STRUCTURAL ANALYSIS OF THE C-TERMINUS OF LUNG SURFACTANT PROTEIN B (SP-B)

By

ANNA KUZNETSOVA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2013

1
To Alexanders – my dear husband and my lovely son
ACKNOWLEDGMENTS

I am very thankful to my advisors Prof. Joanna Long and Prof. Gail Fanucci for their support during my doctorate studying at the University of Florida. With their research guidance and invaluable help I was able to grow on both professional and personal levels. I would like to express my gratitude to Prof. Arthur Edison and Dr. James Rocca for teaching me solution NMR techniques and data processing. I would also like to thank other members of my doctorate committee Prof. Alexander Angerhofer and Prof. Nicole Horenstein for their contribution to my research. I would also like to express my greatest appreciation to my colleagues from the research groups of Dr. Long and Dr. Fanucci for their intensive discussions, exchange of ideas. Joint research with my fellow graduate students from these two labs helped me significantly to achieve my goals. Also, I would like to thank my family and friends, who supported and encouraged me in all of my beginnings.

I also thank Dr. Charles Schwieters, Dr. Andrew Nieuwkoop and Dr. Donghua Zhou for helpful discussions and providing assistance with structure calculations. The NMR instrumentation provided by the National Science Foundation's National High Magnetic Field Laboratory is gratefully acknowledged as well as support from the University of Florida and Gates Foundation funds. The assistance of Dr. Alfred Chung in peptide synthesis is also gratefully acknowledged.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ...................................................................................................................... 4

LIST OF TABLES .............................................................................................................................. 7

LIST OF FIGURES ........................................................................................................................... 8

LIST OF ABBREVIATIONS ................................................................................................................. 12

ABSTRACT ........................................................................................................................................ 15

CHAPTER ........................................................................................................................................ 17

1 INTRODUCTION ............................................................................................................................. 17

   Pulmonary Surfactant as a Surface Active Agent in Lungs ...................................................... 17
   Pulmonary Surfactant in Alveoli ............................................................................................... 17
   Pulmonary Surfactant Cycle .................................................................................................. 17
   Pulmonary Surfactant Chemical Composition .................................................................. 18
   Lipid Trafficking by Lipid-Associated Lung Surfactant Proteins ................................... 20
   Clinical Aspects of Pulmonary Surfactant Deficiency .................................................... 21
   Pulmonary Surfactant Replacement .................................................................................. 22

2 METHODS FOR STUDYING LIPID DYNAMICS AND MEMBRANE ACTIVE PEPTIDE STRUCTURE IN MEMBRANE MIMETIC ENVIRONMENTS ........................................ 29

   Circular Dichroism (CD) ....................................................................................................... 29
   Solid State NMR Spectroscopy ............................................................................................ 32
      NMR of lipid bilayers ....................................................................................................... 33
      $^{31}$P Chemical Shift Anisotropy ..................................................................................... 34
      $^2$H Quadrupolar Coupling .......................................................................................... 39
      DePaking .......................................................................................................................... 44
   Solution State NMR Method of Protein Structure Determination .................................. 47
      Protein Resonance Assignment ....................................................................................... 48
      Restraint Generation for Protein Structure Calculation .............................................. 53
      Protein Structure Determination by Simulated Annealing Approach ...................... 54
      Protein Structure Validation .......................................................................................... 58
   Methods of Peptide Production .......................................................................................... 59
      Solid Phase Peptide Synthesis ....................................................................................... 59
      Heterologous Peptide Expression in E.Coli ................................................................. 61
   Peptide Purification by High Performance Liquid Chromatography .......................... 64

3 INTERACTION OF THE C-TERMINUS OF SP-B WITH LIPID BILAYERS MIMICKING LUNG IS MODULATED BY ACYL CHAIN SATURATION .............................................. 84
Materials and Methods .................................................................................................................. 85
Synthesis of SP-B_{59-80} .............................................................................................................. 85
Heterologous Expression of a SPB_{59-80} Double Mutant (SP-B'_{59-80}) ......................... 85
Preparation of Peptide:Lipid Samples .......................................................................................... 86
CD Experiments ........................................................................................................................... 86
Solid-state NMR Analysis ............................................................................................................ 87

Results ........................................................................................................................................... 88
Secondary Structure ...................................................................................................................... 88
The Interaction of SP-B_{59-80} with Lipid Headgroups .............................................................. 88
^{2}H NMR Studies of DPPC:POPG Lipid Miscibility on Addition of SP-B_{59-80} .................. 90
Effects of SP-B_{59-80} on Lipid Acyl Chains ............................................................................ 92
Discussion ...................................................................................................................................... 98

4 SOLUTION NMR STRUCTURE OF THE C-TERMINUS OF SP-B IN METHANOL AND IN DPC MICELLES ............................................................................................................. 117

Materials and Methods .............................................................................................................. 118
Synthesis of SP-B_{59-80} .............................................................................................................. 118
Heterologous Expression of ^{15}N-labeled SP-B_{59-80} Double Mutant (SP-B'_{59-80}) ...... 118
Circular Dichroism (CD) experiments ......................................................................................... 119
NMR samples preparation ............................................................................................................. 120
NMR experiments ........................................................................................................................ 120
NMR data analysis and structure calculation ............................................................................. 122

Results .......................................................................................................................................... 123
Secondary structure in methanol and in DPC micelles ................................................................. 123
NMR Resonance Assignments for SP-B_{59-80} in methanol .................................................... 123
Three-Dimensional Structure of SP-B_{59-80} in Methanol ......................................................... 125
NMR Resonance Assignment for SP-B'_{59-80} in DPC Micelles ................................................. 127
Three-Dimensional Structure of SP-B'_{59-80} in DPC Micelle ..................................................... 129
Discussion .................................................................................................................................... 132

5 CONCLUSIONS AND FUTURE DIRECTIONS ......................................................................... 163

APPENDIX: SCRIPTS FOR STRUCTURE CALCULATIONS AND REFINEMENT .... 167

LIST OF REFERENCES .................................................................................................................. 185

BIOGRAPHICAL SKETCH ............................................................................................................ 205
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1</td>
<td>$^1$H chemical shifts for SP-B$_{59-80}$ in methanol measured at 300 K.</td>
<td>160</td>
</tr>
<tr>
<td>4-2</td>
<td>$^3$J$<em>{HN,\alpha H}$ couplings for SP-B$</em>{59-80}$ in methanol measured at 300 K.</td>
<td>160</td>
</tr>
<tr>
<td>4-3</td>
<td>A summary of structural statistics for the 10 final structures of SP-B$<em>{59-80}$ in methanol and SP-B$'</em>{59-80}$ bound to DPC micelles.</td>
<td>161</td>
</tr>
<tr>
<td>4-4</td>
<td>$^1$H and $^{15}$N chemical shifts for $^{15}$N labeled SP-B$_{59-80}$ in DPC micelles measured at 315 K.</td>
<td>162</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Particles in the alveolar subphase.</td>
<td>26</td>
</tr>
<tr>
<td>1-2</td>
<td>Pulmonary surfactant renewal cycle.</td>
<td>27</td>
</tr>
<tr>
<td>1-3</td>
<td>Hypothetical in vitro model for specific DPPC adsorption to the air-water interface.</td>
<td>28</td>
</tr>
<tr>
<td>2-1</td>
<td>Far UV CD spectra associated with various types of secondary structure.</td>
<td>65</td>
</tr>
<tr>
<td>2-2</td>
<td>Chemical structures of glycerophospholipids.</td>
<td>66</td>
</tr>
<tr>
<td>2-3</td>
<td>Gel to liquid crystalline phase transition of the bilayer lamellar phase.</td>
<td>66</td>
</tr>
<tr>
<td>2-4</td>
<td>Types of lipid phases.</td>
<td>67</td>
</tr>
<tr>
<td>2-5</td>
<td>Origins of the powder pattern obtained in the presence of chemical shift anisotropy (CSA) interaction.</td>
<td>68</td>
</tr>
<tr>
<td>2-6</td>
<td>Chemical shift anisotropy of the rigid lipid phosphate group is a measure of the degree of order and characterizes the types of the motions that are experienced by the group in the lipid phase.</td>
<td>69</td>
</tr>
<tr>
<td>2-7</td>
<td>Polymorphisms and phosphorus NMR lineshapes.</td>
<td>70</td>
</tr>
<tr>
<td>2-8</td>
<td>Effect of the quadrupole interaction on Zeeman energy levels of nucleus with spin I=1 with axial symmetry.</td>
<td>71</td>
</tr>
<tr>
<td>2-9</td>
<td>The angles $\alpha$, $\gamma$, $\theta'$ used to define the motional averaging in lipid bilayers by relating the principal axis of the EFG tensor (usually the direction of C-$^2$H bond) to the applied magnetic field $H_o$ direction.</td>
<td>72</td>
</tr>
<tr>
<td>2-10</td>
<td>Origin of the powder pattern of $^2$H-NMR spectra observed for deuterium nuclei incorporated into phospholipid tails in membranes in uni- and multilamellar vesicles.</td>
<td>73</td>
</tr>
<tr>
<td>2-11</td>
<td>Solid-state NMR provides residual quadrupolar couplings (RQCs) that directly correspond to ensemble-averaged molecular structure.</td>
<td>74</td>
</tr>
<tr>
<td>2-12</td>
<td>Chemical shift pattern of different amino acids in 2D $^1$H-$^1$H TOCSY spectrum.</td>
<td>75</td>
</tr>
<tr>
<td>2-13</td>
<td>Protein sequence assignment via “backbone walking”.</td>
<td>76</td>
</tr>
<tr>
<td>2-14</td>
<td>A summary of solid phase peptide synthesis with Fmoc N-terminus protection group.</td>
<td>78</td>
</tr>
</tbody>
</table>
2-15 pET-31b expression system designed by Novagen and SPB<sub>C</sub> peptide purification scheme................................................................. 79

2-16 Overexpression of 18 kDa KSI-SPB<sub>59.80</sub>-His<sub>6</sub> fusion construct in BL21(DE3)pLysS cells monitored as function of time after overexpression induction................................................................. 80

2-17 Nickel-affinity chromatogram of purification of KSI-SPB<sub>59.80</sub>-His<sub>6</sub> fusion construct from inclusion bodies in denaturing conditions................................................................. 80

2-18 Cleavage efficiency of KSI-SPB<sub>59.80</sub>-His<sub>6</sub> fusion construct by cyanogen bromide in 80% formic acid was monitored as a function of time ......................... 81

2-19 Pure SPB<sub>59.80</sub> peptide in RP-HPLC fractions after lyophilization.......................... 81

2-20 Mass spectra of expressed SP-B<sub>59.80</sub>, wild type SP-B<sub>59.80</sub> and <sup>15</sup>N-SP-B<sub>59.80</sub> peptides.................................................................................. 82

3-1 CD spectra at 45 °C of SP-B<sub>59.80</sub> at a P:L molar ratio of 1:100 in 4:1 DPPC:POPG and 3:1 POPC:POPG................................................................. 106

3-2 Phosphorous NMR spectra of 4:1 DPPC-d<sub>62</sub>:POPG MLVs with SP-B<sub>59.80</sub> ... 107

3-3 Phosphorous NMR spectra of 3:1 POPC-d<sub>31</sub>:POPG MLVs with SP-B<sub>59.80</sub> ... 108

3-4 DSC scans for 4:1 DPPC-d<sub>62</sub>:POPG LUVs with SP-B<sub>59.80</sub> ...................... 109

3-5 Deuterium NMR spectra as a function of temperature and SP-B<sub>59.80</sub> molar ratio for 4:1 DPPC-d<sub>62</sub>:POPG MLVs and 4:1 DPPC:POPG-d<sub>31</sub> MLVs. ............................ 109

3-6 First moment, M<sub>1</sub>, as a function of temperature, for the <sup>2</sup>H NMR spectra plotted in Fig.3-5 ............................................................................. 110

3-7 Deuterium NMR spectra of 4:1 DPPC-d<sub>62</sub>:POPG MLVs with SP-B<sub>59.80</sub> at the indicated P:L molar ratios. ................................................................. 111

3-8 Deuterium NMR spectra of 3:1 POPC-d<sub>31</sub>:POPG MLVs with SP-B<sub>59.80</sub> at the indicated P:L molar ratios. ................................................................. 112

3-9 Order parameter profiles for the sn-1 chains of DPPC-d<sub>62</sub> and POPG-d<sub>31</sub> in 4:1 DPPC:POPG MLVs at 44 °C with SP-B<sub>59.80</sub> ......................................................... 113

3-10 Order parameter profiles for the sn-1 chains of POPC-d<sub>31</sub> and POPG-d<sub>31</sub> in 3:1 POPC:POPG MLVs at 44 °C with SP-B<sub>59.80</sub> ......................................................... 114

3-11 Changes in d<sub>31</sub>-palmitoyl acid chain order parameters in 4:1 DPPC:POPG MLVs and 3:1 POPC:POPG MLVs on addition of SP-B<sub>59.80</sub> ................................. 115

3-12 Models of SP-B<sub>59.80</sub> interactions with the two lipid environments .................. 116
4-1 CD spectra of SP-B_{59-80} in methanol at 303K and bound to DPC micelles at 318K ........................................................................................................................................ 137
4-2 300ms NOESY spectrum for SP-B_{59-80} in methanol at 300 K ........................................... 138
4-3 Summary of the NOE interactions and CSI measurements for SP-B_{59-80} in methanol .................................................................................................................................................... 140
4-4 Summary of the NOE interactions for SP-B_{59-80} in methanol and for SP-B'_{59-80} in DPC micelles .................................................................................................................................................... 141
4-5 SP-B_{59-80} structure in methanol .............................................................................................. 142
4-6 {\textsuperscript{1}}H-{\textsuperscript{15}}N HSQC spectrum and assignments for {\textsuperscript{15}}N-labeled SP-B'_{59-80} in DPC micelles at 315K .................................................................................................................................................... 143
4-7 Alternating {\textsuperscript{1}}H-{\textsuperscript{1}}H TOCSY and {\textsuperscript{1}}H-{\textsuperscript{1}}H NOESY from {\textsuperscript{15}}N-edited 3D TOCSY-HSQC and NOESY-HSQC spectra of {\textsuperscript{15}}N-labeled SP-B'_{59-80} peptide in DPC micelles .................................................................................................................................................... 143
4-8 Summary of the NOE resonances for SP-B'_{59-80} in DPC micelles .................................... 146
4-9 Three-dimensional structure of SP-B'_{59-80} peptide in DPC micelles. ................................ 147
4-10 SP-B'_{59-80} in DPC micelles with residues colored according to their charge ...... 148
4-11 Three-dimensional surface models of SP-B_{59-80} structure in methanol and in DPC micelle .................................................................................................................................................... 149
4-12 Possible mechanism of DPPC trafficking by SP-B_{59-80}. ................................................. 149
4-13 The Ramachandran plot of ensemble of 10 SP-B_{59-80} structures in methanol .. 150
4-14 The Chi1-Chi2 plots show the chi1-chi2 sidechain torsion angle combinations for 10 structure ensemble of SP-B_{59-80} in methanol. ................................................................. 151
4-15 The Ramachandran plot of ensemble of 10 SP-B_{59-80} structures in methanol with data points colored according to G-factor ................................................................. 153
4-16 Plots of main chain parameters for 10 SP-B_{59-80} structures in methanol compared to the well-refined structures at a similar resolution. .................................................. 154
4-17 The Ramachandran plot of ensemble of 6 SP-B'_{59-80} structures in DPC micelles .................................................................................................................................................... 155
4-18 The Chi1-Chi2 plots show the chi1-chi2 sidechain torsion angle combinations for ensemble of 6 structures of SP-B'_{59-80} in DPC. ................................................................. 157
The Ramachandran plot of ensemble of 6 SP-B\textsubscript{59-80} structures in DPC with data points colored according to G-factor.

Plots of main chain parameters for 6 SP-B\textsubscript{59-80} structures in DPC compared to the well-refined structures at a similar resolution.

Molecular models of SP-B\textsubscript{59-80} directing specific DPPC adsorption to the air-water interface and facilitating tight packing interactions between the bilayers of tubular myelin.
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARDS</td>
<td>Acute respiratory distress syndrome</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>COSY</td>
<td>Correlation spectroscopy</td>
</tr>
<tr>
<td>CSA</td>
<td>Chemical shift anisotropy</td>
</tr>
<tr>
<td>CSI</td>
<td>Chemical shift index</td>
</tr>
<tr>
<td>C-TERMINUS</td>
<td>Carboxy terminus</td>
</tr>
<tr>
<td>DIPSI</td>
<td>Decoupling In the Presence of Scalar Interactions</td>
</tr>
<tr>
<td>DPC</td>
<td>Dodecylphosphatidylcholine</td>
</tr>
<tr>
<td>DPPC</td>
<td>Dipalmitoylphosphatidylcholine</td>
</tr>
<tr>
<td>DPPG</td>
<td>Dipalmitoylphosphatidylglycerol</td>
</tr>
<tr>
<td>DSC</td>
<td>Differential scanning calorimetry</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EFG</td>
<td>Electric field gradient</td>
</tr>
<tr>
<td>EPR</td>
<td>Electron paramagnetic resonance</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier transform infrared</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione-S-transferase</td>
</tr>
<tr>
<td>H,C,N</td>
<td>Proton, Carbon, Nitrogen</td>
</tr>
<tr>
<td>HFIP</td>
<td>Hexafluoro-2-propanol</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-performance liquid chromatography</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single-quantum coherence</td>
</tr>
<tr>
<td>IVM</td>
<td>Internal variable module</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>KSI</td>
<td>Ketosteroid isomerase</td>
</tr>
<tr>
<td>LB</td>
<td>Lamellar body</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>LS</td>
<td>Lung surfactant</td>
</tr>
<tr>
<td>LUV</td>
<td>Large unilamellar vesicle</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>Matrix-assisted laser desorption time-of-flight</td>
</tr>
<tr>
<td>MAS</td>
<td>Magic angle spinning</td>
</tr>
<tr>
<td>MBP</td>
<td>Maltose binding protein</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>MLV</td>
<td>Multilamellar vesicle</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>Nuclear Overhauser effect</td>
</tr>
<tr>
<td>NOESY</td>
<td>Nuclear Overhauser enhancement spectroscopy</td>
</tr>
<tr>
<td>N-TERMINUS</td>
<td>Amino terminus</td>
</tr>
<tr>
<td>PAS</td>
<td>Principal axis system</td>
</tr>
<tr>
<td>PC</td>
<td>Phosphatidylcholine</td>
</tr>
<tr>
<td>PE</td>
<td>Phosphatidylethanolamine</td>
</tr>
<tr>
<td>PG</td>
<td>Phosphatidylglycerol</td>
</tr>
<tr>
<td>POPC</td>
<td>Palmitoyloleoylphosphatidylcholine</td>
</tr>
<tr>
<td>POPG</td>
<td>Palmitoyloleoylphosphatidylglycerol</td>
</tr>
<tr>
<td>PS</td>
<td>Pulmonary surfactant</td>
</tr>
<tr>
<td>RDS</td>
<td>Respiratory distress syndrome</td>
</tr>
<tr>
<td>RF</td>
<td>Radiofrequency</td>
</tr>
<tr>
<td>RMSD</td>
<td>root mean square deviation</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RQC</td>
<td>Residual quadrupolar coupling</td>
</tr>
<tr>
<td>RP-HPLC</td>
<td>Reverse phase high performance liquid chromatography</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>SAPLIPs</td>
<td>Saposin-like proteins</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SP-A</td>
<td>Surfactant protein A</td>
</tr>
<tr>
<td>SP-B</td>
<td>Surfactant protein B</td>
</tr>
<tr>
<td>SP-B&lt;sub&gt;59-80&lt;/sub&gt;</td>
<td>Surfactant protein B residues 59 to 80</td>
</tr>
<tr>
<td>SP-B&lt;sub&gt;63-78&lt;/sub&gt;</td>
<td>Surfactant protein B residues 63 to 78</td>
</tr>
<tr>
<td>SP-B&lt;sub&gt;C&lt;/sub&gt;</td>
<td>Surfactant protein B C-terminus (residues 59 to 80)</td>
</tr>
<tr>
<td>SP-C</td>
<td>Surfactant protein C</td>
</tr>
<tr>
<td>SP-D</td>
<td>Surfactant protein D</td>
</tr>
<tr>
<td>SPPS</td>
<td>Solid phase peptide synthesis</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TFE</td>
<td>Trifluoroethanol</td>
</tr>
<tr>
<td>TM</td>
<td>Tubular myelin</td>
</tr>
<tr>
<td>TOCSY</td>
<td>Total correlation spectroscopy</td>
</tr>
<tr>
<td>TSP</td>
<td>3-(Trimethylsilyl)propionic-2,2,3,3-d&lt;sub&gt;4&lt;/sub&gt; acid sodium salt</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
STRUCTURAL ANALYSIS OF THE C-TERMINUS OF LUNG SURFACTANT PROTEIN B (SP-B)

By
Anna Kuznetsova

May 2013

Chair: Gail E. Fanucci  
Cochair: Joanna R. Long  
Major: Chemistry

Lung surfactant is a complex lipid-protein substance that coats the inner layer of alveoli and facilitates breathing by minimizing surface tension. Surfactant protein B (SP-B) is of particular importance since it may specifically enrich dipalmitoylphosphatidylcholine (DPPC) at the air-water interface by trafficking and sorting of lipids. In vivo and in vitro studies have demonstrated that the C-terminus of SP-B (SP-B_{59-80}) promotes lipid exchange and rapid transfer from bilayer stores into interfacial monolayers. Little is known about the atomic-level structure and dynamics of this peptide in the lipid environment. Solution and solid state NMR were used to probe the structure and dynamics of SP-B_{59-80} in detergent micelles and lipid vesicles to compare the effects of the different lipid environments on peptide helical structure and the effects of the peptide on lipid dynamics.

Using standard TOCSY and NOESY experiments, we were able to assign all resonances for SP-B_{59-80} in methanol and in DPC micelles; CD studies indicate the peptide is helical in both lipid environments. Using XPLOR-NIH, an ensemble of 10 structures with backbone RMSD 0.68Å was calculated for SP-B_{59-80} in methanol and an
ensemble of 6 structures with backbone RMSD 0.67Å was calculated for SP-B_{59-80} in DPC micelles. Both structures are unusual amphipathic structures with a helix-turn-helix motif formed due to the presence of proline at position 67. We also detected hydrophobic contacts in methanol and hydrophilic side chain contacts in micelles due to the two helices positioned with their hydrophilic sides facing outwards in polar methanol solvent and facing inwards in the non-polar environment of the DPC micelle. The interchange between two conformations suggests a possible mechanism for interaction of the C-terminus of SP-B with DPPC and subsequent lipid transfer. The N- and C-terminal helices, with distinct charged and non-polar faces connected by a flexible loop, “roll” around their axes to change orientation of their charged faces relative to each. If a DPPC molecule is bound to SP-B_{59-80} during this “rolling” motion, it could be transferred from the lipid bilayers to the monolayer at the air/water interface or between lipid leaflets. The net result would be destabilization of local bilayer structure or even translocation of the DPPC lipid.
CHAPTER 1
INTRODUCTION

Pulmonary Surfactant as a Surface Active Agent in Lungs

A fundamental problem associated with breathing stems from the high surface area of the lung. Lungs require such a large surface area to carry out gas exchange at a physiologically appropriate level.\(^{(1)}\) This is attained through the alveoli – tiny aqueous lined air sacs with a high radius of curvature and a surface area of \(\sim 200-380 \text{ m}^2\).\(^{(2)}\) However, according to LaPlace’s law \((p = 2\gamma/r)\), the smaller the alveolar radius \((r)\) the higher the pressure \((p)\) required to keep it inflated. There is a general need to minimize surface tension \((\gamma)\) in small, gas-containing, aqueous lined structures by introducing surfactants to stabilize them. With a standard air/water interface it would be impossible to inflate the lungs at atmospheric pressure. The task of reducing surface tension and increasing lung compliance is accomplished by pulmonary surfactant (PS), which covers the respiratory surface of the alveoli. Pulmonary surfactant (PS) is a surface-active lipid–protein complex lining the alveoli that reduces alveolar surface tension to near-zero values at the lung air-liquid interface and allows inflation and oxygen exchange at ambient pressure.

Pulmonary Surfactant in Alveoli

Pulmonary Surfactant Cycle

PS is synthesized by specialized cells of the alveolar epithelium - alveolar epithelial type II cells, also called type II pneumocytes, which make up 5% of the epithelial surface. PS is stored in these cells in the form of tightly packed bilayers in vesicles called lamellar bodies (LBs). Upon secretion into the extracellular space, the LBs form an extensive lipid-rich network of bilayers termed tubular myelin (TM), that
covers rapidly and efficiently the whole respiratory surface (Fig. 1-1, Fig. 1-2). From tubular myelin phospholipids are transferred to the monolayer at the air-water interface to efficiently cover and stabilize this interface for adequate respiration (Fig. 1-2 and Fig. 1-3).

**Pulmonary Surfactant Chemical Composition**

PS is conserved across mammalian species and contains 80% phospholipids, 5–10% neutral lipids (mainly cholesterol), and 10% proteins by weight. The primary phospholipids of surfactant are dipalmitoylphosphatidylcholine - DPPC (~40-50%), palmitoyloleoylphosphatidylcholine - POPC (~30%) and anionic palmitoyloleoylphosphatidylglycerol - POPG (~10%). In an aqueous environment phospholipids self-assemble into bilayers (the structural form of LBs), but at polar/non-polar interfaces (e.g. the air–liquid interface) phospholipids can form monolayers, with the headgroups oriented toward the aqueous phase and the hydrophobic acyl chains pointing toward the air. DPPC, which has melting temperature (T<sub>m</sub>) of 41.5°C, is most likely favored as the major phospholipid of PS at the air-water interface because its fully saturated acyl chains can be very tightly packed and withstand high lateral pressure. The higher the concentration of phospholipid molecules at the interface, the fewer the number of water molecules exposed to air providing the large reductions of surface tensions at the alveolar air–liquid interface. However, to quickly cover the alveolar surface, surfactant lipids also have to be fluid at physiological temperature (37°C). This is attained by mixing DPPC with lower-melting unsaturated phospholipids.

However, lipids alone cannot achieve the near zero surface tension present in lung surfactant. The unique physical properties of PS are determined by low levels of lipid-associated proteins: Surfactant protein B (SP-B) and Surfactant protein C (SP-C).
SP-B and SP-C are extremely hydrophobic, soluble in organic solvents such as chloroform/methanol mixtures, and thus co-purify with lipids in chloroformic extractions of PS. Due to their low concentration and high hydrophobicity, little is known about the atomic-level structure and dynamics of these proteins. The other two surfactant proteins SP-A and SP-D are hydrophilic protein complexes that protect the respiratory surface from pathogens.

Mature SP-B is an extremely hydrophobic, lipid associated homodimer, made up of two 79-81 amino acid monomers that are disulfide-linked. It contains high levels of valine, leucine, isoleucine, alanine, phenylalanine, and tryptophan. The native form of the protein contains six highly conserved cysteines that form three intramolecular disulfide bonds and a seventh cysteine residue that forms an intermolecular disulfide bond. The intramolecular disulfide bonds define a fold that has been conserved for an estimated 300 million years. The unique pattern of intramolecular disulfide bonds, the resulting fold and its hydrophobicity place SP-B in the saposin like family of proteins (SAPLIP). Other proteins in the saposin family include amoebapores from Entamoeba histolytica, acid sphingomyelinase, acyloxyacyl hydrolase, and sphingolipid activator proteins A–D (saposins A–D). Like other members of the saposin superfamily, SP-B is primarily helical in lipid bilayers and structure-promoting solvents. However, SP-B is unique in this family with its high hydrophobicity and integration into lipid assemblies. FTIR and CD spectroscopies indicate that in lipid bilayers the amphipathic helices of SP-B lie in the plane of the lipid lamellae. One of the interesting features of SP-B and a major difference between SP-B and SP–C is that SP-B bears intermittent highly charged residues, 9 of 12 which
are basic and localized in amphipathic helix regions of the N- and C-terminal segments of the protein, imparting a net cationic charge to SP-B of > + 5 at physiologic pH.\(^{(26, 27)}\) A net cationic charge of SP-B at physiologic pH may promote a selective interaction of SP-B with the anionic PG, ample in surfactant.\(^{(28)}\)

**Lipid Trafficking by Lipid-Associated Lung Surfactant Proteins**

Hydrophobic surfactant proteins B and C (SP-B\(^{(29)}\) and SP-C,\(^{(30, 31)}\) promote rapid transfer of phospholipids from bilayers of TM to interfacial monolayers at the lung air-water interface (Fig. 1-3). SP-B and SP-C are required as mediators at relatively low concentration (\(~0.1\ \text{mol}\%)\) to make the transfer of phospholipids through the aqueous hypophase rapid and efficient.\(^{(10)}\) However, their mechanisms of action are still lacking as are atomic-resolution structures of SP-B and SP-C in lipid membranes or lipid-like environments. FTIR indicates that in lipid bilayers SP-B is a helical protein with the helices in the plane of the lipid lamellae.\(^{(32)}\) In contrast, SP-C adopts a helical structure with a transmembrane orientation.\(^{(33)}\) Both SP-B and SP-C distribute preferentially in disordered regions of membranes and interfacial monolayers.\(^{(9, 34)}\)

In contrast with SP-C, SP-B, which directly interacts with the phospholipids in surfactant, is particularly critical to its function.\(^{(35)}\) SP-B is absolutely required for survival and normal respiration. Hereditary SP-B deficiency is lethal in humans\(^{(36)}\) and in SP-B knockout mice;\(^{(37)}\) antibodies against SP-B cause respiratory distress syndrome (RDS) \textit{in vivo}\(^{(38)}\) and insufficient SP-B is a leading cause of respiratory distress in premature infants. SP-B is strictly required not only for the biogenesis and packing of PS into lamellar bodies,\(^{(39, 40)}\) but it also underpins the unique PL packing and dynamics observed in PS. It is also thought that SP-B may specifically enrich DPPC at the air-water interface by trafficking and sorting of lipids in bulk PS,\(^{(41)}\) but the mechanism of
lipid trafficking by SP-B is unknown. SP-B could promote formation and stabilization of certain non-bilayer intermediate structures required for efficient phospholipid transfer.\(^{(42)}\) Monolayer experiments with films containing SP-B have shown that SP-B facilitates ordered and reversible compression-driven structural transitions by formation of associated bilayer patches, which can quickly reinsert into the interfacial monolayer upon expansion.\(^{(43)}\) Compression of pure lipid films usually ends with an irreversible collapse.\(^{(44, 45)}\) Thus respiratory failure associated with absence of SP-B is probably caused not by impaired PS production but rather by the loss of a proper dynamic flow of lipids.\(^{(46)}\)

**Clinical Aspects of Pulmonary Surfactant Deficiency**

Deficient or dysfunctional PS leads to severe respiratory disease. Maturation of type II pneumocytes and PS secretion occur late in gestation and PS is often insufficient in premature infants leading to Respiratory Distress Syndrome (RDS), a major cause of perinatal mortality.\(^{(46)}\) RDS is associated with high alveolar surface tension, increased work of breathing and alveolar collapse.\(^{(47)}\) Impaired gas exchange at the alveolar-capillary barrier causes a decrease in blood oxygen levels (hypoxia) and an increase in carbon dioxide (respiratory acidosis).\(^{(48)}\) Adult patients suffering from extensive lung inflammation, infection or trauma often develop acute RDS (ARDS) due to inactivation of PS by inflammatory by-products and blood components leaked into the airways through a deteriorated alveolar-capillary barrier.\(^{(46)}\) The pivotal role of PS in RDS and ARDS has led to the development of therapies which rely on either synthetic or modified natural PS (extracted from bovine or porcine sources). Current FDA-approved therapy for RDS in premature infants relies on animal-derived PS. Such treatment creates the risk of severe immune response if administered repeatedly, but this is not as crucial for
RDS treatment in preterm babies since they start producing their own surfactant when they start breathing. In contrast, the need for regular administration of surfactant preparations for treatment of ARDS in adults and mature immune function makes natural PS therapy impossible in older populations due to surfactant inhibition and severe immune response.\(^{(49)}\) Consequently, synthetic lung surfactant replacements are an attractive goal for treatment of RDS and ARDS, as they are considerably cheaper to produce, remove the immunologic risks, and may be crucial to effectively treating ARDS. However, although current synthetic PS replacements are less susceptible to degradation by plasma proteins and remove the risk of immune response, they reduce surface tension much less than animal-derived surfactant. To better mimic the natural surfactant we need to elucidate the molecular mechanisms by which surfactant proteins and membranes modulate respiratory physiology.\(^{(50)}\)

**Pulmonary Surfactant Replacement**

The hydrophobicity and disulfide bridges in SP-B have made heterologous expression and purification of the full-length protein in large quantities impractical. Synthetic, peptide-based replacements of SP-B for treatment of RDS and ARDS have received notable attention not only because they are much easier to make with high yield and purity than the full-length protein,\(^{(51, 52)}\) but also because they remove the immunologic risks associated with animal-derived surfactant.\(^{(53-55)}\)

While the entire 80 amino acid SP-B protein is essential for lung surfactant organization, dynamics and respiration, fragments of the native sequence have shown significant biophysical function. Peptides corresponding to N-terminal and C-terminal 20-25 residue fragments of the native sequence have shown surface activity similar to native SP-B,\(^{(17, 56)}\) with maximal activity achieved using a chimeric construct of the two
This could be explained by the fact that both N-terminal and C-terminal peptides form cationic amphipathic helices in SDS micelles and lipid bilayers, which makes them the most critical fragments of SP-B for phospholipid binding and their rapid transfer from bilayer stores into interfacial monolayers, but their individual roles in lipid trafficking are not well understood. Both N-terminal and C-terminal peptides promote lipid exchange in vitro and restore lung function in animal models of respiratory distress. Surfactant preparations containing these synthetic peptides improve oxygenation and lung compliance in surfactant-deficient animal models.

The preponderance of leucines and relative lack of prolines and aromatic sidechains in the C-terminus distinguish it from the N-terminus, which contains four prolines as well as a phenylalanine and a tryptophan. The peptides also have varying spacing between hydrophilic and hydrophobic amino acids. These differences lead to varying secondary structures and penetration into lipid lamellae, and we have documented they have different effects on phospholipid dynamics.

Although to date the N-terminal fragment has been more extensively studied in terms of its molecular properties and structural adaptation in lipid environments, many in vivo activities of SP-B are fulfilled by its C-terminal end, specifically residues 59–80 (SP-B_{59-80}). SP-B_{59-80} has shown efficacy in altering phospholipid properties based on in vitro assays measuring surface tension, such as pulsating bubble surfactometry, as well as in vivo assays demonstrating gain of function in surfactant-deficient fetal rabbits.

The most successful potential synthetic replacement for SP-B named KL₄ peptide, KLLLLKLLLLKLLLLKLLLLK, was designed to mimic the charge distribution
and hydrophilic/hydrophobic ratio of the primary sequence of SP-B_{59–80}. This simple peptide mimetic of SP-B_{59–80} relieves respiratory distress syndrome,\(^{(76)}\) however KL_{4}/lipid preparations are not as efficient in lowering surface tension as native PS. Moreover, an FTIR study of KL_{4} in DPPC/DPPG concluded the peptide can form a helix spanning the bilayers and it has been posited that KL_{4} might more closely mimic SP-C rather than C-terminus of SP-B.\(^{(77)}\) Thus, understanding how SP-B_{59–80} functions in the lipid environment would allow the development of mimetics which are more stable and efficient in lowering surface tension than KL_{4}, but less susceptible to degradation by plasma proteins than native SP-B. The recent solution NMR study of residues 63–78 of the C-terminus of SP-B (SP-B_{63–78}) reconstituted in either SDS micelles or the organic solvent HFIP found the first five residues to be unstructured and established that the rest of the sequence formed a helix in both SDS micelles and organic solvent.\(^{(60)}\) This study and a previous CD study\(^{(59)}\) of SP-B_{59–80} in TFE and SDS micelles are the only documented structural assessments of the C-terminal region of SP-B. SP-B_{59–80} is believed to form an amphipathic helix involved in lipid organization, but direct structural measurements in lipid contexts have not been documented. Detailed study of the interaction of SP-B with phospholipids is necessary to create more effective clinical surfactant preparations. Based on understanding the properties of SP-B_{59–80}, new non-natural peptide analogs of SP-B that exploit the protein's qualities in altering lipid dynamics could be made to extend or enhance artificial lung surfactant therapies. Thus, understanding how SP-B_{59–80} affects the molecular and biophysical properties of lipids is of particular relevance to the treatment of various forms of RDS.
The roles of membrane proteins in many physiological events can be understood only by studying their structure and dynamics in their native lipid environments. Surfactant protein B minimizes surface tension at the alveolar air-water interface by improving dynamics of the lipid redistribution at the interface through trafficking and sorting of LS lipids, especially DPPC lipid enriched at the air-water interface. The C-terminus of SP-B (SP-B$_{59-80}$) may select for DPPC lipid and promote its exchange and rapid transfer from bilayer stores into interfacial monolayers. The synthetic LS mimics in this study utilized binary mixture DPPC:POPG to mirror the interfacial LS phases, whereas POPC:POPG mixture was used as the mimic of bilayer storage phases of LS. Chapter 3 provides more detail on the partitioning of SP-B$_{59-80}$ peptide into these LS mimics and effects of SP-B$_{59-80}$ on lipid dynamics. Chapter 4 discusses the structures of SP-B$_{59-80}$ peptide in methanol solution and in DPC micelles, which mimic DPPC-rich LS environments. Chapter 2 introduces the theory behind the techniques used in this work. Chapter 5 summarizes all results and provides conclusions for the dissertation as well as the incentives for future experiments.
Figure 1-1. Particles in the alveolar subphase. In this electron micrograph section of a rat lung, lamellar bodies (LB) are seen forming tubular myelin (TM) (bar at lower right=1.0 μm). The remaining vesicular structures may represent both used and rejected surfactant materials. Inset: detail of tubular myelin at lower left, showing small projections in the corners, thought to represent SP-A (bar=0.1 μm). This adult rat was exposed to NO$_2$ for 48 h, but no differences were seen from controls. Figure borrowed with permission from: Goerke, J.(1998), Pulmonary surfactant: functions and molecular composition, BBA 1408, 79-89.
Figure 1-2. Pulmonary surfactant renewal cycle: PS is synthesized in alveolar epithelial type II cells, stored there as tightly packed bilayers in lamellar bodies (LBs), and secreted into the extracellular space to form an extensive lipid-rich network of bilayers - tubular myelin (TM). From TM bilayers phospholipids are transferred to the monolayer at the air-water interface to efficiently cover and stabilize this interface for adequate respiration (see also next figure). Figure borrowed and adapted with permission from: Serrano, A.G., and Perez-Gil, J. (2006) Protein-lipid interactions and surface activity in the pulmonary surfactant system, *Chemistry and Physics of Lipids* 141, 105-118.
Figure 1-3. Hypothetical in vitro model for specific DPPC adsorption to the air-water interface. Phospholipids (mainly DPPC) are transferred from tubular myelin bilayers to the monolayer at the air-water interface to efficiently cover and stabilize this interface for adequate respiration. Adapted with permission from: Blanco, O., and Perez-Gil, J. (2007) Biochemical and pharmacological differences between preparations of exogenous natural surfactant used to treat Respiratory Distress Syndrome: Role of the different components in an efficient pulmonary surfactant, *European Journal of Pharmacology* 568, 1-15.
CHAPTER 2
METHODS FOR STUDYING LIPID DYNAMICS AND MEMBRANE ACTIVE PEPTIDE STRUCTURE IN MEMBRANE MIMETIC ENVIRONMENTS

Circular Dichroism (CD)

The circular dichroism (CD) technique allows structural assessment of proteins and provides quick measurement of structural changes, which are often essential to their biological function. This method allows monitoring of structure under a variety of conditions in which proteins actually operate, including interacting with lipid bilayers. In chapters 3 and 4 CD spectra provided prompt assessment of SP-B59-80 peptide structure in different environments (methanol, DPC micelles, lipid bilayers mimicking LS) and showed that it is helical in all of them with some minor possible secondary structure differences due to environment. Consequently, the next chapters have been focused on high resolution structures of SP-B59-80 in methanol and DPC micelles, which could provide insights into its structure and interaction with lipid environment of LS.

Light is an unpolarized electromagnetic wave with randomly oriented electric vectors, which propagates equally in all directions. Plane polarized light is made up of 2 circularly polarised components of equal magnitude, one rotating counter-clockwise (left handed, L) and the other clockwise (right handed, R). Circular dichroism (CD) refers to the differential absorption of these 2 components. If, after passage through the sample being examined, the L and R components are not absorbed or are absorbed to equal extents (for an achiral chromohore), the recombination of L and R would look like circle and would regenerate radiation polarized in the original plane. However, if L and R are absorbed to different extents (for a chiral chromohore), the resulting radiation after recombination of unequal L and R vectors would look like ellipse and would be said to possess elliptical polarization, which gives rise to CD signal. Thus CD
signal is measured as the difference in absorbance between the L and R circularly polarized components \( \Delta A = A_L - A_R \), and like any light absorption value it can be rewritten in terms of Beer’s law:

\[
\Delta A = (\varepsilon_L - \varepsilon_R)C_l
\]

(2-1)

where \( \varepsilon_L \) and \( \varepsilon_R \) are the molar extinction coefficients for the L and R circularly polarized components of light respectively; \( C \) is the molar concentration and \( l \) is the path length in cm. Although \( \Delta A \) is measured, for historical reasons CD signal is always reported in degrees of ellipticity (\( \theta \)). When there is no difference in the absorbance of right- and left-circular polarized light \( (E_R = E_L) \), \( \theta \) is 0° and the resulting combined radiation is linearly polarized. When there is complete absorbance of the circular polarized light in one direction \( (E_R = 0 \text{ or } E_L = 0) \), \( \theta \) is 45° and the resulting radiation is circularly polarized. The case in between is elliptically polarized light with ellipticity defined through its tangent, which is approximately equal to the angle \( \Theta \) itself for small angles:

\[
\tan \theta = (E_R - E_L) / (E_R + E_L) \approx \theta
\]

(2-2)

where \( E_R - E_L \) and \( E_R + E_L \) are the minor and major axes of the ellipse resulting from addition of two light components with different amplitudes and circularly polarized in different directions.

By analogy to the definition of molar circular dichroism \( \Delta \varepsilon \), molar ellipticity \( [\theta] \) is circular dichroism in degrees of ellipticity \( \theta \) corrected for concentration (in mols/L) and cell length (in cm):

\[
\theta = [\theta] [M] l
\]

(2-3)

The CD signal is obtained as ellipticity in units of millidegrees and is normalized according to the protein concentration to yield molar ellipticity \([\theta]\) per amino acid.\(^{(78)}\)
\[ \theta = \frac{\theta_{\text{deg}}}{# \text{AA}[M]l} \]  

(2-4)

The number of amino acids in the protein is denoted as \#AA, [M] is the molar concentration, \( \theta \) deg is the CD signal in millidegrees, and \( l \) is the path length in centimeters of the cuvette.

In proteins, the chromophores of interest include the peptide bond (absorption below 240 nm), aromatic amino acid side chains (absorption in the range of 260 to 320 nm) and disulphide bonds (weak broad absorption bands centered around 260 nm).\(^{(79)}\)

Protein secondary structure is defined by the conformation of the peptide bonds and the CD absorption of peptide bonds in the far UV region (240 nm and below) is characteristic of the different types of regular secondary structure found in proteins (Fig. 2-1).\(^{(79)}\) The spectrum of a polypeptide in random conformation displays a small positive \( n \rightarrow \pi^* \) transition at \( \sim 230 \) nm and a large single \( \pi \rightarrow \pi^* \) transition at approximately 195 nm. The spectrum of the \( \alpha \)-helix vice versa displays a large negative \( n \rightarrow \pi^* \) transition at 222 nm. This usually weak forbidden \( n \rightarrow \pi^* \) transition is enhanced in \( \alpha \)-helix, because even though \( n \) and \( \pi \) orbitals of an isolated amide bond are orthogonal, the orbitals of neighboring chromophores are not necessarily orthogonal and thus the transition is more likely to occur in a polymer between neighboring amide bonds. The strong \( \pi \rightarrow \pi^* \) transition in \( \alpha \)-helix spectrum is split into two transitions because of exciton coupling with negative band at approximately 208 nm and a positive band at approximately 192 nm.\(^{(81)}\) The CD spectrum of the antiparallel \( \beta \)-structure has a negative band at 218 nm and a positive band at 195 nm also showing evidence of exciton coupling in the \( \pi \rightarrow \pi^* \) transition, however, the splitting of the transitions are different than in the case of the \( \alpha \)-helix.
Consequently, CD spectrum of the α-helical conformation (Fig. 2-1) has characteristic shape with two negative peaks at 208 nm and 222 nm of the equal intensity. However, CD spectra of SP-B_{59-80} in different environments showed that the peak at 208 nm was much lower than the peak at 222 nm and suggested helical conformation of SP-B_{59-80} may have some deviations from purely α-helical state.

**Solid State NMR Spectroscopy**

Solid state NMR was used to study the effect of SP-B_{59-80} on lipid dynamics in bilayers, which provided some indirect evidence of the peptide partitioning depth into the lipid bilayers and allowed some assumptions about the structure of SP-B_{59-80} in these lipid bilayers.

In solid state NMR besides the Zeeman interaction of the nuclear spin with the external magnetic field, there are additional perturbations of the total Hamiltonian due to magnetic interactions (dipolar coupling, chemical shift) and electric interactions (quadrupolar coupling). The corresponding 3x3 tensors that describe these couplings in every direction provide information regarding both the structure and dynamics of biomolecular system. The principal values and principal axis systems (PAS) of the various coupling tensors in their motionally averaged form yield structural information about the system and define overall lineshape observed in NMR spectra, whereas their fluctuations are related to the dynamics of the system or the mean-squared amplitudes and rates of the motions and lead to characteristic relaxation behavior. Structural fluctuations of fluid membrane lipids give rise to averaging of the coupling tensors affecting the NMR lineshapes and relaxation times that are determined by the corresponding coupling tensor. Solid-state NMR is thus a unique technique because it
provides information about both structure and dynamics at the level of individual atoms.\textsuperscript{(82)} Studies of biological membranes and other lipid systems by solid-state NMR have been mostly carried out using phosphorus ($^{31}$P) and deuterium ($^2$H) nuclei of lipid molecules. The unique feature of the method is that it is used to study not only macroscopic properties of lipid systems, such as lipid phases in the system and phase transitions between them, but also microscopic properties at the atomic level, such as local environment in a very small submolecular region of the membrane. The physical state of the hydrocarbon chain region of phospholipid bilayer is studied by $^2$H NMR by determining motional anisotropies of C-$^2$H bonds,\textsuperscript{(82)} whereas $^{31}$P NMR is a valuable tool for lipid phase behavior analysis by looking at the motional properties of lipid headgroups via $^{31}$P chemical shift anisotropies (CSA).

**NMR of lipid bilayers**

Representative glycerophospholipids shown in Figure 2-2 possess polar head groups X that differ in their size, capacity for hydrogen bonding, charge and the nonpolar acyl chains, which vary in their length and degree of unsaturation.\textsuperscript{(82)} Biological membranes are phospholipid bilayers that form due to the hydrophobic effect which causes the hydrophobic lipid acyl chains partition within the bilayer interior to avoid contact with water; the polar and/or charged head groups interact with water at the membrane surface. In excess water phospholipid bilayers undergo phase transition from highly ordered lamellar gel phase with tiled chains ($L_\beta$) to the lamellar liquid-crystalline phase ($L_\alpha$) as temperature increases (Fig. 2-3).\textsuperscript{(82)} Other types of phospholipid nanostructures are also possible, including the normal and inverted hexagonal phases, cubic phases and micellar (Fig. 2-4),\textsuperscript{(83, 84)} depending on the structures of the individual lipid molecules. Liquid-crystalline lamellar phase is classified
as smectic A lyotropic phase under liquid crystal classification. In lyotropic liquid crystals solvent molecules fill the space around the compounds to provide additional degree of freedom of concentration and fluidity that enable formation of a variety of different phases. In the smectic A phase, the molecules are oriented along the layer normal, also called director axis, which also represents the axis of motional averaging (Fig.2-3). Phospholipids in the bilayer can be represented as rigid rods with their most probable orientation parallel to the director axis. However, if followed for a long enough time the molecules will constantly execute angular excursions around this preferred orientation. The description of liquid crystals involves an analysis of order: positional order (whether molecules are arranged an ordered lattice) and orientational order (whether molecules are mostly pointing in the same direction), and moreover order can be either short-range or long-range. Orientational order parameter, which measures angular excursions and fluctuations around the director axis, is usually quantitatively defined based on the average of the second Legendre polynomial:

$$S = 1/2 \left\langle 3 \cos^2 \delta - 1 \right\rangle$$  (2-5)

where S is order parameter ranging from S=1 for perfectly aligned sample and S=0 for completely completely random and isotropic sample; $\delta$ is instantaneous angle between the bilayer director axis and local director axis, which usually is C-2H bond vector in a $^2$H NMR of lipid acyl chains.

31P Chemical Shift Anisotropy

31P NMR is mainly applied to membrane studies to detect lipid phase changes and the degree of organization of lipid headgroups. Various phases, shown in Fig. 2-4, can be adopted by hydrated phospholipids depending on the nature of the lipids, level of
hydration, pH, temperature, presence of ions, etc. The technique is based on the large chemical shift anisotropy exhibited by phosphorus in phosphates, the relatively high sensitivity of this nucleus, the fact that it does not require any synthetic labeling as $^{31}$P is a 100% abundant nucleus, and that there is generally only one phosphate group present per lipid molecule.$^{(86, 87)}$ $^{31}$P is a spin 1/2 nucleus, which means it is dipolar with only one transition between two energy levels. Dipolar nuclei are spherical with a uniform charge distribution over their surface, disturbing the external magnetic field independent of direction. This is in contrast to what is seen for quadrupolar, spin 1 nuclei (see below), and results in a much sharper NMR signal.$^{(88)}$ The two anisotropic interactions primarily affecting 1D $^{31}$P solid state NMR spectra and contributing to the line broadening are chemical shift anisotropy (CSA) and heteronuclear dipolar couplings to protons. The latter is removed with high power proton decoupling and thus the CSA is the dominant interaction typically measured.

When a sample is placed in a strong magnetic field, the nuclei of the same NMR-active isotope in different chemical environments experience different magnetic fields due to the shielding effect of the surrounding electrons and therefore resonate at different characteristic frequencies, which are determined as a result of a reduction of the externally applied magnetic field on the nucleus ($B_0$) by the shielding electrons. This reduction is known as the chemical shift $\sigma$ and it is measured in units of parts per million (ppm) relative to the external magnetic field. In general form total Hamiltonian $H$ of interaction between $^{31}$P nucleus and the surrounding magnetic field can be represented by Zeeman interaction of nucleus with $B_0$ applied magnetic field and its chemical shift shielding by the bond electrons from the applied external field: $^{(89, 90)}$
\[ H = -\gamma h(1 - \sigma)B_0 \]  

(2-6)

where \( \sigma \) (chemical shielding tensor) shows the extent of the nucleus screening from the effect of applied field \( B_0 \) by the bonding electrons in each direction in three dimensional space. Consequently, the pattern of bonding electrons around the \( ^{31}\text{P} \) nucleus in the tetrahedral phosphate group greatly influences its chemical shift. This makes \( ^{31}\text{P} \) chemical shift to be highly dependent on relative orientation of the phosphate group with respect to the applied magnetic field \( B_0 \) demonstrating large chemical shift anisotropy (Fig. 2-5). Due to rapid molecular tumbling in liquids, the electronic environments of the \( ^{31}\text{P} \) nuclei are spatially averaged during signal collection and chemical shift tensor \( \sigma \) describing all \( ^{31}\text{P} \) nuclei becomes isotropic (i.e. the same in every direction) and collapses to a number \( \sigma \), thus giving single chemical shift for each type of chemical bonding environment irrespective to molecular orientation. However, in solids in the absence of such spatial averaging lead to the chemical shift remaining anisotropy with its magnitude dependent on both molecular orientation and molecular environment. The anisotropy of the \( ^{31}\text{P} \) chemical shift in lipids is defined by a 3x3 matrix (tensor) \( \sigma \) and leads to the dependence of the resulting \( ^{31}\text{P} \) resonant frequency on the molecular orientation as well giving rise to the characteristic static \( ^{31}\text{P} \) solid state NMR lineshape (Fig. 2-5). In the laboratory reference frame this tensor represents a regular 3x3 matrix with 9 components, but when represented in the nuclear reference frame, also called principle axis system, this tensor becomes diagonal. The diagonal components, labeled \( \sigma_{11}, \sigma_{22}, \sigma_{33} \), which represent eigenvalues of this axis system, are also called the tensor principal values or principal components. The principal axis system for chemical shielding tensor \( \sigma \) of \( ^{31}\text{P} \) in phospholipid phosphate moiety is shown on figures 2-5 and
2-6 with its components experimentally measured to be $\sigma_{11}=80\text{ppm}$, $\sigma_{22}=25\text{ ppm}$, $\sigma_{33}=-110\text{ppm}$. If the principal components and principal axes of the chemical shift tensor $\sigma$ are known, then the chemical shift $\sigma$ can be calculated for any orientation:

$$\sigma = \sigma_{11} \cos^2 \theta_1 + \sigma_{22} \cos^2 \theta_2 + \sigma_{33} \cos^2 \theta_3$$  \hspace{1cm} (2-7)

where $\theta_i$ is the angle between the i-th principal axis of chemical shielding tensor and the direction of applied magnetic field $B_0$.$^{(89,90)}$ The static powder pattern, which is the sum of spectra for all possible orientations, can be calculated using eqn. 2-7. For microcrystalline sample, where all angles between the applied magnetic field and principal axes are populated, the powder pattern demonstrates three different $\sigma_{11}$, $\sigma_{22}$, $\sigma_{33}$ components (Fig.2-6, A). Also vice versa the principal components of chemical shielding tensor of $^{31}\text{P}$ can be estimated directly from the $^{31}\text{P}$ spectrum as shown on the same figure with $\sigma_{11}$ and $\sigma_{33}$ defining its edges and $\sigma_{22}$ defining the central peak. Rapid anisotropic motion can average some of the components of the chemical shift tensor as shown on Fig. 2-6 (B,C) for lamellar and hexagonal phases correspondingly. In membrane lamellar phase (Fig. 2-6, B) lipids rapidly rotate around the normal to the lipid bilayer averaging $\sigma_{22}$, $\sigma_{33}$ components that are perpendicular to the axis of rotation, whereas $\sigma_{11}$ component does not change as it is parallel to the axis of motional averaging. The resulting tensor becomes axially symmetric:

$$\sigma_{\parallel} = \sigma_{11}$$

$$\sigma_{\perp} = (\sigma_{22} + \sigma_{33})/2$$  \hspace{1cm} (2-8)

The values of the perpendicular ($\sigma_{\perp}$) and parallel ($\sigma_{\parallel}$) tensor components correspond to the extremes in frequencies of line shape due to the bilayer normal of the lipid membranes being oriented either parallel or perpendicular relative to the external
magnetic field. This results in a $^{31}$P powder lineshape with a high-field peak and a low-field shoulder, which is a characteristic lineshape observed for fluid phospholipid uni- and multilamellar vesicles in a static solid-state NMR experiment. Its breadth can be described by the CSA ($\Delta \sigma$), the difference between the $\sigma_\perp$ and $\sigma_\parallel$ tensor elements.$^{(91)}$

The lineshape is affected by both motions and partial orientation of the bilayers in the magnetic field. The more motions exist the narrower the lineshape and the smaller is the anisotropy $\Delta \sigma$. Each part of this powder pattern is also related to a certain orientation of the bilayer in a spherical liposome to the external magnetic field. In spherical lipid vesicles there are more perpendicularly oriented lipids relative to the magnetic field axis ($\theta=90^\circ$) than parallel oriented lipids ($\theta=0^\circ$) since the probability of the particular angular orientation of the lipid over the sphere scales as $\sin \theta$. The probability of orientation with particular angle $\theta$ defines the spectral intensity at a particular frequency, that corresponds to this value of angle $\theta$.$^{(92)}$ At high magnetic fields, the plane of the bilayer tends to favor a perpendicular orientation to the magnetic field leading to elongated ellipsoidal liposomes, due to the anisotropy of the magnetic susceptibility of phospholipid molecules being negative.$^{(91)}$

If $\sigma_{11}$ axis of the phosphate moiety is also involved into rapid rotation (Fig. 2-6, C), then there is further averaging of $\sigma_\parallel$ and $\sigma_\perp$ components into $\sigma'_\parallel$ and $\sigma'_\perp$ effective components and the effective tensor still has axial symmetry, but the total chemical shift anisotropy $\Delta \sigma = \sigma'_\parallel - \sigma'_\perp$ is twice smaller compared to the bilayer case. The more motion of $\sigma_{11}$ axis is allowed the more it is reduced with the anisotropy collapsing to
single isotropic peak with $\sigma_{iso} = (\sigma_{11} + \sigma_{22} + \sigma_{33})/3$, if $\sigma_{11}$ axis samples all possible angles fast enough.\(^{(93)}\)

$^{31}$P ssNMR spectroscopy of phospholipids is often used to gather information about lipid polymorphisms as the spectral lineshapes are reflective of the different lipid phases that phospholipids can adopt: gel and liquid-crystalline lamellar (bilayer) phases, the inverted hexagonal phase, and isotropic phases such as micellar and cubic phases (Figure 2-7).\(^{(87)}\) If multiple phases are present, the spectrum will be a superposition of lineshapes. For lamellar phases, the liquid-crystalline phase with more lipid motion gives a much narrower spectrum compared to the gel phase. $^{31}$P spectra of hexagonal phases exhibit twice narrower lineshape with reversed asymmetry compared to lamellar phases due to further averaging of $\sigma_{11}$ axis of the phosphate moiety in comparison to the bilayer phase as described in the paragraph above, and shown on Fig. 2-6, C. The cylinders in a hexagonal phase (H$_{II}$) have a very small radius (Fig. 2-3) and the lateral diffusion about the cylinder axis causes rotation of $\sigma_{11}$ axis of the phosphate moiety around the cylinder axis. In micelles and fluid isotropic phases with rapid reorientation in three dimensions anisotropic interactions are completely averaged out giving isotropic lineshapes with single resonances like those seen in solution NMR spectra.

$^{2}$H Quadrupolar Coupling

An important feature of $^2$H NMR spectroscopy is that it enables both membrane lipids and membrane proteins to be studied by substitution of $^2$H for $^1$H. Substitution allows labeling specific individual C–$^2$H bonds in lipids with $^2$H and thus obtaining site-specific atomically resolved information used for resonance assignment in fully perdeuterated acyl chains in phospholipids.\(^{(82)}\) Specific lipids can also be isotopically
enriched either on one or both acyl chains to study the behavior of a particular lipid and a particular acyl chain in a complex lipid mixture. Static $^2$H solid state NMR probes motion in general and allows to detect changes occurring at the molecular level in a biologically relevant environment such as lipid membranes when specific variables are introduced. For example, lipid melting and other phase changes can be studied as lipid $^2$H lineshapes change with changing temperature, introduction of proteins, etc.\(^{(87)}\)

Interaction with proteins can change the deuterated lipid’s dynamics and assembly also triggering visible spectral changes. Relative protein insertion depth into lipid bilayer is also estimated using profile of lipid dynamic distribution along the acyl chain.

Deuterium has a $I=1$ spin nucleus and as such has a quadrupolar moment and three Zeeman energy levels and two transitions (Fig. 2-8). However, since the three energy levels are equally spaced, the two transitions between them are degenerate and give rise to a single peak with frequency $v_0$ in isotropic solutions. In anisotropic systems this degeneracy is removed due to the coupling of the quadrupole moment of the $^2$H nucleus with the electric field gradient (EFG) created by the surrounding bonding electrons of the C–$^2$H bond. Quadrupolar coupling described by the Hamiltonian $H_Q$ causes the singlet splitting into doublet with the frequency separation between the two resonances defined as quadrupolar splitting $\Delta v_Q$ (Fig. 2-8). Any nucleus with spin $I=1$ has a not spherical shape and thus an uneven charge distribution representing an electric quadrupole. Electric quadrupole interacts with an EFG analogously to the interaction of an electric dipole with an electric field. Even though magnetic couplings, such as dipolar couplings between the $^2$H and $^1$H nuclei, $^2$H chemical shifts, are also
present in $^2$H NMR their contribution is generally neglected because they are much smaller than the dominant electric quadrupolar interaction.\(^{(82)}\)

For symmetric EFG tensor, as in the case of C-$^2$H bond, the quadrupole splitting is defined in general form as:

$$
\Delta \nu_q(\theta) = \frac{3}{2} (e^2 q Q / h) \cdot \frac{3 \cos^2 \theta - 1}{2}
$$

(2-9)

where \((e^2 q Q / h)\) is the quadrupole coupling constant and \(\theta\) is the angle between the principal axis of EFG tensor, usually the direction of C-$^2$H bond, and applied external magnetic field \(H_0\). For saturated C-D bonds, the quadrupolar coupling, is 167 kHz in the static limit.\(^{(94, 95)}\) The static quadrupolar splitting constant (167 kHz) was determined by measuring the splitting values for C-D bonds in several deuterated alkane containing compounds such as ethane and acetonitrile in frozen solids.\(^{(95)}\) In equation 2-9 angle \(\theta\) can be replaced by the combination of three angles \(\alpha, \gamma, \theta'\) which when combined also relate the principal axis of the EFG tensor to the applied magnetic field \(H_0\) direction (Fig. 2-9):\(^{(96)}\)

$$
\Delta \nu_q(\alpha, \gamma, \theta') = \frac{3}{2} (e^2 q Q / h) \cdot \left( \frac{3 \cos^2 \alpha - 1}{2} \right) \left( \frac{3 \cos^2 \gamma - 1}{2} \right) \left( \frac{3 \cos^2 \theta' - 1}{2} \right)
$$

(2-10)

where a bar denotes a time average, \(\gamma\) represents the angle between the C-$^2$H bond vector and the instantaneous segmental chain orientation, \(\alpha\) represents the angle between the instantaneous segmental chain orientation and the director of the motion, taken to be the normal to the bilayer surface, and \(\theta'\) is the angle between the director of the motion and the magnetic field direction \(H_0\). The angles \(\alpha, \gamma\) and \(\theta'\) are shown in Fig. 2-9. There are two different parts that can be separated from each other in the eqn. 2-
the first two time averaged braces represent dynamic averaging and are denoted by the time averaged order parameter \((S_{C\cdot D})\); the last term defines orientational dependence through angle \(\theta'\). The two parts of the equation are then separated and treated differently:\(^{(82)}\)

\[
\Delta \nu_q = \frac{3}{4} \left( e^2 q Q / h \right) S_{CD} \left| 3 \cos^2 \theta' - 1 \right|
\]

(2-11)

where \( S_{CD} = \left( \frac{3 \cos^2 \alpha - 1}{2} \right) \left( \frac{3 \cos^2 \gamma - 1}{2} \right) = \left( \frac{3 \cos^2 \alpha' - 1}{2} \right) \)

(2-12)

\(\alpha'\) is the combination of \(\alpha\) and \(\gamma\) angles represents the time varying angle between the major axis of EFG (C\(^2\)-H bond vector) and the director of the motion (normal to the bilayer surface) and is also shown on Fig. 2-9.

By definition the order parameter \(S_{C\cdot D}\) measures two types of dynamics: fluctuation of the C\(^2\)-H bond about the instantaneous segmental chain orientation (conformational order) and fluctuation of the long molecular axis around the director of the motion (molecular order). Orientational term \((3\cos^2\theta'-1)\) in the eqn 2-10 is responsible for formation of very broad powder lineshape for liposomes (Pake powder pattern), because it defines specific resonance frequency of a deuterium doublet as function of \(\theta'\) orientation of the bilayer normal with respect to magnetic field.\(^{(96)}\) Since in phospholipid vesicle many different orientations are possible ranging from perpendicular (\(\theta'=90^\circ\)) to parallel (\(\theta'=0^\circ\)) with some orientations more probable than the other, then the signal from many lipid vesicles represents a continuum of deuterium doublets covering all values of angle \(\theta'\) and corresponding different frequencies that depend on the orientation.\(^{(97)}\) The shape of the Pake pattern is due to the fact that there are only two possible orientations along the external magnetic field (parallel and antiparallel), but
many orientations perpendicular to the $B_0$-field (Fig. 2-10). Consequently, the frequency for the 90° orientation of the bilayer normal is more intense than the 0° shoulder. This follows simply from the spherical shape of lipid vesicle and thus also applies to the shape of $^{31}$P CSA spectra of lipid vesicles discussed in the previous section.\(^{(98)}\)

Separation between two most intense peaks $\Delta v_q(90^\circ)$ with $\theta' = 90^\circ$ can be easily measured experimentally and used to calculate order parameter $S_{CD}$ from eqn. 2-11, assuming $\theta'$ is 90° and the static quadrupolar splitting constant $(e^2 qQ / \hbar)$ is 167 kHz:

$$S_{CD} = \frac{4}{3} \Delta v_q(90^\circ) / (e^2 qQ / \hbar)$$  \(2-13\)

Consequently, quadrupole splitting between the most intense peaks is effectively reduced by the factor of $S_{CD}$, which is equal to 1 for perfectly ordered systems and approaches 0 for completely disordered systems, so acyl chain order and order parameter $S_{CD}$ decrease as the quadrupolar splitting decreases, and vice versa.\(^{(98)}\) The more motion at a specific acyl carbon the narrower the lineshape becomes and the two peaks for a particular bilayer orientation separated by the quadrupolar splitting will be closer together (Figure 2-11). The powder spectrum of the whole lipid molecule, which is perdeuterated at every acyl chain position, is complex as it represents the superposition of individual powder spectra for every pair of deuterons at a particular carbon along the deuterated acyl chain(s) with a specific quadrupolar splitting assigned to each pair. Once $\Delta v_q(90^\circ)$ is determined for each particular deuterated methylene group, the $S_{CD}$ values can be calculated and plotted against the carbon number of the DPPC acyl chain from 2 to 16 to show an increase or decrease in order in comparing various lipid samples (Fig. 2-11). The deuterons at the most distal carbon position
always have more motion and thus demonstrate the lowest order parameter in the
deuterated lipid acyl chain due to the fast rotation of the methyl group. The signal of the
terminal CD₃ group deuterons is also 1.5 times bigger than for the other (methylene)
positions due to the one additional deuteron. The multiple splittings of perdeuterated
lipid acyl chains and associated line broadening complicate lipid $^2$H spectra because
peaks from the different labeled positions have different order parameters and overlap
with each other. Assignment of specific frequencies to each CD₂ or CD₃ group in the $^2$H
lineshape is very difficult to do with any degree of accuracy without deconvoluting the
whole spectrum, which is known as dePaking and is described in the section below. The
dePaking transforms complicated, broad lineshapes to individual frequencies and allows
order parameters to be calculated using more accurate splitting values. Qualitative
information like phase or structure information can be obtained from non-dePaked
spectra, but quantitative analyses, such as order parameter profile and insertion depth
estimate, require more precise frequency assignments that could be obtained only from
deconvoluting (dePaking) the broad powder lineshapes.

**DePaking**

A broad powdertype NMR spectrum of the random multilamellar dispersion of
perdeuterated phospholipid consists of overlapping Pake powder patterns of deuterons
at individual positions in lipid (Figure 2-11 A,B). In the process of dePaking these
individual Pake patterns are extracted from the original spectrum and deconvoluted to
leave only the peaks that correspond to the most abundant orientation. A dePaked
spectrum now consists of the more highly resolved individual deuteron subspectra
making individual frequencies more easily assigned (Figure 2-11 C). In other words, this
procedure calculates sort of “oriented” spectrum from an unoriented sample. After
dePaking resonances are still a bit broad, but overall the peaks are sharper, better resolved, and more easily assigned. Samples with membrane bilayers (i.e. MLVs or LUVs) demonstrate little change in the average orientation of the molecules relative to the bilayer normal (solid state regime). Consequently, molecular motions do not completely average tensors of the anisotropic interactions (CSA, dipole-dipole, and quadrupolar interactions mentioned above), which retain a spatial dependent component in the form of a second order Legendre polynomial, $3\cos^2\theta'-1$, where $\theta'$ is the angle between the external static magnetic field and the bilayer normal. The spatial component varies between 1 and $-1/2$ (i.e. $\theta' = 0^\circ$ to $90^\circ$). When $\theta'$ is set to the magic angle, $54.7^\circ$, $\cos^2\theta'$ is equal to $1/3$ and the spatial dependence component of these anisotropic NMR interactions becomes equal to zero (Fig. 2-10). Thus anisotropy can be almost completely removed from the spectrum by spinning at the magic angle with respect to the external magnetic field, which is beneficial for the resonance assignment as the anisotropy of the molecular interactions broadens the resonances in unoriented samples and decreases resolution leading to spectra which are broad superpositions from all the contributions of the possible orientations of the lipid molecules in the sample with respect to the external magnetic field. Another way to remove anisotropy is to orient lipid bilayers, but for many samples this is not feasible.

However, in some cases retaining spatial dependence component of the anisotropic NMR interactions is crucial for analysis because it provides information about molecular orientation in the system; this is important for distinguishing between lipid polymorphisms in our samples. However, in order to keep resolution high the orientational distribution for the NMR interactions and the anisotropies that define their
strengths need to be separated. The original broad powder spectrum is described by the following two equations: \(^{(101)}\)

\[
S(\omega) = \int g(x) \left[ \rho(\theta) \frac{\partial \theta}{\partial \omega} \right] dx, \quad \theta = \theta(x, \omega)
\]

or

\[
S(\omega) = \int \rho(\theta) \left[ \rho(x) \frac{\partial x}{\partial \omega} \right] d\theta, \quad x = x(\theta, \omega)
\] (2-14)

where \(g(x)\) is the anisotropy distribution function and \(\rho(\theta)\) is the orientation distribution function. In equations 2-14 \(g(x)\) is a lineshape function for each anisotropy, such as for a single \(^{31}\)P CSA in a pure phospholipid sample; \(\rho(\theta)\), is a superposition of spectra from the individual oriented spectra of a powder pattern and there is one orientation distribution function for each orientation, \(\theta\).

DePaking can be performed in different ways: either extract \(g(x)\) when \(\rho(\theta)\) is known or oppositely calculate \(\rho(\theta)\) from the measured data when \(g(x)\) is known.

For example, in a perfectly spherical lipid vesicle bilayers adopt completely random orientations with respect to the magnetic field, but the function \(\rho(\theta)\) describing their orientation distribution around the sphere is known and it is proportional to \(\sin \theta\), where \(\theta\) is the angle between the bilayer normal and the magnetic field. Since \(\rho(\theta)\) is known for the sphere, the anisotropy distribution function \(g(x)\) can be extracted from the spectrum of lipid vesicle using a standard inversion (dePaking) procedure. \(^{(102)}\) For samples with ellipsoidal deformation of the MLVs, in which the lipid bilayers align to some degree in the magnetic field and are called partially aligned, the orientational or probability distribution function becomes: \(^{(101)}\)
\[ \rho_z(\theta) \propto \sin \theta [1 - (1 - k_e)\cos^2 \theta]^{-2} \]  

(2-15)

where \( k_e \) is the square of the ratio of the long to short axes of the ellipsoids. If \( k_e \) is equal to 1 the lipid vesicle is a sphere. If this value is >1 the shape is more ellipsoidal with its long axis along the external magnetic field. Deconvolution of spectra of MLVs with ellipsoidal deformation is obtained with an iterative procedure which simultaneously determines \( k_e \) and dePakes the spectrum. Since our lipid samples showed some degree of alignment in the magnetic field, this iterative procedure was utilized and dePaking of NMR data was accomplished with previously published algorithms which simultaneously dePake and determine macroscopic ordering in partially aligned lipid spectra using Tikhonov regularization.\(^{101, 102}\) Prior to dePaking \(^{31}\)P NMR spectra were referenced to phosphate buffer, whereas dePaked spectra were quantitated by fitting the two peaks with Lorentzian line shapes. Assignments of \(^2\)H resonances were made based on assignments of individual deuterons in lipid vesicles reported by Petrache, \textit{et al.}\(^{103}\)

**Solution State NMR Method of Protein Structure Determination**

Analysis of the nuclear magnetic resonance spectra of the particular protein requires that most of the NMR resonances observed for this protein in the given conditions are assigned to the specific atoms in the protein sequence. Resonance assignment is said to be completed for the protein if it is determined which chemical shift corresponds to which atom. It can be achieved for proteins using two main types of NMR experiments: one where magnetization is transferred through the chemical bonds, and the other where the transfer is through space, irrespective of the bonding structure. The first type is used to assign the different chemical shifts to specific nuclei, and the second type is mainly used to generate the distance restraints used in the
following structure calculation. In these experiments magnetization is detected for the specific isotope, usually proton, by irradiation of the sample placed into the static magnetic field with pulses of electromagnetic energy that have specific frequency in the radio diapason equal to the frequency of Larmor precession of this isotope. Then the magnetization is transferred between different nuclei using delays determined by intrinsic internuclear couplings. The whole process is controlled by the pulse sequence of the experiment and the change in the pulse sequence allows one to investigate and select specific types of connections between nuclei. The process of assignment of each amino acid in the sequence is also called “sequential backbone walking”. The exact procedure depends on whether the protein is isotopically labelled or not (see below).

Protein Resonance Assignment

The conventional assignment procedure for unlabelled protein involves a set of two dimensional homonuclear experiments: correlation spectroscopy (COSY) or total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). Both COSY and TOCSY experiments transfer magnetization through the chemical bonds between adjacent protons and these transfers are represented by off-diagonal cross peaks between adjacent protons, whereas protons themselves appear on the diagonal (Fig. 2-12, A). This interaction or spin-spin coupling of two nuclear spins $I_1$ and $I_2$ through the chemical bonds connecting them is called J-coupling and can be expressed in the form $I_1 I_2$. The mechanism proposed in 1951 by N.F. Ramsey and E. M. Purcell states that J-coupling can be represented as the sum of the magnetic interaction between each nucleus and the electron spin of its own atom together with the exchange coupling of the electron spins with each other. In COSY magnetization can only be transferred between protons on adjacent atoms, whereas in TOCSY it is
transferred among all protons within a given spin system, even between distant protons as long as there are J-couplings between every intervening proton. All possible COSY magnetization transfers for valine are shown as red peaks on Fig. 2-12, A: HN→Hα, Hα→HN, Hα→Hβ, Hβ→Hα, Hβ→Hγ, Hγ→Hβ. TOCSY connects all protons within valine spin system HN↔Hα↔Hβ↔Hγ (red and green peaks) and thus shows links from the backbone NH group to all side-chain hydrogens of that residue giving a so called “amide trace” characteristic of the valine sidechain structure. Thus a so called spin system shown on Fig 2-12, A is characteristic for the valine structure. Different types of protons have characteristic chemical shifts, which allows amino acids to have distinct patterns in the “amide traces” indicative of their structures (Fig. 2-12, B). Consequently, using a TOCSY spectrum the amino acid type can be identified or narrowed down significantly.(106)

To connect the different spin systems in a sequential order, the NOESY experiment has to be used. In NOESY magnetization is transferred through space, consequently, this experiment will show for each NH group all ¹H resonances which are within about 5-6Å of the NH hydrogen regardless of whether they are in the same spin system or not (Fig. 2-13, A). This transfer of magnetization or spin polarization known as Nuclear Overhauser Effect (NOE) happens directly from one population of nuclear spins to another through space, rather than through electrons spins of chemical bonds connecting them. The nuclear spin polarization is transferred from one nuclear spin population to another via cross-relaxation mechanism. Thus, atoms that are in close proximity to each other can give a NOE, whereas spin coupling is observed only when the atoms are connected by 2–3 chemical bonds. The neighboring residues are
inherently close in space, so the sequential assignment is done on the assumption that the two neighboring NH groups are always visible. Thus two NH groups can be linked because they each have an NOE to the other NH group. The additional presence of Hα(i)-NH(i+1) NOE resonances and NOEs of side-chain protons from i\(^{th}\) residue to NH groups of i+1 residue can further confirm that the residues i and i+1 are sequentially connected (Fig. 2-13, B). Sequential assignment is completed when the spin systems of i\(^{th}\) and i+1 amino acid residues “mapped out” by TOCSY are connected in sequential order through the strong NOESY peaks showing that HN, H\(\alpha\) and H\(\beta\) of i\(^{th}\) amino acid are very close in space to HN\(_{i+1}\) atom of the next amino acid. Helical sections are generally easier to assign, as NOEs from NH(i) are visible not only to NH(i±1), but also to NH(i±2) and sometimes NH(i±3). Having a rough idea of the secondary structure and topology of the protein can thus significantly aid backbone assignment using 2D proton spectra only.\(^{(106)}\) These non sequential NOE crosspeaks show how the protein amino acid sequence is folded in space and are used to derive distance restraints for protein structure calculation after sequential assignment is finished. For example, on Fig. 2-13, NOEs between H\(\alpha\), H\(\delta\) of L27 and the amide proton of G32 are not sequential.

The use of homonuclear 2D \(^{1}\text{H}-^{1}\text{H} \) TOCSY and NOESY for sequence assignment is restricted to very small proteins or peptides because different protons have very similar chemical shifts and are located in the same area of the spectrum thus causing peak overlap. The larger the protein, the more significant is the problem of peak overlap.

For small proteins, less than 10 kDa, the problem of peak overlap can be solved by \(^{15}\text{N}\)-labelling the protein. If magnetization is sent from the amide protons to the other
protons through the $^{15}\text{N}$ nucleus directly attached to the amide protons, then the one- 
bond scalar J-couplings between the amide protons and the attached nitrogens $^1J(^1\text{H-}
^{15}\text{N})$ can be employed to separate 2D $^1\text{H}^1\text{H}$ TOCSY and NOESY traces of different 
amino acids in the 3rd dimension according to the $^{15}\text{N}$ chemical shift of amide nitrogen of 
each amino acid in the process of acquiring 3D$^{15}\text{N}$-NOESY-HSQC and the $^{15}\text{N}$-TOCSY-
HSQC spectra. $^1J(^1\text{H}^{15}\text{N})$ couplings are $\sim$92Hz and are much larger than homonuclear
$^1\text{H}^1\text{H}$ J-couplings of a few Hz, consequently separation between the TOCSY and 
NOESY planes along the $^{15}\text{N}$ dimension is much greater than any distance between the
$^1\text{H}^1\text{H}$ peaks in the planes. Due to the reduction of overlap in 2D $^1\text{H}^1\text{H}$ TOCSY and
NOESY planes more peaks can be detected in each spectrum, which allows one to do 
better backbone assignment and generate more distance restraints for a more precise  
structure calculation. Prior to sequential assignment of the $^{15}\text{N}$-labelled protein it is  
convenient to measure a relatively quick heteronuclear 2D single quantum correlation
(HSQC) spectrum also called “fingerprint” spectrum. In $^{15}\text{N}$-HSQC each $^1\text{H}$ bound to a 
heteronucleus, in this case $^{15}\text{N}$, gives one peak, consequently, there is one signal for 
each amino acid residue. The exceptions to this rule are the proline, which has no 
amide hydrogen, and certain other residues (arginine, lysine, glutamine, asparagine and 
tryptophan) with nitrogen-containing sidechains that give rise to additional signals. $^{15}\text{N}$-
HSQC spectrum allows researchers to determine whether the expected number of
peaks is present and thus to identify possible problems due to multiple conformations or
sample heterogeneity. This simple evaluation predicts the feasibility of doing
subsequent longer, more laborious and expensive experiments. It is not possible to
assign peaks to specific atoms from the heteronuclear single quantum correlation alone,

51
consequently, 3D$^{15}$N-NOESY-HSQC and the $^{15}$N-TOCSY-HSQC spectra have to be acquired for the sequential assignment and structure determination.

Although an assignment process using 3D $^{15}$N experiments (TOCSY-HSQC and NOESY-HSQC) is an improvement over 2D methods for larger peptides/small proteins, the method also has drawbacks, which are obvious for the proteins larger than 10 kDa.$^{(107)}$

1. Still impossible to unambiguously assign all the side chain proton resonances.

2. With increasing protein size the $^{15}$N edited TOCSY-HSQC correlations will become weaker due to the longer correlation time and you may not get a complete transfer along the side chain, thus making amino acid type determination impossible.

3. Assignment relies on HN-X NOEs, which are conformation dependent and degeneracy in the $^1$H dimension becomes a problem, especially for $\alpha$-helical proteins and larger proteins.

The excellent resolution of 3D experiments coupled with efficiency of direct scalar coupling transfers via $^{13}$C/$^{15}$N made the “triple resonance” ($^1$H$^{13}$C$^{15}$N) assignment method the standard way to study proteins larger than 10-15 kD by NMR. The large single bond heteronuclear couplings (e.g. $^{13}$C’O –$^{15}$N, $^1$J$=15$Hz) give efficient transfer through the backbone when compared to small homonuclear $^1$H couplings of several Hz. Also C$\alpha$ and C’ chemical shifts are quite different (C$\alpha$ ≈ 40-65 ppm; C’ ≈ 165-180ppm). This allows the treatment of C$\alpha$ and C’ carbon atoms as separate nuclei, consequently, the particular magnetization transfer along the backbone (from N to CO or to C$\alpha$) can be selected and manipulated independently. This approach removed the
use of NOEs for making sequential assignments, consequently, it has the advantage of being less ambiguous and more reliable.\(^{107}\)

**Restraint Generation for Protein Structure Calculation**

After the protein sequential assignment is achieved, i.e. each chemical shift is assigned to each proton in the protein sequence, each crosspeak in a \(^1\)H-\(^1\)H NOESY spectrum of the protein starts to signify spatial proximity between the two specific protons with the intensity of the crosspeak decaying as the distance between the protons increases. The distance between these specific protons can be determined from the intensity of the crosspeak between them, because the intensity is proportional to the distance to the minus 6th power. Thus each peak in NOESY can be converted into a maximum distance between the nuclei, usually between 1.8 and 6 Å, to generate a list of experimentally determined distance restraints. Distance restraints are the most widely used type of restraints in protein structure calculation. The intensity-distance relationship is not exact, so usually a distance range is used in the distance restraint for structure calculation. For the protein structure calculation correct sequential assignment is of great importance because the NOESY crosspeaks have to be assigned to the correct nuclei based on the chemical shifts in order to generate meaningful distance restraints.\(^{104, 105}\)

In addition to distance restraints, one can generate angular restraints for the protein psi and phi backbone torsion angles using two different experimental techniques. Both approaches utilize the fact that different values of the psi and phi backbone torsion angles dictate certain geometry around the alpha carbon, which in turn affects the coupling constants and chemical shifts of the nearby atoms. Consequently, if the coupling constants \(^3\)J(HA-HN) or the chemical shifts of HN, \(^{15}\)N,
HA, CA, CB are experimentally measured, then the psi and phi backbone torsion angles can be predicted. The first approach uses the Karplus equation\textsuperscript{(108, 109)} to generate angle restraints from coupling constants between the amide and alpha protons:

\[ ^3J(\text{HA-HN}) = A\cos^2(\theta) + B\cos(\theta) + C \]  

(2-16)

where \( \theta \) is the dihedral angle between the two vicinal protons for which the coupling is measured; \( A, B, \) and \( C \) are empirically derived parameters. The superscript "3" indicates that a \(^1\text{H} \) atom is coupled to another \(^1\text{H} \) atom three bonds away. Such hydrogens bonded to neighbouring atoms are termed vicinal. The magnitude of these couplings is generally the smallest when the torsion angle is close to 90° and the largest at angles of 0 and 180°. In the second approach psi and phi torsion angles are searched for in the database of proteins with known structures against measured chemical shifts using the TALOS+ algorithm.\textsuperscript{(110)}

**Protein Structure Determination by Simulated Annealing Approach**

The experimentally determined restraints are organized into lists and converted to the proper text format to be used as input to the computer programs, such as XPLOR-NIH,\textsuperscript{(111, 112)} for the structure calculation process. In general any structure determination program converts the experimental restraints and the known general protein properties (bond lengths, angles, etc.) into separate energy terms \( E_{\text{exp}}^i(X) \) and \( E_{\text{chem}}^k(X) \) correspondingly, which are functions of atomic coordinates:

\[ E_{\text{total}}(X) = \sum_j E_{\text{exp}}^j(X) + \sum_k E_{\text{chem}}^k(X) \]  

(2-17)

This conversion is performed in such a way that energy increases if the atomic coordinates deviate from the ideal positions where the specific type of restraint is completely satisfied, i.e. an energy “penalty” is generated for violation of the particular
restraint. Certain weighting factors $\omega^j$ and $\omega^k$ are assigned to the different energy terms to describe the importance of each term for the structure determination:

$$E_{\text{hybrid}}(X) = \sum_j \omega^j E^{j}_{\text{exp}}(X) + \sum_k \omega^k E^{k}_{\text{chem}}(X)$$  \hspace{1cm} (2-18)

Then the whole hybrid energy $E_{\text{hybrid}}$ is minimized in the process of searching for the ideal set of atomic coordinates, which would satisfy all of the restraints the best possible way, while a delicate balance is maintained in satisfying different restraints due to fine tuning their weighting factors. Once this ideal set of coordinates is found the energy becomes the lowest possible and the global minimum energy is achieved. The optimization process starts from the unfolded structures (extended protein strands) and results in an ensemble of structures that will converge only if the sufficient experimental data were provided to the program to dictate a certain fold. The calculations have to start from extended strands rather than from entirely random structures, which may already be folded, because once the polypeptide chain has folded incorrectly in a local minimum it is unlikely to converge to the correct global minimum region.\(^{(113)}\)

In the XPLOR program the total target function $F_{\text{tot}}$ (potential energy of the system) for which the global minimum region is searched is made up of the following terms:\(^{(112, 113)}\)

$$F_{\text{tot}} = F_{\text{covalent}} + F_{\text{repel}} + F_{\text{NMR}} + F_{\text{RAMA}}$$  \hspace{1cm} (2-19)

$F_{\text{covalent}}$ is the target function for maintaining correct bond lengths, angles and planes, and is given by:\(^{(113)}\)

$$F_{\text{covalent}} = \sum_{\text{bonds}} k_b (r - r_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{impropers}} k_\omega (\omega - \omega_0)^2$$  \hspace{1cm} (2-20)
where the force constants for the bond $(k_b)$, angle $(k_\theta)$ and improper torsions $(k_\omega)$ are set to uniform high values to ensure nearly perfect stereochemistry throughout the calculations. The improper torsion terms serve to maintain planarity and chirality.

$F_{\text{repel}}$ is the target function used to prevent unduly close nonbonded contacts and is given by: $^{(113)}$

$$F_{\text{repel}} = \begin{cases} 0 & \text{if } r \geq s r_{\text{min}} \\ k_r (s^2 r_{\text{min}}^2 - r^2)^2 & \text{if } r < s r_{\text{min}} \end{cases} \quad (2-21)$$

The values of $r_{\text{min}}$ are the standard values of the van der Waals radii as represented by the Lennard—Jones potential used in the CHARMM empirical energy function; $s$ is a van der Waals radius scale factor, and $k_r$ the van der Waals repulsion force constant.

$F_{\text{NMR}}$ is a complex term whose functional form depends on the difference between the calculated and target value of a particular experimentally measured NMR restraint. It measures how well the NMR observables calculated from a molecular structure match their experimental counterparts. $^{(112)}$ An additional term RAMA used in the structure calculations is a multi-dimensional torsion angle database potential term that is derived from the Protein Data Bank and biases structures towards existing features seen in the database. However, it is readily overridden by experimental restraints in cases of conflict. Ideally, only $F_{\text{NMR}}$ class of potential terms would be necessary in structure determination, but NMR restraints alone do not contain sufficient structural information. $^{(112)}$

In the Monte Carlo algorithm, the temperature of the system is temporarily increased in order to search for the global minimum region of the target function. The simulated annealing method in XPLOR achieves the same effect by using Newton's
equations of motion to increase temporarily the temperature of the system in order to search for the global minimum region of a target function. The simulated annealing procedure involves the simultaneous integration of Newton's equations of motion: \[^{(113)}\]

\[
\frac{\partial^2 X_i}{\partial t^2} = -\frac{1}{m_i} \frac{\partial}{\partial X_i} F_{\text{tot}}(X_1, X_2, \ldots, X_n)
\tag{2-22}
\]

for all \( n \) atoms in the system whose temperature is given by: \[^{(113)}\]

\[
(T_{\text{Temp}})_i = \frac{2}{k_B(3n-6)} \left( \sum_{i=1}^{n} m_i V_i^2 \right)
\tag{2-23}
\]

Integration of the classical equations of motion is performed using a Verlet integration algorithm with initial velocities assigned to a Maxwellian distribution at an appropriate temperature. Unbiased sampling of the conformational space consistent with the restraints is achieved by varying the random number seed used to assign the initial velocities. This ensures that the different trajectories diverge during the early stages of the simulations and only converge later as more and more experimental restraints are satisfied. \[^{(113)}\]

In biomolecular NMR structure determination, many internal coordinates are known or presumed to take agreed-upon values and it is generally not desirable that these known coordinates be altered. For instance, bond lengths and angles are taken from high resolution crystal structures. It is also true that the configuration space of internal coordinates to be searched in a structure determination calculation can be much smaller. For instance, the size of torsion angle space is approximately one third that of Cartesian coordinates for proteins. For these reasons, it is desirable to perform molecular dynamics and minimization in arbitrary internal coordinates, such as bond stretching, bending, and torsion angles, which are specified in the internal variable
module (IVM) in XPLOR. IVM can also be used to perform dynamics in the full Cartesian space. The simulated annealing calculations routinely performed in XPLOR IVM usually use torsion angles as internal coordinate system. The common procedure consists of performing molecular dynamics in torsion angle space starting at a high temperature, and then slowly decreasing the temperature in order to find the global minimum region. In Xplor–NIH annealing protocols, potential parameters are generally ramped while the temperature is decreased, such that the potential energy is initially softer, with lower barriers. As simulated annealing progresses and the temperature is decreased, force constants are increased such that the potential takes its desired final form at the end of the annealing protocol. The example of XPLOR-NIH script for simulated annealing calculation is provided in Appendix A.

**Protein Structure Validation**

It is important to acknowledge the fact that the ensemble of structures obtained in a structure calculation is an experimental model, i.e. a representation of experimental data measured by NMR, where all measured restraints have been satisfied in the generated protein structure. However, such experimental models of protein structures can contain errors, and it is very important to detect these errors in a process known as validation. In general the quality of a model will depend on both the quantity and quality of experimental data used to generate it and the correct interpretation of such data. One of the common structure validation programs is PROCHECK. PROCHECK checks the stereochemical quality of a protein structure, producing a number of plots analyzing its overall and residue-by-residue geometry. Examples of such validation plots for the structure of SP-B<sub>59-80</sub> in methanol and in DPC micelles are provided in chapter 4 (Figs. 4-13 to 4-20).
The precision of a model indicates the variance of the measured data set under the same conditions. Precision of the structure ensemble derived using solution NMR is determined with root mean square deviation (RMSD) factor of different structures in the ensemble from the mean structure. The more precise is the structural model the smaller is RMSD of the structure ensemble. The accuracy of the model indicates the degree to which a measurement approaches its true value. In practice, for an NMR structure ensemble it is evaluated by the degree of agreement between the model and a set of experimental data as there is no "standard molecule" against which to compare models of proteins.

**Methods of Peptide Production**

The experiments reported in chapter 3 utilized wild type SPB_{59-80} peptide that was prepared synthetically and purified using acetonitrile gradient and high performance liquid chromatography (HPLC). Then pure SPB_{59-80} peptide was dissolved in methanol and reconstituted with model lung surfactant lipid mixtures 4:1 DPPC:POPG and 3:1 POPC:POPG. Production of {^{15}}\text{N SP-B}_{59-80} uniformly labeled version of the peptide, required for the experiments in chapter 4, can be achieved only via peptide expression in *E.Coli* using minimal media with {^{15}}\text{NH}_{4}\text{Cl} as the sole nitrogen source.

**Solid Phase Peptide Synthesis**

Solid phase peptide synthesis (SPPS), pioneered by Merrifield,\textsuperscript{(115)} is now the most popular technique for production of short peptides. SPPS allows not only the synthesis of natural peptides that are hard to express, but also the incorporation of unnatural amino acids and isotopically enriched residues, modification of the backbone, and the synthesis of D-proteins, which consist of D-amino acids. In SPPS technique the small, insoluble yet porous resin beads (supports) are treated with functional units
('linkers'). Then the carboxyl group of the first amino acid is covalently coupled to the linker, whereas its N-terminus is deprotected and then coupled to the carboxyl group of the next N-protected amino acid unit. Thus the C-terminus of the growing peptide is attached to the solid-phase and the peptide continuously elongates one residue at a time during the cycle of coupling new amino acid, washing away reactant, deprotecting the end of the growing peptide chain, and washing again. Then the cycle coupling-wash-deprotection-wash is repeated again until the desired peptide sequence is produced. Chemical peptide synthesis starts at the C-terminal end of the peptide and ends at the N-terminus, which is the opposite of protein biosynthesis in ribosomes, which starts at the N-terminal end. The peptide remains covalently attached to the bead until cleaved from it by trifluoroacetic acid. The superiority of this technique lies in the ability to keep the peptide immobilized and intact on the solid phase during filtration, whereas liquid-phase by-products and excess reagent are washed away after each reaction and replaced by the reagents to perform the next organic reaction.\(^{116, 117}\) The N-termini of amino acid monomers are protected by Fluorenylmethyloxy carbonyl (Fmoc) groups before they are added onto a deprotected amino acid chain to avoid the possible adverse reactions during synthesis.\(^{116}\) In addition many amino acids also have reactive side chain functional groups, which can interact with free termini or other side chain groups during peptide synthesis adversely affecting final peptide yield and purity. The specific side chain protecting groups used during the peptide synthesis are usually based on the benzyl (Bzl) or tert-buty (tBu) group.\(^{118}\) The side chain protecting groups are known as permanent or semipermanent protecting groups, because they can
withstand the multiple cycles of chemical treatment during synthesis and are only removed with strong acids after peptide synthesis is completed.

Fmoc protection was used during SPPS of SPB\textsubscript{59-80} peptide for experiments in chapter 3 with Fmoc group protecting the \(\alpha\)-amino group and resin-linkage agents. The steps of Fmoc SPPS can be summarized as follows:\textsuperscript{116, 117}

1. The Fmoc protected amino acid is attached to the resin via a linker.
2. The Fmoc protecting group is removed with piperidine in dimethylformamide (DMF) to deprotect the residue.
3. The carboxyl group of the next Fmoc protected amino acid is activated with carbodiimide and coupled to the amino acid linked to the solid phase support at basic pH. The coupling reaction is usually performed twice to insure optimal yield.
4. The deprotection/coupling cycle is repeated to yield the desired amino acid sequence.
5. The linker/resin support and side-chain protecting groups are cleaved with TFA, yielding a free peptide. Then the peptide is purified with HPLC.

The prominent features of the above SPPS reaction steps are outlined in Figure 2-14.

**Heterologous Peptide Expression in *E.Coli***

Basic science studies as well as clinical trials of synthetic LS formulations require large quantities of the peptides. Peptide isolation from natural sources or their chemical synthesis are generally not cost-effective, whereas a relatively cheap recombinant approach provides not only easy scale-up for large-scale peptide production but also the possibility of production of larger peptides and uniform isotopic labeling, which becomes too expensive in traditional solid-state chemical synthesis even of relatively
short peptides. Escherichia coli is the most widely used host for the peptide expression due to its fast growth rate, low cost and the fact that peptides generally do not need post-translational modification to be active. The fact that the most stable and hence abundantly produced proteins in E. coli are usually bacterial in origin and protease resistant is important for the peptide expression as well. Thus to achieve the best yields peptide coding sequences have to be codon optimized for expression in E. Coli and peptides are often expressed as fusion proteins, which protects them from proteolytic degradation. This was an important breakthrough in the biosynthetic peptide production since attempts to produce small polypeptides of less than 100 amino acids generally failed unless these peptides were part of a fusion protein. Soluble hydrophilic carriers, such as thioredoxin, glutathione-S-transferase (GST) and maltose-binding proteins (MBP), allow quick purification from crude lysate by affinity chromatography using carrier protein as an affinity tag to purify the fusion protein. Highly hydrophobic fusion partners, such as ketosteroid isomerase (KSI), enhance the stability of peptides by driving them into the insoluble inclusion bodies, which not only prevents proteolytic degradation of the fusion complex but also allows efficient separation from the soluble host proteins.

Since many recombinant proteins are nonfunctional when tethered to a carrier protein, several cleavage methods have been developed in order to cleave the desired protein from the carrier protein. Site specific proteases such as thrombin or factor X, preserve the protein native structure, but often do not cleave quantitatively at the anticipated site. Chemical cleavage by CNBr, which cleaves quantitatively and specifically at methionine residues, requires strongly acidic conditions, which are
incompatible with the survival of native protein structure. Consequently, chemical cleavage by CNBr can only be used for the production of small polypeptides of less than 100 amino acid residues since their native structure preservation is not required during purification.\textsuperscript{(120)} Also chemical cleavage is not compatible with some fusion carriers, for example GST, because the non-specific cleavage at the carrier will substantially complicate the subsequent purification.\textsuperscript{(119)}

All these concepts have been applied in the method developed by Kuliopulos and Walsh,\textsuperscript{(120)} which utilizes extremely hydrophobic ketosteroid isomerase (KSI) as the carrier protein of fusion constructs that comprise multiple copies of the target peptides in addition to a His-tag sequence. His-tag is used for the affinity purification of the fusion construct on Ni-column under denaturing conditions, which is the only possible way of purification of a highly hydrophobic protein. The highly hydrophobic KSI is used to drive the fusion protein into protease resistant inclusion bodies. This strategy requires that methionine (Met) residues be included onto the sequence as separators of each individual component of the fusion protein (i.e., KSI, target peptides, and His-tag) in order to allow for excision of target peptides by CNBr polypeptide digestion. Final recovery of the target peptides is accomplished through chromatographic separation of the fragments in the cleavage mixture by HPLC.\textsuperscript{(121)} In the commercially available plasmid pET31b\textsuperscript{(122)} designed by Novagen (Fig. 2-15) the transcription of the fusion complex KSI-Met-[(peptide)-Met]\textsubscript{n}-His\textsubscript{6} is set under regulation by a strong bacteriophage T7 RNA polymerase promotor. T7 RNA polymerase gene is placed under lacUV5 control, and T7 RNA polymerase synthesis is induced by IPTG, which in turn induces fusion complex expression in the host cell. T7 RNA polymerase is so selective and
active that almost all of the cell’s resources are converted to target gene expression.\(^{(123)}\)

The scheme of peptide expression and purification using pET31 expression system presented on Fig. 2-15 is followed by the necessary chromatograms and analyses of peptide purity on Figs. 2-16 to 2-20.

**Peptide Purification by High Performance Liquid Chromatography\(^{(124)}\)**

Solid-phase peptide synthesis and recombinant DNA techniques have allowed the production of large quantities of peptides and proteins which need to be highly purified. A method of choice for the isolation of peptides and proteins from a wide variety of synthetic or biological sources is High Performance Liquid Chromatography (HPLC), because it provides the excellent resolution under a wide range of conditions for very closely related molecules, as well as structurally quite distinct molecules. Reverse phase HPLC (RP-HPLC), where peptides are separated by their hydrophobicity, is the most commonly used mode of peptide separation, although other techniques also find application, for example, ion-exchange and size exclusion chromatography, which are based on molecule surface charge or size respectively. In the case of synthetic peptides, RP-HPLC is generally employed both for the initial analysis during synthesis and the final large scale purification. The isolation of peptides and proteins from a biological cocktail, however, often requires a combination of techniques to produce a homogenous sample and HPLC techniques are introduced at the later stages following initial precipitation, clarification and preliminary separations.

The separation in RP-HPLC technique depends on the hydrophobic binding of the solute molecule from the mobile phase to the immobilized hydrophobic ligands attached to the stationary phase, i.e., the sorbent. The solute mixture is initially applied to the sorbent in the presence of aqueous buffers, and the solutes are eluted by the addition of
organic solvent to the mobile phase. Elution can proceed either by isocratic conditions where the concentration of organic solvent is constant, or by gradient elution whereby the amount of organic solvent is increased over a period of time. The solutes are, therefore, eluted in order of increasing molecular hydrophobicity. RP-HPLC is more commonly employed for the isolation of peptides rather than proteins, because its harsh conditions cause the irreversible denaturation of protein and do not allow recovery of the protein in a biologically active form. In RP-HPLC peptides are usually retained on n-alkylsilica-based sorbent and eluted with gradients of increasing concentrations of organic solvent such as acetonitrile containing an ionic modifier such as trifluoroacetic acid (TFA). Separations can be easily manipulated by changing the gradient slope, the operating temperature, the ionic modifier, or the organic solvent composition.\(^{(124)}\)

Figure 2-1. Far UV CD spectra associated with various types of secondary structure. Solid line, \(\alpha\)-helix; long dashed line, anti-parallel \(\beta\)-sheet; dotted line, type I \(\beta\)-turn; cross dashed line, extended 3\(\alpha\)-helix or poly (Pro) II helix; short dashed line, irregular structure. Figure borrowed with permission from: Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism, *Biochimica Et Biophysica Acta-Proteins and Proteomics* 1751, 119-139.
Figure 2-2. Chemical structures of glycerophospholipids. The polar head groups vary in their size and charge: zwitterionic head groups phosphocholine (PC) and phosphoethanolamine (PE), and the anionic head group phosphoglycerol (PG). The nonpolar acyl chains vary in their length and degree and position of unsaturation.

Figure 2-3. Gel to liquid crystalline phase transition of the bilayer lamellar phase. When melting above the temperature $T_m$ phospholipids obtain more degrees of freedom and more motions in their acyl chains as well as more axial rotation compared to the gel phase.
Figure 2-4. Types of lipid phases. Hexagonal phase is shown in its normal and inverted types; bilayer possesses ordered gel phase, which melts into disordered liquid crystalline phase upon heating. Borrowed and adapted with permission from: Koynova, R., and Caffrey, M. (2002) An index of lipid phase diagrams, *Chemistry and Physics of Lipids* 115, 107-219.
Figure 2-5. Origins of the powder pattern obtained in the presence of chemical shift anisotropy (CSA) interaction. The CSA interaction results in a powder pattern due to the distribution of populations of orientations the lipid bilayers adopt. The CSA is defined by a tensor with elements $\sigma_{11}$, $\sigma_{22}$ and $\sigma_{33}$ in lipid bilayers that are time averaged to $\sigma_{\parallel}$ and $\sigma_{\perp}$ components. This is because the average tensor is oriented with respect to the bilayer normal. With solid state NMR you can obtain orientation information that is lost in solution NMR. Each part of the powder pattern is related to a certain orientation of the bilayer and each orientation leads to a different frequency. With spherical vesicles you have more perpendicular than parallel orientations. This is illustrated by the colored spheres showing chances are higher for an angle of rotation of bilayer normal that is $90^\circ$ relative to the external magnetic field ($B_0$). As the angle decreases, chances of the lipid orienting at that angle in the magnetic field leads to a lower intensity. Figure adapted with permission from: Levitt, M. H. (2008) Spin dynamics: Basics of nuclear magnetic resonance, 2 ed., Wiley.
Figure 2-6. Chemical shift anisotropy of the rigid lipid phosphate group is a measure of the degree of order and characterizes the types of the motions that are experienced by the group in the lipid phase. The shape and the width of the pattern are strictly dependent on the orientation of the axis of motional averaging relative to principal components of the chemical shift tensor $\sigma_{11}$, $\sigma_{22}$, $\sigma_{33}$.$^{(89, 90)}$ A: For microcrystalline sample, where all angles between the applied magnetic field and principal axes are populated, the static powder pattern is the sum of spectra for all possible orientations and has three different $\sigma_{11}$, $\sigma_{22}$, $\sigma_{33}$ components. B: Rapid anisotropic motion around normal to the lipid bilayer does not change $\sigma_{11}$ as it is parallel to the axis of motional averaging, but averages the other two components of the chemical shift tensor $\sigma_{22}$, $\sigma_{33}$ that correspond to the axes perpendicular to the axis of rotation. C: If $\sigma_{11}$ axis of the phosphate moiety is also involved into rapid rotation, then there is further averaging of $\sigma_{\|}$ and $\sigma_{\perp}$ components into $\sigma'_{\|}$ and $\sigma'_{\perp}$ effective components and the effective tensor still has axial symmetry, but the total chemical shift anisotropy $\Delta \sigma = \sigma'_{\|} - \sigma'_{\perp}$ is reduced.

Figure 2-7. Polymorphisms and phosphorus NMR lineshapes. The lineshapes for $^{31}$P NMR spectra correspond to different phases or polymorphisms. The phase is shown on the left and its corresponding spectrum on the right. Several phases can result in an isotropic peak. Additional lipid motion further averages $^{31}$P CSA and makes the lineshape narrower. Borrowed and adapted with permission from: Pfeiffer, H., Weichert, H., Klose, G., and Heremans, K. (2012) Hydration behavior of POPC/C$_{12}$-Bet mixtures investigated by sorption gravimetry, $^{31}$P NMR spectroscopy and X-ray diffraction, Chemistry and Physics of Lipids 165, 244-251.
Figure 2-8. Effect of the quadrupole interaction on Zeeman energy levels of nucleus with spin $I=1$ with axial symmetry. The quadrupolar splitting $\Delta \nu$ corresponds to the difference in the transition frequencies of the single quantum transitions, and is due to the perturbing interaction of the $^2\text{H}$ nuclear quadruple moment with the electric field gradient (EFG) of the C–$^2\text{H}$ bond. Adapted with permission from: Brown, M. F., and Nevzorov, A. A. (1999) $^2\text{H}$-NMR in liquid crystals and membranes, *Colloids and Surfaces a-Physicochemical and Engineering Aspects* 158, 281-298.
Figure 2-9. The angles $\alpha$, $\gamma$, $\theta'$ used to define the motional averaging in lipid bilayers by relating the principal axis of the EFG tensor (usually the direction of C-$^2$H bond) to the applied magnetic field $H_0$ direction. $\gamma$ represents the angle between the C-$^2$H bond vector and the instantaneous segmental chain orientation, $\alpha$ - the angle between the instantaneous segmental chain orientation and the director of the motion, taken to be the normal to the bilayer surface and $\theta'$ the angle between the director of the motion and the magnetic field direction. Adapted with permission from: Dufourc, E. J., Smith, I. C. P., and Jarrell, H. C. (1983) A $^2$H-NMR analysis of dihydrosterculoyl-containing lipids in model membranes - structural effects of a cyclopropane ring, *Chemistry and Physics of Lipids* 33, 153-177.
Figure 2-10. Origin of the powder pattern of $^2$H-NMR spectra observed for deuterium nuclei incorporated into phospholipid tails in membranes in uni- and multilamellar vesicles (LUVs and MLVs). Different angles between magnetic field $H_0$ and the axis of motional averaging (dashed arrow) lead to different quadrupole splittings. Since membranes rotate slowly on the timescale of $^2$H-NMR splittings, all these subspectra will contribute. Figure adapted with permission from: Levitt, M. H. (2008) *Spin dynamics: Basics of nuclear magnetic resonance*, 2 ed., Wiley.
Figure 2-11. Solid-state NMR provides residual quadrupolar couplings (RQCs) that directly correspond to ensemble-averaged molecular structure. There are two spin transitions in $^2$H NMR that lead to two powder patterns that overlap resulting in a characteristic pattern for deuterium nucleus in solid state NMR (A). POPC lipid with multiple deuterated sites at sn-1 position and different motions of $^2$H atoms at different carbon positions demonstrates broad powder-pattern spectrum with several overlapping $^2$H lineshapes (B), where each $^2$H lineshape has different quadrupolar splitting $\Delta \nu_Q$, which is affected by lipid mobility at the site. Numerically deconvoluted (de-Paked) $^2$H NMR spectrum (C) reveals quadrupolar splittings $\Delta \nu_Q^{(i)}$ for individual methylene segments ($i$) of acyl chains because dePaking transforms powder lineshapes into individual frequencies. As the quadrupolar splitting increases the acyl chain order increases as well with the terminal methyl being most mobile or disordered. Order parameter profile (D) shows order parameter calculated from $\Delta \nu_Q^{(i)}$ for $^2$H at each carbon position in sn-1 acyl chain.
Figure 2-12. Chemical shift pattern of different amino acids in 2D $^1$H-$^1$H TOCSY spectrum. A: The schematic of the whole TOCSY spectrum for valine amino acid that contains backbone amide, alpha, beta and gamma protons. B: Schematic representation of the amide TOCSY traces for all amino acids. Plotted using chemical shift values from table 2.3 in Wuthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley & Sons, Inc., USA.
Figure 2-13. Protein sequence assignment via “backbone walking”. A: Alternating TOCSY (plain circles) and NOESY (circles with blue edge) amide traces of amino acids show that HN, Hα and Hβ of amino acid i are sequentially connected through space to HN atom of the next amino acid i+1 via strong NOE correlations. B (next page): “backbone walking” for SP-B\textsubscript{59-80} in methanol with initial assignments shown with green arrows. Amino acids give different amide traces in TOCSY spectrum. NOESY spectrum contains the same peaks as in TOCSY and additional NOE peaks between the protons, which are close in space, but not connected by bonds inside one amino acid spin system.
Figure 2-13. Continued
Figure 2-14. A summary of solid phase peptide synthesis with Fmoc N-terminus protection group. 1) Fmoc protected amino acid attached to a resin via linker is deprotected using piperidine. 2) The next Fmoc protected amino acid is activated with carbodiimide and 3) coupled to the growing peptide chain. 4) The deprotection/coupling cycle is repeated. The desired amino acid sequence is deprotected and cleaved from the resin with TFA to yield a free peptide.
Figure 2-15. pET-31b expression system designed by Novagen and SPBc peptide purification scheme.\(^{(122)}\) The pET-31b vector is designed for cloning and high-level expression of peptide sequences fused with the 125 amino acid KSI protein upstream and His\(_6\)-tag downstream of the peptide sequence; peptide is separated by methionine codons from tags.\(^{(123)}\)
Figure 2-16. Overexpression of 18 kDa KSI-SPB<sub>59-80</sub>-His<sub>6</sub> fusion construct in BL21(DE3)pLysS cells monitored as function of time after overexpression induction. Lane #1-sizes of the protein standards in kDa, #2 - 0 hours after induction, #3 - 2 hours, #4 - 4 hours, #5 - 5 hours, #6 - 19 hours.

Figure 2-17. Nickel-affinity chromatogram of purification of KSI-SPB<sub>59-80</sub>-His<sub>6</sub> fusion construct from inclusion bodies in denaturing conditions (A); extremely hydrophobic KSI-SPB<sub>59-80</sub>-His<sub>6</sub> fusion construct after dialysis against water precipitates in the dialysis bag (B).
Figure 2-18. Cleavage efficiency of KSI-SPB<sub>59-80</sub>-His<sub>6</sub> fusion construct by cyanogen bromide in 80% formic acid was monitored as a function of time: t<sub>0</sub> is 0 hours after overexpression induction, t<sub>1</sub> is 2 hours, t<sub>2</sub> is 3 hours, t<sub>3</sub> is 19 hours, t<sub>4</sub> is 22 hours, t<sub>5</sub> is 23 hours. Sizes of the protein standards (PS) are shown on the right side in kDa.

Figure 2-19. Pure SPB<sub>59-80</sub> peptide in RP-HPLC fractions after lyophilization. Peptide was loaded at different concentrations to estimate its quantity.
Figure 2-20. Mass spectra of expressed SP-\textit{B}'_{59-80} (Top), wild type SP-\textit{B}_{59-80} (Bottom) and \textit{^{15}N-SP-\textit{B}'_{59-80}} (next page) peptides with theoretical m/z 2578, 2531 and 2610 respectively.
Figure 2-20. Continued.
CHAPTER 3
INTERACTION OF THE C-TERMINUS OF SP-B WITH LIPID BILAYERS MIMICKING LUNG IS MODULATED BY ACYL CHAIN SATURATION*

To date, KL$_4$ that was developed based on the hydrophilic/hydrophobic pattern in SP-B$_{59-80}$, is the only synthetic replacement for SP-B that has enjoyed the most clinical success.\(^{(71-75)}\) However, their primary sequences have only modest similarity. There are only positively charged lysines in KL$_4$, each separated by four leucines, whereas SP-B$_{59-80}$ additionally has negatively charged terminal amino acids and variable spacing of three or four hydrophobic residues between polar or charged residues. Consequently, a significant question remaining is whether KL$_4$ and SP-B$_{59-80}$ act similarly, despite clear deviations at the primary amino acid level. These differences could lead to varying secondary structures and penetration into lipid lamellae as well as different effects on phospholipid dynamics. One key to understanding how similar KL$_4$ is to SP-B$_{59-80}$ in its behavior is to examine the binding of SP-B$_{59-80}$ to model lipid membrane systems, its effects on their dynamics, and any effects on the Lβ to Lα phase transition in DPPC lipids; these results can then be directly compared to similar previous studies of KL$_4$.

The study of the effects of SP-B$_{59-80}$ on 4:1 DPPC:POPG and 3:1 POPC:POPG MLVs\(^{\text{68}}\) was included in this dissertation since it provided important preliminary results, which determined the course and the main goals of this dissertation. My role was to express and purify SP-B'$_{59-80}$ with isoleucines incorporated in lieu of methionines to demonstrate that the mutations cause no effects on lipid dynamics and the results are

* previously published in full version as: Antharam, V.C., Farver, R.S., Kuznetsova, A., Sippel, K.H., Mills, F.D., Elliott, D.W., Sternin, E., Long, J.R. (2008), Interactions of the C-terminus of lung surfactant protein B with lipid bilayers are modulated by acyl chain saturation, BBA 1778, 2544-2554. The data was also partially published in dissertation by Antharam, V.C. (2008), Biophysical characterization of peptide mimics of lung surfactant protein B.
identical to those of the native sequence, which were obtained by other authors in the paper. Lipid composition of 4:1 DPPC:POPG is similar to formulations commonly used in studying lung surfactant, while 3:1 POPC:POPG is a paradigm lipid system commonly employed to probe peptide/lipid interactions, particularly in studies of cationic, amphipathic helices. Lipid phases of these compositions could also be found in localized areas of the alveoli during the breathing cycle.

Materials and Methods

Synthesis of SP-B_{59-80}

SP-B_{59–80}, (DTLLGRMLPQLVCRLVLRCSMD) was synthesized via solid-phase peptide synthesis on a Wang resin (ABI 430, ICBR, UF) and cleaved from the resin with King's reagent and ether precipitated. The cleaved product was purified via RP-HPLC using a C18 Vydac column with a water/acetonitrile gradient (containing 0.3% TFA). The fractions corresponding to SP-B_{59–80} were collected and the peptide was verified by mass spectrometry with a mass to charge ratio of (m/z) of 2533. Dried peptide was weighed and dissolved in methanol to a stock concentration of approximately 1 mM, and aliquots were analyzed by amino acid analysis for a more accurate determination of concentration (Molecular Structure Facility, UC Davis).

Heterologous Expression of a SPB_{59-80} Double Mutant (SP-B'_{59-80})

For later experiments, SP-B'_{59–80} was expressed using a pET31 construct (EMD Biosciences, Inc., Gibbstown, NJ) incorporating a codon-optimized synthetic gene for SP-B_{59–80} (DNA2.0, Menlo Park, CA) in BL21(DE3) cells with subsequent purification and cleavage using established protocols. The expressed sequence of SP-B'_{59–80} was modified to incorporate isoleucines in lieu of methionines (DTLLGRILPQLVCRLVLRCSID) for compatibility with a cyanogen bromide cleavage
reaction. Following HPLC purification of the final product, fractions corresponding to SP-B<sub>59–80</sub> were collected and the peptide was verified by mass spectrometry with a mass to charge ratio (m/z) of 2497. Dried peptide was weighed and dissolved in methanol to a stock concentration of approximately 1 mM, and aliquots were analyzed by amino acid analysis for a more accurate determination of concentration (Molecular Structure Facility, UC Davis). This peptide was used to obtain results shown in Figs. 3-1, 3-5, 3-6, and 3-11 as well as the POPG-d31 results in Figs. 3-9 and 3-10. Its effects on lipid dynamics were ascertained to be identical to those of the native sequence.

**Preparation of Peptide:Lipid Samples**

POPC, DPPC, POPG, POPC-d31, DPPC-d62 and POPG-d31 were purchased as chloroform solutions (Avanti Polar Lipids, Alabaster, AL) and concentrations were verified by phosphate analysis<sup>1</sup> (Bioassay Systems, Hayward, CA). The lipids were mixed at a molar ratio of 4:1 DPPC:POPG and 3:1 POPC:POPG in chloroform and aliquoted. For samples containing peptide, a methanol solution of SP-B<sub>59–80</sub> was added to lipid solutions with final peptide: lipid (P:L) molar ratios ranging from <1:1000 to >1:50. The samples were dried under a stream of nitrogen with the sample temperature maintained at 42–50°C in a water bath; the resulting films were suspended in cyclohexane, flash-frozen, and lyophilized overnight to remove residual solvent.

**CD Experiments**

3 mg of peptide-lipid powder was solubilized in 1 mL of 10 mM HEPES buffer at pH 7.4, with 140 mM NaCl, to achieve a concentration of 40 μM SP-B<sub>59–80</sub> with 4 mM lipids. Samples were placed in a 50 °C water bath to facilitate solubilization accompanied by 3–5 freeze-thaw cycles with vortexing to achieve equilibration. Peptide-lipid MLVs were extruded through 100 nm filters (Avanti Polar Lipids, Alabaster, AL) 15–
25 times above the Tm of the lipids to form LUVs just prior to CD analysis. CD experiments were performed on an Aviv Model 215 (Lakewood, NJ) at 45°C using a 200–260 nm wavelength range, a 1 nm step size and averaging of 40–50 scans. Background contributions from the buffer and LUVs were removed by subtracting appropriate controls.

**Solid-state NMR Analysis**

For each solid-state NMR sample, 30 mg of peptide-lipid powder was placed in a 5 mm diameter NMR tube and 200 μL of buffer containing 5mM HEPES at pH 7.4, 140mM NaCl, and 1mM EDTA in ²H depleted water (Cambridge Isotopes, Andover MA) was added. Samples were made using 4:1 DPPC-d62:POPG, 4:1 DPPC:POPG-d31, 3:1 POPC-d31:POPG, and 3:1 POPC:POPG-d31 lipid preparations. NMR samples were then subjected to 3–5 freeze-thaw cycles with gentle vortexing to form MLVs. ³¹P and ²H NMR data were collected on 500 and 600 MHz Bruker Avance systems (Billerica, MA) using standard 5mm BBO probes. For ³¹P NMR experiments, 25 kHz proton decoupling was employed during acquisition to remove dipolar couplings. Spectra were acquired at 34, 39, and 44°C to verify sample equilibration with 1024–2048 scans and a 5 s recycle delay between scans to minimize RF sample heating. For the ²H NMR experiments, data were collected using a quad echo sequence (90°-τ-90°-τ-acq with τ=30 μs) with a B1 field of 40 kHz. Spectra were acquired at 34, 39, and 44°C with 1024 or 2048 scans and a 0.5 s recycle delay between scans. To monitor the phase transitions of the lipids, for some samples spectra were also collected over a range of 30–44°C in 2° increments. DePaking of NMR data was accomplished with previously published algorithms which simultaneously dePake and determine macroscopic ordering in partially aligned lipid spectra using Tikhonov regularization.\(^{(101)}\)
$^{31}$P NMR spectra were referenced to phosphate buffer prior to dePaking and dePaked spectra were quantitated by fitting the two peaks with Lorentzian line shapes. Assignments of $^2$H resonances were made based on published values for DPPC.\textsuperscript{(103, 131)}

**Results**

**Secondary Structure**

The CD spectra at 45°C of SP-B\textsubscript{59–80} incorporated into lipid vesicles are shown in Fig. 3-1. The CD spectra are characterized by double minima at 206–208 and 222 nm, as is typically seen for peptide helices. Interestingly, the spectra have a minimum ellipticity at 206–208 significantly lower than the minimum at 222 nm. Similar spectra have been observed for KL\textsubscript{4} in DPPC:POPG LUVs\textsuperscript{(132)} and peptides which are constrained to form π-helices in buffer.\textsuperscript{(133)} Interpretation of the CD spectra in terms of helix content is complicated by the fact that SP-B\textsubscript{59–80} does not form a typical amphipathic α-helix when projected on a helical wheel. Fitting of the CD data with standard deconvolution software\textsuperscript{(134)} led to secondary structure estimates of 75–82% α-helix, 17–23% random coil, and negligible β-sheet for SP-B\textsubscript{59–80} interacting with POPC:POPG LUVs. Secondary structure estimates of 98–100% α-helix with negligible random coil and β-sheet populations were seen for SP-B\textsubscript{59–80} interacting with DPPC:POPG LUVs. However, the quality of the fits was poor due to the nonstandard shape of the CD spectra and poor quality of the data below 200 nm due to light scattering from the lipids.

**The Interaction of SP-B\textsubscript{59–80} with Lipid Headgroups**

$^{31}$P solid-state NMR for 4:1 DPPC:POPG and 3:1 POPC:POPG MLVs with varying levels of SP-B\textsubscript{59–80} were collected to assess the effect of this peptide on lipid phases and the orientation of the phospholipid headgroups. Shown in Fig. 3-2 are $^{31}$P spectra
for the DPPC-d62:POPG samples at 44°C. Also shown are the dePaked spectra allowing a clear determination of the individual, time-averaged chemical shift anisotropies (CSAs) of the PC and PG headgroups. Only lineshapes consistent with lamellar phases are observed and the resonance for the POPG lipids moves with addition of peptide; this is more evident after dePaking the spectra. As seen in other studies, \(^{(132, 135, 136)}\) spontaneous macroscopic lipid alignment occurs in the magnetic field. Due to this phenomenon, normally spherical MLVs undergo a deformation to more ellipsoidal geometries, distorting the powder spectra. The extent of magnetic field alignment was accounted for in the dePaking algorithm. \(^{(101)}\) The difference in PG and PC CSAs is due to the difference in their preferred headgroup orientations relative to the membrane normal. \(^{(89)}\) With increasing levels of SP-B\(_{59-80}\), the PC CSA is invariant, but the PG CSA lessens with increasing peptide levels. The CSA for POPG alone is significantly smaller than observed for POPG in 4:1 DPPC:POPG or 3:1 POPC:POPG mixtures prior to the addition of SP-B\(_{59-80}\). \(^{(132)}\) However, for the binary 4:1 DPPC:POPG mixture containing higher concentrations of SP-B\(_{59-80}\), the \(^{31}\) P CSAs are more comparable to those of the neat lipids. Thus, SP-B\(_{59-80}\) is clearly affecting the interactions of the PC and PG lipids. Shown in Fig. 3-3 are \(^{31}\) P spectra for the POPC-d31:POPG samples at 44°C and their dePaked counterparts. For 3:1 POPC:POPG MLVs, addition of SP-B\(_{59-80}\) leads to a considerably smaller decrease in the PG CSA; in contrast, KL\(_4\) which has similar effects on the PG CSA in both DPPC:POPG and POPC:POPG mixtures. \(^{(132)}\) Based on these results, when POPG is interacting with PC headgroups, electrostatic interactions cause the PG headgroups to reorient leading to a subsequent increase in their averaged CSA values. Addition of surfactant peptide
disrupts this interaction, causing the PG CSAs to move toward values seen for POPG alone; if the surfactant peptide is highly cationic, electrostatic interactions lead to similar decreases in the POPG CSA regardless of whether the PC lipid is fully saturated or monounsaturated. In the case of SP-B_{59-80}, the negatively charged amino acids at the termini result in less electrostatic interaction between the peptide and the PG headgroup. However the partitioning of the peptide into the DPPC:POPG lipids is sufficient to significantly alter the orientation of the PG headgroup due to its overall effects on lipid packing and dynamics; in POPC:POPG lipids it is not. The $^{31}$P NMR spectra for 3:1 POPC-d31:POPG MLVs containing varying concentrations of SP-B_{59-80} are typical for lamellar phases up to 1.7 mol% peptide (Fig. 3-3). However, the spectrum at 2.5 mol% peptide suggests the onset of other bulk dynamics in the lipids, consistent with exchange between a lamellar phase and either a micellar phase or a hexagonal phase. With the onset of a phase transition at 2.5 mol% peptide, no clear distinction between the PC and PG lipids can be made.

$^2$H NMR Studies of DPPC:POPG Lipid Miscibility on Addition of SP-B_{59-80}

DSC thermograms for 4:1 DPPC-d62:POPG LUVs containing varying levels of SP-B_{59-80} indicate SP-B_{59-80} has little effect on DPPC:POPG Lipid Miscibility (Fig. 3-4; data by Antharam, V.C.) with the phase transition moving to slightly higher temperatures with increasing peptide concentration. In contrast, KL$_4$ has been shown by DSC and fluorescence microscopy to promote lipid phase separation and domain formation. The relative differences between KL$_4$ and SP-B$_{59-80}$ are not completely unexpected as KL$_4$ contains solely hydrophilic residues that are cationic while SP-B$_{59-80}$ contains a mixture of cationic and anionic residues. Thus, partitioning of KL$_4$ with POPG
may be more energetically favorable, enhancing phase separation in DPPC/POPG mixtures below the phase transition temperature of DPPC.

To determine more precisely whether addition of SP-B<sub>59–80</sub> causes phase separation of the DPPC and POPG lipids, <sup>2</sup>H NMR spectra were collected over the temperature range of the phase transitions for samples which contained either deuterated DPPC or deuterated POPG (Fig. 3-5). A first moment analysis of the data was used to determine the phase transition temperatures of the individual lipids (Fig. 3-6). From these spectra, it is clear that in the 4:1 DPPC:POPG samples the lipids are fully miscible with the DPPC and POPG melting at similar temperatures. The phase transition seen for deuterated DPPC is at a slightly lower temperature (midpoint of 30.6°C) than for deuterated POPG (32.7 °C). This is in part due to the fact that a larger percentage of the fatty acyl chains are deuterated (80% vs. 10%) (139) but may also reflect slight differences in POPG content between the two samples since they were made from different lipid stock solutions. For the samples that do not contain SP-B<sub>59–80</sub>, the spectra at intermediate melting temperatures are a superposition of gel phase and liquid phase spectra. Addition of SP-B<sub>59–80</sub> affects the phase transition temperature of both lipids with the phase transition temperature for DPPC-d62 increasing to 31.5°C and the POPG-d31 phase transition temperature decreasing to 30.3°C. On addition of SP-B<sub>59–80</sub>, at lower temperatures the POPG-d31 spectra have a larger percentage of the dynamic lipid phase evident and in the DPPC-d62 spectra the dynamic phase is absent, as can be seen by comparing spectra at 30°C and 32°C. The peptide also leads to an increase in temperature at which the lipids are completely melted (34°C vs. 38°C for deuterated DPPC; 36°C vs. 38°C for deuterated POPG), consistent with the DSC data.
Attempts to determine the fractions of liquid phase and gel phase by spectral subtractions with samples containing differing percentages of SP-B<sub>59–80</sub> were unsuccessful due to variations in the lipid alignment and acyl chain order parameters. Thus, addition of SP-B<sub>59–80</sub> leads to some phase separation of the lipids at lower temperatures with a significant fraction of POPG remaining in the gel phase; at higher temperatures, the lipids exhibit similar spectra consistent with the lipids becoming fully miscible. These differences are too subtle to be observed via differential scanning calorimetry. Above the phase transition temperature of the domain with the higher T<sub>m</sub>, the peaks in the <sup>2</sup>H spectra coalesce to a single resonance for each position in the acyl chains suggesting the absence of separate domains in the fluid phase or exchange between domains is fast on the NMR time scale.

**Effects of SP-B<sub>59–80</sub> on Lipid Acyl Chains**

<sup>2</sup>H NMR spectra of DPPC-d<sub>62</sub>:POPG and POPC-d<sub>31</sub>:POPG MLVs at 44°C with varying levels of SP-B<sub>59–80</sub> (Figs. 3-7 and 3-8) allow determination of the effect of this membrane-active peptide on lipid dynamics in the fluid phase and insight into the depth of peptide penetration. For 3:1 POPC-d<sub>31</sub>:POPG MLVs, decreases are seen in the ordering of the POPC sn-1 acyl chains on addition of SP-B<sub>59–80</sub>. This suggests the peptide binds to the interface, increasing the lateral spacing between the lipids, allowing more motion in the acyl chains. In contrast, addition of SP-B<sub>59–80</sub> to 4:1 DPPC-d<sub>62</sub>:POPG MLVs increases the ordering of the DPPC acyl chains, particularly toward the middle of the bilayer. This suggests either 1) insertion of SP-B<sub>59–80</sub> deeply into the DPPC:POPG bilayers, restricting the motional freedom of the acyl chains, or 2) an electrostatic interaction of the peptide with the lipid headgroups resulting in a change in the lipid packing and headgroup conformation. Also shown in Figs. 3-7 and 3-8 are the
dePaked spectra which allow assignment of each C–D bond in the acyl chain and determination of the time-averaged order parameters, \(<S_{CD}>\). For samples in which the lipid bilayers show a tendency to align in the magnetic field, the dePaking must take into account the effects of lipid alignment on the probability distribution function. Assuming the magnetic field leads to an ellipsoidal deformation of the MLVs, the probability distribution becomes:

\[
p_{l(E)}(\theta) \propto \sin(\theta)[1 - (1 - k_E^2)\cos^2 \theta]^{-2}
\]  

(3-1)

where \(k_E\) refers to square of the ratio of the long to short axes of the ellipsoids.

Using Tikhonov regularization, \(k_E\) and the dePaked spectra were determined simultaneously. Lipid acyl chain order parameter profiles\(^{140}\) were generated (Figs. 3-9 and 3-10) by assigning the dePaked spectra, measuring the quadrupolar splitting \((\Delta \nu_q)\) for deuterium atoms at various positions along the acyl chain, and determining their order parameter using Eq. (3-2), where \(\frac{3}{4} \frac{e^2 q Q}{h}\) is the quadrupolar coupling of the deuterium nucleus to the electric field gradient. A static quadrupolar coupling of 167 kHz\(^{94}\) was assumed in calculating the order parameters.

\[
\Delta \nu_q = \frac{3}{4} \frac{e^2 q Q}{h} (3\cos^2 Q - 1)S_{CD}
\]  

(3-2)

Comparison of these order parameter profiles to those for the individual lipids show that the DPPC order parameters for 4:1 DPPC:POPG MLVs with P:L ratios >1:200 are higher than would be expected for DPPC alone.\(^{132}\) Within the resolution of the \(^2\)H NMR experiment, no phase separation is seen. The profiles also yield a more detailed picture of how SP-B\(_{59-80}\) affects the lipid dynamics with carbon positions 9–16
the most affected in the PC lipid acyl chains for both lipid systems. The effects of SP-B<sub>59–80</sub> on the PG lipids were similarly monitored by collecting <sup>2</sup>H NMR spectra for 4:1 DPPC:POPG-d31 and 3:1 POPC:POPG-d31 MLVs, dePaking the spectra, and generating order parameter profiles. The order parameter profiles for the sn-1 chain on POPG in these lipid mixtures are also given in Figs. 3-9 and 3-10. In these experiments an overall decrease in ordering of the POPG acyl chains is seen on addition of SP-B<sub>59–80</sub>. Decreases in the order parameters are seen at all the acyl positions in POPG in 3:1 POPC:POPG MLVs and are slightly larger than the changes for POPC. The POPG acyl chain order parameters in 4:1 DPPC:POPG MLVs do not appear to be as affected by SP-B<sub>59–80</sub>, however a comparison of the results further down the acyl chain for POPG-d31 when mixed with DPPC vs. POPC again demonstrates that the peptide interacts with these lipid systems in a manner dependent on the degree of saturation of the fatty acid chains. Since SP-B<sub>59–80</sub> is most likely interacting with both lipid populations in the fluid phase rather than segregating the lipids and interacting with a single phase above the Lß to Lα phase transition, the changes in POPG-d31 and DPPC-d62 order parameter profiles on addition of SP-B<sub>59–80</sub> to DPPC:POPG MLVs may seem to be contradictory. The observed changes in order parameters can be reconciled when the data are viewed as the change in order parameter at each position along the acyl chain on addition of peptide relative to the order parameters of the lipids without the peptide. Shown in Fig. 3-11 are the changes in order parameters at particular methylene positions in the sn-1 palmitoyl acyl chains for the individual lipids in 3:1 POPC:POPG and 4:1 DPPC:POPG MLVs, respectively, on addition of peptide at a P:L molar ratio on the order of 1:100. From these graphs, the behavior of the PG lipids relative to the PC
lipids is clearly offset, but the trends with respect to acyl chain position are similar. The offset of the profiles for the PG lipids is consistent with $^{31}$P NMR findings described above showing that association of SP-B$_{59-80}$ with the lipids leads to a change in the orientation PG headgroup and an overall decrease in the $^{31}$P CSA for the POPG lipids. Thus, the offset in the PG profiles in Fig. 3-11 relative to the PC profiles does not reflect significantly less internal order in the PG lipid methylene chains per se, but instead reflects a change in the average orientation of the PG lipid director relative to the membrane normal. The degree of offset for the POPG profiles in the DPPC:POPG and POPC:POPG mixtures correlate with the degree to which the POPG $^{31}$P CSA is affected by addition of peptide. Thus the addition of SP-B$_{59-80}$ is most likely equally affecting the internal dynamics of the methylene chains in both lipids in a similar manner. From their profiles, the methylenes in the plateau region are less affected than those further down the acyl chain in the POPC:POPG lipid system. This behavior is similar to changes observed on addition of antimicrobial peptides, which typically cause larger changes in order further down the methylene chain,\(^{(127, 141, 142)}\) but is in contrast to the interaction of KL$_4$\(^{(132)}\) or saposin C\(^{(143, 144)}\) with monounsaturated lipids, which have a larger effect on the plateau region. Of particular interest to the roles of SP-B$_{59-80}$ and KL$_4$ in lung surfactant formulations are their similar effects on 4:1 DPPC:POPG MLVs. Previously we have observed that KL$_4$ increases order along the DPPC acyl chains with its greatest effect at the center of the bilayers, and this same trend is seen with SP-B$_{59-80}$. The profile for POPG shows little change in order toward the center of the bilayers and a decrease in the plateau region. However, as discussed above, these differences reflect a change in the average orientation of the POPG headgroup and if this change in
orientation is taken into account, the internal order at individual methylene positions in
the POPG acyl chain is likely increasing over the entire length of the acyl chain.
Increases in lipid acyl chain ordering have been observed in response to polyelectrolyte
binding to lipid headgroups\(^{(145, 146)}\) as well as in response to the addition of small
molecules, such as cholesterol,\(^{(147, 148)}\) or transmembrane peptide helices which
partition into the acyl chain region of the lipids.\(^{(149, 150)}\) In the case of polyelectrolyte
binding, larger changes are seen for the plateau region of the lipids reflecting a change
in overall orientation of the lipids due to changes in packing of the headgroups. Smaller
changes are also seen near the center of the bilayers and are interpreted as resulting
from closer packing of the lipids on binding of the electrolyte. For transmembrane
helices similar in length to SP-B\(^{59-80}\), increases in order parameters which are similar in
magnitude are seen for both the methylene positions in the plateau region and toward
the center of the bilayers; these trends are observed in both POPC and DPPC bilayers.
This is expected since insertion of a transmembrane peptide of this length primarily
affects the thickness of the bilayer and would interact with both the lipid acyl chains and
headgroups across the entire span of the bilayers. In contrast, on addition of
cholesterol, larger changes in order parameters are observed toward the center of the
bilayer\(^{(147)}\) since cholesterol partitions to the bilayer interior and does not strongly
interact with the lipid head groups. Interestingly, the effects of SP-B\(^{59-80}\) and KL\(_4\) on
order in DPPC:POPG bilayers are most similar to cholesterol, suggesting the bulk of
each peptide is partitioning deep within the lipid bilayers. The higher degree of ordering
toward the center of the bilayer suggests the surfactant peptides are penetrating deeply
into the bilayer and decreasing the mobility of the acyl chains. However, the smaller
changes seen in the plateau region (carbons 3–8) suggest the peptides do not adopt a transmembrane orientation. The thermodynamic penalty imposed by placing SP-B_{59–80} in a transmembrane orientation would be prohibitive since it would place the hydrophilic amino acids at positions 6, 10, 14, and 16 into the hydrophobic core.\(^{(151)}\) Thus, the changes in order parameters indicate SP-B_{59–80} lodges into the hydrophobic region of the bilayer, while maintaining a perpendicular orientation to the bilayer normal. This type of interaction would lead to a negative curvature strain within the lipid bilayers, a phenomenon which has been hypothesized as being important to lung surfactant function.\(^{(152)}\) A recent solution NMR study of SP-B_{63–78} bound to SDS micelles found the peptide forms an amphipathic helix at the water interface with the hydrophilic residues solvent accessible,\(^{(60)}\) consistent with our interpretation. Looking at the same sn-1 C–D positions for POPC-d31 in 3:1 POPC:POPG MLVs (Figs. 10 and 11), it can be seen that SP-B_{59–80} decreases order by as much as 20%, with the largest decrease in ordering seen for position 15. This strongly correlates to SP-B_{59–80} partitioning at the interface of 3:1 POPC:POPG MLVs and increasing the area per lipid molecule leading to more motion in the acyl chains. A striking decrease in the plateau region is seen at 2.5 mol% peptide along with the appearance of an isotropic peak in the \(^{2}\text{H}\) spectrum (Fig. 3-8), consistent with the \(^{31}\text{P}\) NMR data suggesting that this concentration of peptide leads to destabilization of the lamellar phase and exchange of the lipids between a lamellar and nonlamellar phase.

From these results we conclude that the penetration and interaction of SP-B_{59–80} with the lipids in lung surfactant is dependent on the degree of saturation in the lipids. Intriguingly, the peptide penetrates deeply into mixtures containing a high level of
saturated lipids even though the order of the fatty acyl chains in these lipids is higher than those in monounsaturated lipids. When interacting with monounsaturated lipids, SP-B_{59–80} binds at the interface and has effects similar to other amphipathic peptides, including antimicrobial peptides, which increase curvature strain in the lipids and can cause micelle formation at higher concentrations. The interaction of SP-B_{59–80} with these binary lipid systems is on the whole very similar to behavior previously seen for KL_{4}, suggesting the two peptides have very similar mechanisms of binding and similar effects on the dynamics of the lipids.

**Discussion**

The effects of the C-terminal region of SP-B, specifically residues 59–80, on lipid dynamics and acyl chain ordering were explored in two binary lipid systems that differ in their degree of monounsaturation. SP-B_{59–80} has been used as a template for the design of simple molecules for the treatment of respiratory distress syndrome (RDS), the most notable of these being the 21 amino acid peptide KL_{4} which has shown exceptional efficacy in lung surfactant formulations, particularly when compared to more conventional therapies that rely upon exogenous sources of SP-B.\(^{(153)}\) The SP-B_{59–80} peptide is presumed to be helical when interacting with lipids based on the distribution of charged and uncharged residues in its primary amino acid sequence and this is borne out by our CD data. To date, few studies examining the structure of SP-B_{59–80} and its interactions with lipids have been carried out despite its demonstrated surface activity and role in the rational design of KL_{4}. Additionally, subsequent studies of KL_{4} have suggested its properties might be more similar to SP-C, which is a small, highly hydrophobic transmembrane protein in contrast to SP-B, which is found at membrane interfaces. Previously we have examined the structure of KL_{4}\(^{(154)}\) and found it to be
consistent with the peptide partitioning in bilayers with the helix axis perpendicular to the membrane normal, but its depth of penetration is dependent on the degree of saturation in the lipids.\textsuperscript{(132)} Whether SP-B\textsubscript{59–80} binds similarly to lipid interfaces and affects the properties of the lipids in the same manner is the subject of this study. While the sequence of KL\textsubscript{4} is based on SP-B\textsubscript{59–80}, there are some notable differences in their primary amino acid sequences. First, KL\textsubscript{4} contains only two types of amino acids, the hydrophobic leucine and cationic lysine; in SP-B\textsubscript{59–80} the distribution of hydrophobic and hydrophilic residues is similar, but the hydrophilic amino acids are a mixture of anionic (D\textsubscript{59} and D\textsubscript{80}), polar (Q\textsubscript{68}) and cationic (R\textsubscript{64}, R\textsubscript{72}, and R\textsubscript{76}) residues at physiologic pH. The spacing of lysines in KL\textsubscript{4} is based on the five charged amino acids in SP-B\textsubscript{59–80}, but the preponderance of long, cationic sidechains would allow more favorable electrostatic interactions with the anionic POPG lipids and phospholipid interfaces in general. If both peptides form amphipathic helices at lipid interfaces, their secondary structure might also differ since the spacing of the polar residues in SP-B\textsubscript{59–80} is primarily every four residues rather than every five residues, as in KL\textsubscript{4}. We have found that KL\textsubscript{4} has the ability to deeply embed in DPPC-rich bilayers, which we have attributed to the lysine sidechains being able to snorkel up to the phosphates at the lipid interface.\textsuperscript{(132)} While arginine residues would also allow this behavior, the aspartic acid residue sidechains at the ends of SP-B\textsubscript{59–80} would be too short and the electrostatic interactions with the phosphate moieties would be unfavorable. Finally, the occurrence of a proline residue at position 67 in SP-B\textsubscript{59–80} could affect the helical nature of the peptide. Nonetheless, both peptides have been demonstrated to be surface active and effective in lowering surface tension at air/water interfaces. Based on the DSC data, SP-
B\textsubscript{59–80} does not have the same effect on the macroscopic phase properties of DPPC:POPG mixtures as KL\textsubscript{4}. DSC experiments performed by Saenz, et al.\textsuperscript{(137)} and our group\textsuperscript{(132)} as well as epifluorescence studies\textsuperscript{(138)} indicate that KL\textsubscript{4} mediates phase separation in DPPC:POPG environments. In contrast, DSC indicates SP-B\textsubscript{59–80} has little effect on the thermodynamic properties of DPPC:POPG LUVs other than small effects on the cooperativity of the L\textbeta to L\textalpha phase transition. However, the \textsuperscript{2}H NMR data show there is a small degree of phase separation in the gel phase on addition of SP-B\textsubscript{59–80}. In contrast to the calorimetry data, which reports on bulk thermodynamic properties of the lipids, \textsuperscript{31}P and \textsuperscript{2}H NMR data monitoring lipid dynamics above the phase transition temperature show the two peptides interact similarly with lipid vesicles and have similar effects on lipid dynamics in 4:1 DPPC-d62:POPG lipid environments. The \textsuperscript{31}P NMR data indicate that neither peptide affects the DPPC CSA, while both peptides affect the POPG CSA. While an interaction of SP-B\textsubscript{59–80} with the anionic PG headgroups is one plausible explanation of this data, it should be noted that SP-B\textsubscript{59–80} does not possess the distinct cationic charge periodicity found in KL\textsubscript{4}; the sole use of lysines for charged residues in KL\textsubscript{4} could allow for a more enhanced electrostatic interaction with PG compared to SP-B\textsubscript{59–80}. This fact is borne out by quantitative comparison of the data for the two peptides. The changes in the POPG CSA, which report on changes in the average orientation of the POPG headgroups, with increasing concentrations of SP-B\textsubscript{59–80} are less than half those seen for KL\textsubscript{4}. Changes caused by SP-B\textsubscript{59–80} can be attributed to the peptide generally affecting lipid-lipid interactions rather than a direct electrostatic interaction with the POPG headgroups. However, addition of either peptide leads to increases in the time-averaged deuterium order parameters for the DPPC lipids,
particularly at carbons 9–15. These increases suggest equally deep penetration of the peptides into the DPPC:POPG bilayers and are similar in scale for both peptides. For both peptides a transmembrane orientation is unlikely as this would place 3–4 polar amino acids into the bilayer interior. In this regard, our observations support KL₄ and SP-B₅₉₋₈₀ partitioning similarly into DPPC:POPG lipid bilayers. In 3:1 POPC-d₃₁:POPG lipids, the changes in the $^{31}$P CSAs for the lamellar vesicles on addition of SP-B₅₉₋₈₀ are smaller than those seen in 4:1 DPPC:POPG samples or on addition of KL₄ to POPC:POPG mixtures. However, at 2.5 mol% SP-B₅₉₋₈₀, a striking difference is seen in the $^{31}$P powder patterns suggesting the onset of exchange between the lipid lamellar phase and a more dynamic lipid phase. The $^2$H NMR order parameters decrease with increasing peptide concentrations. These increases in fatty acid mobility are similar to the trends found for KL₄, but the changes with SP-B₅₉₋₈₀ are considerably greater in magnitude. These findings indicate that the SP-B₅₉₋₈₀ remains in the interfacial region of the lipid bilayers at lower concentrations. At 2.5 mol% SP-B₅₉₋₈₀, a large coalesced peak in the center of the $^2$H spectrum appears indicating the formation of a second, non-bilayer lipid phase is occurring. The $^{31}$P and $^2$H NMR data are consistent with the formation of small vesicles with motions that are fast on the NMR timescale, and may also implicate SP-B₅₉₋₈₀ in lipid shuttling, lysis or degradation. Lipid lysis and degradation has been suggested as one of the many functions of SP-B that could be necessary for effective surfactant recycling and remediation.$^{155, 156}$ An antimicrobial function has also been attributed to SP-B, and the lipid dynamics seen at 2.5 mol% SP-B₅₉₋₈₀ are similar to behavior seen with higher concentrations of amphipathic antimicrobial peptides such as LL37, pardaxin, magainin, and synthetic derivatives.$^{127, \text{...}}$
This function is of particular interest as well to the saposin protein family, of which SP-B is the most hydrophobic member; this family includes several proteins which bind and traffic lipids for enzymatic modification or degradation.\(^{(22)}\) However, further studies are needed to differentiate between either a role for SP-B\(_{59-80}\) in lipid degradation or in forming nonlamellar phases. Interestingly, we only saw this effect in 3:1 POPC:POPG, and not in DPPC:POPG at 2.5 mol\% SP-B\(_{59-80}\). This suggests the activity of SP-B\(_{59-80}\) is dependent on the saturation level of the acyl chains in a concentration dependent manner and provides a means of discrimination between lipids in the dynamic lung environment. Based on the CD, \(^{31}\)P NMR and \(^{2}\)H NMR data, a model of how SP-B\(_{59-80}\) interacts with these lipid systems can be put forth (Fig. 3-12). It should be noted that in this model we assume that SP-B\(_{59-80}\) is helical in a lipid environment, based on our CD data, and we propose that the peptide assumes an orientation in which the helix axis is perpendicular to the membrane normal due to the occurrence of charged residues throughout the primary sequence. Such an assumption is valid given the findings in the literature pertaining to the helical nature of SP-B and in particular the C-terminal region.\(^{(55, 60, 158, 159)}\) However, from examination of the primary amino acid sequence and our \(^{2}\)H NMR derived order parameters it is conceivable that the type of helical structure SP-B\(_{59-80}\) adopts may not be a canonical \(\alpha\)-helix and the pitch of the helix may change as the peptide more deeply penetrates into the saturated lipid environments. For KL\(_4\), this type of structural transition has been observed by solid-state MAS NMR experiments (Mills, F.D. and Long ,J.R., unpublished data). To answer the question of the type of helix SP-B\(_{59-80}\) forms in a lipid environment, similar MAS NMR experiments on \(^{13}\)C labeled peptide in complex with lipids need to be performed.
Since an increase in DPPC acyl chain ordering is seen on addition of SP-B<sub>59–80</sub> to 4:1 DPPC:POPG lipid vesicles, we conclude the peptide penetrates deeply into the hydrophobic region of the lipids and restricts acyl chain motion. In 3:1 POPC:POPG lipid vesicles, the disordering of the acyl chains indicate that the peptide is at the lipid/buffer interface in this lipid system. This is consistent with the CD data which shows solely a helical conformation for SP-B<sub>59–80</sub> when it is interacting with 4:1 DPPC:POPG vesicles, suggesting complete binding of the peptide, and a mixture of helix and random coil when it is interacting with 3:1 POPC:POPG vesicles, suggesting the peptide might partition between the lipid interface and aqueous phases. The results for addition of SP-B<sub>59–80</sub> to DPPC:POPG suggest the peptide may deeply penetrate into the bilayer with the charged sidechains “snorkeling” to the interface as we have postulated to occur for KL<sub>4</sub>. The snorkeling hypothesis is particularly relevant to KL<sub>4</sub> since all of the charged residues are lysines. However, while SP-B<sub>59–80</sub> also has periodically spaced charged residues, the aspartic acid sidechains are not as long and their interaction with the phosphate moieties would not be as favorable. We are currently pursuing EPR and NMR studies which will allow refinement of our model. This will yield more accurate information in terms of orientation, depth penetration and structure of the peptide at particular amino acids. However, from the current study we can conclude that SP-B<sub>59–80</sub> interactions with lung surfactant lipids and subsequent changes in lipid dynamics are dependent on the level of fatty acid saturation. While SP-B<sub>59–80</sub> and KL<sub>4</sub> behave similarly in this regard, key differences are also evident. In particular, SP-B<sub>59–80</sub> is able to alter bilayer structure in POPC:POPG vesicles at a concentration as low as 2.5 mol% while KL<sub>4</sub> is able to facilitate phase separation in DPPC:POPG vesicles. The phase
separation characteristics of KL₄ may be of key relevance to the peptide’s success in the clinic and can most likely be attributed to the use of solely positively charged residues rather than a mixture of cationic and anionic residues. The properties of SP-B₁₅₉₋₈₀ in POPC:POPG vesicles may be due to the fact that polar and charged residues are spaced every four residues instead of every five residues leading to subtle changes in the structure and interaction of the peptide with the lipid lamellae relative to KL₄; alternatively electrostatic interactions may lead to the peptide being even more peripherally located in the lipid/water interface in comparison to KL₄.

Our CD spectra of SP-B₁₅₉₋₈₀ in 4:1 DPPC:POPG and 3:1 POPC:POPG lipid environments (Fig. 3-1) indicate the peptide is almost entirely helical in both lipid preparations, however the peptide may not form a typical α-helix with similar CD spectra seen for peptides constrained to form π-helices.⁴⁶⁰

³¹P NMR was used to examine lipid headgroup polymorphism and dynamics in 4:1 DPPC:POPG and 3:1 POPC:POPG lipid systems in the presence of SP-B₁₅₉₋₈₀ peptide at different peptide:lipid (P/L) molar ratios (Fig.3-2, 3-3). The difference in PG and PC resonances in ³¹P spectra is due to the difference in their preferred headgroup orientations relative to the membrane normal. In both lipid mixtures, with increasing levels of peptide, the PC Chemical Shift Anisotropy (CSA) is invariant, but the PG CSA decreases towards a value characteristic for the neat POPG lipid. Electrostatic interactions of POPC with POPG reorient anionic PG headgroups which leads to an increase in PG CSA, but addition of positively charged peptide disrupts this interaction and causes the PG CSAs to move back toward neat POPG values. At the highest P/L ratio the POPC/POPG mixtures transition to a nonlamellar phase.
Deuterium NMR spectra of 4:1 DPPC:POPG and 3:1 POPC:POPG lipid samples with varying levels of SP-B_{59-80} allowed us to infer the depth of peptide penetration into the bilayer by looking at lipid acyl chain ordering and using DePaking of the spectra to assign each C–D bond in the acyl chain and determine the time-averaged order parameters $<S_{CD}>$. Order parameters at particular methylene positions in the palmitoyl acyl chains of PC lipids depend on concentration of SP-B_{59-80} (Fig.3-9, 3-10). Peptide addition to POPC:POPG sample decreases ordering of the deuterated POPC acyl chains. This suggests the peptide binds to the interface, increasing the lateral spacing between the lipids and allowing more motion in the acyl chains. In contrast, addition of peptide to DPPC:POPG increases the ordering of the deuterated DPPC acyl chains, particularly toward the middle of the bilayer. This indicates deep peptide penetration which restricts motion of acyl chains. For the second lipid, POPG, in each lipid mixture the changes in the order parameters are affected by the changes in the average orientation of lipid headgroup, but the overall changes in acyl chain order parameters follow those of the PC lipids.

Our results indicate SP-B_{59-80} penetrates deeply into DPPC:POPG bilayers and binds more peripherally to POPC:POPG bilayers (Fig. 3-12). Similar behavior has been observed for KL_{4}, a peptide modeled after SPB_{59-80}. Our laboratory has determined that this is because KL_{4} forms unusual helical structures which are lipid dependent. KL_{4} changes its helical pitch and penetration as a function of lipid composition. The deep penetration of SP-B_{59-80} within DPPC:POPG lipid bilayers would require a similar change in helical structure. The structural plasticity and variable penetration depth of KL_{4} and potentially SP-B_{59-80} can affect the stability and composition of lung surfactant
lipid structures and may provide a mechanism for lipid trafficking from lamellar bodies and tubular myelin to the air-water interface in a manner that selects for DPPC.

![Figure 3-1](image)

Figure 3-1. CD spectra at 45 °C of SP-B<sub>59–80</sub> at a P:L molar ratio of 1:100 in 4:1 DPPC:POPG (-----) and 3:1 POPC:POPG (- - - - - -). The final peptide concentration was 40 μM. Spectra taken with 1 nm step size and averaging of 40–50 scans.
Figure 3-2. Phosphorous NMR spectra of 4:1 DPPC-d82:POPG MLVs with SP-B59-80 at the indicated P:L molar ratios. (Top) Static NMR spectra and (Bottom) DePaked spectra. Spectra were collected with 200 µL of sample at 44 °C, 2048 scans.
Figure 3-3. Phosphorous NMR spectra of 3:1 POPC-d$_{31}$:POPG MLVs with SP-B$_{59-80}$ at the indicated P:L molar ratios. (Top) Static NMR spectra and (Bottom) DePaked spectra. Spectra were collected with 200 µL of sample at 44 °C, 2048 scans.
Figure 3-4. DSC scans for 4:1 DPPC-d$_{62}$:POPG LUVs with SP-B$_{59-80}$ at the indicated P:L molar ratios.

Figure 3-5. Deuterium NMR spectra as a function of temperature for (a) 4:1 DPPC-d$_{62}$:POPG MLVs, (b) 4:1 DPPC-d$_{62}$:POPG MLVs with SP-B$_{59-80}$ at a P:L molar ratio of 1:100, (c) 4:1 DPPC:POPG-d$_{31}$ MLVs, and (d) 4:1 DPPC:POPG-d$_{31}$ MLVs with SP-B$_{59-80}$ at a P:L molar ratio of 1:100. Spectra were collected with 200 µL of sample and 2048 scans.
Figure 3-6. First moment, $M_1$, as a function of temperature, for the $^2$H NMR spectra plotted in Fig.3-5; data for 4:1 DPPC-d$_{62}$:POPG MLVs (open squares), 4:1 DPPC-d$_{62}$:POPG MLVs with SP-B$'_{59-80}$ at a P:L molar ratio of 1:100 (closed squares), 4:1 DPPC:POPG-d$_{31}$ MLVs (open triangles), and 4:1 DPPC:POPG-d$_{31}$ MLVs with SP-B$'_{59-80}$ at a P:L molar ratio of 1:100 (closed triangles) show the variations in the melting temperatures of the lipids on addition of peptide.
Figure 3-7. Deuterium NMR spectra of 4:1 DPPC-d$_{62}$:POPG MLVs with SP-B$_{59-80}$ at the indicated P:L molar ratios. Spectra were taken at 44 °C. (Top) Static NMR spectra; (Bottom) dePaked spectra.
Figure 3-8. Deuterium NMR spectra of 3:1 POPC-d$_{31}$:POPG MLVs with SP-B$_{59-80}$ at the indicated P:L molar ratios. Spectra were taken at 44 °C. (Top) Static NMR spectra and (Bottom) dePaked spectra.
Figure 3-9. Order parameter profiles for the \( sn-1 \) chains of DPPC-\( d_{62} \) (top) and POPG-\( d_{31} \) (bottom) in 4:1 DPPC:POPG MLVs at 44 °C with SP-\( B_{59-80} \) at the indicated P:L molar ratios.
Figure 3-10. Order parameter profiles for the sn-1 chains of POPC-$d_{31}$ (top) and POPG-$d_{31}$ (bottom) in 3:1 POPC:POPG MLVs at 44 °C with SP-B$_{59-80}$ at the indicated P:L molar ratios.
Figure 3-11. Changes in d₃₁-palmitoyl acid chain order parameters in 4:1 DPPC:POPG MLVs (top) and 3:1 POPC:POPG MLVs (bottom) on addition of SP-B₅₉₋₈₀ at a P:L molar ratio of 1:100.
Figure 3-12. Models of SP-B<sub>59--80</sub> interactions with the two lipid environments studied assuming a helical peptide conformation. (Left) Based on <sup>2</sup>H NMR data, SP-B<sub>59--80</sub> deeply penetrates into 4:1 DPPC:POPG lipid bilayers. (Right) In contrast, in 3:1 POPC:POPG MLVs lipids, a more peripheral interaction of SP-B<sub>59--80</sub> with the lipid headgroup region is proposed. A transmembrane orientation of the peptide is unlikely as it would place four polar residues in the hydrophobic interior.
The C-terminal fragment of SP-B, specifically residues 59–80 (SP-B$_{59-80}$), can efficiently transfer phospholipids, mainly DPPC, to air-water interfaces to lower interfacial surface tension as it was demonstrated in in vivo studies where surfactant-deficient fetal rabbits partially restored lung function after application of synthetic SP-B$_{59-80}$.\(^{(17)}\) As shown in the previous chapter, SP-B$_{59-80}$ partitions differently within lipid bilayers with different levels of acyl chain saturation. This differential partitioning within lipid bilayers provides a potential mechanism for transferring of DPPC lipid by SP-B to the air-water interface. However, since direct structural measurements in varying lipid contexts have not been documented, the exact mechanism of such lipid transfer is not known. The C-terminal peptide is believed to form an amphipathic helix involved in lipid organization. Further enhancement of artificial lung surfactant therapies requires knowledge of the high resolution structure of SP-B$_{59-80}$ in the lipid environment of the lung and an understanding of the mechanism of interaction between SP-B$_{59-80}$ and specific lipids.

Since high resolution structural studies in bulky lipid vesicles by solid state NMR are quite expensive, technically challenging and require a lot of time, preliminary measurements were made in various smaller lipid bilayer mimetics (detergent micelles, bicelles, amphipols and nanodiscs) using solution NMR.\(^{(163)}\) A recent solution NMR study of residues 63–78 of the C-terminus of SP-B (SP-B$_{63-78}$) reconstituted in either SDS micelles or the organic solvent HFIP found the first five residues to be unstructured and established that the rest of the sequence formed a helix in both SDS micelles and organic solvent.\(^{(60)}\) However, measurements of surface tension reduction upon mixing of
DPPC/POPG lipids with peptides of different lengths from the C-terminus of SP-B clearly demonstrate that the longer SP-B_{59-80} peptide can further reduce surface tension to 6.7 dynes/cm (vs 10.6 dynes/cm for a shorter SP-B_{64-80} peptide) at a peptide concentration of 1 mol% relative to the lipids.\(^{(17, 55)}\) Moreover, micelles of denaturing detergent SDS are not a good mimic for DPPC lipid bilayer.

In this chapter, solution NMR is used to determine the structure of the C-terminus of human SP-B, residues 59-80 (SP-B\(_{59-80}\)), in methanol and in micelles of the zwitterionic mild detergent \textit{dodecylphosphocholine} (DPC). Structural studies of SP-B_{59-80} in the more physiologically relevant DPC micelles, which mimic the major surfactant lipid DPPC, could provide a better understanding of the mechanism by which SP-B may specifically transfer DPPC in tubular myelin bilayers and to the phospholipid monolayer at the air-water interface.

\textbf{Materials and Methods}

\textbf{Synthesis of SP-B_{59-80}}

SP-B\(_{59–80}\) (DTLLGRMLPQLVCRLVLRCSMD) was synthesized via solid-phase peptide synthesis on a Wang resin (ABI 430, ICBR, UF), cleaved with King's reagent\(^{(164)}\) and purified via RP-HPLC with a C18 Vydac column and a water/acetonitrile gradient (0.3% TFA). Purity of the product was verified by mass spectrometry (m/z = 2533). Peptide was dissolved in methanol to a stock concentration of \(~1\text{mM, and the final concentration was determined by amino acid analysis (Molecular Structure Facility, UC Davis). This peptide was used for structural measurements in methanol.}

\textbf{Heterologous Expression of }^{15}\text{N-labeled SP-B}_{59-80} \text{Double Mutant (SP-B'}_{59-80}\text{)}

\(^{15}\text{N-labeled SP-B'}_{59–80}\) was expressed using a pET31 construct (EMD Biosciences, Inc., Gibbstown, NJ) incorporating a codon-optimized synthetic gene for SP-B\(_{59–80}\).
(DNA2.0, Menlo Park, CA) in BL21(DE3)pLysS cells (Novagen, EMD Biosciences, Inc., Gibbstown, NJ) in minimal media\(^{(165)}\) containing \(^{15}\)N-ammonium chloride as a nitrogen source with subsequent purification and cleavage using established protocols.\(^{(120)}\) The expressed sequence of SP-B'\(^{59-80}\) was modified to incorporate isoleucines in lieu of methionines (DTLLGRILPQLVCRLVLRCSID) for compatibility with a cyanogen bromide cleavage reaction. SP-B'\(^{59-80}\) peptide contained an additional homoserine lactone residue at its C-terminus, resulting from cyanogen bromide cleavage of the peptide from the fusion protein. Following HPLC purification of the final product, the peptide was verified by mass spectrometry (m/z = 2610). Peptide was dissolved in methanol to a stock concentration of \(\approx 1\) mM, and the final concentration was determined by amino acid analysis (Molecular Structure Facility, UC Davis). This peptide was used for structural measurements in DPC micelles, which required \(^{15}\)N-editing to resolve the proton resonances.

**Circular Dichroism (CD) experiments**

DPC-d\(_{38}\) detergent was dissolved in chloroform at 50mg/ml. The detergent solution in chloroform was mixed with a methanol solution of SP-B\(_{59-80}\) to obtain a final peptide: detergent molar ratio of 1:200. The sample was dried under nitrogen at 315–323 K in a water bath, suspended in cyclohexane, flash-frozen, and lyophilized overnight to remove residual solvent. Dried sample was solubilized in 10mM sodium phosphate buffer at pH 7.4 to achieve a concentration of 40\(\mu\)M SP-B\(_{59-80}\) with 8mM detergent. A sample of SP-B\(_{59-80}\) in methanol was also prepared at a concentration of 40\(\mu\)M. SP-B\(_{59-80}\) secondary structure was monitored by far UV CD spectroscopy. CD experiments were performed on an Aviv Model 215 (Lakewood, NJ) at 303 K or 318K over the range of
200–260 nm with a 1nm step size and averaging of 40–50 scans. Baseline scans were obtained using the appropriate peptide-free controls.

**NMR samples preparation**

Methanol stock solutions of 0.7 mM SPB<sub>59-80</sub> and 0.5 mM <sup>15</sup>N-labeled SP-B'<sub>59-80</sub> were incubated for 1-2 hours with 15-fold molar excess TCEP to fully reduce disulfide bonds in the peptide. To prepare a sample of 1.2 mM SPB<sub>59-80</sub>, peptide stock solution was dried under nitrogen gas and redissolved in methanol-d<sub>3</sub> (99.8% - Sigma Aldrich, Inc., St. Louis, MO) before transfer to a 5mm NMR tube (Norell, Inc., Landisville, NJ) for NMR measurements. This sample was used to collect 2D TOCSY and NOESY spectra. To prepare a sample of 1.2 mM <sup>15</sup>N-labeled SP-B'<sub>59-80</sub> in 80 mM DPC-d<sub>38</sub>, peptide stock solution was mixed with a chloroform solution of 50mg/ml DPC-d<sub>38</sub> (98% - Cambridge Isotope Laboratories, Andover, MA) and solvents were evaporated using nitrogen. The resulting film was suspended in cyclohexane, flash-frozen and lyophilized overnight to remove residual solvent. The dry sample was solubilized in 110 µL of 50 mM aqueous sodium phosphate buffer (pH 4.5) with 10% D<sub>2</sub>O (99.96%— Cambridge Isotope Laboratories, Andover, MA) containing 1 mM 3-(Trimethylsilyl)propionic-2,2,3,3-d<sub>4</sub> acid sodium salt (TSP) (98% - Sigma Aldrich, Inc., St. Louis, MO) as an internal reference; this solution was transferred to a 2.5 mm NMR tube (Norell, Inc., Landisville, NJ). The <sup>15</sup>N labeled sample was used to collect 2D <sup>15</sup>N-HSQC, TOCSY and NOESY spectra and 3D <sup>15</sup>N-edited NOESY-HSQC and TOCSY-HSQC spectra.

**NMR experiments**

All NMR spectra were recorded on a Bruker Avance II 600 MHz spectrometer equipped with a 5mm Z-gradient TXI cryo-probe and Topspin 2.1 software. In all experiments WATERGATE<sup>(166, 167)</sup> and States-TPP<sup>(168)</sup> were used for solvent...
suppression and quadrature detection in the t1 dimension, respectively. Sequence assignment for SPB_{59-80} in methanol-d3 solution was achieved using TOCSY (MLEV-17;{(169)} 70 ms mixing time) and NOESY (300 ms mixing time) spectra acquired at 300 K. The spectra were collected with 2048 · 512 data points. The $^1$H spectral width in both dimensions was 11.0 ppm with chemical shifts referenced to TMS using residual CHD$_2$OH signal at 3.31 ppm as a secondary reference.

For multidimensional heteronuclear NMR studies of $^{15}$N-labeled SP-B'_{59–80} in DPC micelles, a set of well-resolved spectra was recorded at 312 K, including $^{15}$N-HSQC (2048 · 128 data points and 13.0 · 80.0 ppm width) and 3D $^{15}$N-edited NOESY-HSQC and TOCSY-HSQC spectra. 3D NOESY-HSQC spectra were collected with the following parameters for each dimension listed in the order of acquisition: $^1$H acquisition 2048 points, 16 scans, 14.0 ppm; $^{15}$N dimension 40 points, 30.0 ppm; $^1$H indirect 256 points, 11.7 ppm. Three different mixing times, 75, 150, and 300 ms, were used for the NOESY-HSQC and spin diffusion was evaluated via NOE buildup curves for several peaks. As the result of finding balance between maximum number of NOEs important for sequential assignment and structure determination and minimum spin diffusion effects we selected 300ms spectrum for sequential assignment and determination of distance constraints. The 3D TOCSY-HSQC was collected using DIPSI-2{(170)} with a 60ms mixing time and the following parameters: $^1$H acquisition 2048 points, 8 scans, 12.0 ppm; $^{15}$N dimension 40 points, 37.0 ppm; $^1$H indirect 128 points, 12.0 ppm. 2D TOCSY (DIPSI-2;{(170)} mixing time 60 ms) and NOESY (75, 150, 200 and 300 ms mixing times) spectra were acquired for this sample under the same conditions to help with sidechain assignments in 3D spectra and to obtain additional distance restraints.
between side chains. All chemical shifts in the DPC-containing sample were referenced using TSP (3-(Trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt) as an internal standard.

**NMR data analysis and structure calculation**

NMR data were processed using NMRPipe\(^{(171)}\) and analyzed with SPARKY\(^{(172)}\). For the SPB\(_{59-80}\) structure in methanol, the major restraints were derived from the 300 msec mixing time 2D NOESY spectrum; cross peak intensities were converted to distance restraints of 2.0–3.5, 2.0–4.5, and 2.0–6.0 Å, corresponding to strong, medium, and weak NOE cross peaks, respectively. For the DPC-containing SP-B\(_{59-80}\) sample cross peak intensities were categorized as strong, medium, weak and very weak to set upper bound distances of 3.8, 4.5, 5.0 and 6.0 Å, respectively; the lower bound distance was set to 2.0 Å. For both the DPC and methanol structures an extended covalent structure was used to generate starting coordinates, and an ensemble of 100 minimized structures was calculated using a standard simulated annealing protocol in Xplor-NIH version 2.32.\(^{(111, 112, 173)}\) The structure was further refined via several rounds of minimization using standard simulated annealing protocol for refinement in Xplor-NIH version 2.32 and minor adjustments to the constraints have been made. Ten structures with minimum total energy were selected out of 3000 structures to meet the following criteria: no NOE-derived distance violations of >0.50 Å, improper angle violations <5°, rmsds from ideal bonds <0.01 Å, and rmsds from ideal bond angles <5°. The structures were viewed and analyzed using VMD viewer in VMD-XPLOR\(^{(173)}\) and quality of structures was evaluated with PROCHECK V 3.4.4\(^{(114, 174)}\) and MOLPROBITY\(^{(175)}\). The structure, \(^1\)H chemical shifts and \(^3\)J\(_{HN,αH}\) coupling constants of SPB\(_{59-80}\) in methanol were deposited in the Protein Data Bank (PDB entry 2m0h) and...
BioMagResBank (BMRB entry 18809). The structure, \(^1\)H and \(^{15}\)N chemical shifts of SP-B\(_{59-80}\) bound to DPC micelles were deposited in the Protein Data Bank (PDB entry 2m1t) and BioMagResBank (BMRB entry 18879).

**Results**

**Secondary structure in methanol and in DPC micelles**

Far-UV CD spectra of SP-B\(_{59-80}\) both in methanol and in DPC micelles are characterized by two negative maxima at 208 nm and at 222 nm as is typically observed for helical peptides (Fig. 4-1). In CD spectra of conventional \(\alpha\)-helices the negative maxima at 208 and 222 nm typically have equal intensity,\(^{(176)}\) but for both spectra in Figure 1 the negative maximum at 208 nm is more intense than the one observed at 222 nm. Previous CD studies of SP-B\(_{59-80}\)\(^{(68)}\) and KL\(_4\)\(^{(132)}\) in lipid bilayers are similar to the spectra in Figure 4-1. These observations clearly demonstrate that the SP-B\(_{59-80}\) is helical in different environments, but the exact helical structure it adopts is determined by the context of the surrounding environment.

**NMR Resonance Assignments for SP-B\(_{59-80}\) in methanol**

The structure of SP-B\(_{59-80}\) in methanol was analyzed by \(^1\)H NMR at 300 K. In the presence of TCEP, signals were sharper and their number decreased, indicating a loss of oligomer formation. Complete sequence-specific proton resonance assignments of SP-B\(_{59-80}\) in methanol were achieved using \(^1\)H-\(^1\)H TOCSY and \(^1\)H-\(^1\)H NOESY spectra.\(^{(104)}\) The NH-\(\alpha\)H region of the NOESY spectrum with assignments is shown in Figure 4-2. A complete list of assignments is deposited in the BMRB and given in Table 4-1.

For SP-B\(_{59-80}\) in methanol, continuous series of \(\alpha\)-H-highfield shifts exceeding by 0.2-0.6 ppm the chemical shifts predicted for random coil sequences were measured for
residues 61-64 and 68-78; the chemical shift index (CSI) is plotted in Figure 4-3 (Bottom) and is suggestive of helix formation for these regions. (177, 178)

The $^3J_{HN, \alpha H}$ coupling constants of 4.8 Hz observed for residues 61-62 and 69-76 are also consistent with an $\alpha$-helical structure (179) (Table 4-2). $^3J_{HN, \alpha H}$ couplings could not be measured for residues 63, 64, 66, 68, 71, 72, 77 and 78 due to overlap. The observed NN($i,i+1$) and $\alpha N(i,i+3)$ NOE patterns (Fig. 4-3) also suggest that residues 61-64 and 68-78 are helical. A nearly complete pattern of $\alpha N(i,i+3)$ connectivities was seen for residues 61-64 and 66-79. In a well-structured helix, the $\alpha N(i,i+3)$ should be much stronger than either the $\alpha N(i,i+2)$ and/or the $\alpha N(i,i+4)$ NOEs (180) and Figure 4-3 shows that this is clearly the case for the SP-B$_{59-80}$ in methanol. However, the intensities of $\alpha N(i,i+3)$ NOEs are approximately the same either for $3_{10}$-helix ($\alpha N(i,i+3) = 3.3$ Å) or an $\alpha$-helix ($\alpha N(i,i+3) = 3.4$ Å). In contrast, $\alpha \beta(i,i+3)$ NOEs should be prominent for an $\alpha$-helix and weak for a $3_{10}$-helix. (104) Also $\alpha N(i,i+2)$ NOEs are unique for $3_{10}$-helix and $\alpha N(i,i+4)$ connectivities are a unique feature of an $\alpha$-helix. (104) Figure 4-3 shows that at the N- and C-terminal ends both $\alpha N(i,i+2)$ and $\alpha N(i,i+4)$ NOEs are present, consequently the helix type of the ends was determined by the relative intensity of the two types of NOEs for each residue. $\alpha N(i,i+4)$ NOEs become quite weak toward the C-terminus suggesting the C-terminal residues 78-80 form a $3_{10}$-helix due to the existence of two $\alpha N(i,i+2)$ NOEs in this region.

Comparison of the $\alpha N(i,i+2)$ and $\alpha N(i,i+4)$ NOEs for N-terminal residues 61-64 suggests a mixture of $3_{10}$ and $\alpha$-helix, however, this interpretation is somewhat ambiguous since the $\alpha N(i,i+2)$ connectivity can also arise from random coil and/or
nascent helix conformations (each of which contains isolated turns giving transient $i \rightarrow i + 3$ hydrogen bonds) without persistent $3_{10}$-helix.\textsuperscript{(181)} Many crystal structures of protein and peptide helices consistently show helix termini are subject to deviations from regular $\alpha$-helix and C-terminal helices often end with a turn of $3_{10}$-helix.\textsuperscript{(182-184)} Numerous NMR studies on helical peptides also report NOEs consistent with mixed populations of $3_{10}$-helix and $\alpha$-helix.\textsuperscript{(180, 185)}

Long-range inter-helical NOEs that position the two helices close in space are consistent with a helical hairpin structure and a few of the inter-helical NOEs observed are reported in Figure 4-3. A close inspection of the per residue NOE distribution for L62-R64 and Q68-R76 residues of SP-B$_{59-80}$ in methanol shows a higher number of NOE contacts than for the residues at the ends of the sequence. Also long range NOEs have been detected for pairs of residues G63-C71 and M65-C71, which form hydrophobic contact between the two helices (Fig. 4-4, A). These observations are consistent with a helical hairpin structure.

**Three-Dimensional Structure of SP-B$_{59-80}$ in Methanol**

The structure of SP-B$_{59-80}$ in methanol was calculated by restrained simulated annealing with subsequent restrained molecular dynamics and energy minimization using XPLOR-NIH.\textsuperscript{(173)} 459 Interproton distance restraints were used, 217 of which were intra-residual, 96 were sequential, 88 were medium-range, 4 were long range and 54 were ambiguous. Dihedral restraints (12) from $^3$J$_{HN,\alpha H}$ -couplings were also used in the calculations. The 10 structures with the lowest total energies and no NOE restraint violations greater than 0.29 Å were used for structural analysis (Table 4-3). Figure 4-5A shows the superposition of the ten structures with rmsd of the backbone atoms (N, Cα, C', O) of 0.68 Å for all residues and 0.28 Å for the more structured residues 61-78
(Table 4-3). A minimized average structure from the ensemble with well-formed secondary structure elements (Fig. 4-5B) can be characterized as a helical L-shaped hairpin motif, where two helices are separated by a turn, which is induced by a trans conformation of P67 (Figure 4-5C side view obtained by 90 rotation of the structure in 4-4B). The N-terminal helix (residues 61-65) is a right-handed 3_{10}-helix containing the achiral G63 residue in the middle, the C-terminal helix is a well-defined right-handed α-helix extending from L69 to C77 ending with a mix of random coil, 3_{10}- and α-character for the last and the most disordered turn of the helix (residues 78 to 80). In the ensemble (Fig. 4-5A) the ends of the peptide are more mobile and slightly less defined than the rest of the peptide. This may be a result of not including hydrogen bonds constraints in the calculation. Residues L66 and P67 form a flexible loop between the two helices. Conformational averaging is more likely for NMR structures of small peptides, especially in the end regions and regions containing aromatics, resulting in a few violations of nonbonded contacts and phi, psi torsion angles.

As can be seen in Figure 4-5 (B and C), SP-B_{59-80} is amphipathic in methanol. The hydrophobic amino acid side-chains shown in green reside along the bottom side of the structure, thus providing an extended, uninterrupted hydrophobic surface. Positively charged residues (blue), negatively charged residues (red) and a polar Q68 residue (grey in the middle of peptide) are all located on the other side of the structure creating a hydrophilic surface. Moreover, the structure shown in Figure 4-5 suggests that hydrophobic interaction between N- and C-terminal helices G63-C71 plays an essential role in stabilizing the bent helical hairpin structure of SP-B_{59-80} peptide in methanol. This distribution of lipophilic/hydrophilic residues may also give rise to the oligomerization
tendency of SP-B\textsubscript{59–80} peptide in methanol, since the lipophilic residues would be buried in the oligomeric form but hydrophilic residues remain exposed to the solvent, stabilizing the oligomeric structures in methanol leading to formation of intermolecular disulfide bonds in the absence of a reducing agent for 1mM methanolic solutions of SP-B\textsubscript{59–80}.

**NMR Resonance Assignment for SP-B'\textsubscript{59–80} in DPC Micelles**

Preliminary 2D \textsuperscript{1}H-\textsuperscript{1}H NOESY and TOCSY spectra (not shown) of unlabeled SP-B\textsubscript{59–80} peptide in DPC micelles had cross-peak line widths that were significantly broader than in the methanol sample. This prevented unambiguous assignment of a greater portion of NOE cross peaks. However, a sample of universally \textsuperscript{15}N-labeled SP-B'\textsubscript{59–80} peptide in DPC micelles showed significant dispersion of signals in the \textsuperscript{15}N dimension of an \textsuperscript{1}H-\textsuperscript{15}N HSQC spectrum (Fig. 4-6) and adequate resolution of almost all \textsuperscript{1}H,\textsuperscript{15}N-cross-peaks. The good resolution in the \textsuperscript{15}N dimension allowed recording of \textsuperscript{15}N-edited NOESY-HSQC and TOCSY-HSQC spectra with \textsuperscript{1}H-\textsuperscript{1}H NOESY and TOCSY traces for each residue resolved in separate \textsuperscript{15}N-planes.

Almost complete backbone assignment of SP-B'\textsubscript{59–80} in DPC micelles was achieved from combined analyses of \textsuperscript{1}H-\textsuperscript{1}H NOESY and TOCSY planes of the corresponding three-dimensional \textsuperscript{15}N-edited NOESY-HSQC and TOCSY-HSQC spectra and identification of the sequential cross signals in a 2D \textsuperscript{1}H-\textsuperscript{1}H NOESY connecting \(\alpha\)H, \(\beta\)H and HN of residue \(i\) with HN of residue \(i+1\) (Fig. 4-7). Sequential NOEs in the \textsuperscript{1}H,\textsuperscript{15}N NOESY-HSQC spectra involving side chain atoms were used to confirm the backbone assignments. Due to the helical nature of the peptide, strong sequential NH-NH cross peaks were observed in the \textsuperscript{1}H,\textsuperscript{15}N NOESY-HSQC spectra, facilitating assignment by extended amid-amide walks. Assignment of aliphatic side-chain protons
for short side chain residues was achieved using only the $^1$H,$^{15}$N TOCSY-HSQC spectrum. However, assignment of aliphatic side chain protons for longer side chain residues (L, I, V, R) required comparison of $^1$H,$^{15}$N TOCSY-HSQC and 2D $^1$H-$^1$H TOCSY spectra since the self-relaxation of the nuclear spins during the mixing (spin-lock) period of the TOCSY pulse sequence did not allow efficient magnetization transfer from the HN to the terminal side chain protons of these residues.

Additional distance restraints between side chains were obtained from the 2D NOESY spectra with 200 and 300 ms mixing times. The chemical shifts of $^{15}$N-labeled SP-B$^{59-80}$ peptide in DPC micelles at 315 K, referenced to TSP, are listed in Table 4-4.

As expected, the analyses of NOE patterns from the $^1$H,$^{15}$N-NOESY-HSQC spectrum and the sidechain region of the 2D $^1$H-$^1$H NOESY spectrum of SP-B$^{59-80}$ peptide in DPC micelles revealed the presence of diagnostic medium-range NOEs connecting the $\alpha$H of residue i with HN protons of i+2, i+3, i+4 residues and i+3 $\beta$H residue (Fig. 4-8). These NOE patterns are similar to NOE patterns obtained for the methanol sample. The L73 and L75 resonances were not resolved (Fig. 4-6) since these residues are located in the center of 7-residue palindromic CRLVLRC sequence and have very similar $^{15}$N and NH chemical shifts. Spectra of a shorter analog, SP-B$^{63-78}$, taken by Booth and coworkers$^{(60)}$ also show that $^{15}$N and NH chemical shifts for L73 and L75 are quite similar to each other in SDS micelles. $\alpha$H-$\beta$H (i,i + 3) connectivities were observed in residues L61-L65 and L66-R76 on a 2D $^1$H-$^1$H NOESY spectrum. This pattern of NOEs suggests that in DPC micelles SP-B$^{59-80}$ acquires a helical conformation similar to the SP-B$^{59-80}$ conformation in methanol, where P67 induces a kink separating two helices. This is in agreement with the CD data (Fig. 4-1).
noteworthy that several long-range NOE interactions between the two helices, primarily involving side chains of charged residues D59, R64, Q68, R72, R76, D80 were observed defining the helical hairpin structure. SP-B'_{59-80} in DPC micelles additionally had αH-NH (i, i+3) and αH-NH (i, i+4) NOEs from T60 to D80. It is noteworthy that αH-NH (i, i+2) NOEs become stronger for the N- and C- terminal residues of the sequence relative to the methanol sample. No $^3J_{\text{HN,\alpha H}}$ proton-proton coupling constants were used in the analysis of SP-B'_{59-80} in DPC because their observation was not possible due to relatively broad line widths in the $^1$H NMR spectra.

**Three-Dimensional Structure of SP-B'_{59-80} in DPC Micelle**

An ensemble of conformations for SP-B'_{59-80} in complex with DPC micelles was determined based on 433 distance constraints including 108 intra-residual, 88 sequential, 142 medium range, 15 long range and 80 ambiguous NOEs (Table 4-3). The DPC-bound structure of SP-B'_{59-80} peptide is well defined. Figure 4-9A shows the superposition of all backbone atoms (N, C$_\alpha$, C',O) for the 6 lowest energy structures of SP-B'_{59-80} in DPC micelles with rmsd of the backbone atoms (N, C$_\alpha$, C', O) of 0.67 Å for all residues and 0.48 Å for the more structured residues 61-78 (Table 4-3). Residues L61–S78 of SP-B'_{59-80} demonstrate the lowest rmsd values for the backbone and side chain atoms. In contrast, residues D59-T60 and I79-D80 have relatively higher rmsd both for the backbone and side-chain atoms. Fig. 4-9B shows a selected representative structure of SP-B'_{59-80} bound to DPC. This structure has a short right-handed helix at the N-terminus (residues D59–R64), showing partial helical character common for nascent helices, and a well-formed right-handed α-helix at the C-terminus (residues Q68–D80); the two helices are connected by a short loop consisting of residues I65, L66, and P67.
The N and C terminal helices of SP-B'59-80 assume a defined orientation relative to each other giving rise to a helical L-shaped hairpin structure (Fig. 4-9B,C).

The L-shaped fold of SP-B'59-80 in DPC micelles is predominantly amphipathic and possesses a continuous hydrophobic face maintained by residues L61, L62, G63, I65, L66, P67, L69, V70, C71, L73, V74, L75, C77 and I79; residues D59, T60, R64, Q68, R72, R76 and D80 are positioned on the hydrophilic face of the peptide (Fig. 4-9B,C). The SP-B'59-80 structure in DPC micelles has increased long range side chain/side chain packing interactions between the charged sidechains of the N- and C-termini helices relative to SP-B59-80 in methanol (Fig. 4-4, B) likely due to the presence of the amphipathic micelles (Figure 4-9 and 4-10). These side chain interactions between D59, T60, R64, Q68, R72, R76 and D80 rotate the two helices so that arginine side chains face each other in DPC-bound structure (Fig. 4-9C) rather than tilt outward from each other like in methanol (Fig. 4-5C). P67 forms the turn of the loop. In other words, side chains of charged residues in SP-B'59-80 appear to be involved in a strong interaction that is locking the two helices into a compact L-shape conformation in the DPC micelle environment (Fig. 4-10).

The views of the DPC-bound structure of SP-B'59-80 (Fig. 4-9 B and C) and structure of SP-B59-80 in methanol (Fig. 4-5 B and C) have been selected to emphasize their amphipathic nature and overall similarity of structures. However, there are several striking differences between the two structures that result from difference in polarity of polar methanol and non-polar hydrophobic core of DPC micelle. In both structures the C-terminal helix (residues 68-80) is a right-handed amphipathic α-helix, whereas the short N-terminal helix (residues 61-64) forms a nascent right-handed 3_{10} helix in
methanol and adopts a nascent right-handed α-helix in DPC micelles. Of particular interest is P67 positioned in the flexible hinge, which breaks the helical conformation and allows rotation of N- and C- terminal helices with respect to each other. Due to flexibility of its proline-induced loop the two helices are either positioned with their hydrophilic sides outwards facing the polar methanol solvent as in methanol structure (Fig. 4-5C) or inwards hiding from the non-polar environment of the DPC-micelle interior (Fig. 4-9C). It is important to point out that in refinement calculations the structure in methanol was obtained if its NMR constraints were applied to a starting structure based on the DPC data and vice versa. This indicates that each set of constraints is unique, complete and sufficient to drive folding into the specific conformer.

Similar changes in orientation of amphipathic helices with environment have been shown for phospholamban,\(^{(186)}\) which adopts an “L-shaped” helix-turn-helix conformation in organic solvents\(^{(187, 188)}\) and in DPC micelles.\(^{(186)}\) The two charged arginine residues (R9 and R13) as well as polar phosphorylation sites S16 and T17 are oriented toward the inside facing domain II in the structures obtained in organic solvents thus hiding from the hydrophobic environment, whereas in DPC micelles they are oriented outward facing the bulk solvent placing the charged side chains toward the bulk solvent and the hydrophobic residues toward the micelle interior \(^{(186)}\). The major difference between phospholamban and SP-B\(_{59-80}\) is the balance between hydrophobic and hydrophilic amino acids in their structures with SP-B\(_{59-80}\) peptide having larger hydrophobic face and smaller hydrophilic face than phospholamban. Consequently, due to this major shift in hydrophobicity SP-B\(_{59-80}\) peptide partitions deeper into DPC micelle hydrophobic core and places charged amino acid side chains closer to each other to
shield them from hydrophobic environment of the micellar core and still allow for their efficient interaction with the DPC headgroups. SP-B$_{59-80}$ also partitions deeply into 4:1 DPPC:POPG lipid bilayer.(68, 69)

**Discussion**

A previous solution NMR study of SP-B$_{63-78}$ reconstituted in either SDS micelles or the organic solvent HFIP found that SP-B$_{63-78}$ adopts the same structure in both environments with the first five residues being unstructured and the rest of the sequence forming an amphipathic α-helix.(60) Our structural studies of SP-B$_{59-80}$ with a longer and functionally more relevant sequence have revealed essential structural differences between SP-B$_{63-78}$ and SP-B$_{59-80}$ that may explain why SP-B$_{59-80}$ reduces surface tension in lipid monolayers more efficiently than peptides with the shorter sequence at comparable concentrations.(17, 55) Using methanol instead of HFIP as well as the milder DPC detergent instead of the denaturant SDS could also account for some of the observed structural differences between our work and previously published data for SP-B$_{63-78}$. DPC micelles most closely mimic the major surfactant lipid DPPC and thus likely provide a better understanding of the mechanism of interaction between SP-B and DPPC in pulmonary surfactant. The four extra residues at the N-terminus and two extra residues at the C-terminus of SP-B$_{59-80}$ compared to SP-B$_{63-78}$ include two negatively charged residues, D59 and D80, which provide not only a counter-balance for the three positive arginine residues but also stabilize the N-terminal helical structure which allowed us to observe the interhelical NOEs between the N- and C-terminal helices and to obtain a helical L-shaped hairpin structure. The helical hairpin structure also highlights the role of P67 residue in the loop between the two helices as a mobile hinge. This loop allows the peptide to change its topology with environment.
The helical hairpin is a known structure for many amphipathic peptides. For example, this fold is adopted by Pa4 pardaxin\(^{(189)}\) and MSI-594\(^{(190)}\) in LPS micelles and by Pa4 pardaxin in DPC micelles.\(^{(142)}\) In these helical hairpin structures the two helices interact through hydrophobic contacts. For example, in the structure of Pa4 in LPS micelles, F15 found in the proline-containing loop between the two helices plays an important role in bringing the nonpolar faces of the two helices together.\(^{(189)}\) In the structure of MSI-594 in LPS micelles, F5 interacts with other hydrophobic sidechains\(^{(190)}\) and its mutation to alanine causes loss of the strong interaction between the two helices; this opens up the helical hairpin structure of MSI-594 and decreases its antibacterial properties dramatically.\(^{(191)}\) It is interesting that Pa4 pardaxin also changes structure with environment: the DPC-bound structure of Pa4\(^{(142)}\) lacks the interhelical packing interactions observed in LPS micelles and thus adopts a loose helical hairpin with a sharp kink rather than a tight turn at the proline-containing loop separating the two helices.

The existence of the P67-based hinge between the N- and C-terminal helices in both the SP-B\(_{59-80}\) and the SP-B'\(_{59-80}\) structures also suggests a “crab claw” structure which can bind specifically to zwitterionic detergent or phospholipids (Fig. 4-11, 4-12), suggesting an interesting mechanism for how it might traffic DPPC and lower surface tension. Of particular interest, the center of the “crab claw” of SP-B\(_{59-80}\) contains the positively charged arginine side chains which electrostatically attract the negatively charged phosphate moiety in DPPC molecule while the tips of the “claw” are made of negatively charged aspartates, which may be important for initial binding to the positively charged choline headgroup of DPPC. The DPPC molecule could then remain
bound to the peptide while it changes conformation leading to movement of the DPPC molecule as well. Both structure ensembles of SP-B\textsubscript{59-80} in methanol and of SP-B\textsuperscript{'}59-80 in DPC have been extensively refined with XPLOR-NIH refine_nordc protocol while being continuously evaluated using good quality structure criteria via PROCHECK. The results of final evaluation shown in Figs. 4-13 to 4-20 demonstrate quite good quality of the final refined structures.

While the structure of SP-B has not been solved, several more soluble members of the saposin and saposin-like family of proteins (SAPLIPs) have been characterized. The saposin fold also assembles amphipathic helical hairpin structures which are important to lipid bilayer perturbation and lipid trafficking. Physiological functions of saposins and SAPLIPs can be classified by the level of lipid packing perturbation as (i) mere membrane binding accompanied by only local disordering of the lipid structures for membrane targeting; (ii) membrane perturbation, but without permeabilization, e.g. lipid extraction and presentation for enzymatic activity; and (iii) permeabilization as a killing principle – the essential activity of defense proteins.\textsuperscript{(192)} Only two amino acid residues with functional groups are strictly conserved between the saposin isoforms: a tyrosine residue and proline residue, both located in the loops connecting two amphipathic helices into a helical hairpin structure inside the saposin fold. Since not only saposins but also many SAPLIPs possess a proline residue at the similar position, it might be of structural benefit for lipid perturbation and extraction.\textsuperscript{(192)} The examples of SAPLIPs with antimicrobial, cytolytic activity that are able to penetrate membranes are NK-lysin\textsuperscript{(193)} and its human homolog granulysin.\textsuperscript{(194)} These proteins have several positively charged amino acids, especially arginine residues, on their surfaces. Their
positive charges presumably form the initial contact with the membrane followed by rotation of the two leaves of the protein towards each other, called “scissoring”, generating hydrophobic patches on the protein surface and enabling interaction with the hydrophobic core of the bilayer.\textsuperscript{(194)} In contrast, the saposins, which extract and present lipids from various membranes for enzymatic processing or antigen presentation, are all primarily negatively charged proteins without pronounced electropositive patches on their surface.\textsuperscript{(192)} For example, saposin C contains three clusters of negatively charged glutamate residues implicated in membrane binding. However, the negatively charged electrostatic surface of saposin C needs to be partially neutralized to trigger membrane binding.\textsuperscript{(195)}

Mature SP-B has been shown to destabilize phospholipid membranes directly, causing eventual fusion between liposome membranes and rapid leakage of their content.\textsuperscript{(196-198)} Destabilization of lipid bilayer packing could also be part of the mechanism by which SP-B promotes formation of a surface-associated multilayer reservoir of surfactant to stabilize the interfacial monolayer when it is laterally compressed during exhalation. It has been shown that upon compression of monolayer films containing SP-B, buckled multilayer patches form and remain associated with the monolayer. Upon expansion of the film at a lower surface pressure, the buckled structures are quickly reinserted into the interfacial film to re-form a flat monolayer.\textsuperscript{(199-206)} In contrast, compression of pure lipid films usually ends with an irreversible collapse.\textsuperscript{(207, 208)} These observations suggest that SP-B is a critical component for lipid trafficking from the aqueous subphase to the lipid monolayer.\textsuperscript{(209, 210)}
Based on our two observed structures for SPB_{59-80} we propose possible mechanism for interaction with DPPC and its trafficking between lipid lamellae by the C-terminus of SP-B (Fig. 4-12): the N- and C-terminal helices with the distinct charged and non-polar faces are connected by a flexible loop, which allows the two helices to “roll” around their axes to change orientation of their charged faces relative to each. If a DPPC molecule is bound to SP-B_{59-80} during this “rolling” motion, it could be “squeezed out” via concerted “rolling” motion of SP-B_{59-80} helices resembling the process of squeezing water out of clothing by two rollers in an antique laundry mangle machine. The net result would be destabilization of local bilayer structure or even translocation of the DPPC lipid.

In summary, we have determined the three-dimensional structure of SP-B'_{59-80} in DPC detergent micelles by three-dimensional heteronuclear NMR spectroscopy and the three-dimensional structure of SP-B_{59-80} in methanol using two-dimensional NMR techniques. The overall similarity of these structures and key differences in tertiary contacts provide novel insight into the potential mechanism of DPPC trafficking by the C-terminus of SP-B. Our structures and the proposed mechanism for DPPC trafficking may facilitate the engineering of novel peptide analogs of SP-B that will be more efficient in surface tension reduction in lungs.
Figure 4-1. CD spectra of SP-B<sub>59-80</sub> in methanol at 303K (red) and bound to DPC micelles in 10 mM sodium phosphate, pH 7.4, at 318K (green). Peptide concentration in each sample is 40μM. In the DPC-containing sample the peptide:detergent molar ratio is 1:200.
Figure 4-2. 300ms NOESY spectrum for SP-B<sub>59–80</sub> in methanol at 300 K. A: NH-αH region. The continuous line represents the assignment pathway with sequential assignments in black and other assignments in blue. By default the assignments have the form: #residue HA-#residue HN (if present, other protons are indicated by letter and number). B: sidechain region. NOEs of the proline system relevant to sequential assignment are shown.
Figure 4-2. Continued.
Figure 4-3. Summary of the NOE interactions and CSI measurements for SP-B\textsubscript{59--80} in methanol. A: Sequential i, i+3; $\alpha$H-NH; $\alpha$H-βH signals and i, i+4 NOE interactions for residues 61-64 and 67-77 are characteristic of an $\alpha$-helix. Black bars indicate unambiguous NOEs and gray bars indicate ambiguous NOEs. Strong, medium and weak classifications are indicated by the height of the bar for the sequential NOEs. Some NOEs are not marked because they could not be clearly identified due to resonance overlap. B: chemical shift index (CSI) plotted as the difference between random coil and observed $\alpha$H chemical shifts.\textsuperscript{(211)} Significantly negative CSI values are indicative of a helical conformation. The number of distance constraints per residue is presented in Fig. 4-12, A.
Figure 4-4. Summary of the NOE interactions for SP-B<sub>59–80</sub> in methanol (A) and for SP-B<sub>59–80</sub>' in DPC micelles (B). The number of distance constraints per residue is presented, with white bars representing intra-residue constraints, pale gray bars representing sequential constraints, dark gray bars representing medium-range (i - j < 5) constraints, and black bars representing long-range (i - j ≥ 5) constraints. An increased number of long range interactions is observed for residues essential to tertiary structure formation: hydrophobic interaction between G63 and V70 in methanol (A) and interaction between charged side chains of D59, T60, R64, Q68, R72, R76 in DPC micelles (B). Interactions of pairs of charged residues T60-R64, R64-Q68, Q68-R72, R76-D80, which are not long-range but are essential for tertiary structure, contribute to an increased number of medium-range interactions observed for these residues.
Figure 4-5. A: Superpositions of the final set of 10 refined SP-B\textsubscript{59–80} structures in methanol with respect to the backbone atoms (N, C\textsuperscript{\alpha}, C',O). B and C: Representative views of SP-B\textsubscript{59–80} in methanol illustrating the amphipathic nature of SP-B\textsubscript{59–80}. Green: hydrophobic residues L, M, C, V, P; Red: negatively charged aspartate residues; Blue: positively charged arginine residues; Grey: polar Q68 residue in the middle of peptide, hydroxyl-containing residues T60, S78 at the ends and nonpolar G63 adjacent to blue R64.
Figure 4-6. $^1$H-$^{15}$N HSQC spectrum and assignments for $^{15}$N-labeled SP-B'$_{59-80}$ in DPC micelles at 315K. Residues L73 and L75, located in the center of the 7-residue palindromic sequence CRLVLRC are too overlapped to be assigned as separate peaks.

Figure 4-7. Alternating $^1$H-$^1$H TOCSY (red) and $^1$H-$^1$H NOESY (blue) from $^{15}$N-edited 3D TOCSY-HSQC and NOESY-HSQC spectra of $^{15}$N-labeled SP-B'$_{59-80}$ peptide in DPC micelles. Assignments are indicated by the residue on top of each strip. Sequential NOEs of the backbone walk connecting protons of residue $i$ with HN of residue $i+1$ are shown by red arrows and other NOEs by black arrows of decreasing boldness indicating increasing distance. A: Assignments for residues at the ends T60-R64 and C77-D80. B: Assignments for residues in the middle I65-R76.
Figure 4-7. Continued.
Figure 4-7. Continued.
Figure 4-8. Summary of the NOE resonances for SP-B’59-80 in DPC micelles. A: The series of i, i+3 αH-NH and αH-βH signals and i, i+4 αH-NH from residues 60-64 and 67-77 is characteristic of an α-helix formation for these residues. Black bars indicate strong unambiguous NOEs and dotted bars indicate weak NOEs. Strong, medium and weak classifications are indicated by the height of the bar for the sequential NOEs. Some NOEs are not marked because they could not be clearly identified due to resonance overlap. B: CSI plotted as the difference between random coil and observed αH chemical shifts. The number of distance constraints per residue is presented in Supplementary information Fig. S1, B.
Figure 4-9. Three-dimensional structure of SP-B\textsubscript{59-80} peptide in DPC micelles. A: Superposition of backbone atoms (N, Ca, C',O) of the 6 lowest energy structures of SP-B\textsubscript{59-80} in DPC micelles. B and C: A representative DPC-bound conformation of SP-B\textsubscript{59-80} showing the distribution of charged and nonpolar residues characteristic of an amphipathic peptide from the top view (B) and a side view (C) including all the side chains. Extensive side chain to side chain interactions involving D59, R64, Q68, R72, R76 and D80 are present in the NMR data. Green: hydrophobic residues L, I, V, C, P; Red: negatively charged D residues; Blue: positively charged R residues; Grey: polar Q68 residue in the center of peptide, hydroxyl-containing residues T60, S78 at the ends and non-polar G63 adjacent to blue R64.
Figure 4-10. SP-B$_{59-80}$ in DPC micelles with residues colored according to their charge to show pairs of electrostatic interaction of positively charged R residues (blue) with negatively charged phosphate moiety (red) and of negatively charged D residues (red) with positively charged choline moiety (blue) of DPC detergent. Green: hydrophobic residues L, I, V, C, P; Grey: polar Q68 residue in the center of peptide and hydroxyl-containing residues T60, S78 at the ends and non-polar G63 adjacent to blue R64. DPC micelle coordinates in PDB format were obtained from MD simulations of 65 DPC molecules with 6305 water molecules for 1100 ps.\textsuperscript{(213)} Coordinates obtained with permission from: Tieleman, D. P., van der Spoel, D., and Berendsen, H. J. C. (2000) Molecular dynamics simulations of dodecylphosphocholine micelles at three different aggregate sizes: Micellar structure and chain relaxation, *Journal of Physical Chemistry B* 104, 6380-6388.
Figure 4-11. Three-dimensional surface models of SP-B_{59-80} structure in methanol interacting with a DPC molecule (Left) and in DPC micelle (Right). Residues are colored according to their charge to show electrostatic interaction between positively charged (blue) arginine or choline moiety of DPC detergent and negatively charged (red) aspartate or phosphate moiety of DPC detergent. Green: hydrophobic residues L, I, V, C, P; Grey: polar Q68 residue in the middle of peptide and hydroxyl-containing residues T60, S78 at the ends and non-polar G63 adjacent to blue R64.

Figure 4-12. Possible mechanism of DPPC trafficking by SP-B_{59-80}. Structures are colored according to charge of residues: positively charged arginines (blue) and negatively charged aspartates (red) and hydrophobic residues (green).
Figure 4-13. The Ramachandran plot of ensemble of 10 SP-B$_{59-80}$ structures in methanol generated by PROCHECK \cite{114} shows the phi-psi torsion angles for all residues in the ensemble of 10 structures (except those at the chain termini). Glycine residues are separately identified by triangles as these are not restricted to the regions of the plot appropriate to the other sidechain types. The colouring/shading represents the different regions described in \cite{174}: Additional Ramachandran plots were generated for each of the 20 different amino acid types (see next page). The darkest areas (red or green) correspond to the "core" regions representing the most favorable combinations of phi-psi backbone torsion angles.

<table>
<thead>
<tr>
<th>Plot statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Residues in most favoured regions [A,B,L]</td>
<td>130  72.2%</td>
</tr>
<tr>
<td>Residues in additional allowed regions [a,b,l,p]</td>
<td>40   22.2%</td>
</tr>
<tr>
<td>Residues in generously allowed regions [-a,-b,-l,-p]</td>
<td>10   5.6%</td>
</tr>
<tr>
<td>Residues in disallowed regions</td>
<td>0   0.0%</td>
</tr>
<tr>
<td>Number of non-glycine and non-proline residues</td>
<td>180 100.0%</td>
</tr>
<tr>
<td>Number of end-residues (excl. Gly and Pro)</td>
<td>20</td>
</tr>
<tr>
<td>Number of glycine residues (shown as triangles)</td>
<td>10</td>
</tr>
<tr>
<td>Number of proline residues</td>
<td>10</td>
</tr>
<tr>
<td>Total number of residues</td>
<td>220</td>
</tr>
</tbody>
</table>
Figure 4-14. The Chi1-Chi2 plots show the chi1-chi2 sidechain torsion angle combinations for 10 structure ensemble of SP-B_{59-80} in methanol for all residue types whose sidechains are long enough to have both these angles. The darkest areas (green) correspond to the "core" regions representing the most favorable combinations of chi1-chi2 sidechain torsion angles.
Figure 4-14. Continued
Figure 4-15. The Ramachandran plot of ensemble of 10 SP-B$_{59-80}$ structures in methanol generated by PROCHECK (114) with data points colored according to G-factor with yellow as favorable and red as unfavorable.
Figure 4-16. Plots of main chain parameters for 10 SP-B59-80 structures in methanol show how the structures (represented by the solid square) compare with well-refined structures at a similar resolution. The dark band in each graph represents the results from the well-refined structures; the central line is a least-squares fit to the mean trend as a function of resolution, while the width of the band on either side of it corresponds to a variation of one standard deviation about the mean. In some cases, the trend is dependent on the resolution, and in other cases it is not.
Figure 4-17. The Ramachandran plot of ensemble of 6 SP-B\textsubscript{59,80} structures in DPC micelles generated by PROCHECK (114) shows the phi-psi torsion angles for all residues in the ensemble of 6 structures (except those at the chain termini). Glycine residues are separately identified by triangles as these are not restricted to the regions of the plot appropriate to the other sidechain types. The colouring/shading represents the different regions described in (174): Additional Ramachandran plots were generated for each of the 20 different amino acid types (see next page). The darkest areas (red or green) correspond to the "core" regions representing the most favorable combinations of phi-psi backbone torsion angles.
Figure 4-17. Continued.
Figure 4-18. The Chi1-Chi2 plots show the chi1-chi2 sidechain torsion angle combinations for ensemble of 6 structures of SP-B'$_{59-80}$ in DPC for all residue types whose sidechains are long enough to have both these angles. The darkest areas (green) correspond to the "core" regions representing the most favorable combinations of chi1-chi2 sidechain torsion angles.
Figure 4-19. The Ramachandran plot of ensemble of 6 SP-B'_{59-80} structures in DPC generated by PROCHECK\(^{(114)}\) with data points colored according to G-factor with yellow as favorable and red as unfavorable.
Figure 4-20. Plots of main chain parameters for 6 SP-B'_{59-80} structures in DPC show how the structures (represented by the solid square) compare with well-refined structures at a similar resolution. The dark band in each graph represents the results from the well-refined structures; the central line is a least-squares fit to the mean trend as a function of resolution, while the width of the band on either side of it corresponds to a variation of one standard deviation about the mean. In some cases, the trend is dependent on the resolution, and in other cases it is not.
Table 4-1. $^1$H chemical shifts for SP-B$_{59-80}$ in methanol measured at 300 K.

<table>
<thead>
<tr>
<th>Residue</th>
<th>HN</th>
<th>HA</th>
<th>HB</th>
<th>HG</th>
<th>HD</th>
<th>HE</th>
</tr>
</thead>
<tbody>
<tr>
<td>D59</td>
<td>4.435</td>
<td>3.066</td>
<td>2.962</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T60</td>
<td>8.552</td>
<td>4.521</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L61</td>
<td>8.498</td>
<td>4.178</td>
<td>1.736</td>
<td>1.708</td>
<td>1.708</td>
<td>0.991</td>
</tr>
<tr>
<td>L62</td>
<td>8.050</td>
<td>4.135</td>
<td>1.668</td>
<td>1.668</td>
<td>1.642</td>
<td>0.930</td>
</tr>
<tr>
<td>G63</td>
<td>8.166</td>
<td>3.765</td>
<td>3.971</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R64</td>
<td>7.874</td>
<td>4.270</td>
<td>1.967</td>
<td>1.967</td>
<td>1.824</td>
<td>1.738</td>
</tr>
<tr>
<td>M65</td>
<td>7.997</td>
<td>4.521</td>
<td>2.169</td>
<td>2.169</td>
<td>2.586</td>
<td>2.706</td>
</tr>
<tr>
<td>L66</td>
<td>8.110</td>
<td>4.380</td>
<td>1.914</td>
<td>1.769</td>
<td>1.796</td>
<td>0.995</td>
</tr>
<tr>
<td>P67</td>
<td>4.269</td>
<td></td>
<td>2.355</td>
<td>1.897</td>
<td>1.996</td>
<td>2.200</td>
</tr>
<tr>
<td>Q68</td>
<td>7.869</td>
<td>4.023</td>
<td>2.288</td>
<td>2.151</td>
<td>2.480</td>
<td>2.446</td>
</tr>
<tr>
<td>L69</td>
<td>8.025</td>
<td>4.074</td>
<td>1.896</td>
<td>1.794</td>
<td>1.640</td>
<td>0.924</td>
</tr>
<tr>
<td>V70</td>
<td>8.246</td>
<td>3.551</td>
<td>2.221</td>
<td></td>
<td>0.964</td>
<td>1.060</td>
</tr>
<tr>
<td>C71</td>
<td>8.148</td>
<td>3.998</td>
<td>3.120</td>
<td>2.856</td>
<td>2.502</td>
<td></td>
</tr>
<tr>
<td>R72</td>
<td>8.078</td>
<td>3.960</td>
<td>2.042</td>
<td>1.990</td>
<td>1.684</td>
<td>1.684</td>
</tr>
<tr>
<td>L73</td>
<td>8.174</td>
<td>4.090</td>
<td>1.935</td>
<td>1.750</td>
<td>1.821</td>
<td>0.890</td>
</tr>
<tr>
<td>V74</td>
<td>8.403</td>
<td>3.590</td>
<td>2.248</td>
<td></td>
<td>0.950</td>
<td>1.093</td>
</tr>
<tr>
<td>L75</td>
<td>8.273</td>
<td>4.083</td>
<td>1.910</td>
<td>1.634</td>
<td>1.869</td>
<td>0.907</td>
</tr>
<tr>
<td>R76</td>
<td>8.310</td>
<td>4.106</td>
<td>2.080</td>
<td>2.080</td>
<td>1.856</td>
<td>1.773</td>
</tr>
<tr>
<td>C77</td>
<td>8.138</td>
<td>4.313</td>
<td>3.080</td>
<td>2.932</td>
<td>2.670</td>
<td></td>
</tr>
<tr>
<td>S78</td>
<td>7.898</td>
<td>4.370</td>
<td>4.005</td>
<td>3.920</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M79</td>
<td>7.665</td>
<td>4.517</td>
<td>2.184</td>
<td>2.070</td>
<td>2.705</td>
<td>2.564</td>
</tr>
<tr>
<td>D80</td>
<td>7.944</td>
<td>4.763</td>
<td>2.936</td>
<td>2.850</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-2. $^3$J$_{HN,αH}$ couplings for SP-B$_{59-80}$ in methanol measured at 300 K. $^3$J$_{HN,αH}$ couplings could not be measured for residues 63, 64, 66, 68, 71, 72, 77 and 78 due to overlap.

<table>
<thead>
<tr>
<th>$^3$J$_{HN,αH}$ coupling constant</th>
<th>Mean value</th>
<th>Min value</th>
<th>Max value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>T60HN-HA</td>
<td>7.4</td>
<td>6.4</td>
<td>8.4</td>
<td>1</td>
</tr>
<tr>
<td>L61HN-HA</td>
<td>4.8</td>
<td>3.8</td>
<td>5.8</td>
<td>1</td>
</tr>
<tr>
<td>L62HN-HA</td>
<td>4.8</td>
<td>3.8</td>
<td>5.8</td>
<td>1</td>
</tr>
<tr>
<td>M65HN-HA</td>
<td>7.35</td>
<td>6.35</td>
<td>8.35</td>
<td>1</td>
</tr>
<tr>
<td>L69HN-HA</td>
<td>5.0</td>
<td>4.0</td>
<td>6.0</td>
<td>1</td>
</tr>
<tr>
<td>V70HN-HA</td>
<td>4.8</td>
<td>3.8</td>
<td>5.8</td>
<td>1</td>
</tr>
<tr>
<td>L73HN-HA</td>
<td>4.88</td>
<td>3.88</td>
<td>5.88</td>
<td>1</td>
</tr>
<tr>
<td>V74HN-HA</td>
<td>4.8</td>
<td>3.8</td>
<td>5.8</td>
<td>1</td>
</tr>
<tr>
<td>L75HN-HA</td>
<td>4.45</td>
<td>3.45</td>
<td>5.45</td>
<td>1</td>
</tr>
<tr>
<td>R76HN-HA</td>
<td>4.45</td>
<td>3.45</td>
<td>5.45</td>
<td>1</td>
</tr>
<tr>
<td>M79HN-HA</td>
<td>7.73</td>
<td>6.73</td>
<td>8.73</td>
<td>1</td>
</tr>
<tr>
<td>D80HN-HA</td>
<td>8.1</td>
<td>7.1</td>
<td>9.1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4-3. A summary of structural statistics for the 10 final structures of SP-B_{59-80} in methanol and SP-B'_{59-80} bound to DPC micelles.

<table>
<thead>
<tr>
<th>restraints for structure calculation</th>
<th>SP-B_{59-80} in methanol</th>
<th>SP-B'_{59-80} bound to DPC micelles</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE-based distance restraints (total)</td>
<td>459</td>
<td>433</td>
</tr>
<tr>
<td>Intra-residue (</td>
<td>i-j</td>
<td>= 0)</td>
</tr>
<tr>
<td>Sequential (</td>
<td>i-j</td>
<td>= 1)</td>
</tr>
<tr>
<td>Medium range (2 ≤</td>
<td>i-j</td>
<td>≤ 4)</td>
</tr>
<tr>
<td>Long range (</td>
<td>i-j</td>
<td>≥ 5)</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>54</td>
<td>80</td>
</tr>
<tr>
<td>(^3)(J_{\text{HN,\alpha H}}) coupling-based backbone angle restraints (total)</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Restraints violations (Å)(^a,b)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Distance violations &gt; 0.29Å</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Distance violations &gt; 0.5Å</td>
<td>0.050 ± 0.001</td>
<td>0.060 ± 0.012</td>
</tr>
<tr>
<td>rmsd of distance violation (Å)</td>
<td>0.286</td>
<td>0.5</td>
</tr>
<tr>
<td>Max. distance violation (Å)</td>
<td>0.250 ± 0.025</td>
<td>-</td>
</tr>
<tr>
<td>rmsd of (J)-coupling violation (Hz)</td>
<td>0.003 ± 0.000</td>
<td>0.005 ± 0.000</td>
</tr>
<tr>
<td>rmsd from covalent geometry(^b)</td>
<td>0.482 ± 0.006</td>
<td>0.673 ± 0.016</td>
</tr>
<tr>
<td>bonds (Å)</td>
<td>0.332 ± 0.008</td>
<td>0.412 ± 0.031</td>
</tr>
<tr>
<td>angles (deg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>impropers (deg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rmsd from the mean structure(^a)</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>backbone atoms of all residues (Å)</td>
<td>1.32</td>
<td>1.44</td>
</tr>
<tr>
<td>total rmsd (Å)</td>
<td>0.28</td>
<td>0.48</td>
</tr>
<tr>
<td>backbone atoms of residues (61-78) (Å)</td>
<td>1.15</td>
<td>1.36</td>
</tr>
<tr>
<td>total rmsd of residues (61-78) (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot(^c)</td>
<td>130 (72.2%)</td>
<td>59 (54.6%)</td>
</tr>
<tr>
<td>residues in the most favored region</td>
<td>40 (22.2%)</td>
<td>43 (39.8%)</td>
</tr>
<tr>
<td>residues in the additional allowed region</td>
<td>10 (5.6%)</td>
<td>6 (5.6%)</td>
</tr>
<tr>
<td>residues in the generously allowed region</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>outliers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Calculated with VMD\(^{173}\) when the backbone atoms of residues 61–78 of the accepted ensemble of structures are superimposed. \(^b\)Calculated by XPLOR-NIH. \(^{111, 112, 173}\) \(^c\)Calculated with Procheck V3.4.4\(^{114, 174}\) for the accepted ensemble of structures.
<table>
<thead>
<tr>
<th>Residue</th>
<th>N</th>
<th>HN</th>
<th>HA</th>
<th>HB</th>
<th>HG</th>
<th>HD</th>
<th>Ne</th>
<th>HE</th>
</tr>
</thead>
<tbody>
<tr>
<td>D59</td>
<td>-</td>
<td>-</td>
<td>4.650</td>
<td>3.045</td>
<td>3.045</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T60</td>
<td>115.1</td>
<td>8.767</td>
<td>4.510</td>
<td>4.480</td>
<td>1.275</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L61</td>
<td>124.1</td>
<td>8.908</td>
<td>4.156</td>
<td>1.800</td>
<td>1.875</td>
<td>1.666</td>
<td>0.918</td>
<td>0.965</td>
</tr>
<tr>
<td>L62</td>
<td>117.9</td>
<td>8.505</td>
<td>4.020</td>
<td>1.740</td>
<td>1.770</td>
<td>1.527</td>
<td>0.865</td>
<td>0.943</td>
</tr>
<tr>
<td>G63</td>
<td>104.7</td>
<td>8.060</td>
<td>3.775</td>
<td>3.935</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R64</td>
<td>117.2</td>
<td>7.670</td>
<td>4.400</td>
<td>1.760</td>
<td>1.952</td>
<td>1.650</td>
<td>3.204</td>
<td>84.300</td>
</tr>
<tr>
<td>I65</td>
<td>116.5</td>
<td>8.005</td>
<td>4.092</td>
<td>1.940</td>
<td>1.255</td>
<td>0.910</td>
<td>0.840</td>
<td>1.300</td>
</tr>
<tr>
<td>L66</td>
<td>119.7</td>
<td>8.367</td>
<td>4.245</td>
<td>1.800</td>
<td>1.910</td>
<td>1.615</td>
<td>0.873</td>
<td>0.920</td>
</tr>
<tr>
<td>P67</td>
<td></td>
<td>1.930</td>
<td>3.600</td>
<td>3.655</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q68</td>
<td>114.4</td>
<td>7.245</td>
<td>4.055</td>
<td>2.130</td>
<td>2.220</td>
<td>2.445</td>
<td>2.477</td>
<td>110.900</td>
</tr>
<tr>
<td>L69</td>
<td>119.1</td>
<td>8.060</td>
<td>4.070</td>
<td>1.770</td>
<td>1.860</td>
<td>1.650</td>
<td>0.830</td>
<td>0.870</td>
</tr>
<tr>
<td>V70</td>
<td>117.2</td>
<td>8.328</td>
<td>3.535</td>
<td>2.140</td>
<td>1.020</td>
<td>0.930</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C71</td>
<td>116.8</td>
<td>8.058</td>
<td>4.660</td>
<td>3.465</td>
<td>3.140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R72</td>
<td>117.2</td>
<td>7.771</td>
<td>4.370</td>
<td>1.940</td>
<td>1.760</td>
<td>3.193</td>
<td>84.300</td>
<td>7.582</td>
</tr>
<tr>
<td>L73</td>
<td>115.3</td>
<td>7.853</td>
<td>4.445</td>
<td>1.870</td>
<td>1.760</td>
<td>1.566</td>
<td>0.877</td>
<td></td>
</tr>
<tr>
<td>V74</td>
<td>112.2</td>
<td>7.706</td>
<td>4.350</td>
<td>2.335</td>
<td>1.055</td>
<td>0.936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L75</td>
<td>115.3</td>
<td>7.853</td>
<td>4.445</td>
<td>1.870</td>
<td>1.760</td>
<td>1.566</td>
<td>0.877</td>
<td></td>
</tr>
<tr>
<td>R76</td>
<td>118.2</td>
<td>7.913</td>
<td>4.475</td>
<td>1.850</td>
<td>1.710</td>
<td>1.630</td>
<td>1.560</td>
<td>3.178</td>
</tr>
<tr>
<td>C77</td>
<td>120.6</td>
<td>8.440</td>
<td>4.560</td>
<td>3.415</td>
<td>3.170</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S78</td>
<td>116.3</td>
<td>8.460</td>
<td>4.455</td>
<td>3.840</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I79</td>
<td>121.0</td>
<td>8.056</td>
<td>4.170</td>
<td>1.884</td>
<td>1.435</td>
<td>0.893</td>
<td>0.845</td>
<td>1.165</td>
</tr>
<tr>
<td>D80</td>
<td>121.0</td>
<td>8.353</td>
<td>4.720</td>
<td>2.930</td>
<td>2.780</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 5
CONCLUSIONS AND FUTURE DIRECTIONS

Many lung surfactant deficiency and dysfunction pathologies could potentially be treated by therapies utilizing combinations of synthetic LS protein mimics and lipids as LS replacements. This could ameliorate or prevent an immune response, thus expanding the range of treatable respiratory distress syndromes. Presented here were insights into the molecular level behavior of LS constituents and an attempt to elucidate a potential mechanism of interaction between LS protein mimics and LS lipids. This work is a continuation of research aimed at understanding LS on a molecular level to guide development of its synthetic replacements.

In Chapters 3 and 4 we showed that SP-B_{59-80} demonstrates helical structure in methanol and in DPC micelles solutions as well as when associated with DPPC/POPG and POPC/POPG lipid systems as determined by CD measurements. However, we found some minor differences due to the environment, which could be interesting with respect to determining peptide function and required elucidation of the high resolution structures of the peptide in these environments. The lipid systems retain a lamellar organization, even with 3 mol% addition of SP-B_{59-80}. This is in contrast to CD data for the N-terminal SP-B peptide, SP-B_{1-25}, studied by my colleagues, which retains a constant secondary structure when associated with DPPC/POPG and POPC/POPG lipid systems and causes the formation of fluid isotropic lipid phases.\(^{69}\)

In Chapter 3 we showed that SP-B_{59-80} penetrates deeply into DPPC:POPG bilayers and binds more peripherally to POPC:POPG bilayers allowing it to distinguish between fully saturated and monounsaturated lipids.\(^{161}\) Similar behavior has been observed for KL_{4}, a peptide mimic of SPB_{59-80}, which forms unusual lipid dependent
helical structures.\(^{(162)}\) KL\(_4\) changes its helical pitch and penetration as a function of lipid composition. The deep penetration of SP-B\(_{59-80}\) within DPPC:POPG lipid bilayers could require a similar change in helical structure. However, as described in Chapter 4, SP-B\(_{59-80}\) alters its structure by rotating two shorter helices around a proline hinge rather than having a single longer helix with variable pitch. The structural plasticity and variable penetration depth of KL\(_4\) and SP-B\(_{59-80}\) can potentially affect the stability and composition of lung surfactant lipid structures and may provide a mechanism for lipid trafficking from lamellar bodies and tubular myelin to the air-water interface in a manner that selects for DPPC.

With the findings presented in Chapter 4, a more thorough molecular model is established that provides insights into how the C-terminal end of SP-B may interact with DPPC lipid and selectively transfer it to the alveolar air-water interface. The proposed mechanism can drive the development of future SP-B mimetics.

Taken together, these studies highlight that development of successful synthetic LS replacement formulations probably requires one to specifically target peptidomimetics that affect DPPC dynamics. The structural study also highlighted the unique role of the proline in the middle of SP-B\(_{59-80}\) sequence, which acts as a hinge allowing two helices to roll around their axes and adapt to the surrounding environment. This is in contrast to KL\(_4\), a functional mimic of SP-B\(_{59-80}\), which changes helical pitch to adapt to different lipid environments through exchange between i,i+4 H-bonds and i,i+5 H-bonds. The rotation about the proline hinge in SP-B\(_{59-80}\) requires change in a smaller number of bonds compared to helical pitch changes of KL\(_4\) and thus is probably more energetically favorable. Consequently, introduction of the proline residue in the middle
of KL4 peptide could make its structural adaptation to the environment much more energetically favorable, which could in turn make it more efficient in RDS treatment. While we established the effect of SP-B59-80 on lipid dynamics is modulated by the degree of lipid saturation in LS mimics, much work is left if we want to have a better understanding of SP-B59-80 interaction with lipids, particularly study of the high resolution structures and dynamics of SP-B59-80 peptide in the DPPC:POPG and POPC:POPG lipid bilayers. The solution NMR structures of SP-B59-80 in methanol and DPC micelles discussed in this dissertation provide an important model for how the peptide might structurally adapt in a lipid environment, which prompted us to create a model of its interaction with DPPC lipid (Fig. 5-1). However, other structural and dynamics measurements of the peptide in DPPC:POPG and POPC:POPG lipid bilayers are needed to establish the true mechanism for the peptide interaction with and remodeling of lipid assemblies.

Direct measurements of the actual depth of SP-B59-90 penetration into lipid bilayers via power saturation electron paramagnetic resonance (EPR) measurements could confirm our findings at amino acid resolution. Experiments of this type would complement our existing deuterium data, confirming conclusions made from our determination of relative order parameters in the acyl chains of the individual lipid species. Backbone measurements of SP-B59-90 in the DPPC:POPG and POPC:POPG lipid bilayers via solid state NMR will provide high resolution structures of the peptide in these bilayers. This would expand our understanding of its mechanism of interaction with lipids, which would ultimately allow understanding of how lipid trafficking happens in lung surfactant.
A model of how SP-B\textsubscript{59-80} could be involved in directing DPPC adsorption to the air-water interface and facilitating tight packing interactions between the bilayers of tubular myelin has been proposed in Fig. 5-1 based on the structures of SP-B\textsubscript{59-80} in methanol and DPC micelles obtained in this dissertation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{model.png}
\caption{Molecular models of SP-B\textsubscript{59-80} directing specific DPPC adsorption to the air-water interface and facilitating tight packing interactions between the bilayers of tubular myelin. Phospholipids (mainly DPPC) are transferred from tubular myelin bilayers to the monolayer at the air-water interface to efficiently cover and stabilize this interface for adequate respiration. The two structures of SP-B\textsubscript{59-80} in DPC micelle are positioned at the fusion neck to demonstrate how they can stabilize the neck and participate in DPPC lipid trafficking. The structure of SP-B\textsubscript{59-80} in methanol is placed at the bilayer interface to show how it can be involved into interaction between the two bilayers to facilitate their tight packing in tubular myelin.}
\end{figure}
APPENDIX
SCRIPTS FOR STRUCTURE CALCULATIONS AND REFINEMENT

Script modified from anneal_nordc.py - example script in XPLOR-NIH eginput directory,\(^{(179)}\) which was used for initial structure calculation:

```python
xplor.requireVersion("2.24")

# slow cooling protocol in torsion angle space for protein G. Uses NOE, RDC, J-coupling restraints.
# this script performs annealing from an extended structure.
# It is faster than the original anneal.py
# CDS 2009/07/24
#
# this checks for typos on the command-line. User-customized arguments can also be specified.
# xplor.parseArguments()

# filename for output structures. This string must contain the STRUCTURE literal so that each calculated structure has a unique name. The SCRIPT literal is replaced by this filename (or stdin if redirected using <), but it is optional.
outFilename = "SPB5980_SCRIPT_STRUCTURE.sa"
numberOfStructures=100  #usually you want to create at least 20

# protocol module has many high-level helper functions.
import protocol
protocol.initRandomSeed()  #set random seed - by time
command = xplor.command

# generate PSF data from sequence and initialize the correct parameters.
# from psfGen import seqToPSF
seqToPSF('spbc_sequence_xplor.seq', startResid=59)
# generate random extended initial structure with correct covalent geometry
# protocol.genExtendedStructure()

xplor.command("write struct output=spbc.psf end")
protocol.writePDB("spbc_extended.pdb")
```
a PotList contains a list of potential terms. This is used to specify which terms are active during refinement.

```python
from potList import PotList

potList = PotList()
```

parameters to ramp up during the simulated annealing protocol

```python
from simulationTools import MultRa, StaticRamp, InitialParams

rampedParams=[]
highTempParams=[]
```

compare atomic Cartesian rmsd with a reference structure backbone and heavy atom RMSDs will be printed in the output structure files

```python
# from posDiffPotTools import create_PosDiffPot
refRMSD = create_PosDiffPot("refRMSD","name CA or name C or name N",
                           pdbFile='sa_5.pdb',
                           cmpSel="not name H*")
```

set up NOE potential

```python
noe=PotList('noe')
potList.append(noe)
from noePotTools import create_NOEPot
for (name,scale,file) in
[("all",1,"spbc_TCEP_hocd3_HeiVolNote_heightsorted_duplicates_commented_classified_noindexed.tbl"),
   #add entries for additional tables ]:
    pot = create_NOEPot(name,file)
    # pot.setPotType("soft") # if you think there may be bad NOEs
    pot.setScale(scale)
    noe.append(pot)
rampedParams.append( MultRamp(2,30, "noe.setScale( VALUE )") )
```

set up J coupling with Karplus coefficients

```python
# from jCoupPotTools import create_JCoupPot
#jCoup = create_JCoupPot("jcoupl","jna_coup.tbl",
#                        A=6.98,B=-1.38,C=1.72,phase=-60.0)
#potList.append(jCoup)
```

set up dihedral angles

```python
# from xplorPot import XplorPot
dihedralRestraintFilename="ubiquitin_1g6j_noA_dihe.tbl"
#protocol.initDihedrals(dihedralRestraintFilename,
   #useDefaults=False # by default, symmetric
sidechain
   # restraints are included
```
# gyration volume term
# from gyrPotTools import create_GyrPot
gyr = create_GyrPot("Vgyr",
    "resid 1:56") # selection should exclude disordered tails
potList.append(gyr)
rampedParams.append( MultRamp(.002,1,"gyr.setScale(VALUE)" ) )

# hbda - distance/angle bb hbond term
# protocol.initHBDA('hbda.tbl')
potList.append( XplorPot('HBDA') )

# Rama torsion angle database
# protocol.initRamaDatabase()
potList.append( XplorPot('RAMA') )
rampedParams.append( MultRamp(.002,1,"potList['RAMA'].setScale(VALUE)" ) )

# setup parameters for atom-atom repulsive term. (van der Waals-like term)
potList.append( XplorPot('VDW') )
rampedParams.append( StaticRamp("protocol.initNBond()" ) )
rampedParams.append( MultRamp(0.9,0.8, 
    "command('param nbonds repel VALUE end end')" ) )
rampedParams.append( MultRamp(0.004,4, 
    "command('param nbonds rcon VALUE end end')" ) )

# nonbonded interaction only between CA atoms
highTempParams.append( StaticRamp("""protocol.initNBond(cutnb=100, 
    rcon=0.004, 
    tolerance=45, 
    repel=1.2, 
    onlyCA=1)""") )

# bond
# angle
# improper

potList.append( Xplor("BOND") )
potList.append( Xplor("ANGL") )
potList["ANGL"].setThreshold( 5 )
rampedParams.append( MultRamp(0.4,1,"potList['ANGL'].setScale(VALUE)" ) )
potList.append( XplorPot("IMPR") )
potList['IMPR'].setThreshold( 5 )
rampedParams.append( MultRamp(0.1,1,"potList['IMPR'].setScale(VALUE)") )

# Give atoms uniform weights, configure bath/molecule friction coeff.
# protocol.massSetup()

# IVM setup
# the IVM is used for performing dynamics and minimization in torsion-angle
# space, and in Cartesian space.
# from ivm import IVM
dyn = IVM()

# configure ivm topology for torsion-angle dynamics
# protocol.torsionTopology(dyn)

# minc used for final cartesian minimization
# minc = IVM()
protocol.initMinimize(minc)
protocol.cartesianTopology(minc)

# object which performs simulated annealing
# from simulationTools import AnnealIVM
init_t  = 3500.    # Need high temp and slow annealing to converge
cool = AnnealIVM(initTemp =init_t,
    finalTemp=25,
    tempStep =12.5,
    ivm=dyn,
    rampedParams = rampedParams)

def calcOneStructure(loopInfo):
    """this function calculates a single structure, performs analysis on
the
structure, and then writes out a pdb file, with remarks.
""

    # generate a new structure with randomized torsion angles
    # from monteCarlo import randomizeTorsions
    randomizeTorsions(dyn)
    protocol.fixupCovalentGeom(maxIters=100,useVDW=1)
# set torsion angles from restraints
#
# from torsionTools import setTorsionsFromTable
    setTorsionsFromTable(dihedralRestraintFilename)
protocol.writePDB(loopInfo.filename() + ".init")

# calc. initial tensor orientation
#
# initialize parameters for high temp dynamics.
InitialParams( rampedParams )
# high-temp dynamics setup - only need to specify parameters which
#   differ from initial values in rampedParams
InitialParams( highTempParams )

# high temp dynamics
#
protocol.initDynamics(dyn,
    potList=potList,  # potential terms to use
    bathTemp=init_t,
    initVelocities=1,
    finalTime=200,   # stops at 800ps or 8000 steps
    numSteps=2000,   # whichever comes first
    printInterval=200)  # was 100

dyn.setETolerance( init_t/100 )  # used to det. stepsize. default: t/1000
    dyn.run()

# initialize parameters for cooling loop
InitialParams( rampedParams )

# initialize integrator for simulated annealing
#
protocol.initDynamics(dyn,
    potList=potList,
    numSteps=100,  # at each temp: 100 steps or
    finalTime=.2 ,  # .2ps, whichever is less
    numSteps=2000, # whichever comes first
    printInterval=200)

# perform simulated annealing
#
    cool.run()

# final torsion angle minimization
#
protocol.initMinimize(dyn,
    printInterval=50)
    dyn.run()

# final all-atomic degrees of freedom minimization
#
protocol.initMinimize(minc,
    potList=potList,
    dEPred=10)

minc.run()

# do analysis and write structure
loopInfo.writeStructure(potList)
pass

from simulationTools import StructureLoop, FinalParams

StructureLoop(numStructures=numberOfStructures,
     pdbTemplate=outFilename,
     structLoopAction=calcOneStructure,
     genViolationStats=1,
     averageTopFraction=0.1, # report stats on best 50% of structs
     averageContext=FinalParams(rampedParams),
     #
     averageCrossTerms=refRMSD,
     averageSortPots=[potList['BOND'], potList['ANGL'], potList['IMPR'], noe],
     averagePotList=potList).run()
Script modified from refine_nordc.py - example script in XPLOR-NIH eginput directory, which was used for refinement of ensemble of SP-B$_{59-80}$ structures in methanol.

xplor.requireVersion("2.18")

# slow cooling protocol in torsion angle space for protein G. Uses NOE, J-coupling restraints.
# this version refines from a reasonable model structure.
# CDS 2005/05/10
#

(opts,args) = xplor.parseArguments(["quick"]) # check for command-line typos

quick=False
for opt in opts:
    if opt[0]=="quick":  #specify -quick to just test that the script runs
        quick=True
        pass
    pass

outFilename = "SPB5980_SCRIPT_STRUCTURE.sa"
numberOfStructures=3000

if quick:
    numberOfStructures=3
    pass

# protocol module has many high-level helper functions.
# import protocol
protocol.initRandomSeed(8870649)   #explicitly set random seed

# annealing settings
#
command = xplor.command

protocol.initParams("protein")

# generate PSF data from sequence and initialize the correct parameters.
# from psfGen import seqToPSF
#seqToPSF('1g6j_sequence.seq')
#protocol.initStruct("g_new.psf") # - or from file
# generate a random extended structure with correct covalent geometry
# saves the generated structure in the indicated file for faster startup
# next time.
#
protocol.genExtendedStructure("gb1_extended_%d.pdb" %
    protocol.initialRandomSeed())

# or read an existing model
#
protocol.loadPDB("SPBC_refine_nordc_95.sa")

xplor.simulation.deleteAtoms("not known")

protocol.fixupCovalentGeom(maxIters=100, useVDW=1)

# a PotList contains a list of potential terms. This is used to specify
# which terms are active during refinement.
# from potList import PotList

potList = PotList()

# parameters to ramp up during the simulated annealing protocol
# from simulationTools import MultRamp, StaticRamp, InitialParams

rampedParams=[]
highTempParams=[]

# compare atomic Cartesian rmsd with a reference structure
# backbone and heavy atom RMSDs will be printed in the output
# structure files
# from posDiffPotTools import create_PosDiffPot
#refRMSD = create_PosDiffPot("refRMSD","name CA or name C or name N",
#    pdbFile='g_xray.pdb',
#    cmpSel="not name H*")

# set up NOE potential
noe=PotList('noe')

potList.append(noe)

from noePotTools import create_NOEPot

for (name, scale, file) in
    [(['all',1,"spbc_TCEP_hocd3_HeiVolNote_heightsorted_duplicates_commented_cl
        assified_noindex11ambi.tbl"),
        # add entries for additional tables ]:
        
        pot = create_NOEPot(name, file)
        pot.setPotType("soft") #-- if you think there may be bad NOEs
        pot.setScale(scale)
        noe.append(pot)

rampedParams.append( MultRamp(2,30, "noe.setScale( VALUE )") )

174
# set up J coupling - with Karplus coefficients
from jCoupPotTools import create_JCoupPot
cjCoup = create_JCoupPot("jcoup","jna_coup_spbc_in_methanol9.tbl",
A=6.98,B=-1.38,C=1.72,phase=-60.0)
rampedParams.append(MultRamp(1,3,"cjCoup.setScale(VALUE)"))
potList.append(cjCoup)

# Set up dihedral angles
from xplorPot import XplorPot
#protocol.initDihedrals("ubiquitin_1g6j_noA_dihe.tbl",
#                       #useDefaults=False # by default, symmetric
sidechain
#                                           # restraints are included
#
#potList.append( XplorPot('CDIH') )
#highTempParams.append( StaticRamp("potList['CDIH'].setScale(10)"))
#rampedParams.append( StaticRamp("potList['CDIH'].setScale(200)"))
# set custom values of threshold values for violation calculation
#
#potList['CDIH'].setThreshold( 5 ) #5 degrees is the default value, though

# gyration volume term
#
#gyration volume term
#
#from gyrPotTools import create_GyrPot
#gyr = create_GyrPot("Vgyr",
#                    "resid 1:56") # selection should exclude disordered
tails
#potList.append(gyr)
rampedParams.append(MultRamp(.002,1,"gyr.setScale(VALUE)"))

# hbda - distance/angle bb hbond term
#
#protocol.initHBDA('hbda.tbl')
potList.append(XplorPot('HBDA'))

# hbdb - distance/angle bb hbond term
#
#protocol.initHBDB()
potList.append(XplorPot('HBDB'))

#Rama torsion angle database
#
#from torsionDBPotTools import create_TorsionDBPot
torsionDB = create_TorsionDBPot('torsionDB')
potList.append(torsionDB)
rampedParams.append(MultRamp(.002,2,"torsionDB.setScale(VALUE)"))

# setup parameters for atom-atom repulsive term. (van der Waals-like term)
#
potList.append( XplorPot('VDW') )
rampedParams.append( StaticRamp("protocol.initNBond()") )
rampedParams.append( MultRamp(0.9,0.8,
    "command('param nbonds repel VALUE end end')") )
rampedParams.append( MultRamp(.004,4,
    "command('param nbonds rcon VALUE end end')") )
# nonbonded interaction only between CA atoms
highTempParams.append( StaticRamp("protocol.initNBond(cutnb=100,
    rcon=0.004,
    tolerance=45,
    repel=1.2,
    onlyCA=1)""") )

potList.append( XplorPot('BOND') )
potList.append( XplorPot('ANGL') )
potList['ANGL'].setThreshold( 5 )
rampedParams.append( MultRamp(0.4,1,"potList['ANGL'].setScale(VALUE)" )
potList.append( XplorPot('IMPR') )
potList['IMPR'].setThreshold( 5 )
rampedParams.append( MultRamp(0.1,1,"potList['IMPR'].setScale(VALUE)" )

# Give atoms uniform weights, except for the anisotropy axis
# protocol.massSetup()

# IVM setup
#   the IVM is used for performing dynamics and minimization in torsion-angle
#   space, and in Cartesian space.
# from ivm import IVM
dyn = IVM()

# reset ivm topology for torsion-angle dynamics
# dyn.reset()
protocol.torsionTopology(dyn)

# minc used for final cartesian minimization
# minc = IVM()
protocol.initMinimize(minc)
protocol.cartesianTopology(minc)
# object which performs simulated annealing

from simulationTools import AnnealIVM

init_t = 1500.   # Need high temp and slow annealing to converge
cool = AnnealIVM(initTemp =init_t,
                  finalTemp =25,
                  tempStep =3.25,
                  ivm=dyn,
                  rampedParams = rampedParams)

def accept(potList):
    
    return True if current structure meets acceptance criteria

    if potList['noe'].violations()>0:
        return False
    if potList['CDIH'].violations()>0:
        return False
    if potList['BOND'].violations()>0:
        return False
    if potList['ANGL'].violations()>0:
        return False
    if potList['IMPR'].violations()>1:
        return False

    return True

def calcOneStructure(loopInfo):
    
    this function calculates a single structure, performs analysis on
    the
    structure, and then writes out a pdb file, with remarks.

    # initialize parameters for high temp dynamics.
    InitialParams( rampedParams )
    # high-temp dynamics setup - only need to specify parameters which
    # differ from initial values in rampedParams
    InitialParams( highTempParams )

    # high temp dynamics
    #
    protocol.initDynamics(dyn,
        potList=potList, # potential terms to use
        bathTemp=init_t,
        initVelocities=1,
        finalTime=10,    # stops at 10ps or 5000 steps
        numSteps=5000,   # whichever comes first
        printInterval=100)

    dyn.setETolerance( init_t/100 )  # used to det. stepsize. default: 
t/1000
    dyn.run()

    # initialize parameters for cooling loop
InitialParams( rampedParams )

# initialize integrator for simulated annealing
# protocol.initDynamics(dyn,
    potList=potList,
    numSteps=100,       # at each temp: 100 steps or
    finalTime=.2 ,       # .2ps, whichever is less
    printInterval=100)

# perform simulated annealing
# cool.run()

# final torsion angle minimization
# protocol.initMinimize(dyn,
    printInterval=50)
    dyn.run()

# final all-atom minimization
# protocol.initMinimize(minc,
    potList=potList,
    dEPred=10)
    minc.run()

# do analysis and write structure
loopInfo.writeStructure(potList)
pass

from simulationTools import StructureLoop, FinalParams
StructureLoop(numStructures=numberOfStructures,
    pdbTemplate=outFilename,
    calcMissingStructs=True,
    structLoopAction=calcOneStructure,
    genViolationStats=1,
    averagePotList=potList,
    averageSortPots=[potList['BOND'],potList['ANGL'],potList['IMPR'],jCoup,
    noe],
    averageCrossTerms=refRMSD,
    averageTopFraction=0.005, #report only on best 50% of structs
    averageAccept=accept,   #only use structures which pass
    averageContext=FinalParams(rampedParams),
    averageFilename="SPB5980_SCRIPT_ave.pdb",       #generate
    averageFitSel="name CA",
    averageCompSel="not resname ANI and not name H*" ).run()
Script modified from refine_nordc.py - example script in XPLOR-NIH eginput directory, (179) which was used for refinement of ensemble of SP-B'₅₉₋₈₀ structures in DPC micelles.

```python
xplor.requireVersion("2.18")
#
# slow cooling protocol in torsion angle space for protein G. Uses
# NOE, J-coupling restraints.
#
# this version refines from a reasonable model structure.
#
# CDS 2005/05/10
#

(opts, args) = xplor.parseArguments(["quick"])  # check for command-line
typos
quick=False
for opt in opts:
    if opt[0]=="quick":  #specify -quick to just test that the script runs
        quick=True
    pass
pass

outFilename = "2ISPB5980_SCRIPT_STRUCTURE.sa"
numberOfStructures=600
if quick:
    numberOfStructures=3
    pass

# protocol module has many high-level helper functions.
#
import protocol
protocol.initRandomSeed(3421)  #explicitly set random seed
#
# annealing settings
#
command = xplor.command
protocol.initParams("protein")

# generate PSF data from sequence and initialize the correct parameters.
#
#from psfGen import seqToPSF
#seqToPSF('1g6j_sequence.seq')
#protocol.initStruct("g_new.psf") # - or from file
```
# generate a random extended structure with correct covalent geometry
# saves the generated structure in the indicated file for faster startup
# next time.
#

protocol.genExtendedStructure("gb1_extended_%d.pdb" %
    protocol.initialRandomSeed())

# or read an existing model
# protocol.loadPDB("2ISPBC_refine_nordc_108.sa")
xplor.simulation.deleteAtoms("not known")

protocol.fixupCovalentGeom(maxIters=100, useVDW=1)

# a PotList contains a list of potential terms. This is used to specify
# terms are active during refinement.
# from potList import PotList
potList = PotList()

# parameters to ramp up during the simulated annealing protocol
# from simulationTools import MultRamp, StaticRamp, InitialParams
rampedParams = []
highTempParams = []

# compare atomic Cartesian rmsd with a reference structure
# backbone and heavy atom RMSDs will be printed in the output
# structure files
# from posDiffPotTools import create_PosDiffPot
# refRMSD = create_PosDiffPot("refRMSD","name CA or name C or name N",
#    pdbFile='g_xray.pdb',
#    cmpSel="not name H")

# set up NOE potential
noe = PotList('noe')
potList.append(noe)
from noePotTools import create_NOEPot
for (name, scale, file) in
    [('all', 1, "15Nspbc_noesy3D_dpc_1Dec11_04042012_noe_sorted2_ambi17.tbl"),
     # add entries for additional tables
    ]:
    pot = create_NOEPot(name, file)
    pot.setPotType("soft")  # if you think there may be bad NOEs
    noe.append(pot)
    pot.setScale(scale)
noe.append(pot)
rampedParams.append( MultRamp(2, 30, "noe.setScale( VALUE )") )
# set up J coupling - with Karplus coefficients
# from jCoupPotTools import create_JCoupPot
jCoup = create_JCoupPot("jcoup","jna_coup_spbc_in_methanol7.tbl",
# A=6.98,B=-1.38,C=1.72,phase=-60.0)
potList.append(jCoup)

# Set up dihedral angles
from xplorPot import XplorPot
#protocol.initDihedrals("ubiquitin_1g6j_noA_dihe.tbl",
#                       # useDefaults=False   # by default, symmetric
#                       # sidechain
#                       # restraints are included
#                       )
potList.append( XplorPot('CDIH') )
# highTempParams.append( StaticRamp("potList['CDIH'].setScale(10)") )
# rampedParams.append( StaticRamp("potList['CDIH'].setScale(200)") )
# set custom values of threshold values for violation calculation
# potList['CDIH'].setThreshold( 5 ) # 5 degrees is the default value, though

# gyration volume term
# gyration volume term
# from gyrPotTools import create_GyrPot
# gyr = create_GyrPot("Vgyr",
#                     "resid 1:56") # selection should exclude disordered tails
potList.append(gyr)
rampedParams.append( MultRamp(.002,1,"gyr.setScale(VALUE)") )

# hbda - distance/angle bb hbond term
# protocol.initHBDA('hbda.tbl')
# potList.append( XplorPot('HBDA') )

# hbdb - distance/angle bb hbond term
# protocol.initHBDB()
# potList.append( XplorPot('HBDB') )

# Rama torsion angle database
# from torsionDBPotTools import create_TorsionDBPot
torsionDB = create_TorsionDBPot('torsionDB')
potList.append( torsionDB )
rampedParams.append( MultRamp(.002,2,"torsionDB.setScale(VALUE)") )

# setup parameters for atom-atom repulsive term. (van der Waals-like term)
potList.append( XplorPot('VDW') )
rampedParams.append( StaticRamp("protocol.initNBond()") )
rampedParams.append( MultRamp(0.9,0.8,
    "command('param nbonds repel VALUE end
    end')") )
rampedParams.append( MultRamp(.004,4,
    "command('param nbonds rcon VALUE end
    end')") )
# nonbonded interaction only between CA atoms
highTempParams.append( StaticRamp("protocol.initNBond(cutnb=100,
    rcon=0.004,
    tolerance=45,
    repel=1.2,
    onlyCA=1)""")

potList.append( XplorPot("BOND") )
potList.append( XplorPot("ANGL") )
potList['ANGL'].setThreshold( 5 )
rampedParams.append( MultRamp(0.4,1,"potList['ANGL'].setScale(VALUE)") )
potList.append( XplorPot("IMPR") )
potList['IMPR'].setThreshold( 5 )
rampedParams.append( MultRamp(0.1,1,"potList['IMPR'].setScale(VALUE)") )

# Give atoms uniform weights, except for the anisotropy axis
# protocol.massSetup()

# IVM setup
#   the IVM is used for performing dynamics and minimization in torsion-angle
#   space, and in Cartesian space.
# from ivm import IVM
dyn = IVM()

# reset ivm topology for torsion-angle dynamics
# dyn.reset()
protocol.torsionTopology(dyn)

# minc used for final cartesian minimization
# minc = IVM()
protocol.initMinimize(minc)
protocol.cartesianTopology(minc)

# object which performs simulated annealing
# from simulationTools import AnnealIVM
init_t = 2000.  # Need high temp and slow annealing to converge
cool = AnnealIVM(initTemp = init_t,
    finalTemp=25,  
    tempStep =3.25, 
    ivm=dyn,  
    rampedParams = rampedParams)

def accept(potList):
    """
    return True if current structure meets acceptance criteria
    """
    if potList['noe'].violations()>0:
        return False
    if potList['CDIH'].violations()>0:
        return False
    if potList['BOND'].violations()>0:
        return False
    if potList['ANGL'].violations()>0:
        return False
    if potList['IMPR'].violations()>1:
        return False
    return True

def calcOneStructure(loopInfo):
    """
    this function calculates a single structure, performs analysis on
    the structure, and then writes out a pdb file, with remarks.
    """
    # initialize parameters for high temp dynamics.
    InitialParams( rampedParams )
    # high-temp dynamics setup - only need to specify parameters which
    # differ from initial values in rampedParams
    InitialParams( highTempParams )

    # high temp dynamics
    #
    protocol.initDynamics(dyn,
        potList=potList,  # potential terms to use
        bathTemp=init_t,  
        initVelocities=1,
        finalTime=10,  # stops at 10ps or 5000 steps
        numSteps=5000,  # whichever comes first
        printInterval=100)

dyn.setETolerance( init_t/100 )  # used to det. stepsize. default: t/1000
    dyn.run()

    # initialize parameters for cooling loop
    InitialParams( rampedParams )
# initialize integrator for simulated annealing
#
protocol.initDynamics(dyn,
potList=potList,
numSteps=100, # at each temp: 100 steps or
finalTime=.2 , # .2ps, whichever is less
printInterval=100)

# perform simulated annealing
#
cool.run()

# final torsion angle minimization
#
protocol.initMinimize(dyn,
printInterval=50)
dyn.run()

# final all- atom minimization
#
protocol.initMinimize(minc,
potList=potList,
dEPrd=10)

minc.run()

# do analysis and write structure
loopInfo.writeStructure(potList)
pass

from simulationTools import StructureLoop, FinalParams
StructureLoop(numStructures=numberOfStructures,
pdbTemplate=outFilename,
#calcMissingStructs=True,
structLoopAction=calcOneStructure,
genViolationStats=1,
averagePotList=potList,
averageSortPots=[potList['BOND'],potList['ANGL'],potList['IMPR'],
    noe],
#    averageCrossTerms=refRMSD,
    averageTopFraction=0.017, # report only on best 50% of structs
    # averageAccept=accept, # only use structures which pass accept()
averageContext=FinalParams(rampedParams),
averageFilename="2ISPB5980_SCRIPT_ave.pdb", # generate regularized ave structure
averageFitSel="name CA",
averageCompSel="not resname ANI and not name H*" ).run()
LIST OF REFERENCES


89. Seelig, J. (1978) 31P nuclear magnetic resonance and the head group structure of phospholipids in membranes, *Biochim Biophys Acta* 515, 105-140.


BIOGRAPHICAL SKETCH

Anna Kuznetsova was born in 1984 in Stupino, Moscow region, Russia. She received her bachelor’s degree in chemistry from Moscow State University in 2006. In 2007, Anna Kuznetsova joined the graduate program in chemistry department at the University of Florida. She received her Ph.D. from the University of Florida in the spring of 2013. Anna Kuznetsova is the author of several scientific papers. She was also a TA in General Chemistry (I, II) and Biochemistry Laboratory classes.