteins by detergent capture in a microfluidic device. *Journal of the American Chemical Society* **130,** 14324–14328 (2008).

55. Jean-Louis, R. & Levy, D. Reconstitution of membrane proteins into liposomes. *Methods in Enzymology* **372,** 65–86 (2003).

56. Patist, A., Oh, S. G., Leung, R. & Shah, D. O. Kinetics of micellization: Its significance to technological processes. *Colloids and Surfaces A-Physicochemical and Engineering Aspects* **176,** 3–16 (2001).

57. Patist, A., Kanicky, J. R., Shukla, P. K. & Shah, D. O. Importance of micellar kinetics in relation to technological processes. *Journal of Colloid and Interface Science* **245,** 1–15 (2002).

58. Olmeda, B., Garca-Alvarez, B. & Perez-Gil, J. Structure-function correlations of pulmonary surfactant protein SP-B and the saposin-like family of proteins. *European Biophysics Journal* **42,** 209–222 (2013).

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

60. Wustneck, N., Wustneck, R., Perez-Gil, J. & Pison, U. Effects of oligomerization and secondary structure on the surface behavior of pulmonary surfactant proteins SP-B and SP-C. *Biophysical Journal* **84,** 1940–1949 (2003).

61. Olmeda, B., Garcia-Alvarez, B., Gomez, M. J., Martinez-Calle, M., Cruz, A. & Perez-Gil, J. A model for the structure and mechanism of action of pulmonary surfactant protein B. *The FASEB Journal* **29,** 1–12 (2015).

62. Yu, Y., Wang, J., Shao, Q., Shi, J. & Zhu, W. The effects of organic solvents on the folding pathway and associated thermodynamics of proteins: A microscopic view. *Scientific reports* **6,** 1–12 (2016).

4

The Function of rSP-B Tested Using the Langmuir-Wilhelmy Balance

4.1 Statement of Authorship

Tadiwos Getachew prepared the samples used for analysis, measured, analyzed and interpreted the data, and also wrote the thesis.

4.2 Background

This chapter presents the functional testing of rSP-B with the Langmuir-Wilhelmy Balance (LWB) using DPPC/DPPG (4:1) and DPPC/POPG (7:3) membrane models. LWB was used to measure the pressure-area compression isotherms of major lung surfactant phospholipids with and without rSP-B. The two dimensional isothermal bulk compressibility modulus was then computed from the slope of the measured pressure-area isotherms. The ternary lipid mix of DPPC/POPG/PA (68:22:8 wt/wt/wt) is also discussed as it relates to synthetic exogenous surfactant preparation. There are many techniques for measuring the surface properties of lung surfactants. In order to understand the strengths and weaknesses of these methods, it is important to have some background on how surface activity can be quantified. The main techniques for assessing the surface properties of monomolecular films are the Langmuir-Wilhelmy balance, pulsating bubble surfactometry (PBS) and the captive bubble surfactometry (CBS).

LWB (Figure 4.1) is the classical model standardized in the late 1950s by John Clements when he was working on his pioneering study on lung surfactant [1]. Since its first use, the technique has been frequently used to study the rheological behavior of amphiphile molecules spread at the air-water interface. The advantage of it is that it can generate precise surface pressure (Π)-surface molecular area (A) isotherms, as one can control the number of molecules spread on a known surface area. What is more, LWB, can easily be coupled with other techniques such as microscopy and spectroscopy [2]. As such, LWB is practically useful to investigate the topography of water insoluble monolayers *in situ* or via depositing a film on solid substrates using the Langmuir-Blodgett (LB) technique. Over the years, combination of LWB with different techniques has provided invaluable information on monolayer topography, domain formation and segregation, molecular orientation and composition, as well as electrical surface potential. With respect to surfactant science, for example, LWB provided important insight into how multilayers protrude from the surface layer in a pressure dependent manner [2, 3].

Figure 4.1: A typical depiction of the Langmuir-Wilhelmey surface balance. The Langmuir-Wilhelmey surface balance is characterized by a trough (white area) onto which the film spread, two moving barriers (the two white rods at each end), the pressure sensor at the center of the trough and a temperature sensor located at the far right. The box covering ensures resistance to shake and fluctuation to temperature.

The LWB employs a Teflon trough onto which the film is spread, adjustable moving barriers at either end of the trough to compress the film, and a transducer to measure the surface pressure of the monolayer. The surface area available to a monolayer of an insoluble amphiphile molecule at the air-water interface is reduced by moving the barriers so they are closer together, normally at a constant temperature.

Compression leads to an increase in the surface pressure, Π, and reduces the normal surface tension of water (72.8 mN/m), Υ_o, down by an amount equivalent to Υ, which is the minimal surface tension achieved in the presence of surface active molecules [4], according to

$$\Pi = \Upsilon_0 - \Upsilon, \tag{4.1}$$

where Π is the surface pressure, Υ_o is the surface tension of pure water, and Υ is the surface tension attained in the presence of surface active amphiphile molecule.

The surface pressure can therefore be defined as the change in surface tension as a function of mean molecular area available to the molecules spread at the air-water interface. The interfacial tension is dependent upon the composition of the bulk solution, the surface pressure, and temperature [5]. Surface tension can be precisely quantified by measuring the change in weight when a thin, smooth, plate or rod comes in contact with a liquid. The contributing forces are the combination of buoyancy and the force of wetting (f), with gravity remaining constant [6]. The wetting force, f, is given as

$$f = \gamma P cos\theta, \tag{4.2}$$

and the total contributing force, F, thus becomes

$$F = \gamma P cos\theta - V\Delta\rho g, \tag{4.3}$$

where γ is the surface tension, P is the perimeter of the contact line, θ is the contact angle, V is the displaced volume, $\Delta\rho$ is the density difference between the air and the liquid and g is the gravitational acceleration.

At a contact angle of zero, cos(0) = 1, therefore, complete wetting of the plate/rod happens and the surface tension becomes directly correlated with the measured force. The Wilhelmy balance is an ideal system to quantify liquid surface tension using a rod of a known perimeter.

LWB might suffer from film leakage, which may decrease the calculated mean molecular area by some margin and may also prevent the film from reaching minimal tension. A number of studies reported that leakage starts to occur at a surface tension, γ of ≈18 mN/m, which corresponds to a surface pressure of ≈54 mN/m [2, 7, 8]. According to Prokop *et al.* [7] at such high surface pressure, surface active amphiphile molecules start to escape from the interface and spread onto the trough, wetting the walls and the barriers, thereby stabilizing the interfacial film. Leakage can occur both above and below the air-water interface and at the barriers. Some counteractive measures to prevent leakage are the continuous dam type ribbon trough and the use of tightly fitted barriers [9]. A study by Tabak *et al.* [10] reported that the use of ribbon barriers is freer from leakage artifacts and could give better Π-A cycling isotherms. Another challenge is that due to its open set up and larger film area than alveoli, LWB may not truly mimic *in vivo* conditions of respiration and realistic fast dynamic compression and expansion cycles could be difficult to attain. To this end pulsating bubble surfactometry (PBS) and captive bubble surfactometry (CBS) could better mimic an alveolus as both of these techniques uses an air bubble to closely simulate the air sacs.

PBS uses a capillary tube inserted into a chamber that contains ≈20 µl of surfactant to be tested. The capillary tube is used to create an air bubble inside the chamber by drawing air from the atmosphere [11]. CBS, on the other hand, seals a bubble inside a chamber filled with surfactant liquid by using a hydrophilic sealing material hence CBS alleviates leakage by preventing surfactant from escaping [8]. Although both methods better mimic an alveolus and provide fast cycling isotherms with minimal sample, these techniques also face important drawbacks. First, like LWB, PBS is also an open set up, which implies that the system is exposed to atmospheric pressure and hence can also suffers from film leakage, i.e. surfactants escape from the bubble into the capillary tube. Secondly, the surface tension (Γ) in PBS, is calculated from the Laplace equation using equation 4.4 by assuming spherical bubble. The Laplace equation is

$$P = 2\Gamma/R, \tag{4.4}$$

where P is the pressure and R is the radius of the bubble during oscillation. Comparing the P at the maximum and minimum radii provides ΔP, the pressure gradient.

The shape of the bubble, however, may not always be spherical, for example, when the Γ gets near 1 mN/m, the gravitational force becomes important and compromises the bubble shape [12, 13]. In addition, during rapid pulsating and at low surface tension, the shape of the bubble is further deformed [14, 15]. To avoid shape deformation, film compression in PBS must be fixed to only about 50% area reduction of the bubble [2], which makes it difficult to fully characterize surface rheology. CBS, on the contrary, allows a wider range of film areas and prevents surfactant leakage due to its closed setup [16]. Nonetheless, it too has some disadvantages. Unlike LWB, surface tension and the bubble area are coupled in CBS. During compression, the bubble area diminishes, which reduces the surface tension. At the same time, the smaller surface tension also tend to increase bubble area. Another disadvantage of CBS is that data analysis relies on monitoring the change in shape of the bubble that is freely diffusing in the chamber, which is cumbersome. CBS is also concentration dependent, i.e. surfactant concentration must not exceed ≈3 mg/ml when studying adsorption kinetics as that hampers measurement due to optical limitations [2]. Finally, unlike LWB, coupling CBS or PBS with other techniques such as microscopy is difficult, and hence, the content and concentration of surfactant recruited at the air-water interface can not be known precisely [17]. Hence, the phase composition or topology of surfactant cannot be studied by these methods. That being said, however, all of the *in vitro* methods discussed here are suitable to characterize the three prominent biophysical features of LS, i.e. rapid adsorption, low surface tension during compression, and fast spreadability of the film upon expansion. It is a matter of desired features and degree of accuracy required that motivate the choice of one over the other. With these caveats in mind, we found LWB to be one informative technique to test the function of rSP-B *in vitro*. To fully appreciate and interpret LWB data, in the context of lung surfactant, it is also necessary to further discuss the mean molecular area (MMA), Π-A isotherm and the two dimensional isothermal bulk compressibility modulus (β) of Langmuir monolayers.

The MMA, also known as area per molecule, is defined as the total effective area of the spreading surface divided by the number of molecules that are spread on the surface [18]. For systems that follows the additive rule, at a given surface pressure, the MMA occupied by two immiscible molecules are additive as follows:

$$A_{12} = A_1\chi_1 + A_2\chi_2 \ , \tag{4.5}$$

where A and χ are the area and mole fractions of each pure component in the system.

Deviation from the additive rule indicates miscibility and non-ideal behavior between the constituent molecules. Immiscible films exhibit distinct collapse pressures, specific to each pure component in the system. On the other hand, miscible monolayers have one collapse pressure. Thus, non-micible films can be distinguished from miscible ones by the different collapse pressures they exhibit, that is, the surface pressure at which each component in the system collapses irreversibly.

The interaction effect of mixed monolayers can also be deduced by calculating the excess free energy of mixing (ΔG_{excess}) by using equation 4.6, at a constant surface pressure and temperature, where negative values indicate an attractive interaction [19]. This free energy is given by

$$\Delta G_{ex} = \int_0^{\Pi} [A_{12} - (A_1\chi_1 + A_2\chi_2)]\delta\Pi. \tag{4.6}$$

The Π-A isotherm can be expressed in different units. Figure 4.2 shows a DPPC isotherm expressed as the area occupied by the chains or molecules of DPPC. It can also be expressed as an effective surface area of the trough or as percent area. The area occupied by the molecule is twice the area occupied by the chain for a two chain lipid.

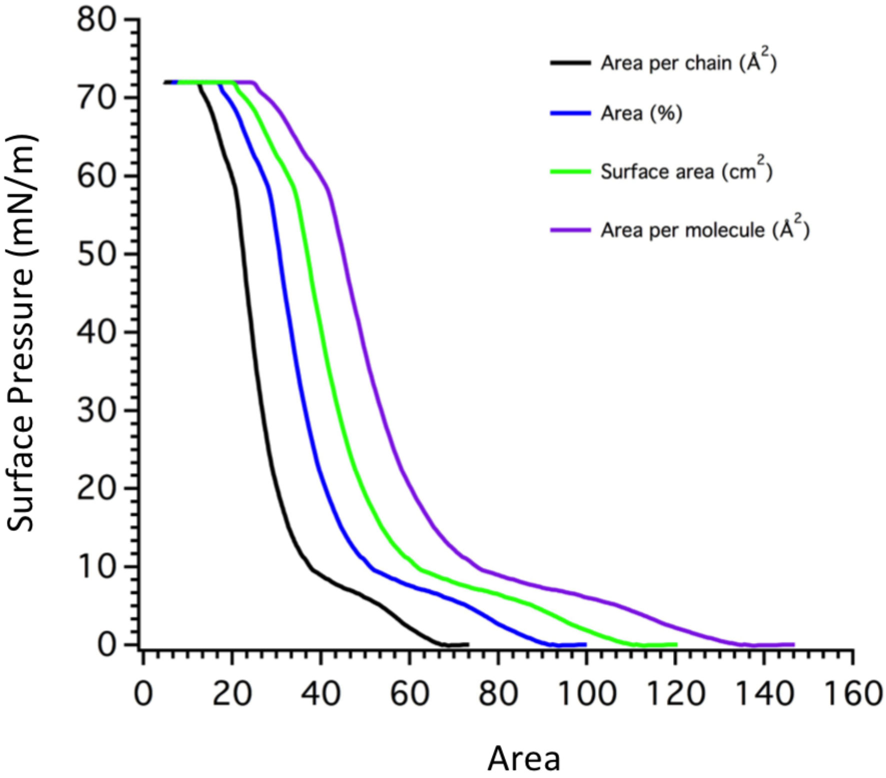


Figure 4.2: Pressure-area isotherms can be expressed in different units. In this case, DPPC isotherm is expressed in terms of area occupied by DPPC chains (black) or molecules (purple) and percent area (blue) or surface area of the trough (green). The DPPC film (1 mg/ml) was spread on an ultra pure water from 1:1 chloroform:methanol (v/v) solution. The film was allowed to stabilize for 10 minutes and the pressure-area isotherm was recorded at room temperature. The isotherm was measured by me.

The mean molecular area (MMA) can be computed according to the equations below:

$$\text{MMA} = \frac{\text{effective area of the trough } (cm^2)}{\text{number of molecules spread}}, \tag{4.7}$$

$$\text{number of molecules spread} = \frac{\text{weight spread } \ (g)}{\text{molecular weight of sample } (g/mole)} \times N_\text{A}, \tag{4.8}$$

$$\text{weight spread} = \text{volume spread } (liters) \times \text{concentration } (g/liters), \tag{4.9}$$

where N_A is the Avogadro's number (6.02×10^{23} molecules/mole). In case of binary or ternary lipid mixtures, equation 4.5 should be considered in the calculation.

The two dimensional compressibility modulus is the ability of the monolayer to withstand mechanical stress. While bulk fluids are only slightly compressible, interfaces, on the contrary, are quite compressible. The two dimensional isothermal bulk modulus, β, of a monolayer can be computed by equation 4.10 [20],

$$\beta = -A(\frac{\delta\Pi}{\delta A})_T = A(\frac{\delta\gamma}{\delta A})_T = 1/C, \tag{4.10}$$

where β is the 2D isothermal bulk compressibility modulus of the monolayer, Π is the surface pressure, A is the surface area, γ is the surface tension and C is the monolayer compressibility. Large values of C or small values of β can therefore mean that the monolayer is relatively easy to compress.

At a surface pressure or area where $\beta \to 0$ or $C \to \infty$ there will be a dip in the β-Π or β-A isotherms [20, 21]. The minimum indicates a first order phase transition [22]. In the β-Π or β-A plots, a minimum can therefore be seen when $\beta \to 0$ or $C \to \infty$. Practically, surfactant with low compressibility, C, in the alveoli reduces the surface tension without much area reduction of the air sacs. This allows a large volume of air to remain in the lungs for gas exchange [17]. Low compressibility is therefore a desirable feature of a good lung surfactant. As such, the compressibility modulus of a film can be a good indication of the quality of a given surfactant.

To illustrate the surface activity data, as well as the relationship between the surface activity and *in vivo* activity, it is useful to look at a DPPC monolayer. At a very large area per lipid molecule, the monolayer spread on the trough is in a gaseous form where the molecules are highly spread apart and have a high degree of freedom (Figure 1.7 and 4.2). In this state, the hydrophobic acyl chains of the lipid(s) make extensive contact with water molecules, which is an entropically disfavored phenomenon. This corresponds to the extreme right hand side of the isotherms in Figure 4.2, with a surface pressure of 0. As the molecules are compressed, however, the randomness decreases and the hydrophobic portion of the lipid molecules start to lift-off the water surface and the surface pressure starts to increase. As such, this state is called a liquid expanded (LE) phase, where the degree of freedom is more constricted and the molecules are closer to each other than in the gaseous phase. The LE phase has a fluid character and remains disordered for the most part. Further compression, therefore, leads to the first order phase transition to the so called liquid condensed phase (LC). Here, the molecules are tilted towards the water surface, and hence, this state is also referred as the tilted condensed phase.

The LC phase is more ordered and less compressible than the LE phase. The plateau at the low surface pressure of the isotherm on Figure 4.2 of DPPC Π-A isotherm represents the first order phase transition and LE-LC coexistent region. At sufficiently high surface pressure, the LC phase transition results in the more ordered solid (S) phase, which is a metastable state. In the S phase, also known as the untilted condensed phase, the acyl chains are perpendicular to the interface as they are closely packed to each other. This state is accompanied by a steep increase in the surface pressure (Figure 4.2). Further compression beyond this pressure leads to the collapse pressure (Πc), above which, the monolayer subjected to compression cannot stand the mounting pressure and collapses. This is marked by the plateau at the top of the Π-A isotherm of DPPC at $\approx$72 mN/m. The Πc, therefore, determines the minimum surface tension for a given monolayer. The higher is the Πc the better is the surface activity of the monolayer [4, 23, 24].

The simple DPPC/POPG (7:3) membrane system is a useful model to effectively mimic some of the important biophysical characteristics of lung surfactant [25]. The mixture contains zwitterionic PC and anionic PG head groups, as well as the disaturated, palmitoyl and the unsaturated, palmitoyl-oleoyl acyl chains. DPPC films make islands of LC lipid domains surrounded by LE phase at low surface pressure. At moderate and high surface pressure, the LC phase predominates, which makes DPPC film quite incompressible [26, 27]. The addition of unsaturated phospholipids to DPPC results in coexistence of LC and LE phases even at a relatively high surface pressure [28]. In the DPPC/POPG system, for instance, above 20°C, POPG exists in the LE phase up to its point of collapse, 48 mN/m [23, 29]. This makes the lipid mixture readily compressible at physiological temperature. On another note, incorporating palmitic acid (PA) to DPPC/POPG film has been shown to modify the mechanical properties of the DPPC/POPG monolayer [30]. Studies have shown that PA increases the order of DPPC domains in the condensed phase hence increasing the solid portion of the monomolecular film at increasing surface pressure [31]. As such, PA may have the ability to regulate the viscosity of the monomolecular film.

One parameter of particular interest to LS function is the equilibrium spreading surface pressure (Π_e). Π_e is the surface pressure when phospholipid vesicles in the aqueous hypophase reach equilibrium with the interfacial monolayer [2, 32]. In the alveolus, this happens at a Π of $\approx$45 mN/m or γ of $\approx$25 mN/m [33]. *De novo* adsorption and spreading of secreted surfactant or clinical surfactants occurs at the Π_e. It is therefore imperative that surfactant rapidly adsorbs and spreads to reach equilibrium within a couple of seconds. A plateau at the Π_e in the pressure-area isotherm is a clear mark of this. Above Π_e, monolayers are in a state of supersaturation and only lateral compression brings about higher Π.

In general, lipid films containing surfactant proteins exhibit four distinct regions on their Π-A compression isotherm (Figure 4.3): I) monolayer regions, $0 \leqslant \Pi \leqslant 40$ mN/m, II) monolayer to multilayer transition region, $40 \leqslant \Pi \leqslant 50$ mN/m, III) multilayer region, $50 \leqslant \Pi \leqslant 70$ mN/m and IV) the collapse region, $70 \leqslant \Pi \leqslant 72$ mN/m [25].

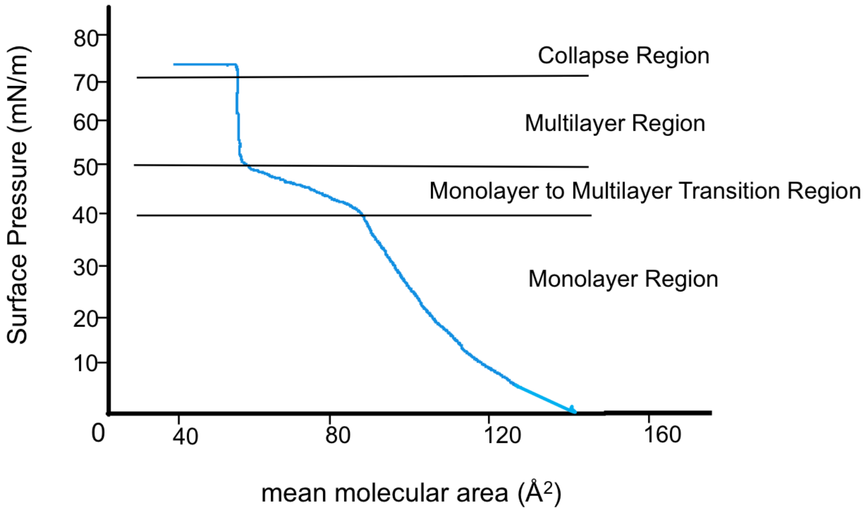


Figure 4.3: A generalized sketch of the pressure-area isotherm of lung surfactant lipids plus the small hydrophobic proteins. The pressure-area isotherms of lung surfactant lipids can generally be divided into four main different regions in the presence of the small hydrophobic proteins *in vitro*. The monolayer region, the monolayer to multilayer transition region, the multilayer region, and the collapse region.

The metastable region ($50 \leqslant \Pi \leqslant 70$ mN/m) of LS, is probably the physiological relevant range. This was corroborated by Schurch *et al.* [34], who measured surface tension in excised rat alveoli at total lung capacity and demonstrated that the surface tension in alveoli cycles between 0 to 30 mN/m, i.e. Π of 40 to 70 mN/m. In contrast, the collapse region probably is non-ideal in mammalian lungs as it is not energetically favourable but if it occurs, it is probably transient at end-expiration [25]. The detailed mechanism for how this is achieved *in vivo* is not clearly understood. A large body of data indicates that SP-B and SP-C contribute to the stabilization and maintenance of lipid films at the interface in the metastable region. It has been shown, by *in vitro* and *ex vivo* experiments, that the two proteins interact with lipid molecules at the interface to promote lipid rearrangement and phase changes. SP-B in particular, as discussed in the introduction, happens to be essential to life. The effect of SP-B *in vivo* has also been explained with respect to the line tension and dipole density of the monomolecular film.

SP-B increases the line tension and dipole density difference between lipid phases within the monomolecular film [20]. Line tension, which is the energy per unit length, arises when the edge of a monolayer or a liquid droplet contributes an extra positive or negative energy to the total surface energy. The dipole density, on the other hand, is the charge accumulated in surfactant or lipid headgroups of a monolayer within domains that leads to a difference in surface potential between domains of coexisting phases [35, 36]. The line tension between domains and the dipole density difference between phases are two opposing forces that govern the size and distribution of lipid domains at equilibrium [37, 38]. Because SP-B increases the line tension and dipole density, lipid domains that form in the presence of SP-B are smaller in size and monodisperse [39]. Consequently, monolayers with SP-B present are able to reach Π_e faster and with lesser area reduction than films without SP-B. The 2D bulk compressibility modulus also increases in the presence of SP-B up to the squeeze-out pressure [20].

The interaction of the proteins with the lipids have also been studied *in vitro* extensively. In mixed PC/PG layers, SP-B and SP-C were found to preferentially interact with and order the acyl chains of PG when used in the physiological amount, that is, 1-2 wt%. No preferential interaction between SP-B/C and PC were observed. PC and PG thus appear to be affected differently by the two proteins [40]. SP-B, however, is superior to SP-C in helping the layer to reach the equilibrium spreading surface pressure quickly as well as in recruiting and packing of LS lipids at the interface effectively [41]. The combined use of SP-B and SP-C was found to be more efficient than either one of the proteins alone. This suggests synergism between the two hydrophobic proteins. The composition of the monomolecular film, therefore, reflects on the ability to sustain the film on repeated cycles as it would happen in the alveoli.

Overall, the currently accepted explanation is that monolayers containing surfactant protein(s) undergo reversible folding, where squeezed multilayer patches form surfactant reservoirs that replenishe the interface upon film expansion [42]. Readsorption of the protrusion mass into the monolayer has been shown to occur below the LC-LE phase transition pressure of the lipid(s) at the interface on the expansion cycle, for example ≈30 mN/m in case of DPPG film at 30°C [43]. Apparently, this expansion plateau was not observed for films containing only SP-C in captive bubble experiments [41], which further underlines the necessity of SP-B for promoting film replenishment. This observation also further emphasizes that the content of the film matters. As clinical surfactant preparations differ in composition (i.e. surfactant proteins (SP-B or SP-C), phospholipid, cholesterol, and additives) they also vary in dose and frequency of administration.

Nonetheless, it was apparent that, during *in vitro* characterization of clinical surfactants, Curosurf, Survanta, Infasurf and BLES all exhibit a plateau region between surface pressures of 40 to 50 mN/m on compression, indicating a 2D to 3D transition that stems from the LE phase [25]. Moreover, all but Survanta were found to support near zero surface tension. Survanta, however, collapses at 62 mN/m. In essence, a good LS should be less compressible and able to reach minimal surface tension quickly with a small area reduction (≈15 to 20%, *in vivo*) leaving a large volume of air and enough space for gas exchange [17]. In this chapter the objective is to assess whether rSP-B possesses the requisite biophysical features of good LS protein when incorporated into simple LS lipids *in vitro*.

4.3 Materials and Methods

Surface pressure-area isotherms were recorded using a Kibron MicroTrough XS (59 mm X 208 mm, Helsinki, Finland) using both a ribbon-trough and micro-trough plates that were controlled by software version 3.6.1. All lipids used were purchased from Avanti Polar Lipids. The model lipids were mixed weight by volume to a final concentration of 1 mg/ml in 1:1 chloroform:methanol (v/v). 10 µl of the lipids in chloroform:methanol (1 mg/ml) were added drop-wise to ultra pure water (18.2 MΩ) using a microsyringe (Hamilton, NV, USA) from the air side, unless stated otherwise.

Recombinant SP-B was prepared as per section 2.2.4 with exchange into the lipid mix of choice (lipid concentration 1 mg/ml) while the protein was on the second stage of IMAC resin. The protein was eluted by pH gradient and was quantified using the Bradford assay. An appropriate volume of the eluted protein and lipid sample was taken from the stock to make the desired final rSP-B concentrations (2, 4, 6, 8 and 10 wt% of the protein per 1 mg/ml of the lipids: as the physiological *in vivo* SP-B concentration is close to 2%, measuring a range of rSP-B concentrations will help assess the functionality of the protein) in 1:1 chloroform:methanol (v/v) solution. To ensure proper mixing, the aliquots were gently and briefly shaken. Ten µl of the sample (protein with the lipids) was spread from the air side drop-wise using a microsyringe in the same manner as was done for lipids alone. A minimum of 10 minutes was given for the organic solvent to evaporate and the film to stabilize. The change in surface pressure was monitored by the sensor (0.2 microgram resolution and sensitivity 0.01 mN/m). To ensure a contact angle of zero, the dyneprobe (perimeter of 1.60 mm and mass of 10 mg) was cleaned with ethanol and flamed in between measurements. All measurements were taken at ambient temperature and the rate of film compression was adjusted as required. Filtering and smoothing of the two dimensional isothermal bulk compressibility modulus of the Π-A isotherms were done using the built-in functions of IGOR Pro 7 (Wavemetrics, Portland, Oregon).

4.4 Results

Before testing rSP-B in lipid models, I measured the Π-A isotherms of lung surfactant phospholipids and confirmed that the isotherms behave as reported in the literature. However, it is also imperative to note that without the added stabilization of the film material due especially by the two small hydrophobic proteins, SP-B and SP-C, lung surfactant lipids lose film material on repeated compression. The isotherms presented herein are only the first Π-A compression isotherms. All of the first Π-A compression isotherms recorded were found to be reproducible between repeated measurements i.e for freshly applied sample.

The Π-A isotherms of the LS lipids are shown in Figure 4.4. The results indicate that DPPC reached a surface pressure of 72 mN/m while DPPG reached ≈60 mN/m, whereas the unsaturated lipids, POPC and POPG, both reached a surface pressure of ≈48 mN/m. The lift-off, which is the area where the surface pressure starts to increase, for both saturated and unsaturated PC, occurs at a larger mean molecular area than for PG. DPPC lifts-off at a MMA of 130 $Å^2$ and POPC at 140 $Å^2$. Likewise, DPPG lifts-off at a MMA of 80 $Å^2$ and POPG at 110 $Å^2$. At ambient temperature, only DPPC exhibits a plateau below a surface pressure of 10 mN/m, between MMA of ≈110 to 80 $Å^2$. According to the literature, the plateau corresponds to the LE-LC co-existent phase [21, 22, 44].

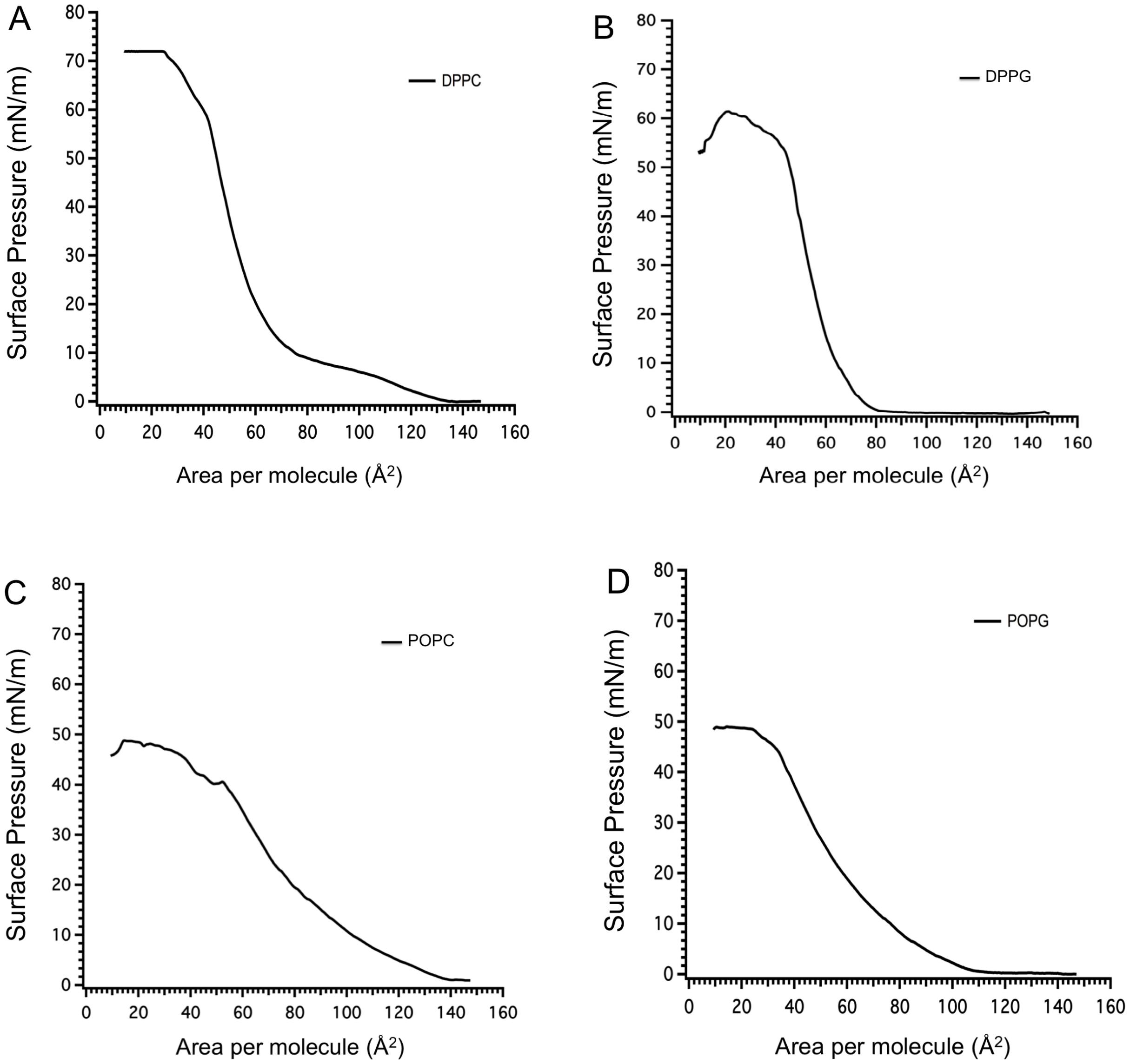


Figure 4.4: Lung surfactant phospholipids exhibit distinct pressure-area isotherms. Panel A to D are isotherms of DPPC, DPPG, POPC and POPG respectively. Films were compressed at room temperature and on ultra-pure water sub-phase at a speed of 2 cm^2/min. Measurements were repeated and the first measurements were found to be reproducible between samples and across batches. The first and second isotherms of all lipids samples however show hysteresis, consistent with irreversible loss of sample.

The same data can also be presented as the 2D bulk compression modulus (β), which is a measure of the incompressibility of the film and emphasizes regions of the compression isotherms where the surface pressure isotherms change quickly. The 2D bulk modulus was calculated from the slope of the pressure-area isotherms for each LS lipid presented in Figure 4.4 as per equation 4.10. The 2D bulk moduli of the LS lipids are shown in Figure 4.5 as a function of mean molecular area and in Figure 4.6 as a function of surface pressure.

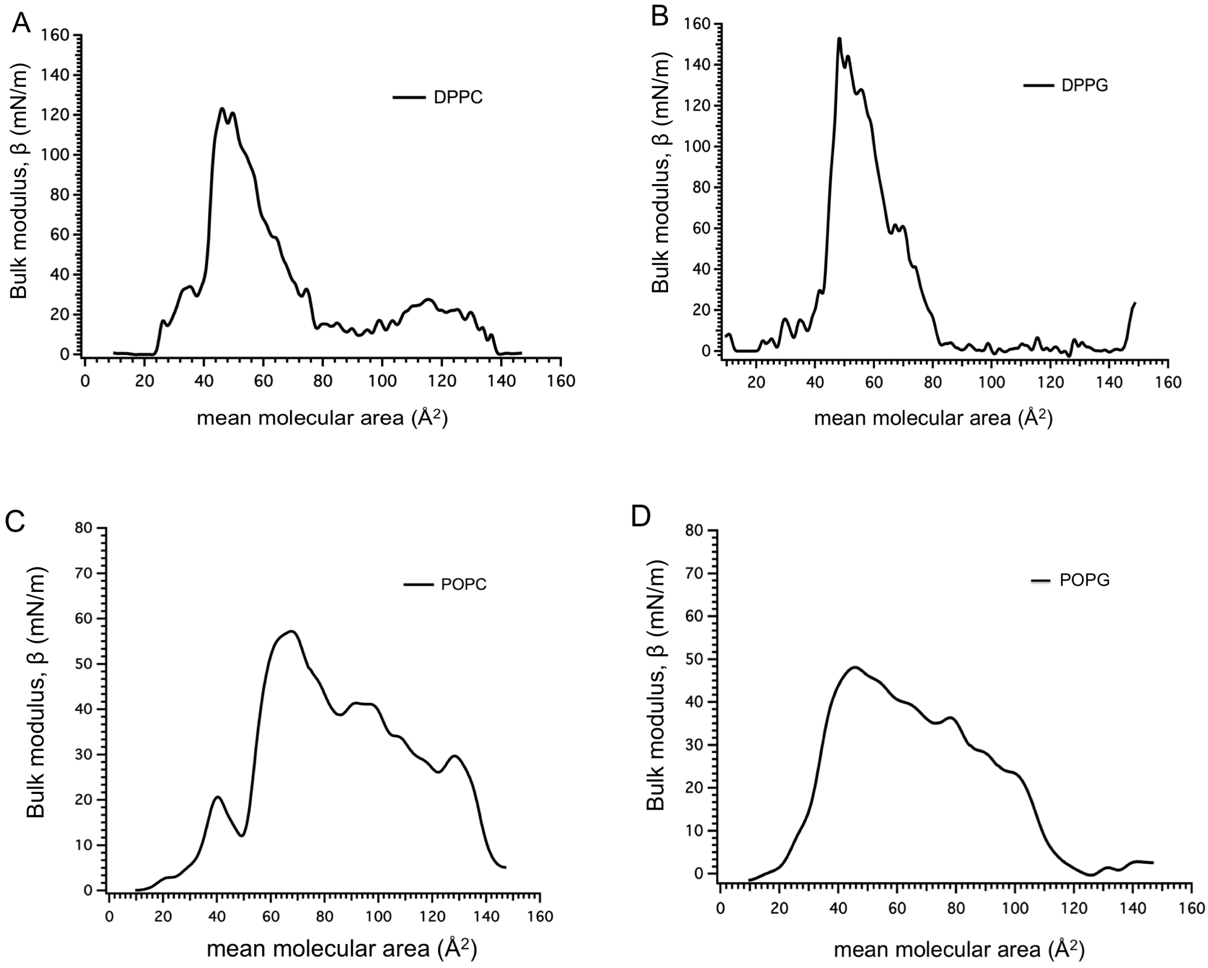


Figure 4.5: Lung surfactant phospholipids exhibit characteristic isothermal surface bulk moduli. The 2D bulk modulus of each isotherm shown in Figure 4.4 is plotted against the mean molecular area. The highest bulk moduli are observed for the saturated lipids, DPPG and DPPC. The unsaturated lipids, POPC and POPG, on the contrary, display lower bulk moduli. The minimum between MMA of ≈110 to 80 Å^2 of DPPC corresponds to LE-LC co-existence plateau observed in the pressure-area isotherm.

The co-existence region, which is the surface pressure or the molecular area where the LE-LC phases of DPPC molecules co-exist together, is apparent in the local minimum on the DPPC β-A graph (Figure 4.5A) between ≈110 to 80 Å^2, as well as by the dip at a surface pressure of ≈8 mN/m in the β-Π plot (the green line) in Figure 4.6. DPPG with a maximum β of ≈160 mN/m is stiffer than DPPC with a β of ≈120 mN/m. Both of these lipids reached their maximal compression moduli at a surface pressure of ≈43 mN/m or MMA of ≈50 Å^2. As the surface area was further decreased, the monolayers of these lipids relax. The phenomenon is apparent in the bulk modulus plot as the β decreases on further compression. Such relaxation is attributed to the transition of the film from a 2D monolayer to 3D aggregates possibly via a nucleation and growth mechanism. Both films eventually collapse at a sufficiently high surface pressure. On the other hand, POPC reached its β_{max} at a larger MMA of ≈65 Å^2 than POPG, which reached the peak of its bulk modulus at a MMA of ≈45 Å^2. Nonetheless, both POPC and POPG are most incompressible at about the same surface pressure, ≈30 mN/m.

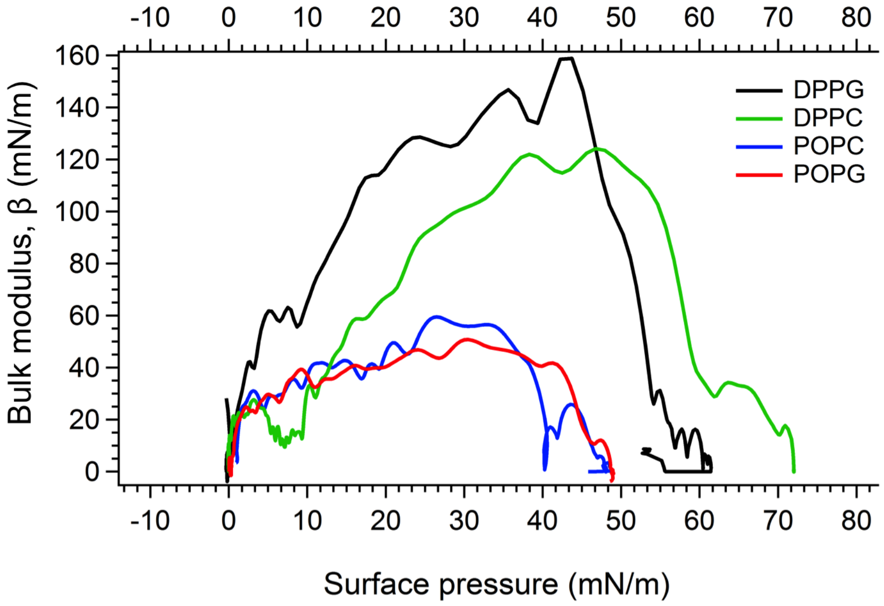


Figure 4.6: The 2D bulk modulus of each isotherm shown in Figure 4.4 plotted against surface pressure. DPPG has the highest bulk modulus even though it has a smaller collapse pressure than DPPC.

As one of the model lipid mixtures used to test the function of rSP-B in this study is DPPC/DPPG at a 4:1 molar ratio, I also measured the Π-A isotherm and calculated the β of that lipid mixture. It was found that the DPPC/DPPG film also achieves near zero surface tension similar to DPPC. Nonetheless, the film was observed to be incompressible at a larger MMA of ≈75 Å^2 as is shown in Figure 4.7B, as opposed to the films of DPPC or DPPG alone that reached the same state at ≈50Å^2 as shown in Figures 4.5A and B respectively.

The β_{max} of DPPC/DPPG, 140 mN/m, is below that for pure DPPG and above that of pure DPPC. All three films, however, reached β_{max} at a similar surface pressure, ≈43 mN/m. Next, I added rSP-B to DPPC/DPPG lipid mix and probed the changes due to the presence of the protein.

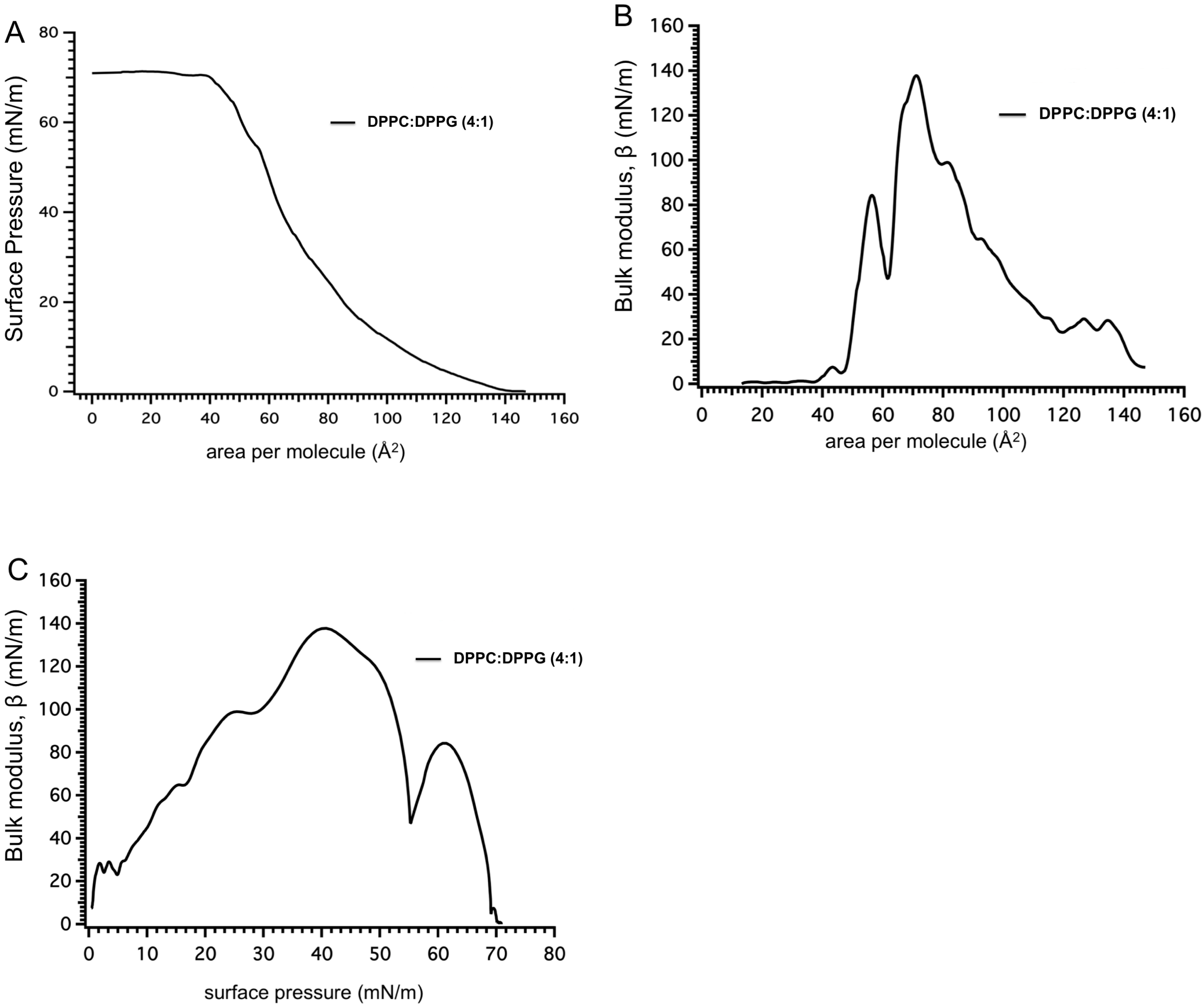


Figure 4.7: Pressure-area isotherm and the 2D bulk modulus of DPPC/DPPG (4:1) lipid mixture. Panel A: Π-A isotherm of DPPC/DPPG (4:1). Panel B and C are the 2D bulk modulus plotted against the area per molecule and the surface pressure respectively.

For the samples with the protein, rSP-B was exchanged into the desired lipid mix and eluted with a lower pH. Before adding the lipids to the column, it was necessary to subject them to five freeze-thaw cycles. Repeated freezing and thawing of the lipids will help to disintegrate the large unilamellar vesicles into smaller vesicles so that the lipids do not plug up the column. After exchanging into the desired mix on the column, the protein was eluted with the same lipid-mix with the pH adjusted from 7.5 to a pH of 5.

The addition of 2 wt% rSP-B to DPPC/DPPG shifts the isotherm to the right (Figure 4.8A) as compared to the control, DPPC/DPPG (Figure 4.8B). The shape of the isotherm with rSP-B is very different from the shape with DPPC/DPPG alone.

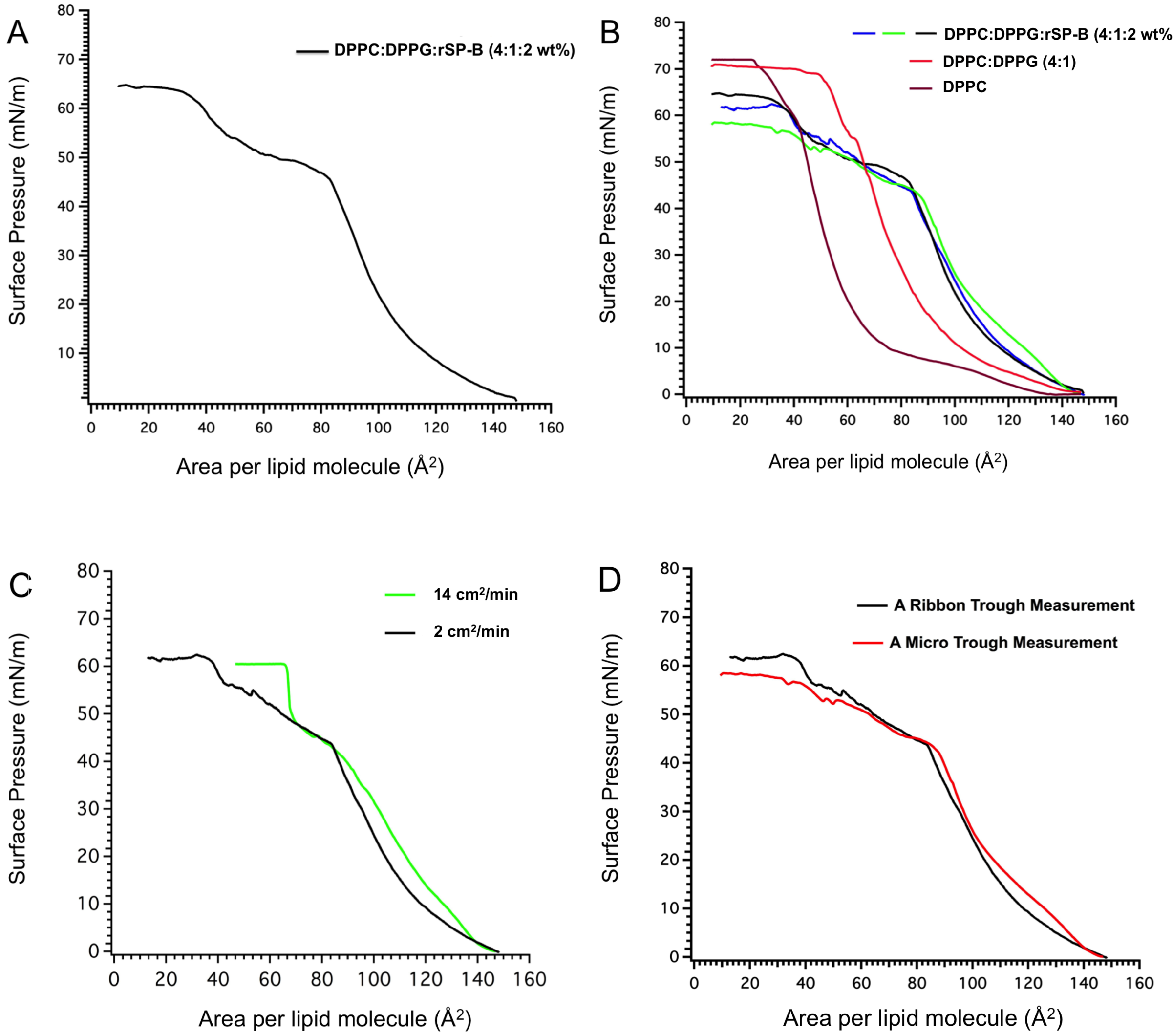


Figure 4.8: rSP-B induces a reproducible alteration of the DPPC/DPPG isotherm in between surface pressures of 40 to 50 mN/m. Panel A: pressure-area isotherm of DPPC/DPPG/rSP-B (4:1:2 wt%), Panel B: overlay of isotherms with and without rSP-B, Panel C: overlay of isotherms compressed at different compression speeds and Panel D: overplot of a ribbon-trough and a micro-trough measurement.

Firstly, the mixture with the protein added, lifts off quickly and reaches the Π_e at a large area per lipid molecule when compared to either one of the DPPC/DPPG or DPPC films. Secondly, with rSP-B present, there is an obvious inflection point at ≈90 $Å^2$. Above this point, the surface pressure largely levels off, with a linear, slight increase in surface pressure as the area is decreased. This high pressure plateau is not seen in the lipid only scans at all. Excitingly, as will be discussed in the Discussion section, 2 wt% of rSP-B may be facilitating the formation of multilayer lipid structures beneath the interfacial layer *in vitro* [29, 39, 45]. The plateau was observed for repeated measurements (Figure 4.8B), i.e. for the first compression isotherm of different protein preparations. The plateau was also apparent when the film was compressed at different compression speeds. Figure 4.8C demonstrates the Π-A isotherm for compression at rates of 14 cm^2/min and 2 cm^2/min with both displaying an inflection point at ≈90 Å^2 and a plateau at smaller areas. At the higher compression speed, there is a jump up to the maximum surface pressure at ≈70 Å^2. Moreover, the results were also repeatable with different troughs, for instance, with a ribbon and a micro-trough of equal dimensions (both from Kibron) as shown in (Figure 4.8D). These initial experiments showed very promising behaviour for recombinant SP-B. The next step was to look at the same data from the point of view of the 2D bulk modulus.

DPPC/DPPG/rSP-B (Figure 4.9B and C) reached the maximum β at a larger MMA of 100 Å^2 or at a lower surface pressure of ≈35 mN/m when compared with the lipid mixture that reached the β_{max} at a smaller MMA of 70 Å^2 or at a higher Π of 45 mN/m. Already at 45 mN/m, the monolayer with rSP-B has released the in-plane stress, perhaps due to growth of 3D aggregates. Put another way, the inverse of β, the compressibility (C), of the film in the presence of rSP-B is high at 45 mN/m. *In vivo*, this is the surface pressure at which the film at the interface and in the bulk phase reach equilibrium and starts to spread [33, 46]. Conversely, the lipid mixture without rSP-B, at the same surface pressure, is at its least compressible state i.e. the peak of its β. Also evident in Figure 4.9D is that the film compressed at 14 cm^2/min (green line) did not show the same dip at 120 Å^2 when compared to the film compressed at 0.4 cm^2/min (black line). This may indicate the that the rate of film compression may affect the phase behaviour of lung surfactant. To this end, an AFM study has shown that BLES compressed at different rate shows different domain morphologies. The film compressed at high speed shows many small nanodomains compared to the film compressed at a slow rate [47]. Phase equilibration in monolayers occurs at a slow rate and hence faster compressions tends to disturb domain morphologies more, before equilibrium is reached, than slow compressions [47]. The small kink (marked by an asterisk) on the Π-A isotherm at ≈56 mN/m (Figure 4.9A, red line) or the sharp dip (also shown in asterisk) in the bulk modulus of the lipid alone mixtures (Figure 4.9B, C and D, red lines) is likely due to well characterized film leakage [2, 7, 8, 48]. To minimize this effect, a ribbon dam type trough was used for measuring cycling isotherms.

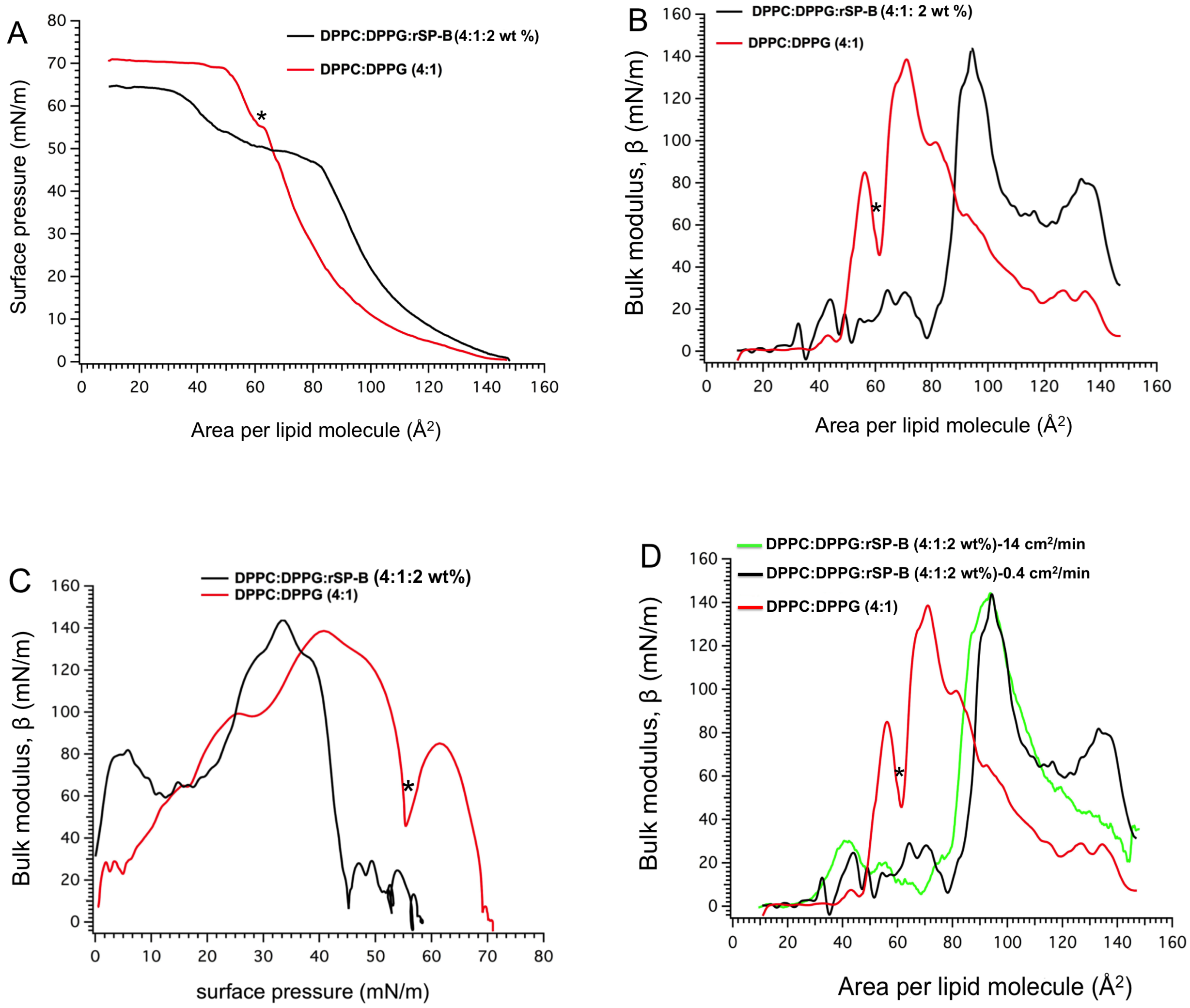


Figure 4.9: The effect of 2 wt% of rSP-B on the phase behaviour and lateral integrity of DPPC/DPPG (4:1) film. Panel A indicates the pressure-area isotherm shifts to larger MMA until the Π_e with rSP-B present. Likewise, Panel B shows the film with rSP-B reach the incompressibility limit at the larger MMA and before the Π_e than DPPC/DPPG. Panel C displays the monolayer in the presence of the protein (black) undergoes a phase change at a $\Pi \approx 13$ mN/m (marked by the minimum of the dip at $\beta \approx 60$ mN/m), which suggests that rSP-B insert into DPPC/DPPG film and affect its packing order. Panel D demonstrates the rate of film compression may have an effect on the phase behaviour of LS lipids. The dip observed at the MMA of $\approx$120 Å^2(black) was not observed when the monolayer was compressed at high compression speed (green line).

The second lipid model I used to test the function of rSP-B, DPPC/POPG simulates some of the biophysical features of LS better than DPPC/DPPG. Due to POPG's fluid nature, DPPC/POPG system is more suitable to study film refinement and reinsertion of LS [25]. Like DPPC/DPPG, it was observed that DPPC/POPG (7:3) also supports a surface pressure of ≈72 mN/m (Figure 4.10A). The β-A/Π plots exhibits a plateau region between MMA values of 100-80 Å^2 (Panel B) or Π of 13 to 23 mN/m (Panel C) respectively.

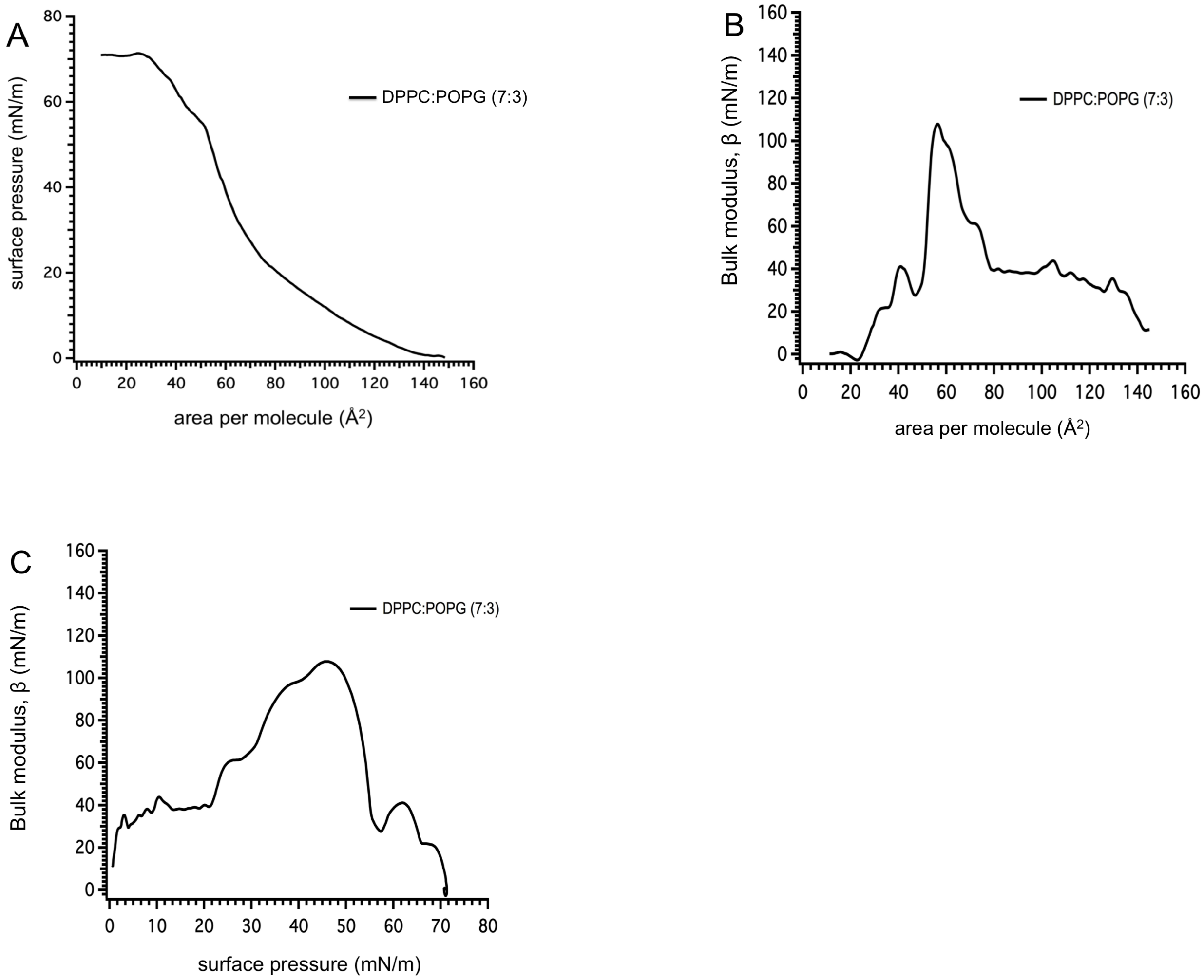


Figure 4.10: The pressure-area isotherm and bulk moduli of DPPC/POPG (7:3) lipid mixture. The β-A and the β-Π plots depict a plateau that is not clearly reflected in the Π-A isotherm. In addition the data indicate that DPPC/POPG (7:3) exhibits a lower maximum 2D bulk modulus than DPPC/DPPG (4:1), which suggests that the former is more fluid and therefore more compressible than the latter.

It has been demonstrated that LS films compressed to a surface pressure ≥30 mN/m show phase segregation, with a continuous LE phase surrounding islands of LC lipid domains [45]. The plateau observed here in the β-Π plot suggests that squeeze-out of fluid PG starts to occurs as early as a surface pressure of 13 mN/m and continues up to a Π of ≈23 mN/m. The β_{max} (110 mN/m) recorded for the DPPC/POPG film was at a MMA of ≈60 Å^2 or surface pressure of ≈48 mN/m.

To test the hypothesis that rSP-B stabilizes and promotes film reinsertion over repeated compression and expansion isotherms, the cyclic pressure-area isotherms of DPPC/POPG film were recorded with and without rSP-B. In the whole lung inspiration result in increase in total volume and surface area of the alveoli and expiration causes a decrease of these. To derive useful information from such changes in surface area, cyclic pressure-area isotherms are measured *in vitro*. In this study, the same film spread was subjected to a number of cycling isotherms, i.e. both the compression as well as the expansion Π-A isotherms were measured. The cycles were then compared to each other and to that of the lipid-only control. For the film with rSP-B, the plateau in the expansion curve at surface pressure ≈25 mN/m, as marked by the arrow in Figure 4.11A, is particularly intriguing as it likely denotes that rSP-B promoted readsorption of film material back into the interface.

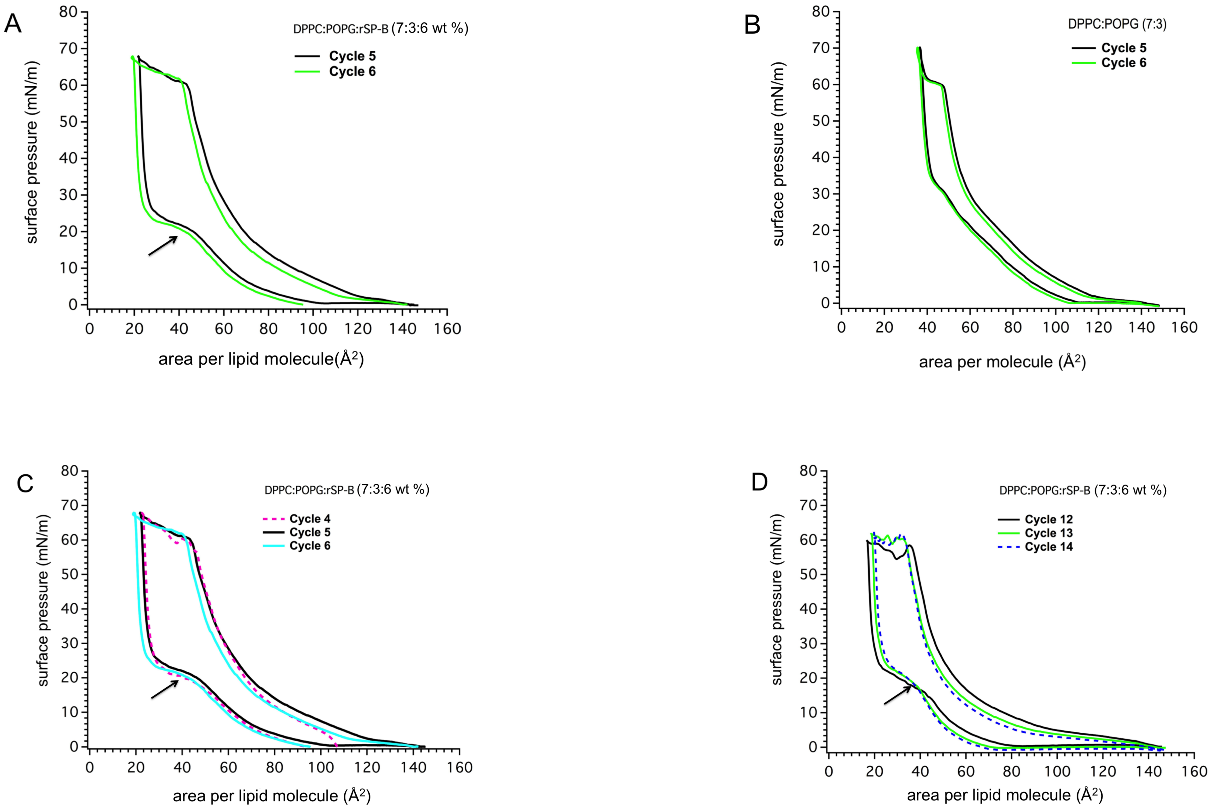


Figure 4.11: Six wt% of rSP-B facilitates film readsorption when incorporated in DPPC/POPG (7:3) film. Repeated compression and expansion pressure-area isotherms of DPPC/POPG/rSP-B film (Panel A) and the lipid-alone film, DPPC/POPG (Panel B). Panel C and D indicates the reproducibility of the film (DPPC/POPG/rSP-B) replenishment when subjected to more repeated cycles. The protein sample was exchanged into the same lipid mixture (1 mg/ml) used as control. Both the protein containing sample and the lipid-alone sample were spread at the air-water interface only once before recording the repeated compression and expansion cycles.

The plateau was not prominent in the lipid alone film (Figure 4.11B). More importantly, the isotherms of rSP-B containing films does not show evidence of aging after repeated cycles. (Figure 4.11C and D). On the contrary, the control displayed aging of the film, i.e. between the first and last cycles measured (data not shown). Hence, rSP-B (6 wt%) must be facilitating film readsorption when incorporated in DPPC/POPG (7:3) lipid membrane *in vitro*. The behaviour of the rSP-B containing lipid films under cycling is very encouraging as SP-B has been proposed to be critical in LS film replenishment during the expansion and compression breathing cycles [41].

Another way of interpreting the same data is looking at the 2D isothermal bulk modulus of the compression and expansion cycles. The 2D bulk modulus of cycle 5 of DPPC/POPG/rSP-B is shown in Figure 4.12B and C.

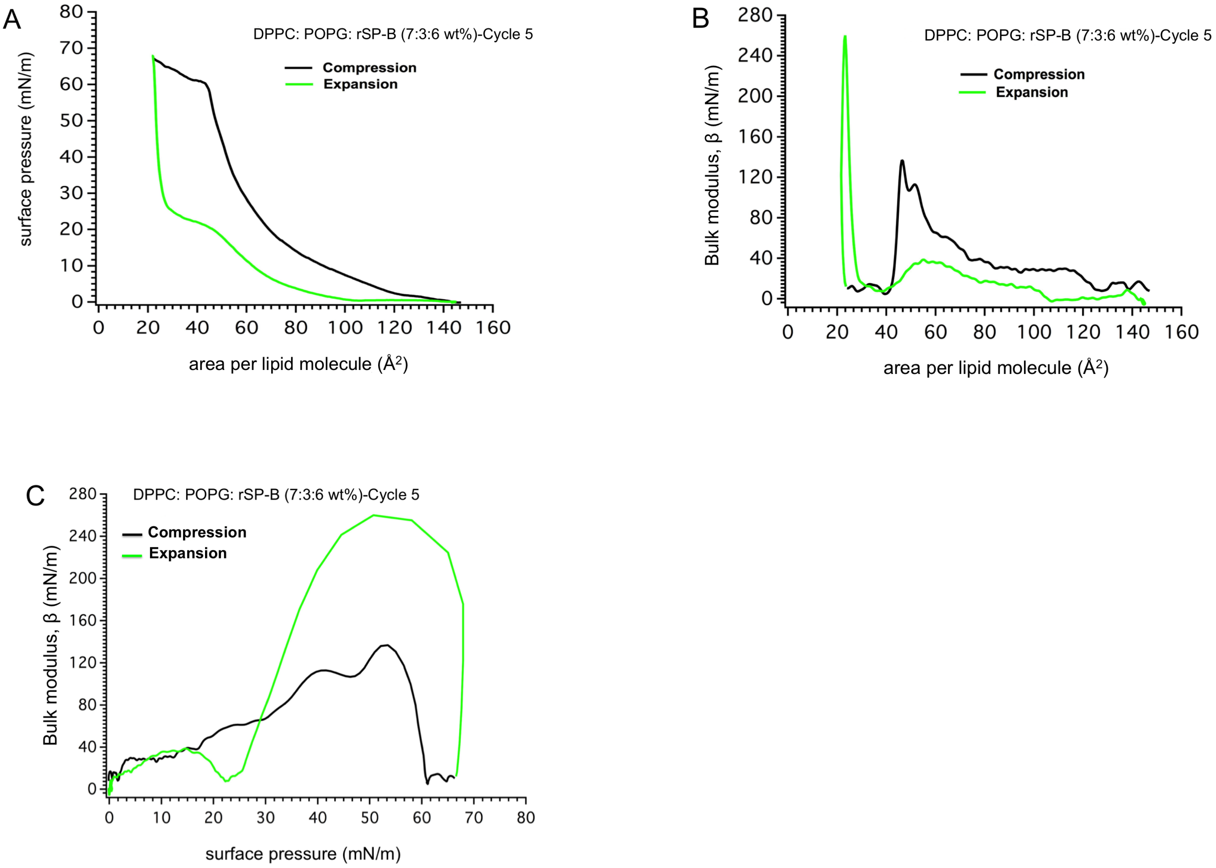


Figure 4.12: Cyclic pressure-area isotherm (Panel A) and bulk moduli of DPPC/POPG/rSP-B (7:3:6 wt%) film expressed as a function of mean molecular area (Panel B) and surface pressure (Panel C). At a mean molecular area of ≈50 Å^2 and bulk modulus of ≈130 mN/m (LC phase), the monolayer, as can be seen from the Π-A isotherm, reaches a collapse pressure of ≈61 mN/m. At this surface pressure the film collapses, possibly by a reversible folding mechanism. Further compression increases the surface pressure to ≈68 mN/m. On the expansion cycle, for a brief moment, the packing density is high with the 2D bulk modulus as high as ≈260 mN/m, which is characteristic to solid (S) phase.

The plateau with rSP-B present begins at a larger molecular area of 120 $\mathrm{\AA}^2$ in Panel B or smaller surface pressure of about 5 mN/m in Panel C. This is in contrast to the lipid-only film with a plateau starting at 100 $\mathrm{\AA}^2$ and a surface pressure of 13 mN/m (Figure 4.10B and C). This illustrates that with rSP-B present, the squeeze-out starts early during compression and also occurs over a large range of molecular area. Moreover, the observation could suggest that more film material leaves the interface when there is rSP-B and hence indicate an ability of rSP-B to refine lipid composition of the layer. At the surface pressure of 48 mN/m, which is the collapse pressure of pure POPG, the bulk modulus in the presence of rSP-B (Figure 4.10C, black line) indicates a slight dip. The increase in the β value from 108 to 130 mN/m in between the surface pressure 48 to 55 mN/m could be due to the increase in the order of the acyl chains of the lipid(s) at the interface on further compression.

At a surface pressure of ≈55 mN/m, the film reaches high packing order. The monolayer at this point is incompressible and thus a further adjustment is necessary to counteract the increase in surface pressure. Between the surface pressure of 56 and 61 mN/m, it can be seen that the compressibility of the film increases, which demonstrates a modification within the film that allowed relaxation of the monolayer. One possible explanation as to why the compressibility increases with increasing surface pressure could be that the film buckles. A similar explanation was given for DPPC monolayers where undulation of the film was reported to occur at a surface pressure of 55 mN/m [49]. At ≈61 mN/m, the monolayer collapses reversibly (Figure 4.12A and C, black lines). Although the surface pressure gradually increases to ≈68 mN/m upon further compression, the film compressibility between the onset of the reversible collapse at 61 mN/m and at 68 mN/m, before film expansion begins, remains nearly constant and high. To see if the process is reversible, the 2D bulk modulus of the expansion cycle was also interpreted.

On the expansion cycle (Figure 4.12B and C, green lines), the bulk modulus increases very rapidly to a β of ≈260 mN/m, which is characteristic of a solid phase. Immediately after, the 2D bulk modulus drops to a β of ≈10-15 mN/m, which is characteristic of a liquid expanded phase. The surface pressure also drops quickly from 68 mN/m to ≈25 mN/m over a small area change (Figure 4.12A and B, green lines). The rise and drop in the bulk modulus and the decrease in the surface pressure are coupled and occurs very quickly. In laboratory settings, it happens within 2-3 seconds. Importantly, the highest 2D bulk modulus recorded on film expansion is nearly at the same surface pressure that the film reaches the highest packing density on compression, ≈55 mN/m. The later and slower event is the reinsertion process of the protrusions that are either associated with the interface or are in close proximity to the surface layer. This process commences at a surface pressure of ≈25 mN/m (Figure 4.12C, green lines) or mean molecular area of 42-58 $Å^2$ (Figure 4.12B, green lines) and covers a wide area change. Evidently, this event occurs when the monolayer is in the LE phase (β of ≈10-15 mN/m), which is in agreement with the literature data. Likewise, this process also seems to unfold at a similar surface pressure to that for which the squeeze-out of the lipid(s) was observed in the β-A and β-Π plots of the lipid-only film as shown in Figure 4.10B and C. Overall, the results indicates that DPPC/POPG/rSP-B (7:3:6 wt%) film releases lateral strain by two mechanisms: formation of protrusion before the Π_e and with increasing surface pressure, by undergoing a reversible folding collapse after the Π_e.

Also of interest was the mixture of DPPC/POPG/PA (68:22:8 wt/wt/wt). Palmitate, which is the salt of palmitic acid is a common additive of some clinical exogenous surfactants. The two hydroxyl groups of PA hydrolyse to give an anionic head group, hence, electrostatic and other weak interactions might underlie some of the behaviour of PA in LS. The ternary lipid mixtures of DPPC/POPG/PA, when spread from four times more volume (40 µl at 1 mg/ml) than DPPC/POPG/rSP-B (10 µl also at 1 mg/ml), demonstrates a slight inflection point at ≈45 mN/m on compression and consistent film replenishment plateau at ≈30 mN/m on film expansion as displayed in Figure 4.13A and B.

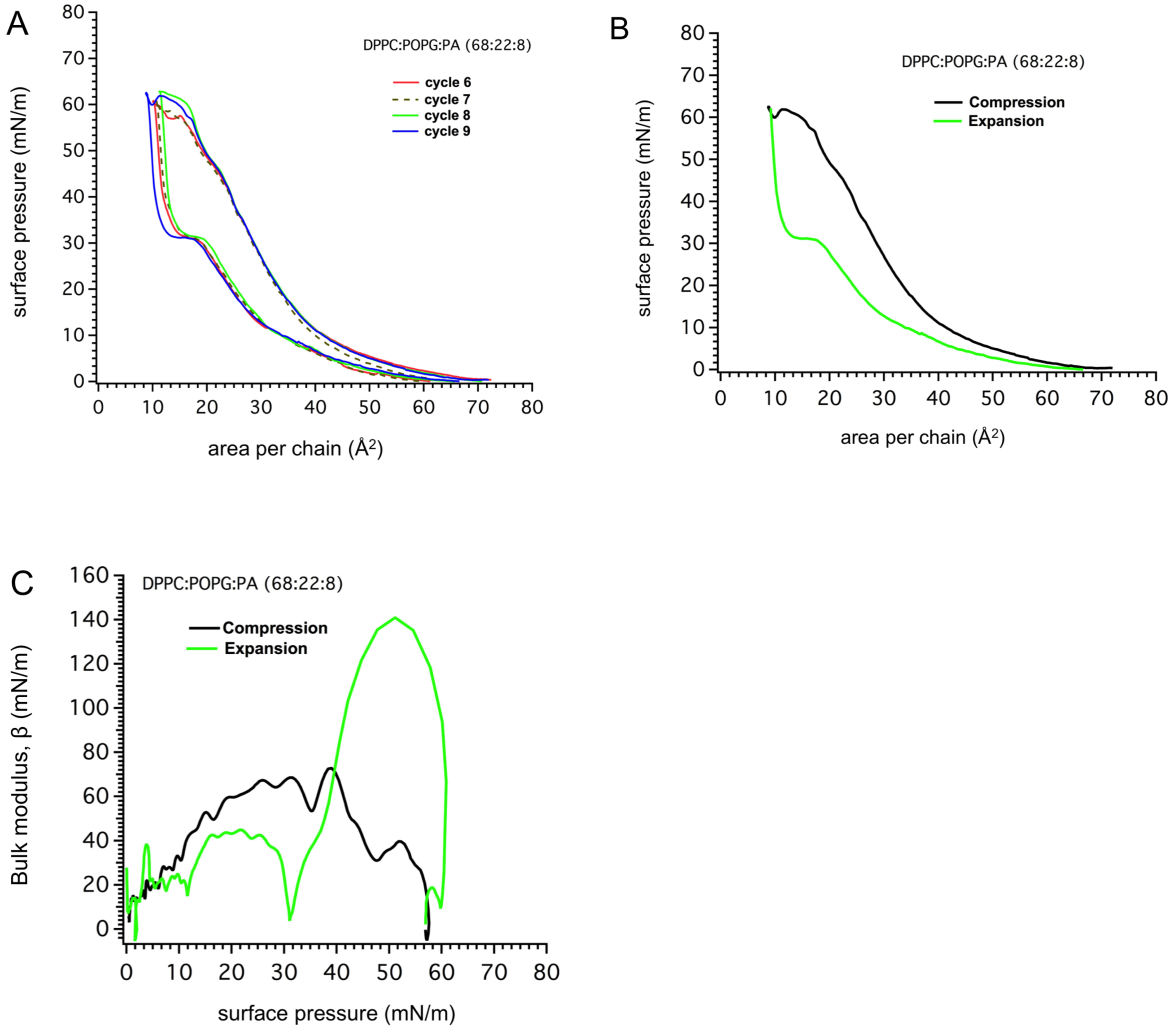


Figure 4.13: Palmitic acid, at large volume spread (40 µl), forces a slight kink at the Π_e and film replenishment plateau at ≈30 mN/m upon film expansion in a DPPC/POPG lipid membrane. Panel A: overlay of repeated Π-A cycling isotherms of DPPC/POPG/PA (68:22:8 wt/wt/wt). Panel B: cycling isotherm of one of the cycle shown in panel A. Panel C: the 2D bulk modulus of the cycle shown in Panel B. PA also exhibits reversible folding collapse at the surface pressure of ≈58 mN/m.

On comparison, only 10 µl of DPPC/POPG/rSP-B, was enough to get to a β_{max} of 130 mN/m on compression and 260 mN/m on expansion. At 40 µl film spread, DPPC/POPG/PA only reached to a β_{max} of 70 mN/m on compression and 140 mN/m on the expansion cycles (Figure 4.13C). Hence, relative to the lipids, it can be said that a small amount of rSP-B is enough to make the DPPC/POPG film less compressible or more sturdy than high amounts of PA. The reinsertion of film material is also evident in the β-Π plot (green line) between surface pressure of 12 mN/m and 32 mN/m, where the monolayer incompressibility rises again to a β of ≈40 mN/m. This observation is indeed suggestive of some sort of network between PA and the DPPC/POPG film.

4.5 Discussion

The Langmuir-Wilhelmy surface balance was used to test the function of rSP-B. Two simple lipid models were assayed with and without rSP-B: DPPC/DPPG (4:1) and DPPC/POPG (7:3). These two lipid mixtures have been utilized successfully to characterize the function of SP-B and SP-C from animals as well as synthetic protein variants [50–53]. The lipid mixtures, in particular, DPPC/POPG, are well suited to characterize the main biophysical features of LS, which are low surface tension on compression and fast spreading and readsorption upon film expansion. First, I have shown that both lipid mixtures can reduce the surface tension to a near zero value on compression *in vitro*, as expected from the literature [50, 51, 53, 54]. Next, rSP-B was added to the lipid mixtures and the compression and expansion isotherms were recorded under similar conditions. Moreover, to elucidate the role of rSP-B and further our understanding of how rSP-B might affect the behaviour of the model lipids, the two dimensional isothermal bulk compressibility modulus was also computed from the pressure-area isotherms. To compare the behaviour of palmitic acid and rSP-B in LS, I also investigated DPPC/POPG/PA lipid mixture.

rSP-B alters the packing order and the phase behaviour of DPPC/DPPG (4:1) lipid membrane. I have seen that DPPC/DPPG/rSP-B (4:1:2 wt%) exhibits a minimum in the β-Π plot at a surface pressure of ≈13 mN/m or in the β-A plot at ≈120 $Å^2$. Such a dip has been demonstrated to indicate a first order phase transition from the LE to LC phase [20–22]. However, the dip was not apparent when the film was compressed at high speed (Figure 4.9D, green line). Nonetheless, it can also be seen that the films compressed at slow (black line) and high (green line) rates, reached the β_{max} at a similar surface pressure as opposed to the lipid-alone film. The fact that the film compressed at a faster rate does not show a similar dip as the film compressed at a slow rate is intriguing, as it suggests that the rate of film compression affects the phase change of LS. This observation is consistent with previous studies that have shown that domain morphology and the sizes of lipid domains are affected by the rate of film compression [47, 55]. In addition, the fact that the DPPC/DPPG-only film does not display a phase transition dip demonstrates that rSP-B inserts into the DPPC/DPPG monolayer and alters the packing order and the phase behaviour of the lipid mixture. This argument is supported by the β-Π plots that reflect the monolayer with rSP-B is highly packed at a surface pressure of 35 mN/m compared to the lipid-only film that reached the highest packing density at a surface pressure of 45 mN/m. The same interpretation holds true when looking at the β-A plots. The film with rSP-B gets to a β_{max} of 142 mN/m at the larger MMA of 100 $Å^2$ than the lipid-alone film that reached a β_{max}

of 139 mN/m at ≈70 $\mathring{A}^2$. A similar behaviour was observed for porcine lung extract by Nag *et al.* [55]. In addition to this, the isotherms with rSP-B lifts off faster and reach the Π_e at a larger MMA than either DPPC/DPPG or DPPC films compressed at the same rate. These attributes are all characteristic of a good quality lung surfactant [17]. Clearly, these observations indicate that rSP-B makes DPPC/DPPG more resistant to compression. There are literature data to suggest that SP-B increases the line tension and dipole density of the monomolecular film [20]. By the same token, rSP-B could be doing the same and hence giving the DPPC/DPPG film an efficient interfacial activity such as reaching the spreading surface pressure faster and with less compression.

Another way of testing the function of rSP-B is to look at whether the protein induces formation of multilayered stacks during compression. Interestingly, at the Π_e, DPPC/DPPG/rSP-B (4:1:2 wt%), demonstrates an inflection point between 90 to 50 $\mathring{A}^2$ (Figure 4.8A and B), which is in agreement with other studies done with the same lipid mixtures with SP-B extracted from animals [50, 51]. This strongly suggests that 2 wt% of rSP-B forces the DPPC/DPPG monolayer to undergo a 2D to 3D transition to release the in-plane stress of the already saturated film. Convincingly, at the inflection point, the compressibility of the DPPC/DPPG/rSP-B film is high, which is in contrast to the lipids alone at a similar surface pressure. The increase in the compressibility after the Π_e indicates that most of the rSP-B is removed by monolayer refinement. However, given the saturated nature of both lipids in the system, it is difficult to identify which of the two lipid might be removed.

Clues as to which lipid might be excluded are given by some literature data. A study by Taneva *et al.* [56] demonstrated that SP-B has a better ability to refine DPPG out of the interfacial layer compared to DPPC. Conversely, Seifert *et al.* [57] showed that in DPPC/DPPG monolayers, SP-B is more likely to associate with the less condensed DPPC to form the protrusion, while the more condensed DPPG remains at the interface. Likewise, a TOF-SIMS study of DPPC/DPPG/SP-C films has also reported some DPPC content in the multilayer stacks [58]. On the other hand, results also show that SP-B does not cause significant ordering of PC acyl chains [40]. Hence, it could also be that PG and SP-B make some form of network likely driven by electrostatic and hydrophobic interactions. Electrostatic interactions are of interest given the highly cationic nature of SP-B and the acidic nature of PG [29]. Nevertheless, it is also important to consider that such interaction might only be possible in the presence of fluid PGs as these lipids provides more space for the protein and increase its solubility [57]. Looking at the Π-A isotherm (Figure 4.8B), it can be noticed that after the plateau region, the isotherms with rSP-B measured using a similar trough reached the same surface pressure as the pure DPPC isotherm. Thus, one could also argue that it is DPPG that is being removed along with rSP-B. The slight increase in the 2D bulk modulus between surface pressures of 45 and 60 in the β-Π plot (Figure 4.9C) could be due to the increase in the order of the acyl chains of the lipid(s) that remain in the interface after the initial expulsion of the lipid(s) and the protein at the Π_e. Finally, I compared the collapse pressures of the of the lipid mixture with the film containing rSP-B.

The monolayer with rSP-B collapses at ≈62 mN/m, which is a lower surface pressure than for DPPC/DPPG. This could happen if the film is depleted of DPPC and collapses at the maximum surface pressure for DPPG. Alternately, due to a lack of fluidizing agent(s) such as POPG and POPC, the lipids might simply lack the flexibility to support zero interfacial tension *in vitro* with rSP-B present. To this end, determining if all of rSP-B is expelled or if some rSP-B still remains at the interface beyond the Π_e is crucial and is a matter that needs more in-depth investigation. If there remains some rSP-B in the monolayer after the Π_e, it will most likely affect the packing order of the lipids and hence how and when the film collapses. In any case, the results herein show that 2 wt% rSP-B has the ability to induce lipid rearrangement in DPPC/DPPG (4:1) lipid mixtures. To investigate if there might be film replenishment, I also looked into the DPPC/POPG system.

rSP-B seems to promote film replenishment when in a DPPC/POPG (7:3) lipid monolayer. Due to the fluid nature of POPG, the lipid mixture of DPPC/POPG provides further opportunity to probe the function of rSP-B, in addition to what can be observed from DPPC/DPPG model. While the β-Π plot of DPPC/POPG-only film (Figure 4.10C) shows a plateau region between surface pressures of 13-23 mN/m, the plateau with DPPC/POPG/rSP-B (Figure 4.12C, black line) starts as low as surface pressure of 5 mN/m and continues to a surface pressure of 20 mN/m. These plateaus are distinct from the dip observed in DPPC/DPPG/rSP-B case (Figure 4.9B and C, black lines). Such observation is consistent with squeeze-out of either lipid, possibly POPG in the earlier case or a protein-lipid network in the later. Although in this particular study (my study), no plateau was observed on the Π-A plot at 23 mN/m during compression in both of the model membranes used, a study by Diemel *et al.* [50] have observed a kink point (a slight turn in the isotherm) at 23 mN/m and a pronounced plateau at ≈40 mN/m on the Π-A isotherm in DPPC/DPPG/SP-B (4:1:4 wt%) film. Using atomic force microscopy, the Diemel group showed that although protrusions form in the DPPC/DPPG film, the height is only one bilayer thick.

This observation was in contrast to the more varied heights of multilayer patches that the same group reported for DPPC/POPG at a similar protein to lipid ratio. It was also noticed that the height of the protrusion depends on whether the unsaturation was on the PC or PG acyl chains. Similar observations were also reported by Krol *et al.* [51] for DPPC/DPPG/SP-B (4:1:1 mol%) film. That group demonstrated globular protrusions and disc-like structures that stem from the LE phase up to a surface pressure of 30 mN/m. In my study, however, the plateau was only evident on the β-Π/A plots of DPPC/POPG model membranes (Figure 4.12B and C). This is striking because, given POPG's fluid nature, it is more likely that the lipid readily squeezes-out of the interface than DPPG. More importantly, with rSP-B present, the plateau was found to span a large range area (120 to 75 Å^2) compared to the lipid-only film (100 to 80 Å^2), which is an indication that more film was ejected from the surface layer. It thus appears that either rSP-B has a better ability to refine fluid PG out of the interfacial layer or it could be that most of the protein itself is removed together with POPG. Nonetheless, the inflection noticed in the DPPC/DPPG/rSP-B film at the spreading surface pressure was not clearly apparent in the DPPC/POPG/rSP-B case, although a very slight kink point can be noticed at ≈45 mN/m. This could be due to protrusion that may occur at the intermediate surface pressure, it could be that the film is more compressible, or it could be simply that more protein is required to see an obvious inflection. To see whether the squeezed-out film material is replenished back into the interface, repeated compression and expansion cycles were compared.

Six wt% of rSP-B facilitates reinsertion of the squeezed-out material during repeated cycling of the same film. The change in the slope at the surface pressure of ≈25 mN/m, as indicated by the arrows in Figure 4.11 on the expansion cycle, is a clear mark of film replenishment. My results are in fair agreement with the work done on SP-B derived peptides and KL4, the synthetic SP-B mimetic peptide, in the same model lipid [52, 53]. However, what was striking was that, with rSP-B, less protein concentration was needed to see consistent and more prominent film reinsertion in contrast to both KL4 and the SP-B derived peptide. The reinsertion plateau observed with rSP-B present could indicate that the protrusion material is retained in association with the monomolecular film. The observation that DPPC/POPG alone does not show the pronounced reinsertion plateau points to the formation of a network between rSP-B and fluid PG. Still at this point, it is difficult to speculate whether all of the rSP-B squeezes-out or if some remains attached to the interface to bridge between the interface and the protrusions. This is a potential future research direction. On the other hand, the bulk 2D compression modulus I observed with rSP-B is also consistent with the work done on mini-B on the same model lipid by

Chakraborty *et al.* [54].

In addition to the 2D bulk modulus of the compression cycle, the 2D bulk modulus of the expansion cycle recorded in this study tells a couple of interesting facts. On the expansion cycle, the surface pressure drops very fast to ≈25 mN/m. We anticipate that this is the surface pressure at which the film reaches equilibrium with the interface on the expansion cycle. This is evident on Figure 4.12A, B and C, green lines, where there is not much area change (Panel A and B) but the surface pressure drops quickly from 68 mN/m to a ≈25 mN/m (Panel A and C). In this step, the 2D bulk modulus increases very rapidly to β of ≈260 mN/m, characteristic of solid or semi-crystalline state [59, 60]. It could be that the film compressed beyond the Π_e is metastable that it is intially resisting expansion until equilibrium is reached. Later on presumably after equilibration is reached, in this case at a surface pressure of 25 mN/m, the β decreases to ≈10-15 mN/m, which is characteristic of LE phase [61]. The monolayer at this point will therefore have enough room to accommodate the squeezed-out film back into the interfacial film. Previous studies have demonstrated that SP-B partitions into the LE phase and then squeezes-out with fluid lipids during film compression upto the Π_e [20, 33, 62, 63]. Hence, it is reasonable to assume that the squeezed-out film also reincorporates when the monolayer is in the LE phase on the reverse cycle.

In the presence of rSP-B, the reinsertion of the protrusions occurs over a wide area change as the surface area enlarges because of film expansion. As said, this process is marked by the start of the plateau at a surface pressure of 25 mN/m on the Π-A plot (Figure 4.12A) or by the local maximum in β at a mean molecular area of 58 Å^2 (Figure 4.12B, green line) or surface pressure between 25 to 10 mN/m (Figure 4.12C, green line) of the β-A and β-Π plots respectively. These observations support the suggestion that rSP-B makes some form of network with the expelled lipid(s) and also suggest that the network structure reinserts when the monolayer is fluid enough. In our experimental set up, at an expansion speed of 14 cm^2/min, the sudden drop in pressure takes 2-3 seconds while the later reinsertion process roughly takes twice that time. By using a LB trough and flourescence micrograph, Warriner *et al.* [64] also demonstrated that the collapse structures of LS lipid films do not immediately reincorporate on film expansion. Instead, the film respreads initially without incorporating the multilayer structures and later incorporates the multilayer patch starting at surface pressure of ≈20-15 mN/m. Reincorporation continues to as low a surface pressure as 5 mN/m.

Overall, the evidences suggest that the DPPC/POPG/rSP-B (7:3:6 wt%) system eases lateral strain initially by squeezing out fluid PG before the Π_e and then possibly buckles reversibly at a high surface pressure prior to the onset of irreversible collapse. The film seems to start to effectively spread back to the interface starting at a surface pressure of ≈25 mN/m.

To enhance the surface activity, spreading kinetics, and readsorption rate of exogenous lung surfactant, different additives are also used either together with SP-B/SP-C or in place of the small proteins. Of particular note are PA and its derivative alcohol, hexadecanol (HD). PA is used in Survanta, a formulation that contains both SP-B and SP-C, whereas HD was successfully used in Exosurf, a protein free synthetic surfactant that is discontinued from the market. To elucidate the role of PA, therefore, I also investigated DPPC/POPG/PA (68:22:8 wt/wt/wt) model membrane in spread films.

DPPC/POPG/PA film shows a slight kink at the Π_e and film reinsertion at ≈30 mN/m (Figure 4.13). The kink point has also been reported by other groups [30, 65]. The change in slope on the expansion cycle at ≈30 mN/m indicates some form of interaction between the constituent molecules. To this end, grazing incidence X-ray diffraction studies have shown that PA/HD specifically interacts with DPPC in lipid monolayers [31, 66, 67] leading to tight packing of DPPC chains and increasing the solid portion of the film at a particular surface pressure. A separate study also reported a hydrogen bond between PA and DPPC in mixed monolayres [68]. The Π-A isotherm of DPPC/POPG/PA film collapses earlier (65 mN/m, data not shown) than the DPPC/POPG (72 mN/m) film. This indicates that PA prevents the DPPC/POPG film from reaching zero interfacial tension. On the contrary, the β_{max} observed in the presence of PA (Figure 4.13C, black line) is less than the β_{max} recorded for either the DPPC/POPG/rSP-B (Figure 4.12C, black line) or the DPPC/POPG (Figure 4.10C) case. Hence, PA also seems to increase the compressibility of DPPC/POPG film. The Bringezu group [31] have shown that the compressibility of DPPC/POPG/PA increases with an increase in PA concentration. It thus looks like that PA and rSP-B, when used alone, have antagonistic effect. That is, while small amount of rSP-B is enough to make DPPC/POPG less compressible, PA may rather have a fluidizing effect.

On another note, when used in conjunction, synergistic behaviour of SP-B and PA are also well documented [29, 30, 69, 70]. This may underlie the necessity to use PA and SP-B/SP-C together in replacement surfactant. The fact that, Surfaxin, a formula that contains PA and KL4, is discontinued from use suggests the need for SP-B/SP-C in exogenous surfactant. As of yet there is no alternative to using animal derived SP-B. As far as we are aware, this is the first study to use rSP-B. That said, however, DPPC/POPG is just a simple model of LS lipids, and perhaps DPPC/POPG/PA/rSP-B could be worth investigating.

Finally, despite the encouraging results that are consistent with the data in the literature, the author would still like to suggest the use of mini-B and/or synthetic clinical surfactant as a positive control. In depth analysis would be required to solidify the use of recombinant SP-B as a replacement clinical surfactant and a start would be to demonstrate that recombinant SP-B compares positively with synthetic surfactant that are in the market. Otherwise, results presented here are repeated a minimum of three times and were found to be consistent with both different isolates of protein or aliquotes taken from the same preparation.

4.6 References

1. Clements, J. A. Surface tension of lung extracts. *Proceedings of the Society for Experimental Biology and Medicine* **95,** 170–2 (1957).

2. Zuo, Y. Y., Veldhuizen, R. A. W., Neumann, A. W., Petersen, N. O. & Possmayer, F. Current perspectives in pulmonary surfactant - Inhibition, enhancement and evaluation. *Biochimica et Biophysica Acta* **1778,** 1947–1977 (2008).

3. Veldhuizen, E. J. A. & Haagsman, H. P. Role of pulmonary surfactant components in surface film formation and dynamics. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1467,** 255–270 (2000).

4. Lipp, M. M., Lee, K. Y. C., Takamoto, D. Y., Zasadzinski, J. A. & Waring, A. J. Coexistence of buckled and flat monolayers. *Physical Review Letters* **81,** 1650–1653 (1998).

5. Verwijlen, T., Imperiali, L. & Vermant, J. Separating viscoelastic and compressibility contributions in pressure-area isotherm measurements. *Advances in Colloid and Interface Science* **206,** 428–436 (2014).

6. Yuan, Y. & Lee, T. R. *Contact angle and wetting properties* in *Surface science techniques* (eds Bracco, G. & Holst, B.) 3–34 (Springer, Berlin, Heidelberg, 2013).

7. Prokop, R. M. & Neumann, A. W. Measurement of the interfacial properties of lung surfactant. *Current Opinion in Colloid & Interface Science* **1,** 677–681 (1996).

8. Schurch, S., Bachofen, H., Goerke, J. & Possmayer, F. A captive bubble method reproduces the in situ behavior of lung surfactant monolayers. *Journal of Applied Physiology* **67,** 2389–2396 (1989).

9. Notter, R. H. & Morrow, P. E. Pulmonary surfactant: A surface chemistry viewpoint. *Annals of Biomedical Engineering* **3,** 119–159 (1975).

10. Tabak, S. A. & Notter, R. H. Modified technique for dynamic surface pressure and relaxation measurements at the air-water interface. *Review of Scientific Instruments* **48,** 1196–1201 (1977).

11. Enhorning, G. Pulsating bubble technique for evaluating pulmonary surfactant. *Journal of Applied Physiology* **43,** 198–203 (1977).

12. Hall, S. B., Bermel, M. S., Ko, Y. T., Palmer, H. J., Enhorning, G. & Notter, R. H. Approximations in the measurement of surface tension on the oscillating bubble surfactometer. *Journal of Applied Physiology* **75,** 468–477 (1993).

13. Chang, C.-H. & Franses, E. An analysis of the factors affecting dynamic tension measurements with the pulsating bubble surfactometer. *Journal of Colloid and Interface Science* **164,** 107–113 (1994).

14. Liao, Y.-C., Basaran, O. A. & Franses, E. I. Effects of dynamic surface tension and fluid flow on the oscillations of a supported bubble. *Colloids and Surfaces A-Physicochemical and Engineering Aspects* **282-283,** 183–202 (2006).

15. Liao, Y.-C., Basaran, O. A. & Franses, E. I. Hydrodynamic effects on the oscillations of supported bubbles: implications for accurate measurements of surface properties. *Colloids and Surfaces A: Physicochem. Eng. Aspects* **250,** 367–384 (2004).

16. Schurch, S., Schurch, D., Curstedt, T. & Robertson, B. Surface activity of lipid extract surfactant in relation to film area compression and collapse. *Journal of Applied Physiology* **77,** 974–986 (1994).

17. Banerjee, R. Surface chemistry of the lung surfactant system: Techniques for in vitro evaluation. *Current Science* **82,** 420–428 (2002).

18. Mansour, H. M. & Zografi, G. Relationships between equilibrium spreading pressure and phase equilibria of phospholipid bilayers and monolayers at the air-water interface. *Langmuir* **23,** 3809–19 (2007).

19. Adams, E. M. & Allen, H. C. Palmitic acid on salt subphases and in mixed monolayers of cerebrosides: Application to atmospheric aerosol chemistry. *Atmosphere* **4,** 315–336 (2013).

20. Dhar, P., Eck, E., Israelachvili, J. N., Lee, D. W., Min, Y., Ramachandran, A., Waring, A. J. & Zasadzinski, J. A. Lipid-protein interactions alter line tensions and domain size distributions in lung surfactant monolayers. *Biophys Journal* **102,** 56–65 (2012).

21. Ghazvini, S., Ricke, B., Zasadzinski, J. A. & Dhar, P. Monitoring phases and phase transitions in phosphatidylethanolamine monolayers using active interfacial microrheology. *Soft Matter* **11,** 3313–3321 (2015).

22. Yu, Z.-W., Jin, J. & Cao, Y. Characterization of the liquid-expanded to liquid-condensed phase transition of monolayers by means of compressibility. *Langmuir* **18,** 4530–4531 (2002).

23. Takamoto, D. Y., Lipp, M. M., von Nahmen, A., Lee, K. Y., Waring, A. J. & Zasadzinski, J. A. Interaction of lung surfactant proteins with anionic phospholipids. *Biophysical journal* **81,** 153–69 (2001).

24. Zuo, Y. Y., Chen, R., Wang, X., Yang, J., Policova, Z. & Neumann, A. W. Phase transitions in dipalmitoylphosphatidylcholine monolayers. *Langmuir* **32,** 8501–8506 (2016).

25. Zhang, H., Fan, Q., Wang, Y. E., Neal, C. R. & Zuo, Y. Y. Comparative study of clinical pulmonary surfactants using atomic force microscopy. *Biochimica et Biophysica Acta* **1808,** 1832–1842 (2011).

26. McConlogue, C. W. & Vanderlick, T. K. A close look at domain formation in DPPC monolayers. *Langmuir* **13,** 7158–7164 (1997).

27. Kim, K., Choi, S. Q., Zasadzinski, J. A. & Squires, T. M. Interfacial microrheology of DPPC monolayers at the air-water interface. *Soft Matter* **7,** 7782 (2011).

28. Grunder, R., Gehr, P., Bachofen, H., Schurch, S. & Siegenthaler, H. Structures of surfactant films: A scanning force microscopy study. *European Respiratory Journal* **14,** 1290–1296 (1999).

29. Ding, J., Doudevski, I., Warriner, H. E., Alig, T., Zasadzinski, J. A., Waring, A. J. & Sherman, M. A. Nanostructure changes in lung surfactant monolayers induced by interactions between palmitoyloleoylphosphatidylglycerol and surfactant protein B. *Langmuir* **19,** 1539–1550 (2003).

30. Ding, J., Takamoto, D. Y., Von Nahmen, A., Lipp, M. M., Lee, K. Y. C., Waring, A. J. & Zasadzinski, J. A. Effects of lung surfactant proteins, SP-B and SP-C, and palmitic acid on monolayer stability. *Biophysical Journal* **80,** 2262–2272 (2001).

31. Bringezu, F., Ding, J., Brezesinski, G. & Zasadzinski, J. A. Changes in model lung surfactant monolayers induced by palmitic acid. *Langmuir* **17,** 4641–4648 (2001).

32. Yan, W., Biswas, S. C., Laderas, T. G. & Hall, S. B. The melting of pulmonary surfactant monolayers. *Journal of Applied Physiology* **102,** 1739–1745 (2007).

33. Keating, E., Zuo, Y. Y., Tadayyon, S. M., Petersen, N. O., Possmayer, F. & Veldhuizen, R. A. W. A modified squeeze-out mechanism for generating high surface pressures with pulmonary surfactant. *Biochimica et Biophysica Acta* **1818,** 1225–1234 (2012).

34. Schurch, S., Goerke, J. & Clements, J. a. Direct determination of volume- and time-dependence of alveolar surface tension in excised lungs. *Proceedings of the National Academy of Sciences* **75,** 3417–3421 (1978).

35. Israelachvili, J. N. *Thermodynamic principles of self-assembly* in *Intermolecular and Surface Forces* (ed Israelachvili, J. N.) 3rd ed., 503–534 (Acadamic press, Boston, 2011).

36. Rufeil-Fiori, E. & Banchio, A. J. Domain size polydispersity effects on the structural and dynamical properties in lipid monolayers with phase coexistence. *Soft Matter* **14,** 1870–1878 (2018).

37. Lee, D. W., Min, Y., Dhar, P., Ramachandran, A., Israelachvili, J. N. & Zasadzinski, J. A. Relating domain size distribution to line tension and molecular dipole density in model cytoplasmic myelin lipid monolayers. *Proceedings of the National Academy of Sciences* **108,** 9425–9430 (2011).

38. Min, Y. Phase dynamics and domain interactions in biological membranes. *Current Opinion in Chemical Engineering* **15,** 76–83 (2017).

39. Cruz, A., Vázquez, L., Vélez, M. & Pérez-Gil, J. Effect of pulmonary surfactant protein SP-B on the micro- and nanostructure of phospholipid films. *Biophysical Journal* **86,** 308–320 (2004).

40. Brockman, J. M., Wang, Z., Notter, R. H. & Dluhy, R. A. Effect of hydrophobic surfactant proteins SP-B and SP-C on binary phospholipid monolayers II: Infrared external reflectance-absorption spectroscopy. *Biophysical journal* **84,** 326–40 (2003).

41. Schürch, D., Ospina, O. L., Cruz, A. & 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,** 3290–3299 (2010).

42. Lee, K. Y. C. Collapse mechanisms of Langmuir monolayers. *Annual Review of Physical Chemistry* **59,** 771–791 (2008).

43. Alig, T. F., Warriner, H. E., Lee, L. & Zasadzinski, J. A. Electrostatic barrier to recovery of dipalmitoylphosphatidylglycerol monolayers after collapse. *Biophys Journal* **86,** 897–904 (2004).

44. Crane, J. M., Putz, G. & Hall, S. B. Persistence of phase coexistence in disaturated phosphatidylcholine monolayers at high surface pressures. *Biophysical Journal* **77,** 3134–3143 (1999).

45. Sibug-Aga, R. & Dunn, R. C. High-resolution studies of lung surfactant collapse. *Photochemistry and Photobiology* **80,** 471–476 (2004).

46. Schürch, S., Green, F. H. & Bachofen, H. Formation and structure of surface films: Captive bubble surfactometry. *Biochimica et Biophysica Acta* **1408,** 180–202 (1998).

47. Zuo, Y. Y., Keating, E., Zhao, L., Tadayyon, S. M., Veldhuizen, R. A., Petersen, N. O. & Possmayer, F. Atomic force microscopy studies of functional and dysfunctional pulmonary surfactant films. I. Micro- and nanostructures of functional pulmonary surfactant films and the effect of SP-A. *Biophys J* **94,** 3549–3564 (2008).

48. Schürch, S. Surface tension properties of surfactant. *Clinics in Perinatology* **20,** 669–682 (1993).

49. Yang, Y. P. & Tsay, R. Y. Morphology of compressed dipalmitoyl phosphatidylcholine monolayers investigated by atomic force microscopy. *Journal of Luminescence* **127,** 186–191 (2007).

50. Diemel, R. V., Snel, M. M. E., Waring, A. J., Walther, F. J., Van Golde, L. M. G., Putz, G., Haagsman, H. P. & Batenburg, J. J. Multilayer formation upon compression of surfactant monolayers depends on protein concentration as well as lipid composition. An atomic force microscopy study. *Journal of Biological Chemistry* **277,** 21179–21188 (2002).

51. Krol, S., Ross, M., Sieber, M., Künneke, S., Galla, H.-J. & Janshoff, A. Formation of three-dimensional protein-lipid aggregates in monolayer films induced by surfactant protein B. *Biophysical Journal* **79,** 904–918 (2000).

52. Holten-Andersen, N., Michael Henderson, J., Walther, F. J., Waring, A. J., Ruchala, P., Notter, R. H. & Lee, K. Y. KL4 peptide induces reversible collapse structures on multiple length scales in model lung surfactant. *Biophys Journal* **101,** 2957–2965 (2011).

53. Serrano, A. G., Ryan, M., Weaver, T. E. & Pérez-Gil, J. Critical structure-function determinants within the N-terminal region of pulmonary surfactant protein SP-B. *Biophysical Journal* **90,** 238–249 (2006).

54. Chakraborty, A., Hui, E., Waring, A. J. & Dhar, P. Combined effect of synthetic protein, Mini-B, and cholesterol on a model lung surfactant mixture at the air-water interface. *Biochimica et Biophysica Acta* **1858,** 904–912 (2016).

55. Nag, K., Perez-Gil, J., Ruano, M. L., Worthman, L. A., Stewart, J., Casals, C. & Keough, K. M. Phase transitions in films of lung surfactant at the air-water interface. *Biophysical Journal* **74,** 2983–2995 (1998).

56. Taneva, S. & Keough, K. M. Pulmonary surfactant proteins SP-B and SP-C in spread monolayers at the air-water interface: III. Proteins SP-B plus SP-C with phospholipids in spread monolayers. *Biophysical Journal* **66,** 1158–1166 (1994).

57. Seifert, M., Breitenstein, D., Klenz, U., Meyer, M. & Galla, H.-J. Solubility versus electrostatics: What determines lipid/protein interaction in lung surfactant. *Biophysical Journal* **93,** 1192–1203 (2007).

58. Bourdos, N., Kollmer, F., Benninghoven, A., Ross, M., Sieber, M. & Galla, H.-J. Analysis of lung surfactant model systems with time-of-flight secondary ion mass spectrometry. *Biophysical Journal* **79,** 357–369 (2000).

59. Xu, G., Hao, C., Zhang, L. & Sun, R. Investigation of surface behavior of DPPC and curcumin in Langmuir monolayers at the air-water interface. *Scanning* **2017,** 1–12 (2017).

60. Krajewska, B., Wydro, P., Kyzio, A. & Wydro, P. Chitosan as a subphase disturbant of membrane lipid monolayers. The effect of temperature at varying pH: II. DPPC and cholesterol. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **434,** 359–364 (2013).

61. Davies, J. & Rideal, E. *Interfacial phenomena* 2nd ed., 474 (New York : Academic Press, 1963).

62. Keating, E., Waring, A. J., Walther, F. J., Possmayer, F., Veldhuizen, R. A. W. & Petersen, N. O. A ToF-SIMS study of the lateral organization of lipids and proteins in pulmonary surfactant systems. *Biochimica et Biophysica Acta* **1808,** 614–621 (2011).

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

64. Warriner, H. E., Ding, J., Waring, A. J. & Zasadzinski, J. A. A concentration-dependent mechanism by which serum albumin inactivates replacement lung surfactants. *Biophysical Journal* **82,** 835–842 (2002).

65. Alonso, C., Waring, A. & Zasadzinski, J. A. Keeping lung surfactant where it belongs: Protein regulation of two-dimensional viscosity. *Biophysical journal* **89,** 266–273 (2005).

66. Alonso, C., Bringezu, F., Brezesinski, G., Waring, A. J. & Zasadzinski, J. A. Modifying calf lung surfactant by hexadecanol. *Langmuir* **21,** 1028–1035 (2005).

67. Lee, K. Y. C., Gopal, A., von Nahmen, A., Zasadzinski, J. A., Majewski, J., Smith, G. S., Howes, P. B. & Kjaer, K. Influence of palmitic acid and hexadecanol on the phase transition temperature and molecular packing of dipalmitoylphosphatidyl-choline monolayers at the air-water interface. *The Journal of Chemical Physics* **116,** 774–783 (2002).

68. Ma, G. & Allen, H. C. Condensing effect of palmitic acid on DPPC in mixed Langmuir monolayers. *Langmuir* **23,** 589–597 (2007).

69. Longo, M. L., Bisagno, A. M., Zasadzinski, J. A., Bruni, R. & Waring, A. J. A function of lung surfactant protein SP-B. *Science* **261,** 453–456 (1993).

70. Lipp, M. M., Lee, K. Y. C., Zasadzinski, J. A. & Waring, A. J. Phase and morphology changes in lipid monolayers induced by SP-B protein and its amino-terminal peptide. *Science* **273,** 1196–1199 (1996).

5

Summary

5.1 Summary and Conclusions

Although RDS is considered fully treatable in neonates nowadays, treatment is especially expensive in adult patients and the need to support patient with ventilation is still a drive force to develop a formulation that can be delivered the non-invasive or the less invasive way. On the other hand, many people still succumb to ARDS and the mortality rate is yet unacceptable. If there is going to be hope to develop drug for ALI/ARDS, it is going to need a lot of surfactant material and a formula that resist inhibition [1]. The need is thus to develop synthetic surfactant as effective as clinical surfactant in excess and at a lesser cost.

During the last 30 to 40 years synthesizing and characterizing SP-B have proven to be difficult. On the other hand, SP-C was produced both by chemical synthesis and by heterologous DNA expression. Nonetheless, SP-C produced this way has been found to be prone to amyloid fibril formation and/or β-sheet aggregation, which render the protein ineffective or less active [1–3]. The post-translational modification of SP-C, i.e. palmitoylation of the cysteines has proofed to be crucial and is impossible to introduce by chemical synthesis or by expressing the protein in E-coli. Alteration of SP-C's nature since have led to many variants with desired traits. The variant rSP-C33Leu for example demonstrated better *in vitro* and *in vivo* efficacy and showed resistance to inactivation by serum lipoproteins [4]. Although rSP-C33Leu was produced by peptide synthesis initially, it was recombinant DNA expression that excelled at producing the protein at a larger scale and cheaper cost [4, 5]. These capabilities make rSP-C33Leu the choice for commercial success. Although SP-B's nature and properties has been a challenge to explore such path thus far, the work performed here tried to follow SP-C's suite and find a way to produce and alter SP-B's feature.

Chemical synthesis of SP-B can be complicated by its size, hydrophobicity, disulfide linkages and complex structure. With respect to size, the truncated peptides, mini-B and super mini-B were produced by peptide synthesis [6]. And yet, no technique is available to produce these peptides by recombinant gene expression. Other compounds that replaced SP-B's function with synthetic products have not been quite successful so far. KL4 used in Surfaxin for instance is one good example. The compound makes the drug viscous at room temperature. Surfaxin therefore needs to be heated to 44 degree Celsius and then to be shaken vigorously prior administration, which complicates its ease of handle and use [2, 6]. Nonetheless, due to its gel formation tendency, attempts are currently underway to develop KL4 as an aerosol formulation, Aerosurf [1, 7]. Another area that gained research attention to replicate SP-B and SP-C use are non-natural helical polypeptides with poly-N subsituted glycines (peptoids). Some peptoids have been developed to emulate the function of SP-B and SP-C *in vitro* and *in vivo* [2, 8]. These compounds are resistant to protease and reduce immunogenic response but there is lack of data of clearance and catabolism of peptoids in the lung. Thus, it may take a while before these products are in use, assuming that all other requirements are met. Modification of SP-B's nature is hence an important consideration. Mini-BLeu is one such example and a result of this is the novel synthetic surfactant, CHF5633 (Chiesi Farmaceutici, Italy). The drug combined rSP-C33Leu (1.5%) and mini-BLeu (0.2%) with 1:1 DPPC/POPG and showed resistance to dysfunction better than animal derived surfactant for the first time in history. CHF5633 is in current phase two clinical trial [9–13].

Although SP-C33Leu has been deemed a success, it is believed that there is still room to further modify SP-B for improved biophysical features. Recombinant gene expression is one robust technique to introduce mutation(s) to probe the structure-activity relation or to modify proteins in ways possible the target protein acquires the desired biological function. Importantly, the technique is amenable to scale-up production at cheaper cost than peptide synthesis, which makes it commercially feasible. This thesis work has successfully shown that SP-B can be expressed in bacteria using recombinant DNA methodology, purified with IMAC, renatured on-column, and stabilized using different membrane mimetic agents. Such technology may increase the availability of human SP-B and perhaps may offer patients with respiratory distress the flexibility of larger and more frequently doses. The study also provides the chance to produce SP-B that is labeled with NMR active nuclei for structural studies, which is impossible to do for animal derived SP-B. Of special importance is site directed mutagenesis, which is mandatory to select a variant that may stand lung surfactant inactivation. Likewise, the technique also makes it possible to probe a variety of mutants to investigate the structure-function relation of SP-B.

The second study well characterizes the rSP-B/DPC/SDS system. In the process, the thesis work confirmed that DPC/SDS micelles are removed by dialysis using a pore size as small as 1 kDa. The work also successfully showed the presence of larger rSP-B/DPC/SDS complexes that demonstrate native-like secondary structure with circular dichroism. The secondary/tertiary structure of the protein was also found to be similar in methanol and in DPC/SDS. Apparently, the sizes of the protein in these two surroundings differ to great extent. The protein in the micelles is much larger. These observations gives a hint to us that as with most other SAPLIPs, SP-B do also have a propensity to self-assemble according to its environment [14]. It could be that the tertiary structure is stabilized by covalent disulphide bonds while the quaternary structure is maintained by weak non-covalent interactions that may be destabilized by harsh solvents like methanol. Thus, the traditional way of extracting SP-B with organic solvent may be disrupting the native arrangements of the protein. The size of the protein in the detergent complexes suggests, rSP-B may be adopting supramolecular arrangements in part via weak non-covalent protein-protein/protein-detergent interactions. It could be that such assemblies promote a more efficient way of facilitating the apposition and stabilization of lipid monolayers and the multilayered stacks *in vivo* than the dimer or monomer of SP-B. The study therefore opens a new perspective and dimension for future studies.

From structure determination aspect, the size of the protein in 0.2% DPC/SDS is a limiting factor for liquid state NMR. This has indicated to us that either further optimization is required or other techniques such as solid state NMR and cryo-electron microscopy must be pursued. From optimization point of view, for example, the study have found that DPC/SDS (9:1) micelles tend to form heterogeneous micelles at a concentration between 0.2 to 1%, whereas at concentrations between 1 to 2%, DPC/SDS (9:1) forms homogeneous micelles that demonstrate the right size for solution NMR. It could be that suspending the protein in these micelles, that is, micelles formed at concentrations between 1 to 2% of DPC/SDS (9:1), may prevent the protein from assembling into higher forms to finally lead to success with solution NMR. This is awaiting to be seen and could possibly have implication for other proteins suspended in these detergent mix as well.

The third study characterizes the function of rSP-B with the Langmuir-Wilhelmy balance. The functions of rSP-B in simple lipid membranes tested are also in agreement with rSP-B promoting the large scale reorganization of lipids *in vitro*. The data demonstrates that the function of rSP-B compare favorably to the data reported in literature for mini-B and other truncated SP-B peptides [15, 16] and to that of native SP-B and KL4 [17–19]. It is therefore possible to conclude that human SP-B variant produced in bacteria is functional. These findings, if future functional studies go as planned, are prospective of rSP-B as a potential future alternative to animal derived SP-B. As such, by optimizing the content, nature and concentrations of the lipids and rSP-B carefully a model that satisfies all the necessary biophysical features of lung surfactant could be possible.

As a whole, the work undertaken in this study have used different techniques to meet objectives put in the hypothesis: production of this difficult molecule in bacteria. The thesis work have demonstrates robust protocol for production of SP-B in laboratory setting. In characterizing the structure-function relation of recombinant SP-B, invaluable data have been gathered. Foremost is the concentration of the detergents used. The heterogeneous micelles that form at a rather low concentration of detergent can be dialysed to give a homogeneous sample but the protein remains too big for standard solution NMR. As such the work is also partly a success beyond going towards achieving the goal of the thesis. In functional tests of rSP-B, the protein has been found to perform well in lipid environment *in vitro*. Nonetheless, there are some limitations to the method used; LB trough is not a leak proof method and the dynamics achieved in the LB-trough is not comparable to the true *in vivo* dynamics. Also, the shape of the instrument does not emulate the air-sacs. These limitations could be addressed by studying the protein function using leak proof techniques such as captive bubble and constrained drop surfactometry. In summary, although the work has been largely a success; there is yet much that remains to be learned.

5.2 Future Directions

Reconstituting of rSP-B into phospholipid environments will prove to be important to compare the size and the secondary structure of rSP-B with that of rSP-B suspended in detergent micelles. Using techniques such as analytical ultra-high centrifugation and size exclusion chromatography, the different sub-populations of rSP-B/DPC/SDS could be separated and analysed individually. Optimization of more conditions for structural study by NMR may be important. Native-gel electrophoresis of rSP-B solubilized with different detergent(s) could give important insight on the behaviour, size and oligomeric state of SP-B in different surroundings. The number of subunits that make a complex could also be studied by subjecting the native gel to a second dimension and to denaturing electrophoresis. The function of the protein can further be probed with leak proof techniques such as captive bubble surfactometry and constrained drop surfactometry. A multitude of functional assays in a variety of lipid mixtures and protein concentrations, however, could be necessary. For instance, the combined use of palmitic acid and rSP-B in different lipid mixtures and concentrations could be studied. If preliminary data are promising, animal model studies are also worth thinking about in the long run. Site directed mutagenesis will be invaluable to select SP-B variant(s) that are resistant to inactivation by serum lipoproteins and reactive oxygen species. This has a potential outcome to design treatment for patient with ARDS/ALI. Similarly, the structure-to-function relation of SP-B can be probed by studying different SP-B variants. For safety reasons, it is also imperative to consider introducing an enzymatic cleavage site between rSN and rSP-B instead of the chemical cleavage site. Important to consider also is that due to adsorption and spreading properties, lung surfactant has the potential to be used as a vehicle to deliver other drugs to the lung to areas otherwise are difficult to target provided that surfactant that resist deactivation in inflamed or injured lung comes available [20]. Bacterial pneumonia and ARDS/ALI for instance are two prominent areas that benefit from developing surfactant as a carrier.

5.3 References

1. Hentschel, R., Bohlin, K., van Kaam, A., Fuchs, H. & Danhaive, O. Surfactant replacement therapy: from biological basis to current clinical practice. *Pediatric Research* **88,** 176–183 (2020).

2. Walther, F. J., Gordon, L. M. & Waring, A. J. Advances in synthetic lung surfactant protein technology. *Expert Review of Respiratory Medicine* **13,** 499–501 (2019).

3. Sardesai, S., Biniwale, M., Wertheimer, F., Garingo, A. & Ramanathan, R. Evolution of surfactant therapy for respiratory distress syndrome: Past, present, and future. *Pediatric Research* **81,** 240–248 (2017).

4. Johansson, J. & Curstedt, T. Synthetic surfactants with SP-B and SP-C analogues to enable worldwide treatment of neonatal respiratory distress syndrome and other lung diseases. *Journal of Internal Medicine* **285,** 165–186 (2019).

5. Kronqvist, N., Sarr, M., Lindqvist, A., Nordling, K., Otikovs, M., Venturi, L., Pioselli, B., Purhonen, P., Landreh, M., Biverstål, H., Toleikis, Z., Sjöberg, L., Robinson, C. V., Pelizzi, N., Jörnvall, H., Hebert, H., Jaudzems, K., Curstedt, T., Rising, A. & Johansson, J. Efficient protein production inspired by how spiders make silk. *Nature Communications* **8,** 1–15 (2017).

6. Walther, F. J., Gordon, L. M. & Waring, A. J. Design of Surfactant Protein B Peptide Mimics Based on the Saposin Fold for Synthetic Lung Surfactants. *Biomedicine Hub* **1,** 1–21 (2016).

7. Finer, N. N., Merritt, T. A., Bernstein, G., Job, L., Mazela, J. & Segal, R. An Open Label, Pilot Study of Aerosurf Combined with nCPAP to Prevent RDS in Preterm Neonates. *Journal of Aerosol Medicine and Pulmonary Drug Delivery* **23, Number,** 303–309 (2010).

8. Czyzewski, A. M., McCaig, L. M., Dohm, M. T., Broering, L. A., Yao, L. J., Brown, N. J., Didwania, M. K., Lin, J. S., Lewis, J. F., Veldhuizen, R. & Barron, A. E. Effective in vivo treatment of acute lung injury with helical, amphipathic peptoid mimics of pulmonary surfactant proteins. *Scientific Reports* **8,** 1–9 (2018).

9. Seehase, M., Collins, J. J. P., Kuypers, E., Jellema, R. K., Ophelders, D. R. M. G., Ospina, O. L., Perez-Gil, J., Bianco, F., Garzia, R., Razzetti, R. & Kramer, B. W. New surfactant with SP-B and C analogs gives survival benefit after inactivation in preterm lambs. *PLoS ONE* **7,** 1–10 (2012).

10. Johansson, J. & Curstedt, T. Synthetic surfactants with SP-B and SP-C analogues to enable worldwide treatment of neonatal respiratory distress syndrome and other lung diseases. *Journal of Internal Medicine* **285,** 165–186 (2018).

11. Echaide, M., Autilio, C., López-Rodríguez, E., Cruz, A. & Pérez-Gil, J. In Vitro Functional and Structural Characterization of A Synthetic Clinical Pulmonary Surfactant with Enhanced Resistance to Inhibition. *Scientific Reports* **10,** 1–10 (2020).

12. Sweet, D. G., Turner, M. A., Stranák, Z., Plavka, R., Clarke, P., Stenson, B. J., Singer, D., Goelz, R., Fabbri, L., Varoli, G., Piccinno, A., Santoro, D. & Speer, C. P. A first-in-human clinical study of a new SP-B and SP-C enriched synthetic surfactant (CHF5633) in preterm babies with respiratory distress syndrome. *Archives of Disease in Childhood: Fetal and Neonatal Edition* **102,** F497–F503 (2017).

13. Ronda, L., Faggiano, S., Paredi, G., Michielon, A., Sartor, F., Raschini, A. S., Cavatorta, V., Sgarbi, E., Bettati, S. & Mozzarelli, A. SP-B and SP-C analogues within CHF5633 synthetic surfactant probed by fluorescence labeling. *Journal of Molecular Liquids* **298,** 111983 (2020).

14. Hill, C. H., Read, R. J. & Deane, J. E. Structure of human saposin A at lysosomal pH. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **71,** 895–900 (2015).

15. Chakraborty, A., Hui, E., Waring, A. J. & Dhar, P. Combined effect of synthetic protein, Mini-B, and cholesterol on a model lung surfactant mixture at the air-water interface. *Biochimica et Biophysica Acta* **1858,** 904–912 (2016).

16. Serrano, A. G., Ryan, M., Weaver, T. E. & Pérez-Gil, J. Critical structure-function determinants within the N-terminal region of pulmonary surfactant protein SP-B. *Biophysical Journal* **90,** 238–249 (2006).

17. Diemel, R. V., Snel, M. M. E., Waring, A. J., Walther, F. J., Van Golde, L. M. G., Putz, G., Haagsman, H. P. & Batenburg, J. J. Multilayer formation upon compression of surfactant monolayers depends on protein concentration as well as lipid composition. An atomic force microscopy study. *Journal of Biological Chemistry* **277,** 21179–21188 (2002).

18. Krol, S., Ross, M., Sieber, M., Künneke, S., Galla, H.-J. & Janshoff, A. Formation of three-dimensional protein-lipid aggregates in monolayer films induced by surfactant protein B. *Biophysical Journal* **79,** 904–918 (2000).

19. Holten-Andersen, N., Michael Henderson, J., Walther, F. J., Waring, A. J., Ruchala, P., Notter, R. H. & Lee, K. Y. KL4 peptide induces reversible collapse structures on multiple length scales in model lung surfactant. *Biophys Journal* **101,** 2957–2965 (2011).

20. Baer, B., Souza, L. M. P., Pimentel, A. S. & Veldhuizen, R. A. New insights into exogenous surfactant as a carrier of pulmonary therapeutics. *Biochemical Pharmacology* **164,** 64–73 (2019).