

Conformational Dynamics of the Membrane Enzyme Lipoprotein Signal
Peptidase (LspA) and Characterization of Bicelle Membrane Mimetics

Tracy Anne Caldwell
Downingtown, PA

B.S. James Madison University, 2015

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Molecular Physiology and Biological Physics

University of Virginia
March, 2020

© Copyright by
Tracy Anne Caldwell
All rights reserved
March 2020

Abstract

With the rapid growth of antibiotic resistance, it is imperative that new drugs are developed which target novel pathways. The lipoprotein processing pathway is a novel pathway for antibiotic drug targeting as the enzymes involved, Lgt, LspA, and Lnt, are essential in some organisms including *E. coli*, *S. enterica*, *M. tuberculosis*, and *S. coelicolor*, and have no mammalian homologs. Lipoproteins are characterized by an N-terminal lipid moiety that serves as a membrane anchor, and serve a wide range of functions including in signal transduction, stress sensing, virulence, cell division, sporulation, nutrient uptake, antibiotic resistance, adhesion, and trigger the activation of host innate immune responses. If lipoproteins are not processed correctly they cannot serve these vital functions and the bacteria will be compromised. Lipoprotein signal peptidase (LspA) is an aspartyl protease that carries out the second step in the lipoprotein processing pathway - cleaving the transmembrane helical signal peptide of lipoproteins after lipidation by Lgt.

Lipoprotein signal peptidase (LspA) has been identified as an antibiotic drug target as it targets a novel pathway, is essential in some bacteria, and does not have a mammalian homolog. The crystal structure of LspA has been determined in complex with the antibiotic globomycin. However, the apo and lipoprotein substrate bound structures of LspA have remained elusive. We propose that there are conformational dynamics of LspA in which the β -cradle and PH domains “open” to allow the substrate to enter the active site, or “close” to hide the charged residues of the active site from the surrounding hydrophobic membrane. This hypothesis is investigated using electron paramagnetic resonance (EPR) studies and molecular dynamics (MD) simulations. This hybrid approach allows for visualization of structures consistent with experimental EPR restraints.

In order to efficiently develop drugs to target novel proteins, it is essential to have a biological, quantitative, reproducible, and high-throughput activity assay to test the effectiveness of the developed therapeutics. An activity assay is currently lacking for LspA. Here, a LspA activity assay is sought that will be used to gain a deeper understanding of the protein's mechanism, test requirements for LspA activity, and ultimately be used to test the efficiency of inhibitors in future antibiotic development.

Membrane proteins, including LspA, hold a wide variety of essential functions and comprise a large percentage of drug targets. However, in order to study membrane proteins *in vitro*, a membrane mimetic must be used. This membrane mimic must shield the hydrophobic transmembrane residues of the membrane protein so that it is stable and able to be studied in an aqueous environment. Here, work to better characterize bicelle systems used to study membrane proteins will be described. The classically described bicelle contains a central disk-shaped lipid bilayer encircled by a rim of detergents which screen the hydrophobic lipid tails from water. Characterization of DMPC/DHPC bicelles, including bicelle shape and lipid/detergent mixing, is investigated via SAXS, SANS, MD and fluorescence anisotropy.

Table of Contents

Copyright Page.....	ii
Abstract.....	iii
Table of Contents.....	v
List of Figures.....	xi
List of Tables.....	xvi
CHATER 1. INTRODUCTION.....	1
1.1 <i>Introduction to the Study of Proteins</i>	1
1.1.1 <u>Proteins</u>	1
1.1.2 <u>Protein Dynamics</u>	2
1.1.2.1 Protein Fluctuations.....	4
1.1.2.2 Conformational Dynamics.....	5
1.1.3 <u>Membrane Proteins</u>	5
1.1.4 <u>Membrane Mimetics</u>	9
1.1.4.1 Micelles.....	11
1.1.4.2 Bicelles.....	15
1.1.4.3 Other membrane mimetics.....	18
1.2 <i>Lipoproteins and the Lipoprotein Processing Pathway</i>	19
1.2.1 <u>Lipoproteins</u>	19
1.2.2 <u>Processing of Lipoproteins</u>	23
1.2.3 <u>LspA Structure and Function</u>	27
1.3 <i>Significance of Research</i>	31

	vi
1.3.1 <u>Rise of Antibiotic Resistance</u>	31
1.3.2 <u>Lipoprotein Processing Pathway as an Antibiotic Drug Target</u>	33
1.3.3 <u>Elucidating the Mechanism of LspA</u>	34
1.3.4 <u>Effects of Membrane Mimetic</u>	38
1.4 <i>Thesis Overview</i>	39
1.5 <i>References</i>	39
CHAPTER 2. ELECTRON PARAMAGNETIC RESONANCE SPECTROSCOPY.....	56
2.1 <i>Introduction</i>	56
2.2 <i>Biophysical Techniques used to Study Membrane Proteins</i>	56
2.2.1 <u>Common Biophysical Techniques used to Study Membrane Proteins</u>	58
2.2.1.1 X-ray Crystallography.....	58
2.2.1.2 Cryo-Electron Microscopy.....	58
2.2.1.3 Small Angle Scattering.....	59
2.2.1.4 Nuclear Magnetic Resonance.....	59
2.2.1.5 Förster Resonance Energy Transfer.....	60
2.2.2 Advantages of using EPR to Study Membrane Proteins.....	61
2.3 <i>Primer on EPR Theory</i>	64
2.3.1 <u>Free Electron in a Magnetic Field</u>	64
2.3.2 <u>Hyperfine Interaction</u>	67
2.3.3 <u>Anisotropy</u>	67
2.3.4 <u>Relaxation</u>	68
2.3.5 <u>Correlation Time</u>	69

	vii
2.4 <i>Applications of EPR</i>	70
2.4.1 <u>Site-Directed Spin Labeling</u>	70
2.4.2 <u>Continuous Wave EPR</u>	73
2.4.2.4 How to “Read” a CW Spectrum.....	73
2.4.3 <u>Double Electron Electron EPR</u>	79
2.4.3.1 Dipolar Coupling.....	79
2.4.3.2 Pulsed EPR.....	82
2.4.3.3 Distance Analysis.....	85
2.5 <i>Practical Considerations for EPR Experimentation</i>	90
2.5.1 <u>Sample Preparation</u>	90
2.5.2 <u>CW Spectral Analysis</u>	95
2.5.3 <u>DEER Spectral Analysis</u>	96
2.6 <i>References</i>	100
CHAPTER 3. CONFORMATIONAL DYNAMICS OF LspA.....	108
3.1 <i>Introduction</i>	108
3.1.1 <u>Project Motivation</u>	108
3.1.2 <u>Support of Proposed Conformational Dynamics</u>	112
3.2 <i>Electron Paramagnetic Resonance Studies</i>	116
3.2.1 <u>Continuous Wave EPR Studies</u>	116
3.2.1.1 CW in FC12 Micelles.....	116
3.2.1.2 CW with the Addition of Lipids.....	122
3.2.2 <u>Double Electron Electron Resonance Studies</u>	128

	viii
3.2.2.1 I43R1/A63R1 in FC12 Micelles.....	128
3.2.2.2 M36R1/A63R1 in FC12 Micelles.....	133
3.2.2.3 I43R1/A63R1 in DMPC/FC12 Bicelles.....	137
3.3 <i>Molecular Dynamics Simulations</i>	142
3.4 <i>Comparison of EPR and MD Studies</i>	146
3.4.1 <u>Comparing distance distributions</u>	146
3.4.2 <u>MMM analysis of MD conformations</u>	147
3.4.3 <u>Generating model structures</u>	157
3.5 <i>Conclusions</i>	166
3.6 <i>Methods</i>	166
3.6.1 <u>Expression, purification, and spin labeling of LspA</u>	166
3.6.2 <u>Continuous-wave EPR</u>	168
3.6.3 <u>Pulsed-EPR</u>	168
3.6.4 <u>Molecular Dynamics Simulations</u>	169
3.6.5 <u>Evolutionary Couplings</u>	170
3.7 <i>References</i>	170
CHAPTER 4. LspA ACTIVITY ASSAY.....	174
4.1 <i>Introduction</i>	174
4.1.1 <u>Project Motivation</u>	174
4.1.2 <u>Previous Work</u>	174
4.1.2.1 Initial LspA Studies.....	174
4.1.2.2 Gel Based Assay.....	176

4.1.2.3 Förster Resonance Energy Transfer Assay.....	181
4.2 <i>Attempts at Optimizing Activity Assay</i>	187
4.2.1 <u>Repeating Published Assays</u>	187
4.2.1.1 Gel and Western Assay.....	187
4.2.1.2 Förster Resonance Energy Transfer Assay.....	195
4.2.2 <u>Developing a New LspA Assay</u>	201
4.2.2.1 Fluorescence Polarization Assay.....	201
4.2.2.2 Förster Resonance Energy Transfer Assay with Signal Peptide...	207
4.2.2.3 Electron Paramagnetic Resonance Assay.....	209
4.3 <i>Conclusions</i>	214
4.4 <i>Methods</i>	214
4.4.1 Protein Expression and Purification.....	214
4.4.2 Gel-Based Assay.....	216
4.4.3 FRET Assay.....	217
4.4.4 New FRET Assay.....	218
4.4.5 EPR Assay.....	218
4.5 <i>References</i>	220
CHAPTER 5. BICELLE PROPERTIES.....	228
5.1 <i>Introduction</i>	228
5.1.1 <u>Project Motivation</u>	228
5.1.2 <u>Previous Work: SAXS</u>	232
5.2 <i>SANS</i>	234

	x
5.2.1 <u>SANS Methods</u>	234
5.2.2 <u>SANS Contrast Variation</u>	234
5.2.3 <u>SANS Fitting Approach</u>	237
5.2.4 <u>SANS Results</u>	243
5.3 <i>Collaborations</i>	249
5.3.1 <u>Molecular Dynamics</u>	249
5.3.1.1 MD Methods.....	249
5.3.1.2 MD Results.....	250
5.3.2 <u>Fluorescence Anisotropy</u>	254
5.3.2.1 Anisotropy Methods.....	254
5.3.2.2 Anisotropy Results.....	255
5.4 <i>Conclusions</i>	258
5.4.1 <u>Bicelle Shape</u>	258
5.4.2 <u>Lipid-Detergent Mixing</u>	260
5.4.3 <u>Overall Conclusions</u>	260
5.5 <i>References</i>	261

List of FiguresChapter 1: Introduction

1.1	Timescale of protein motions.....	3
1.2	Membrane protein classification.....	8
1.3	Schematic of micelles and bicelles.....	10
1.4	Types of detergents.....	13
1.5	Micelle shapes.....	14
1.6	Schematic of bicelle formation.....	17
1.7	Lipoprotein localization.....	21
1.8	Schematic of pre-lipoprotein signal peptide.....	22
1.9	Lipoprotein processing pathway.....	26
1.10	Structure of LspA.....	29
1.11	Hypothesized aspartyl-protease cleavage mechanism of LspA.....	30
1.12	Classification of membrane proteins of known structure.....	36
1.13	iCLiP membrane topology compared to LspA.....	37

Chapter 2: EPR

2.1	Biophysical techniques used to study membrane proteins.....	57
2.2	Energy diagram for nitroxide spin label in a magnetic field.....	66
2.3	Nitroxide spin label reaction.....	72
2.4	Mobile, semi-restricted, and immobile CW spectra.....	76
2.5	Characterization of CW lineshape.....	77
2.6	Two-component CW spectra resulting from conformational dynamics.....	78

	xii
2.7 Orientation of spins.....	81
2.8 Four-pulse DEER experiment.....	84
2.9 Effect of average radius and distribution width on DEER signal.....	87
2.10 Multi-component DEER.....	88
2.11 Tikhonov regularization L-curve.....	89
2.12 DMSO addition causes significant CW lineshape changes.....	94
2.13 Spin label rotamers.....	99

Chapter 3: LspA Conformational Dynamics

3.1 Hypothesized LspA conformational dynamics.....	111
3.2 LspA evolutionary couplings.....	113
3.3 Evolutionary couplings between the β -cradle and PH.....	114
3.4 LspA conformational dynamics is supported by crystal structures.....	115
3.5 CW spectra.....	118
3.6 A63R1 CW spectra.....	119
3.7 Fit of A63R1 CW spectra.....	120
3.8 Location of spin label site in <i>S. aureus</i> LspA.....	121
3.9 CW spectra with the addition of lipid.....	124
3.10 DMPC/FC12 CW titration.....	125
3.11 DMPC/FC12 SAXS data.....	126
3.12 CW of DMPC/FC12 LspA with globomycin.....	127
3.13 I43R1/A63R1 DEER analysis.....	130
3.14 I43R1/A63R1 LspA MMM simulated spectra.....	131

	xiii
3.15 M36R1/A63R1 DEER analysis.....	135
3.16 M36R1/A63R1 MMM simulated spectra.....	136
3.17 I43R1/A63R1 DEER analysis in DMPC/FC12 bicelles.....	139
3.18 I43R1/A63R1 DEER in micelles compared to bicelles.....	140
3.19 I43R1/A63R1 MMM simulated spectra compared to DEER in DMPC/FC12 bicelles.....	141
3.20 MD distance measurements.....	143
3.21 MD distance histograms.....	144
3.22 Combined MD histograms.....	145
3.23 MD analysis workflow.....	150
3.24 I43R1/A63R1 apo MD MMM analysis compared to micelle DEER data.....	151
3.25 I43R1/A63R1 apo MD MMM analysis compared to bicelle DEER data.....	152
3.26 I43R1/A63R1 globomycin MD MMM analysis compared to micelle DEER data.....	153
3.27 I43R1/A63R1 globomycin MD MMM analysis compared to bicelle DEER data.....	154
3.28 MD MMM distributions fit to bicelle DEER data.....	155
3.29 M36R1/A63R1 MD MMM analysis.....	156
3.30 MD derived model structures.....	161
3.31 Surface representation of structures.....	162
3.32 Protein electrostatics.....	163
3.33 Cavity fill models.....	164
3.34 Spin labeled LspA.....	165

Chapter 4: LspA Activity Assay

4.1	Schematic of gel-based LspA assay.....	180
4.2	Schematic of FRET-based LspA assay.....	184
4.3	Gel-based LspA assay.....	191
4.4	Gel-based LspA assay western blot.....	192
4.5	Schematic of ICP constructs.....	193
4.6	Fluorescent gel-based assay with FP-ICP construct.....	194
4.7	Emission and excitation spectra of FRET probe.....	198
4.8	FRET assay in FC12 and DDM.....	199
4.9	Analyzing DDM FRET assay.....	200
4.10	FIAsH-EDT ₂ coordination.....	203
4.11	Schematic of FP assay.....	204
4.12	Emission and excitation spectra of FP probe.....	205
4.13	FP assay results.....	206
4.14	Schematic of new FRET assay.....	208
4.15	Schematic of DEER assay.....	212
4.16	Unnatural amino acid incorporation.....	213

Chapter 5: Bicelle Properties

5.1	Detergent, lipid, and bicelle structures.....	231
5.2	SAXS measurements of q=0.1 to 1 bicelles.....	233
5.3	Schematic of SANS contrast variation.....	236
5.4	Schematic of core-shell bicelle model.....	239

	xv
5.5 SANS core-shell bicelle model fits for $q=0.7$	241
5.6 SANS core-shell bicelle model fits for $q=0.3$	242
5.7 Snapshots from MD bicelle simulations.....	252
5.8 Fluorescence anisotropy of bicelles.....	257

List of TablesChapter 1: Introduction

1.1 Projected deaths attributable to antimicrobial resistance per year by 2050.....	32
---	----

Chapter 2: EPR

2.1 Structure and dynamics probed by EPR.....	63
---	----

Chapter 3: LspA Conformational Dynamics

3.1 Percent of rotamers in fit MMM distributions.....	132
---	-----

Chapter 4: LspA Activity Assay

4.1 LspA activity values determined from assays.....	185
4.2 IC ₅₀ values determined from assays.....	186

Chapter 5: Bicelle Properties

5.1 Scattering length density for different bicelle components.....	240
5.2 Fit parameters for the SANS profiles of q=0.7 bicelles.....	245
5.3 Fit parameters for the SANS profiles of q=0.3 bicelles.....	246
5.4 Comparison of q=0.3 and q=0.7 bicelles.....	247
5.5 Match point comparison for tested bicelles.....	248
5.6 Bicelle properties from MD simulations.....	253
5.7 Comparison of bicelle methods.....	259

CHAPTER 1. INTRODUCTION

1.1 Introduction to the Study of Proteins

1.1.1 Proteins

Proteins are highly diverse biomolecules which sustain all living organisms. They have a broad range of functions including maintaining structural stability, producing energy, catalyzing reactions, and sending and receiving signals in the cell. Thus, many drug therapies target specific proteins. The human proteome contains approximately 20,000 proteins, but accounting for alternative splicing and polymorphisms, scientists estimate that there may be up to 6 million unique human protein species.¹ In humans, each cell contains a specific variety of approximately 2 billion protein molecules.² Thus, to understand the human body and to target human diseases, it is imperative for researchers to determine the structure and function of each of these proteins.

Proteins are made up of a specific combination of amino acids connected through peptide bonds. The amino acid sequence, or primary structure, of a protein is dictated by a DNA sequence encoded in the organism's genome. Sequential amino acids in a protein can fold into secondary structural elements called α -helices and β -sheets. The subsequent folding of these structural elements determines the proteins overall three-dimensional structure or tertiary structure. The orientation of multiple bound proteins defines the quaternary structure. Understanding a protein's structure is critical to understanding the protein's biological function. X-ray crystallography, nuclear magnetic resonance (NMR), and more recently cryo-electron microscopy (cryo-EM) are common techniques used to determine high resolution protein structure.

1.1.2 Protein Dynamics

In addition to protein structure, it is essential to understand a protein's movements or dynamics to understand its function and mechanism of action. Protein dynamics are dictated by the probability of distinct bond orientations or conformational states and the energy barriers between these states. These phenomena are understood by thermodynamics and kinetics, respectively. Protein dynamics can be divided into two classes: protein fluctuations and conformational dynamics (Figure 1.1).³

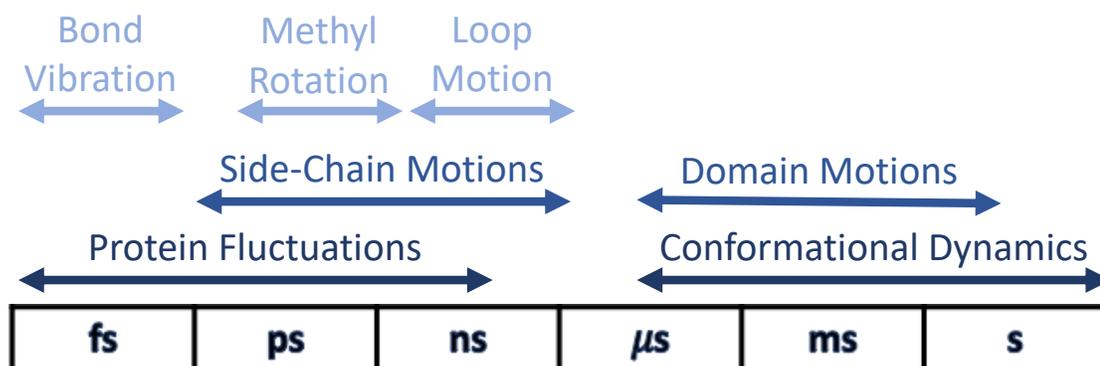


Figure 1.1. Timescale of protein motions. Adapted from Henzler-Wildman and Kern (2007), with permission.³

1.1.2.1 Protein Fluctuations

Protein motions that occur on a ‘fast’ timescale and result in structurally similar states define protein fluctuations. These movements have low ($1 kT$) energy barriers and occur within the energy wells of each conformational dynamic state. Backbone and loop motions occur on a nanosecond timescale, sidechain rotations occur on a picosecond timescale, and bond vibrations occur on a femtosecond timescale (Figure 1.1).³ The overall size and structure of a protein, as well as its conformational dynamics, influence the timescale on which these motions occur.⁴ Protein fluctuations are important for the overall stability of the protein and for ligand and protein interactions.⁴ Fluctuations in a protein binding pocket allow the protein to complement the shape of the binding partner and facilitates efficient binding.^{3,4}

Atomic displacement can be analyzed through X-ray diffraction data and is expressed as the B factor.^{3,5} However, NMR relaxation data is most commonly used to probe dynamics on these timescales and is reported as an order parameter (S^2).⁶⁻⁹ Molecular dynamics simulations are also a strong technique for studying protein fluctuations as they are usually completed at this timescale and can give insight to the local energies and interactions in the protein.³ Continuous wave electron paramagnetic resonance (CW-EPR) can be used to measure local motions by reporting the rotational correlation time of the spin label in the ps to ns timescale.¹⁰⁻¹³ CW-EPR will be further discussed in Chapter 2.

1.1.2.2 Conformational Dynamics

Conformational dynamics are structural rearrangements of a protein into different distinct states. These states are separated by large energy barriers (several kT), and are therefore relatively rare and occur on a ‘slow’ timescale from microseconds to seconds (Figure 1.1).³ Most conformational dynamics occur between protein domains and are described as hinge or shear motions.¹⁴ Many important biological processes occur by conformational dynamics including enzyme catalysis, signal transduction, and protein interactions.³

Different conformational dynamic states can be determined by X-ray crystallography or cryogenic electron microscopy (cryo-EM) by trapping one state at a time. Small-angle X-ray scattering (SAXS) or small-angle neutron scattering (SANS) can give information about the ensemble of states, although in lower resolution. NMR gives both structural information about different states as well as information about the timescale of dynamics through relaxation experiments. Probe based experiments such as fluorescence resonance energy transfer (FRET) and double electron electron resonance (DEER) can also give information about various conformational states by determining a distance distribution between different points in the protein.^{3,10,14,15} DEER spectroscopy will be further discussed in Chapter 2.

1.1.3 Membrane Proteins

Proteins are classified as soluble or membrane proteins. Soluble proteins reside in the aqueous environment of the cell and fold such that their hydrophobic residues are buried in the core of the protein excluded from the polar environment. Membrane proteins,

on the other hand, bury their hydrophobic residues within a lipid bilayer. This thesis is focused on the structure and function of membrane proteins.

Membrane proteins make up about 20-30% of an organism's coding genome and serve a wide range of functions including as receptors, transporters, and enzymes.¹⁶ Over half of current pharmaceuticals target membrane proteins, yet they constitute less than 2% of known protein structures in the protein data bank (PDB).^{17,18} This disparity comes from the difficulty in studying membrane proteins due to the need to reconstitute the protein in a membrane mimetic.¹⁹ For example, the first soluble protein structure of myoglobin was determined by X-ray crystallography in 1958, but it took until 1985 for the first membrane protein structure to be determined.^{20,21} However, recent advances in membrane protein expression and purification, crystallography, NMR, and cryo-EM techniques have resulted in an accelerated pace for membrane protein structural and functional characterization in recent years.²²⁻²⁴

Membrane proteins can be further classified as 1) integral membrane proteins or 2) membrane associated or peripheral membrane proteins (Figure 1.2).

Integral membrane proteins completely transverse the lipid bilayer and serve a variety of functions including in ion transport, cell signaling, nutrient intake, and as membrane bound enzymes. β -barrel proteins are integral membrane proteins made up of 8-26 antiparallel β -strands that twist to form a closed barrel structure, most with hydrophobic residues positioned toward the membrane and hydrophilic residues toward the interior.^{18,25-27} They are only found in the outer-membranes of gram-negative bacteria, mitochondria, and chloroplasts.²⁵ The majority of integral membrane proteins (~83%) are made of one (bitopic) or more (polytopic) α -helices. These proteins generally have

hydrophobic residues in the transmembrane region and hydrophilic residues outside of the membrane. Lipoprotein signal peptidase (LspA), the subject of this thesis described in Chapters 3 and 4, is an example of a polytopic α -helical integral membrane protein.²⁸

Peripheral membrane proteins are proteins which interact with the membrane, but do not completely transverse it. Such membrane interactions could be through an amphipathic α -helix perpendicular the to the membrane, an electrostatic or ionic interaction with the membrane headgroups, or by an interaction with a covalently bound lipid. Peripheral membrane proteins play important physiological roles including in bacterial virulence, nutrient uptake, cell signaling, antibiotic resistance, and triggering innate immune responses.^{29,30} Lipoproteins, the substrates of LspA described in greater detail in the section 1.2.1, are peripheral membrane proteins anchored to the membrane by a lipidation modification.³¹

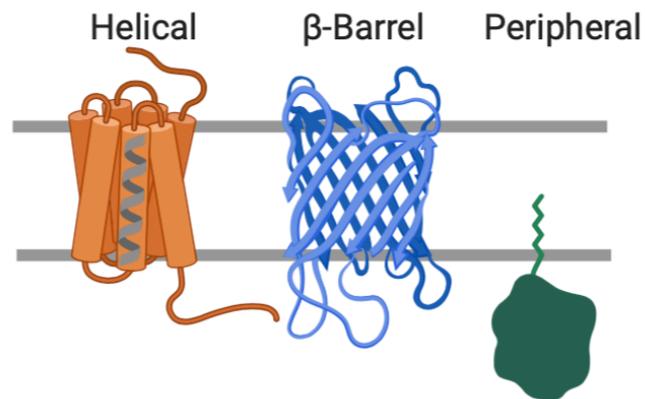


Figure 1.2. Membrane protein classification. Membrane proteins can be classified as integral membrane proteins which span the membrane (depicted by gray lines) and can be helical (left, orange) or β -barrel (middle, blue), or peripheral membrane proteins which are associated with the membrane, but do not span it (right, green).

1.1.4 Membrane Mimetics

When membrane proteins are studied, a membrane mimetic must be used to recapitulate the membrane environment. This membrane mimic must shield the hydrophobic transmembrane residues of the membrane protein so that it is stable and able to be studied in an aqueous environment. Membrane mimetics include micelles, bicelles, nanodiscs, amphipols, lipodisc (or SMALPs), and liposomes.³² Each mimetic has their own advantages and disadvantages, but the consequences of the mimetic on protein fold and function is still largely unknown. We will focus on micelles and bicelles as they were used in this work and will be discussed in Chapters 3-5 (Figure 1.3).

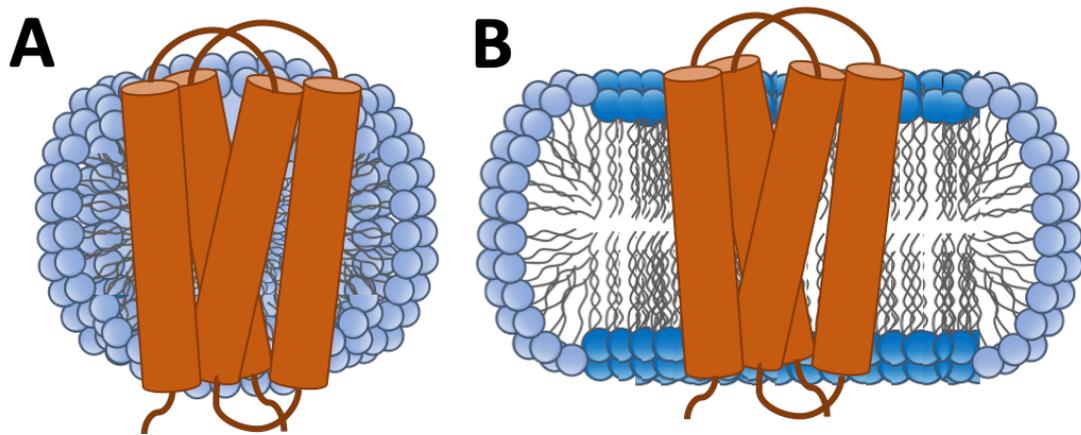


Figure 1.3. Schematic of micelles and bicelles. Micelles (A) are made of detergents (light blue headgroups and gray acyl tails) while bicelles (B) are made up of a bilayer of lipids (dark blue headgroups and gray acyl tails) surrounded by detergent. Representative membrane protein shown in orange.

1.1.4.1 Micelles

The simplest, though not simple at all, membrane mimetic is a detergent micelle. A detergent molecule is amphipathic with a polar headgroup and a hydrophobic hydrocarbon tail. Detergents self-assemble into micelles when their concentration is at or above the detergent-specific critical micelle concentration (CMC). The hydrophilic headgroups form the solvent-exposed outer shell of the micelle while the hydrophobic tails form the middle core of the micelle (Figure 1.3A). When membrane proteins are solubilized in detergent micelles their hydrophobic transmembrane regions reside in the micelle surrounded by the detergent tails, and their hydrophilic loops extend into the surrounding aqueous environment (Figure 1.3A).^{32,33}

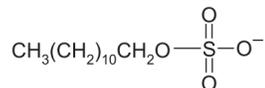
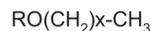
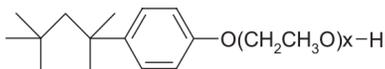
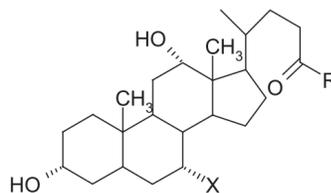
Detergents can be categorized as ionic, bile acid salts, nonionic, or zwitterionic (Figure 1.4). Ionic detergents have a charged headgroup which can be either positive (cationic) or negative (anionic). Ionic detergents in general are good at solubilizing membrane proteins but can be denaturing.¹⁹ Bile acid salts are also ionic, but they are composed of rigid steroidal groups. Nonionic detergents have uncharged headgroups and are thought to be a more mild detergent. Lastly, zwitterionic detergents contain both a positive and negative charge in their headgroups for a net neutral charge.¹⁹ Additionally, detergents can have various numbers of carbons in their tails. The micelle used most in this work is composed of dodecylphosphocholine (DPC, FC12) detergent which has a zwitterionic phosphocholine headgroup and a single 12-carbon tail.

Characterization of detergent aggregates was pioneered by Charles Tanford in the 1970s.³⁴ The CMC value, aggregation number (N, number of detergent monomers per micelle), and properties of the micelle are dependent upon the properties of the detergent

headgroup and tail.^{19,33,35,36} For example, for the same headgroup, detergents with longer tails have smaller CMC values, larger N values, longer headgroup to headgroup distances, larger volumes, and form more stable micelles.^{32,34,35} Specifically, the volume (V) increases with the number of alkyl chain carbons (n_c) according to Tanford's formula: $V=27.4+26.9*n_c$.³⁴ Properties of the headgroup dictate micelle shape and ellipticity. Negatively charged head groups such as lysophosphatidyl glycerols form spherical micelles, non-ionic headgroups such as glucosides and maltosides form oblate micelles, and zwitterionic headgroups such as phosphocholines form prolate micelles (Figure 1.5).³⁵ Mixed micelles, which contain two or more detergents, can also be used. Micelle properties such as size scale linearly with detergent mole ratio, so in theory micelles can be made for specific requirements.³⁷

a.**Ionic detergents**

Sodium dodecyl sulfate (SDS)

**c.****Non-Ionic Detergents**R = glucose, x = 7, n-octyl- β -D-glucopyranosideR = maltose, x = 9, decyl- β -D-maltosidex = 11, dodecyl- β -D-maltosidex = 9, Triton[®] X-100x = 7-8, Triton[®] X-114**b.****Bile Acid Salts**X=H, R = O-Na⁺, sodium deoxycholateX=OH, R = O-Na⁺, sodium cholate**d.****Zwitterionic Detergents**

x = H, CHAPS

x = OH, CHAPSO

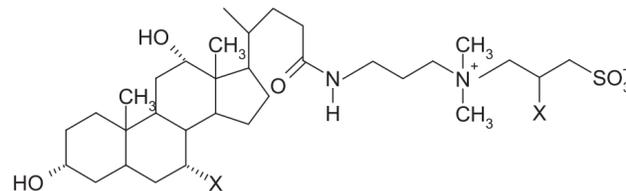


Figure 1.4. Types of detergents. A) Ionic detergents contain a charge. B) Bile acid salts also are charged and have rigid steroid groups. C) Non-ionic detergents are not charged. D) Zwitterionic detergents contain both a positive and negative charge and are thus net neutral. Figure reprinted from Seddon *et al.* (2004), with permission.

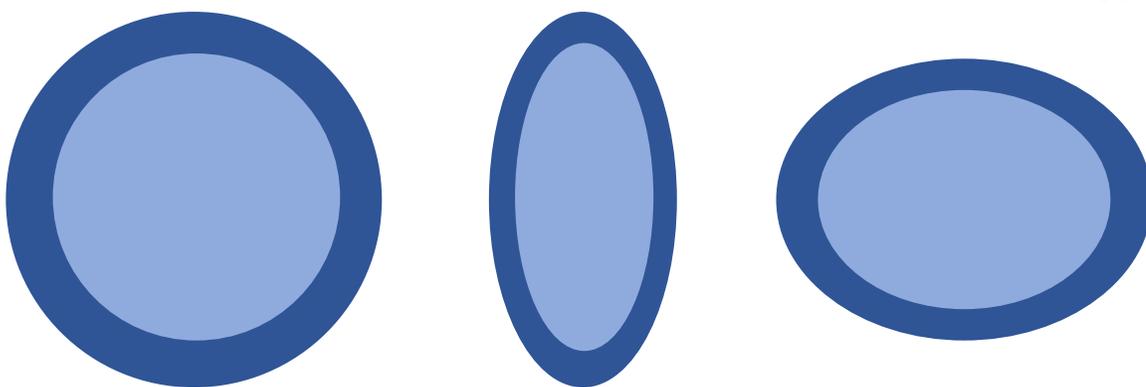


Figure 1.5. Micelle shapes. Micelles can form spherical (right), prolate (middle), or oblate (left) structures. The darker blue represents the outer shell where the headgroups reside and the light blue represents the core where the tails reside.

1.1.4.2 Bicelles

Bicelles are formed by a mixture of lipid and detergent molecules. Lipids occupy a cylindrical volume while detergents occupy a conical volume (Figure 1.6). Thus, when lipids and detergents self-associate they form a disk-like planer core of lipid surrounded by a rim of detergent molecules (Figure 1.3). This phenomena was first reported in 1984,^{38,39} and since bicelles have been structurally characterized by methods such as small angle neutron scattering and NMR.^{40,41,50–54,42–49} One such investigation will be presented in Chapter 5.

Bicelles are characterized by the ratio of lipid to detergent monomers which is called the q-value. The size and shape of bicelles is dependent on the q-value, structure of the detergent and lipid components, concentration of amphiphiles, and experimental temperature.^{42,44,46,47,54,55} The most commonly used bicelle is made of dimyristoylphosphatidylcholine (DMPC) long chain lipid, and dihexanoylphosphatidylcholine (DHPC) short chain lipid (which mimics a detergent in its conical shape).^{32,33,56} Other detergents commonly used to form bicelles include CHAPS, CHAPSO, and Cyclofos-6.^{32,47,57,58} PEGylated lipids or cholesterol have also been added to modulate lipid diffusion in the bicelle.^{57,59}

Bicelles with low q-values (< 0.7 ; also known as fast-tumbling “isotropic” bicelles) have been used to study polytopic integral membrane proteins by NMR.^{41,45,50,60} Several of these studies suggest that the stabilization of membrane protein fold is due to the more “bilayer” nature of bicelles compared to micelles. That is, the segregated lipid core in bicelles is more similar in structure to the native membrane. The bilayer nature of bicelles is still under debate,⁴⁰ but bicelles are seen to improve sample stability of membrane

proteins and in some cases increased protein activity compared to micelles.^{32,61,62} Detergent
– lipid segregation of low-q DMPC/DHPC bicelles is examined in detail in Chapter 5 to
determine if these bicelles are truly “bilayer like”.⁴⁰

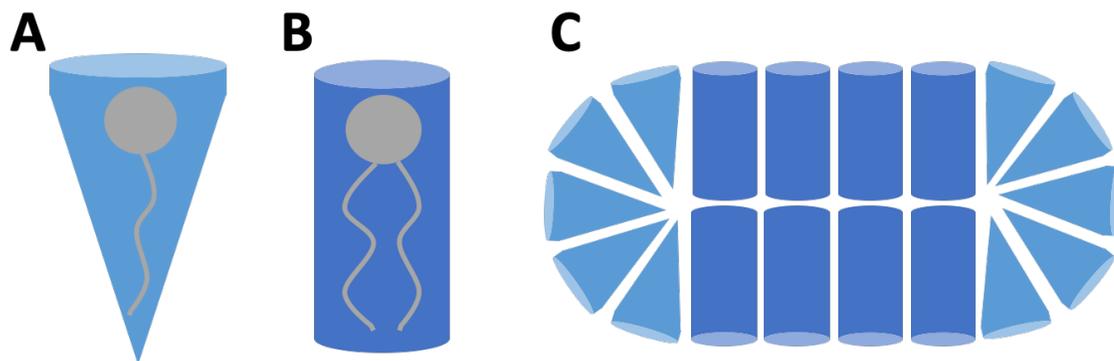


Figure 1.6. Schematic of bicelle formation. Detergents (A) occupy a conical volume while lipids (B) occupy a cylindrical volume. When detergents and lipids self-associate into an ideal bicelle (C) they form a planer core of lipids surrounded by a ring of detergent molecules.

1.1.4.3 Other membrane mimetics

Nanodiscs are made up of phospholipids which assemble into a bilayer-like domain, surrounded by two membrane scaffolding proteins (MSPs) which wrap around and shield the hydrophobic regions of the lipids.^{63–65} The choice of MSP used dictates the size (diameter) of the nanodisc which ranges from 8 to 16 nm, with the most commonly used around 10 nm.^{32,65,66} Various lipids or lipid mixtures can be used including DMPC, *E.coli* lipids, and egg phosphocholine lipids.⁶⁵ Nanodiscs are used in a wide variety of membrane protein studies including in NMR, fluorescence polarization studies, and other assays such as pull-downs or immunoprecipitation.⁶⁶

Amphipols are amphiphathic polymers that surround a membrane protein's hydrophobic transmembrane region.^{32,33,67–69} Different amphipols have various attached hydrophobic moieties including carboxylate, sulfonate, and glucose groups.⁶⁷ The advantage to using amphipols is that detergents are not used which can be denaturing to some proteins. Thus, amphipols are thought to be more stabilizing for membrane proteins and may allow for the retainment of bound native lipids.^{32,33,67,68} Amphipols are useful in biological applications where detergents are otherwise problematic such as in isothermal titration calorimetry (ITC) or surface plasmon resonance (SPR), and are also useful for many other techniques used to determine structural and functional information about membrane proteins.^{67,70}

The most recent development in membrane protein mimetics are styrene-maleic acid (SMA) copolymers. SMA copolymers surround and stabilize patches of lipid bilayers to form lipodisks or SMA lipid particles (SMALPs). SMA copolymers are similar to MSPs, but they are advantageous because they do not require the use of detergents and can extract

membrane proteins directly from biological membranes.^{32,68,71,72} As SMALPs are quite new, their applications are still being explored.

Liposomes are spherical lipid bilayer vesicles. Different types of lipids can be used to modulate the charge, hydrophobic thickness, and size of the liposome.⁷³ Extrusion can also be used to control for size homogeneity.^{19,32} While liposomes are a good mimic of the native membrane, even the smallest liposomes (called small unilamellar vesicles, SUVs) are very large compared to the other mimetics, with a diameter around 50 nm and a weight of around 17 MDa. Therefore, liposomes cannot be successfully used with techniques such as solution NMR.^{32,73} Liposomes are widely used in a variety of biophysical techniques to study membrane proteins including in EPR and fluorescence experiments.¹⁰ Liposomes can also be used as a drug delivery system.⁷⁴

1.2 Lipoproteins and the Lipoprotein Processing Pathway

1.2.1 Lipoproteins

Lipoproteins are characterized by an N-terminal lipid modification that serves as a membrane anchor.⁷⁵ Lipoproteins in gram-positive bacteria reside in the cytoplasmic membrane facing the exterior, while lipoproteins in gram-negative bacteria can reside in the inner membrane facing the periplasm or the outer membrane facing the periplasm or on the cell surface (Figure 1.7).⁷⁵ When lipoproteins are synthesized (termed prelipoproteins prior to processing) they contain a signal peptide consisting of a positive n-region, a hydrophobic h-region, and a four amino-acid lipobox (or c-region) with the consensus amino acid sequence [LVI][ASTVI][GAS][C] (Figure 1.8). The last cysteine residue of the lipobox becomes dagylated during lipoprotein processing resulting in the

mature lipoprotein (see section 1.2.2). The amino acids directly following the lipobox (+2 to +4 with the lipobox cysteine being +1) act as a sorting signal for the final destination of the lipoprotein. Lipoproteins serve a wide range of vital functions including in bacterial envelope biogenesis and maintenance, signal transduction, stress sensing, virulence, cell division, sporulation, nutrient uptake, antibiotic resistance, and adhesion.^{29,75} Lipoproteins also act in triggering the activation of host innate immune responses.²⁹

Pseudomonas aeruginosa, specifically, has approximately 175 lipoproteins with a wide variety of functions, many of which contribute to its pathogenicity.⁷⁶ Interestingly, during *P. aeruginosa* mucoid conversion in the lungs of cystic fibrosis patients (the state which causes chronic infections), there is a significant induction of lipoprotein-encoding genes, suggesting that lipoproteins may play a significant role in the progression of cystic fibrosis infection.⁷⁶ The lipoprotein and lipoprotein processing proteins in this work are from *P. aeruginosa*.

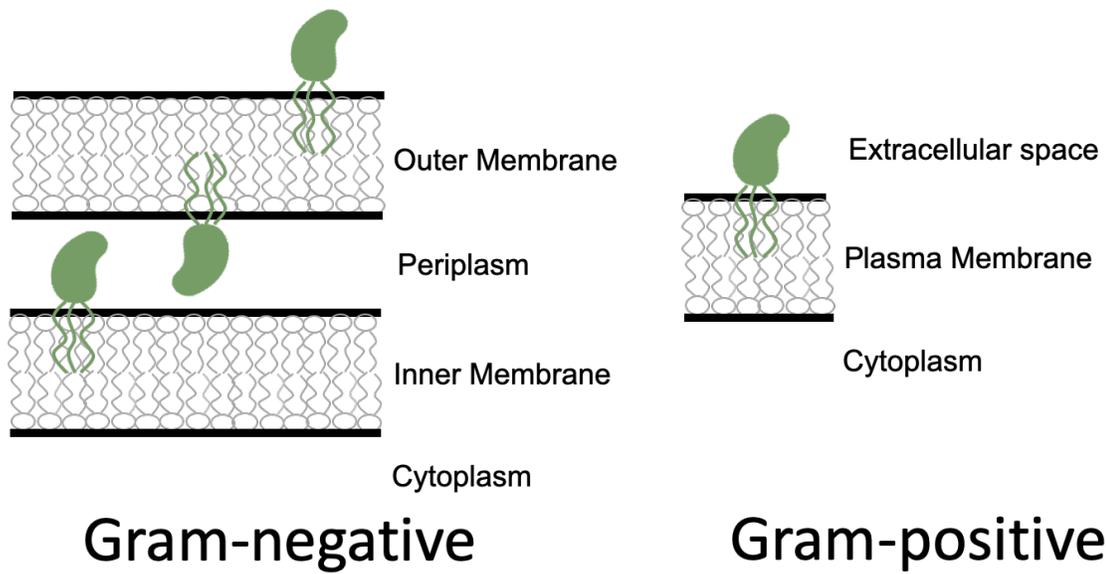


Figure 1.7. Lipoprotein localization. Lipoproteins (green) can reside in either face of the outer membrane or the periplasmic face of the inner membrane in gram-negative bacteria (left). In gram-positive bacteria they are retained in the extracellular face of the inner membrane (right).

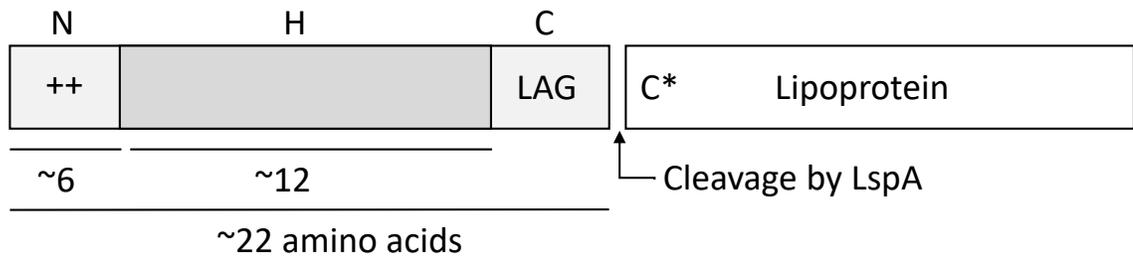


Figure 1.8. Schematic of pre-lipoprotein signal peptide. Lipoprotein signal peptides contain a positive N-region, a hydrophobic H-region, and a conserved lipobox or C-region. The dagylated cysteine is represented by C*.

1.2.2 Processing of Lipoproteins

Preprolipoproteins (lipoproteins with their signal peptides) are synthesized in the cytoplasm and then translocated across the cytoplasmic membrane via the general secretory (Sec) or twin-arginine translocation (TAT) pathways.^{75,77-79} They then undergo modifications by three membrane enzymes in the inner membrane to become mature lipoproteins (Figure 1.9).

First, preprolipoprotein diacylglyceryl transferase (Lgt) catalyzes the attachment of a diacylglycerol molecule (DAG), derived from phosphatidylglycerol (PG), to the cysteine of the lipobox through a thioester bond converting the preprolipoprotein to a prolipoprotein.^{31,75,79-83} The lipid modification secures the prolipoprotein to the membrane. The crystal structure of Lgt from *Escherichia coli* has been determined both in the presence of an inhibitor (palmitic acid) and the substrate (PG). The active site residues were determined by mutagenesis studies, and a front and side cleft through which the substrates PG and the lipobox-containing lipoprotein enter Lgt were identified.⁸⁴ Various donor lipids were tested using a GFP-lipoprotein probe. Dipalmitoyl phosphatidylglycerol (DPPG) was the most effective donor followed by dipalmitoyl phosphatidic acid (DPPA) and dipalmitoyl phosphatidyl-serine (DPPS), which are all negatively charged lipids. Neutral phospholipids were not Lgt substrates, suggesting the polar headgroup plays a role in Lgt recognition.^{84,85}

Second, prolipoprotein signal peptidase (also known as signal peptidase II, Lsp, or LspA) cleaves the signal peptide at the dagylated cysteine to form an apolipoprotein.^{28,80,86,87} Initial biochemical studies of LspA were performed in the early 1980s, but the crystal structure of LspA from *Pseudomonas aeruginosa* and

Staphylococcus aureus were recently determined in 2016 and 2020, respectively.^{28,31,88-91}

LspA is the main focus of this thesis work described in Chapters 3 and 4.

Finally, some apolipoproteins are further processed by N-acyltransferase (Lnt) which N-acylates the dagylated cysteine using phosphatidylethanolamine (PE) as substrate, and forms a mature triacylated lipoprotein.⁹² Gram-positive bacteria lack Lnt. The crystal structure of Lnt from both *Pseudomonas aeruginosa* and *Escherichia coli* (which have 39% sequence identity) were determined and are very similar in structure (RMSD of 1.2 Å).⁹² Active site residues have been identified and reside above the membrane where the membrane domain and periplasmic nitrilase-like domains meet. Additionally, lipids bound in the crystal structure suggest the substrate binding site.⁹² A ping-pong mechanism has been proposed in which the acyl chain is transferred to the catalytic cysteine of Lnt and subsequently transferred to the dagylated N-terminal cysteine of the lipoprotein, with the second step of the reaction being faster than the first.^{93,94} Thus, the first product leaves the active site before the second substrate enters as suggested by one proposed binding site.⁹²⁻⁹⁵ At this point, it is unknown why some lipoproteins are substrates for Lnt while others are not, as lipoproteins can be either di- or tri-acylated.⁹⁶ Lipoproteins in a small subset of bacteria including Mycobacteria are often also glycosylated by Ppm synthase, and the actions of Lnt and Ppm synthase are believed to be tightly correlated.⁹⁷⁻⁹⁹ It has been suggested that gene expression and thus activity of lipoprotein-modifying enzymes (Lgt, LspA, Lnt, and Ppm synthase) are regulated depending on the environmental conditions or developmental stages of the bacteria, as various lipid modifications are seen depending on these factors.⁸¹

The majority of lipoproteins reside in the inner leaflet of the outer membrane. Lipoproteins destined for the outer membrane are transported via the localization of lipoproteins (Lol) pathway as signaled by amino acids following the lipobox (residues +2 to +4 with the lipobox cysteine being +1).^{83,100} The Lol pathway is composed of five proteins: LolA, a periplasmic chaperone; LolB, an outer membrane receptor, LolC, a cytoplasmic membrane protein; LolD, a cytoplasmic ATP-ase; and LolE, a cytoplasmic membrane protein.^{75,81,100} LolE is hypothesized to first interact with the lipoprotein to initiate the transport reaction, the lipoprotein is transferred to LolC and further to LolA through ATP hydrolysis by LolD. However, exactly how the lipoprotein is transferred to the membrane is still unknown.^{81,101,102}

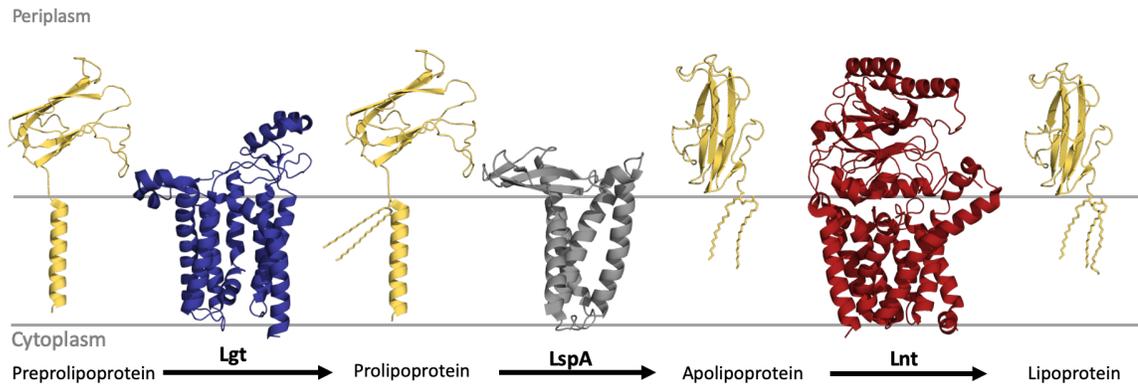


Figure 1.9. Lipoprotein processing pathway. Lipoprotein processing pathway in gram-negative bacteria. Gram-positive bacteria lack Lnt. The membrane is symbolized by gray lines and the lipoprotein shown is ICP (yellow). PDB IDs: Lgt: 5azc, LspA: 5dir, Lnt: 5n6m, ICP: 2wgn

1.2.3 LspA Structure and Function

LspA is an inner membrane protein which cleaves the signal peptide of prolipoproteins. Daglyation of the lipoprotein by Lgt is a prerequisite to cleavage by LspA.^{89,90} The crystal structure of *P. aeruginosa* LspA was determined in complex with the antibiotic globomycin with the *in meso*, or lipid cubic phase method, to a resolution of 2.8 Å (Figure 1.10).²⁸ Crystal structures have also been determined of *S. aureus* LspA in complex with the antibiotics globomycin and myxovirescin to a resolution of 1.9 and 2.3 Å, respectively.¹⁰³ Globomycin and myxovirescin are both cyclic peptide inhibitors of LspA.¹⁰³ Globomycin is commonly used as a negative control in functional studies.^{28,104,105}

LspA consists of four transmembrane helices, a periplasmic helix (PH), and a four-stranded β-sheet, called the β-cradle, which rests on the membrane (Figure 1.10). In the *S. aureus* structures, that authors termed the PH the extracellular loop (EL) as it is from a gram-positive bacteria which does not contain a periplasm. LspA is believed to act as an aspartyl protease, and the catalytic dyad residues have been identified by mutagenesis studies (D124 and D143 in *P. aeruginosa* LspA, and D118 and D136 in *S. aureus* LspA) (Figure 1.11).²⁸ Fourteen additional residues were identified as being strictly conserved. The positioning of globomycin and the crystallization host lipid monoacylglycerol (MAG) in the crystal structure of LspA allowed the authors to propose the orientation of the signal peptide and DAG in the active site. From the proposed orientation, the shape complementarity of LspA to the trigonal feature of the substrate where the signal peptide, lipoprotein, and DAG converge is hypothesized to be what allows LspA to specifically cleave a diverse set of substrates.²⁸ In this position the lipobox of the signal peptide would sit directly in the active site between the catalytic dyad.²⁸ There are still many remaining

questions about the mechanism of LspA, including how the substrate enters and exits the active site and the conformational dynamics of the protein. These questions will be addressed in detail in Chapters 3 and 4.

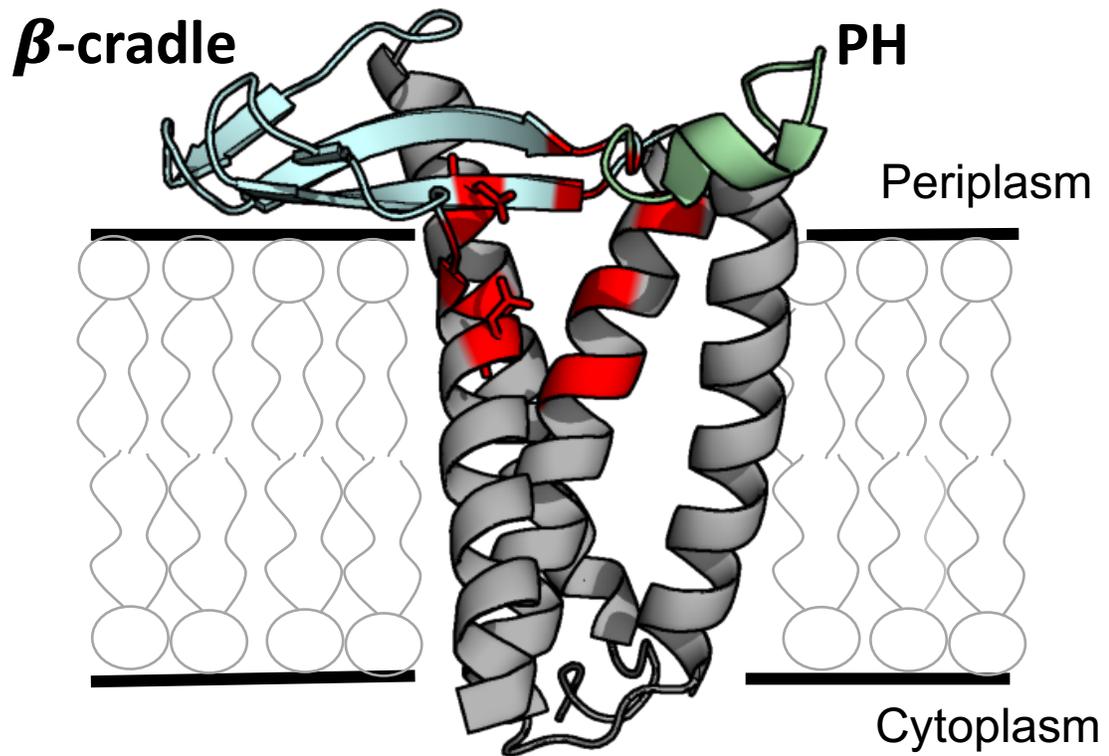


Figure 1.10. Structure of LspA. Crystal structure of *P. aeruginosa* LspA (PDB ID 5dir).²⁸

The transmembrane domain (gray), the PH (green), and the β -cradle (blue) are colored.

Conserved residues are colored red and the catalytic dyad residues are shown as sticks.

Membrane lipids shown as cartoons.

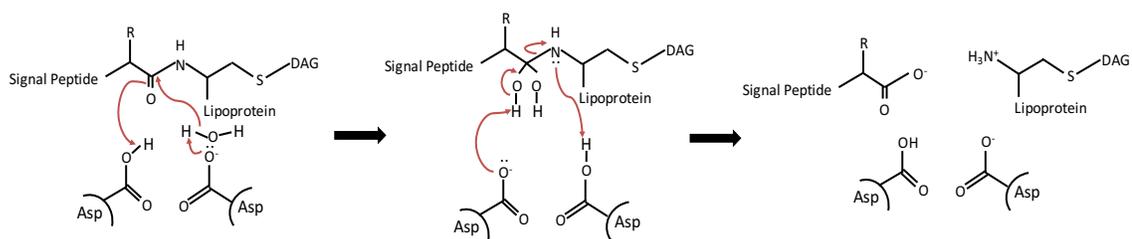


Figure 1.11. Hypothesized aspartyl-protease cleavage mechanism of LspA. Aspartic acid residues shown are the catalytic dyad: D124 and D143 for *P. aeruginosa* LspA, or D118 and D136 for *S. aureus* LspA.

1.3 Significance of Research

1.3.1 Rise of Antibiotic Resistance

Antibiotics have a wide variety of clinical applications including in cancer treatments, organ transplants, surgery, autoimmune diseases, and general sickness.¹⁰⁶ Unfortunately, antibiotic, and more widely antimicrobial (including antibiotic, antiviral, and antifungal), resistance is an exponentially rising medical concern. The World Health Organization states that the growth of antibiotic resistant bacteria *'means that commonplace medical procedures once previously taken for granted could be conceivably consigned to medical limbo. The repercussions are almost unimaginable'*.^{107,108} Over 10 million people are projected to die annually from antimicrobial resistance by the year 2050 (Table 1.1).¹⁰⁹ In addition to mortality, antimicrobial resistance also results in more frequent doctor visits and longer, more expensive treatments and hospital stays.¹⁰⁶ In fact, antimicrobial resistance is predicted to cost \$100.2 trillion worldwide by 2050.¹⁰⁹

After the first widely used antibiotic, penicillin, was developed in the mid-1940s, 20 new classes of antibiotics were developed by 1962. However, since then only four new classes of antibiotics have been developed and reached the market, while the use of antibiotics and the rise of antibiotic resistance has increased dramatically.¹⁰⁶ This decrease in antibiotic development is partially due to many pharmaceutical companies closing their antibiotic research divisions, leaving the development of antibiotics to academic institutions.¹¹⁰ As of March 2019, there were 43 new antibiotics in global clinical development, but only 6 have novel targets.¹¹¹ With the rapid growth of antibiotic resistance it is imperative that new drugs are developed which target novel pathways.

Continent	Deaths
Asia	4,730,000
Africa	4,150,000
Latin America	392,000
Europe	390,000
North America	317,000
Oceania	22,000
Total	10,001,000

Table 1.1. Projected deaths attributable to antimicrobial resistance per year by 2050.¹⁰⁹

1.3.2 Lipoprotein Processing Pathway as an Antibiotic Drug Target

Targeting the lipoprotein processing pathway is a novel pathway for antibiotic development. Lgt, LspA, and/or Lnt are essential in organisms including *E. coli*, *S. enterica*, *M. tuberculosis*, and *S. coelicolor*.^{28,81} Therefore, understanding the molecular mechanisms of these enzymes could lead to the development of antibacterial drugs. All three of the lipoprotein processing enzymes are potential antibiotic drug candidates, though the active site of Lnt, a catalytic triad of the nitrilase type, is also found in mammals making it a challenging target.⁹² Aspartic proteases are also found in mammals. However, the high degree of sequence conservation in residues surrounding the active site of LspA, as well as potential essential membrane interactions provide opportunities for drugs to be developed that will be specific to LspA.

Critically, LspA is essential in many human pathogens including *E. coli* and *S. coelicolor*, and does not have a mammalian homolog. These qualities make LspA an excellent candidate for antibiotic drug development. Additionally, conserved amino acids not only reside in, but surround the active site of LspA. Therefore, any mutation that arose to block antibiotic binding in the active site, as occurs in antibiotic resistance, would also likely interfere with the binding and cleavage of substrate.²⁸ Thus, LspA is a powerful target to combat the development of antibiotic resistance. Here, LspA from *P. aeruginosa* is studied as a model of LspA from all bacterial species, as it has >50% sequence similarity to other biologically significant LspA, such as those found in *E. coli*, *S. aureus*, *P. difficile*, *M. tuberculosis*, and *C. trachomatis*, and a crystal structure is available.^{28,81}

P. aeruginosa is an opportunistic pathogen which primarily infects the immunocompromised, and is the predominant pathogen in the airways of cystic fibrosis

patients.^{76,112} *P. aeruginosa* are also commonly found in hospitalized pneumonia patients (44.6% of cases), and in bloodstream infections (27.9% of cases).¹¹³ In 2013 the Centers for Disease Control and Prevention listed *P. aeruginosa* as a serious threat for drug-resistance¹¹⁴ and in 2017 the World Health Organization listed *P. aeruginosa* as a “Priority 1: Critical” pathogen for research and development of new antibiotics.^{115,116} Multi-drug-resistant *P. aeruginosa* are resistant to carbapenems and other β -lactams commonly used to handle infection, and are difficult to treat due to multiple resistance mechanisms including an increase in expression of efflux pumps, changes in outer membrane permeability (due to a lack of the OprD porin), and enzymes targeting the drug (such as β -lactamases).^{110,112,113,117} Thus, it is imperative that antibiotics are developed to target a novel pathway in *P. aeruginosa*.

1.3.3 Elucidating the Mechanism of LspA

The molecular mechanism and dynamics associated with cleavage of lipoprotein signal peptide cleavage is largely unknown. While the molecular mechanisms of Lgt and Lnt are somewhat well characterized, the mechanism of LspA is speculated.^{28,84,92,118} The active site residues of LspA have been determined by mutagenesis and the signal peptide orientation has been hypothesized (see Section 1.2.3).²⁸ However, understanding the specific molecular mechanism of signal peptide cleavage will aid in future structure-guided drug development.

Additionally, understanding the mechanism of LspA can give insight into the mechanisms of other membrane bound enzymes. Membrane proteins make up about 30% of all proteins in living organisms and are the drug target of about 60% of current

pharmaceuticals, yet they represent less than 2% of the total number of known protein structures. Specifically, there are structures for only ten intramembrane-cleaving proteases (I-CLiPs) (Figure 1.12).¹⁸ I-CLiPs can be categorized into four types: rhomboid serine protease (GlpG), site-2 metalloprotease (S2P), Rce1-type glutamyl protease, and aspartyl protease (such as signal peptide peptidase (SPP) and γ -secretase).¹¹⁹⁻¹²¹ While LspA has never before been characterized as an iCLiP specifically, it is proposed to be an aspartyl protease and functions within the membrane, so therefore falls under this category.²⁸ However, LspA does not show sequence or structural similarity to these proteases (Figure 1.13). Much is still unknown about how iCLiPs recognize their substrates or the molecular mechanism of cleavage.^{120,121} Even less is known about the protein dynamics during substrate binding and cleavage and the role of the lipid environment.¹²²⁻¹²⁴ Here, LspA is studied as a novel iCLiP and could elucidate information applicable to the class of proteins as a whole.

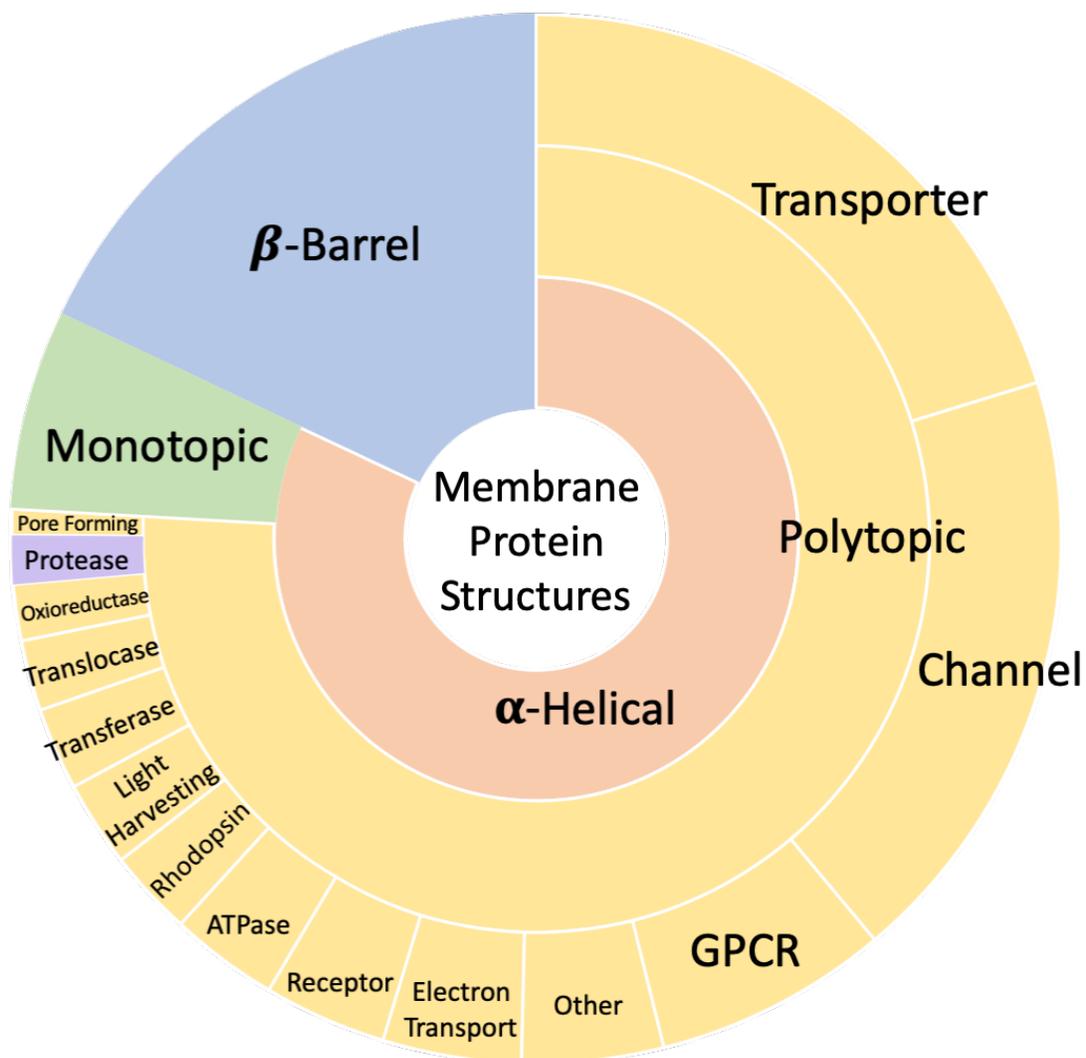


Figure 1.12. Classification of membrane proteins of known structure. There are currently only ten determined membrane protease structures (purple).¹⁸

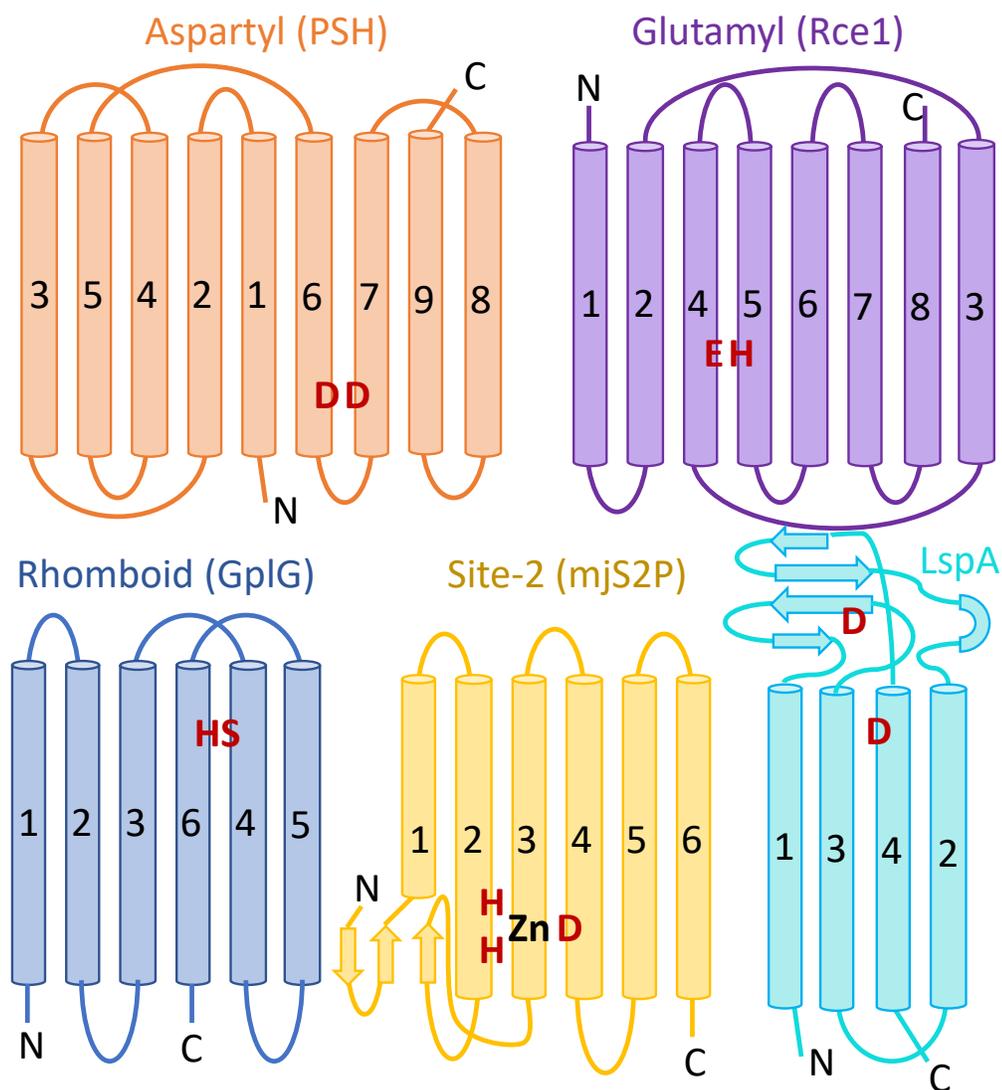


Figure 1.13. iCLiP membrane topology compared to LspA. Catalytic residues shown in red. Adapted from Sun *et al.* (2016), with permission.¹²¹

1.3.4 Effects of Membrane Mimetics

Membrane proteins are very challenging to study because they do not express readily as recombinant proteins (compared to soluble proteins) and they require a membrane mimetic (see section 1.1.4). Experimentally, a membrane mimetic that is suitable for stabilizing a specific membrane protein can be hard to find as usually this is a trial and error process.¹⁹ Additionally, there may be experimental constraints that dictate the mimetic chosen. For example, larger systems such as liposomes or large bicelles tumble slowly are not conducive to study by NMR.^{32,33}

The membrane mimetic used ideally should solubilize the membrane protein in its native conformation and, if enzymatic, have activity similar to that in the native membrane. However, for many membrane proteins the native conformation and activity is unknown. Membrane protein activity and protein backbone dynamics are seen to differ in different mimetic systems.^{19,125} Thus, it is important to consider the effects of the membrane mimetic system used when interpreting all results.

Here, LspA is studied in two membrane mimetics experimentally (micelles and bicelles) and compared to molecular dynamics simulations conducted in a bilayer system (Chapter 3). It was important to take these experimental factors into consideration when interpreting the data from these studies. Additionally, bicelles were characterized to provide more information about the shape and segregation of this widely used system (Chapter 5). These results are crucial to understanding the environment surrounding membrane proteins studied using this mimetic.

1.4 Thesis Overview

Based on the essential function of LspA in many bacterial species, LspA has been proposed as a good candidate for antibiotic drug targeting.^{28,81} While the crystal structure of LspA aided in a proposed signal peptide orientation, a detailed understanding of the molecular mechanism and conformational dynamics associated with cleavage is sought.²⁸ Therefore, the overarching goal of this thesis is to better understand the mechanism of LspA cleavage. To that end, the *primary objectives* of the thesis are the following: 1) determine the conformational dynamics of LspA (**Chapter 3**) using electron paramagnetic resonance (EPR) spectroscopy (**Chapter 2**), and 2) develop a biological activity assay to measure LspA function (**Chapter 4**). Work to better characterize membrane mimetic bicelle systems used to study membrane proteins will also be described (**Chapter 5**).

1.5 References

- 1 Ponomarenko EA, Poverennaya E V., Ilgisonis E V., Pyatnitskiy MA, Kopylov AT, Zgoda VG, *et al.* The Size of the Human Proteome: The Width and Depth. *Int J Anal Chem* 2016. <https://doi.org/10.1155/2016/7436849>.
- 2 Milo R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* 2013;**35**:1050–5. <https://doi.org/10.1002/bies.201300066>.
- 3 Henzler-Wildman K, Kern D. Dynamic personalities of proteins. *Nature* 2007;**450**:964–72. <https://doi.org/10.1038/nature06522>.
- 4 Stank A, Kokh DB, Fuller JC, Wade RC. Protein Binding Pocket Dynamics. *Acc Chem Res* 2016;**49**:809–15. <https://doi.org/10.1021/acs.accounts.5b00516>.

- 5 Frauenfelder H, Petsko G, Tsernoglou D. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* 1979;**280**:558–63.
- 6 Kleckner IR, Foster MP. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta* 2011;**1814**:942–68.
<https://doi.org/10.1016/j.bbapap.2010.10.012>.
- 7 Sapienza PJ, Lee AL. Using NMR to study fast dynamics in proteins: Methods and applications. *Curr Opin Pharmacol* 2010;**10**:723–30.
<https://doi.org/10.1016/j.coph.2010.09.006>.
- 8 Kovermann M, Rogne P, Wolf-watz M. Protein dynamics and function from solution state NMR spectroscopy. *Q Rev Biophys* 2016;**49**:1–43.
<https://doi.org/10.1017/S0033583516000019>.
- 9 Mittermaier A, Kay LE. New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Science (80-)* 2006;**312**..
- 10 Claxton DP, Kazmier K, Mishra S, McHaourab HS. *Navigating Membrane Protein Structure, Dynamics, and Energy Landscapes Using Spin Labeling and EPR Spectroscopy*. vol. 564. 1st ed. Elsevier Inc.; 2015.
- 11 Altenbach C, López CJ, Hideg K, Hubbell WL. Exploring Structure, Dynamics, and Topology of Nitroxide Spin-Labeled Proteins Using Continuous-Wave Electron Paramagnetic Resonance Spectroscopy. *Methods Enzymol* 2015;**564**:59–100. <https://doi.org/10.1016/bs.mie.2015.08.006>.
- 12 Columbus L, Hubbell WL. A new spin on protein dynamics. *Trends Biochem Sci* 2002;**27**:288–95. [https://doi.org/10.1016/S0968-0004\(02\)02095-9](https://doi.org/10.1016/S0968-0004(02)02095-9).
- 13 Columbus L, Hubbell WL. Mapping backbone dynamics in solution with site-

- directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* 2004;**43**:7273–87. <https://doi.org/10.1021/bi0497906>.
- 14 Biehl R, Richter D. Slow internal protein dynamics in solution. *J Phys Condens Matter* 2014;**26**:. <https://doi.org/10.1088/0953-8984/26/50/503103>.
- 15 McHaourab HS, Steed PR, Kazmier K. Toward the fourth dimension of membrane protein structure: Insight into dynamics from spin-labeling EPR spectroscopy. *Structure* 2011;**19**:1549–61. <https://doi.org/10.1016/j.str.2011.10.009>.
- 16 Almén M, Nordström KJ, Fredriksson R, Schiöth HB. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol* 2009;**7**:50. <https://doi.org/10.1186/1741-7007-7-50>.
- 17 Yin H, Flynn AD. Drugging Membrane Protein Interactions. *Annu Rev Biomed Eng* 2016;**18**:51–76. <https://doi.org/10.1146/annurev-bioeng-092115-025322>.
- 18 White S. *Membrane Proteins of Known Structure*. 2019.
- 19 Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: Not just a soap opera. *Biochim Biophys Acta - Biomembr* 2004;**1666**:105–17. <https://doi.org/10.1016/j.bbamem.2004.04.011>.
- 20 Deisenhofer J, Epp O, Huber R, Michel H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3Å resolution. *Nature* 1985;**318**:618–24.
- 21 Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* 1958;**181**:662–6.

- 22 Thonghin N, Kargas V, Clews J, Ford RC. Cryo-electron microscopy of membrane proteins. *Methods* 2018;**147**:176–86. <https://doi.org/10.1016/j.ymeth.2018.04.018>.
- 23 Data P, Pdb B, Single T, Burley SK, Berman HM, Kleywegt GJ, *et al.* *Protein Crystallography*. vol. 1607. 2017.
- 24 Kaptein R, Wagner G. NMR studies of membrane proteins. *J Biomol NMR* 2015;**0**:181–4. <https://doi.org/10.1126/scisignal.274pe36>.Insulin.
- 25 Schulz GE. B-Barrel Membrane Proteins. *Curr Opin Struct Biol* 2000;**10**:443–7. [https://doi.org/10.1016/S0959-440X\(00\)00120-2](https://doi.org/10.1016/S0959-440X(00)00120-2).
- 26 Fairman JW, Noinaj N, Buchanan SK. The structural biology of beta-barrel membrane proteins: a summary of recent reports. *Curr Opin Struct Biol* 2011;**21**:523–31. <https://doi.org/10.1016/j.sbi.2011.05.005>.The.
- 27 Nanda V. Building bigger beta-barrles. *Elife* 2019;**8**:. <https://doi.org/10.7554/eLife.40308>.
- 28 Vogeley L, El Arnaout T, Bailey J, Stansfeld PJ, Boland C, Caffrey M. Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. *Science (80-)* 2016;**351**:876–80. <https://doi.org/10.1126/science.aad3747>.
- 29 Nakayama H, Kurokawa K, Lee BL. Lipoproteins in bacteria: Structures and biosynthetic pathways. *FEBS J* 2012;**279**:4247–68. <https://doi.org/10.1111/febs.12041>.
- 30 Morozova D, Guigas G, Weiss M. Dynamic structure formation of peripheral membrane proteins. *PLoS Comput Biol* 2011;**7**:. <https://doi.org/10.1371/journal.pcbi.1002067>.

- 31 Wu HC, Tokunaga M, Tokunaga H, Hayashi S, Giam C-Z. Posttranslational modification and processing of membrane lipoproteins in bacteria. *J Cell Biochem* 1983;**22**:161–71. <https://doi.org/10.1002/jcb.240220305>.
- 32 Liang B, Tamm LK. Solution NMR of SNAREs, complexin and α -synuclein in association with membrane-mimetics. *Prog Nucl Magn Reson Spectrosc* 2018;**105**:41–53. <https://doi.org/10.1016/j.pnmrs.2018.02.001>.
- 33 Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand É, Marcotte I. Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim Biophys Acta - Biomembr* 2011;**1808**:1957–74. <https://doi.org/10.1016/j.bbamem.2011.03.016>.
- 34 Tanford C. *The hydrophobic effect: formation of micelles and biological membranes*. 1973.
- 35 Oliver RC, Lipfert J, Fox DA, Lo RH, Doniach S, Columbus L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS One* 2013;**8**:. <https://doi.org/10.1371/journal.pone.0062488>.
- 36 Columbus L, Lipfert J, Jambunathan K, Fox DA, Sim AYL, Doniach S, *et al*. Mixing and matching detergents for membrane protein NMR structure determination. *J Am Chem Soc* 2009;**131**:7320–6. <https://doi.org/10.1021/ja808776j>.
- 37 Oliver RC, Lipfert J, Fox DA, Lo RH, Kim JJ, Doniach S, *et al*. Tuning micelle dimensions and properties with binary surfactant mixtures. *Langmuir* 2014;**30**:13353–61. <https://doi.org/10.1021/la503458n>.
- 38 Gabriel NE, Roberts MF. Interaction of Short-Chain Lecithin with Long-Chain

- Phospholipids: Characterization of Vesicles That Form Spontaneously.
Biochemistry 1986;**25**:2812–21. <https://doi.org/10.1021/bi00358a012>.
- 39 Gabriel NE, Roberts MF. Spontaneous Formation of Stable Unilamellar Vesicles.
Biochemistry 1984;**23**:4011–5. <https://doi.org/10.1021/bi00313a001>.
- 40 Caldwell TA, Baoukina S, Brock AT, Oliver RO, Root KT, Krueger JK, *et al*.
Low-q Bicelles are Mixed Micelles. *J Phys Chem Lett* 2018;**9**:4469–73.
- 41 Luchette PA, Vetman TN, Prosser RS, Hancock REW, Nieh MP, Glinka CJ, *et al*.
Morphology of fast-tumbling bicelles: A small angle neutron scattering and NMR
study. *Biochim Biophys Acta - Biomembr* 2001;**1513**:83–94.
[https://doi.org/10.1016/S0005-2736\(01\)00358-3](https://doi.org/10.1016/S0005-2736(01)00358-3).
- 42 Beaugrand M, Arnold AA, Hénin J, Warschawski DE, Williamson PTF, Marcotte
I. Lipid concentration and molar ratio boundaries for the use of isotropic bicelles.
Langmuir 2014;**30**:6162–70. <https://doi.org/10.1021/la5004353>.
- 43 Triba MN, Warschawski DE, Devaux PF. Reinvestigation by phosphorus NMR of
lipid distribution in bicelles. *Biophys J* 2005;**88**:1887–901.
<https://doi.org/10.1529/biophysj.104.055061>.
- 44 Shintani M, Matubayasi N. Morphology study of DMPC/DHPC mixtures by
solution-state ¹H, ³¹P NMR, and NOE measurements. *J Mol Liq* 2015;**217**..
- 45 Piai A, Ru Q, Dev J, Chou JJ. Optimal Bicelle q for Solution NMR Studies of
Protein Transmembrane Partition. *Chemistry (Easton)* 2017;**23**:1361–7.
<https://doi.org/10.1183/09031936.00063810>.The.
- 46 Glover KJ, Whiles JA, Wu G, Yu NJ, Deems R, Struppe JO, *et al*. Structural
evaluation of phospholipid bicelles for solution-state studies of membrane-

- associated biomolecules. *Biophys J* 2001;**81**:2163–71.
[https://doi.org/10.1016/S0006-3495\(01\)75864-X](https://doi.org/10.1016/S0006-3495(01)75864-X).
- 47 Mineev KS, Nadezhdin KD, Goncharuk SA, Arseniev AS. Characterization of Small Isotropic Bicelles with Various Compositions. *Langmuir* 2016;**32**:6624–37.
<https://doi.org/10.1021/acs.langmuir.6b00867>.
- 48 Kot EF, Goncharuk SA, Arseniev AS, Mineev KS. Phase Transitions in Small Isotropic Bicelles. *Langmuir* 2018;**34**:3426–37.
<https://doi.org/10.1021/acs.langmuir.7b03610>.
- 49 Lin TL, Liu CC, Roberts MF, Chen SH. Structure of mixed short-chain lecithin/long-chain lecithin aggregates studied by small-angle neutron scattering. *J Phys Chem* 1991;**95**:6020–7. <https://doi.org/10.1021/j100168a057>.
- 50 Vold RR, Prosser RS, Deese AJ. Isotropic solutions of phospholipid bicelles: A new membrane mimetic for high-resolution NMR studies of polypeptides. *J Biomol NMR* 1997;**9**:329–35. <https://doi.org/10.1023/A:1018643312309>.
- 51 Sanders CR, Schwonek JP. Characterization of Magnetically Orientable Bilayers in Mixtures of Dihexanoylphosphatidylcholine and Dimyristoylphosphatidylcholine by Solid-State NMR. *Biochemistry* 1992;**31**:8898–905. <https://doi.org/10.1021/bi00152a029>.
- 52 Sanders CR, Prestegard JH. Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO. *Biophys J* 1990;**58**:447–60. [https://doi.org/10.1016/S0006-3495\(90\)82390-0](https://doi.org/10.1016/S0006-3495(90)82390-0).
- 53 Ram P, Prestegard JH. Magnetic field induced ordering of bile salt/phospholipid micelles: new media for NMR structural investigations. *BBA - Biomembr*

- 1988;**940**:289–94. [https://doi.org/10.1016/0005-2736\(88\)90203-9](https://doi.org/10.1016/0005-2736(88)90203-9).
- 54 Triba MN, Warschawski DE, Devaux PF. Reinvestigation by phosphorus NMR of lipid distribution in bicelles. *Biophys J* 2005;**88**:1887–901. <https://doi.org/10.1529/biophysj.104.055061>.
- 55 Vold RR, Prosser RS. Magnetically Oriented Phospholipid Bilayered Micelles for Structural Studies of Polypeptides. Does the Ideal Bicelle Exist? *J Magn Reson* 1996;**113**:267–71.
- 56 Marcotte I, Auger M. Bicelles as model membranes for solid-and solution-state NMR studies of membrane peptides and proteins. *Concepts Magn Reson Part A Bridg Educ Res* 2005;**24**:17–35. <https://doi.org/10.1002/cmr.a.20025>.
- 57 Dürr UHN, Soong R, Ramamoorthy A. When detergent meets bilayer: Birth and coming of age of lipid bicelles. *Prog Nucl Magn Reson Spectrosc* 2013;**69**:1–22. <https://doi.org/10.1016/j.pnmrs.2013.01.001>.
- 58 Poulos S, Morgan JLW, Zimmer J, Faham S. *Bicelles coming of age: An empirical approach to bicelle crystallization*. vol. 557. 1st ed. Elsevier Inc.; 2015.
- 59 King V, Parker M, Howard KP. Pegylation of magnetically oriented lipid bilayers. *J Magn Reson* 2000;**142**:177–82.
- 60 Sanders CR, Landis GC. Facile Acquisition and Assignment of Oriented Sample NMR Spectra for Bilayer Surface-Associated Proteins. *J Am Chem Soc* 1994;**116**:6470–1. <https://doi.org/10.1021/ja00093a072>.
- 61 Morrison EA, Henzler-Wildman KA. Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile. *Biochim Biophys Acta - Biomembr* 2012;**1818**:814–20.

- <https://doi.org/10.1016/j.bbamem.2011.12.020>.
- 62 Sanders CR, Landis GC. Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. *Biochemistry* 1995;**34**:4030–40.
- 63 Inagaki S, Ghirlando R, Grisshammer R. Biophysical characterization of membrane proteins in nanodiscs. *Methods* 2013;**59**:287–300.
<https://doi.org/10.1016/j.ymeth.2012.11.006>.
- 64 Laguerre A, Löhr F, Henrich E, Hoffmann B, Abdul-Manan N, Connolly PJ, *et al.* From Nanodiscs to Isotropic Bicelles: A Procedure for Solution Nuclear Magnetic Resonance Studies of Detergent-Sensitive Integral Membrane Proteins. *Structure* 2016:1830–41. <https://doi.org/10.1016/j.str.2016.07.017>.
- 65 Bayburt TH, Sligar SG. Membrane Protein Assembly into Nanodiscs. *FEBS Lett* 2010;**584**:1721–7. <https://doi.org/10.1016/B978-0-12-386043-9.00005-0>.New.
- 66 Schuler MA, Denisov IG, Sligar SG. Nanodiscs as a New Tool to Examine Lipid-Protein Interactions. *Lipid-Protein Interact. Methods Protoc.*, vol. 974. 2013.
- 67 Planchard N, Point É, Dahmane T, Giusti F, Renault M, Le Bon C, *et al.* The Use of Amphipols for Solution NMR Studies of Membrane Proteins: Advantages and Constraints as Compared to Other Solubilizing Media. *J Membr Biol* 2014;**247**:827–42. <https://doi.org/10.1007/s00232-014-9654-z>.
- 68 Klöpfer K, Hagn F. Beyond detergent micelles: The advantages and applications of non-micellar and lipid-based membrane mimetics for solution-state NMR. *Prog Nucl Magn Reson Spectrosc* 2019;**115**:271–83.
<https://doi.org/10.1016/j.pnmrs.2019.08.001>.
- 69 Tribet C, Audebert R, Popot JL. Amphipols: Polymers that keep membrane

- proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A* 1996;**93**:15047–50. <https://doi.org/10.1073/pnas.93.26.15047>.
- 70 Popot JL, Berry EA, Charvolin D, Creuzenet C, Ebel C, Engelman DM, *et al.* Amphipols: Polymeric surfactants for membrane biology research. *Cell Mol Life Sci* 2003;**60**:1559–74. <https://doi.org/10.1007/s00018-003-3169-6>.
- 71 Knowles TJ, Finka R, Smith C, Lin YP, Dafforn T, Overduin M. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc* 2009;**131**:7484–5. <https://doi.org/10.1021/ja810046q>.
- 72 Orwick MC, Judge PJ, Procek J, Lindholm L, Graziadei A, Engel A, *et al.* Detergent-free formation and physicochemical characterization of nanosized lipid-polymer complexes: Lipodisq. *Angew Chemie - Int Ed* 2012;**51**:4653–7. <https://doi.org/10.1002/anie.201201355>.
- 73 Manzo G, Carboni M, Rinaldi AC, Casu M, Scorciapino MA. Characterization of sodium dodecylsulphate and dodecylphosphocholine mixed micelles through NMR and dynamic light scattering. *Magn Reson Chem* 2013;**51**:176–83. <https://doi.org/10.1002/mrc.3930>.
- 74 Li M, Du C, Guo N, Teng Y, Meng X, Sun H, *et al.* Composition design and medical application of liposomes. *Eur J Med Chem* 2019;**164**:640–53. <https://doi.org/10.1016/j.ejmech.2019.01.007>.
- 75 Szewczyk J, Collet JF. *The Journey of Lipoproteins Through the Cell: One Birthplace, Multiple Destinations*. vol. 69. 1st ed. Elsevier Ltd.; 2016.
- 76 Remans K, Vercammen K, Bodilis J, Cornelis P. Genome-wide analysis and

- literature-based survey of lipoproteins in *Pseudomonas aeruginosa*. *Microbiology* 2010;**156**:2597–607. <https://doi.org/10.1099/mic.0.040659-0>.
- 77 Mori H, Ito K. The Sec protein-translocation pathway. *Trends Microbiol* 2001;**9**:494–500. [https://doi.org/10.1016/S0966-842X\(01\)02174-6](https://doi.org/10.1016/S0966-842X(01)02174-6).
- 78 Natale P, Brüser T, Driessen AJM. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms. *Biochim Biophys Acta* 2008;**1778**:1735–56. <https://doi.org/10.1016/j.bbamem.2007.07.015>.
- 79 Tokuda H, Narita S-I. Biogenesis and Membrane Targeting of Lipoproteins. *EcoSal Plus* 2010;**1**:. <https://doi.org/10.1128/ecosalplus.4.3.7>.
- 80 Paetzel M, Karla A, Strynadka NCJ, Dalbey RE. Signal peptidases. *Chem Rev* 2002;**102**:4549–79. <https://doi.org/10.1021/cr010166y>.
- 81 Buddelmeijer N. The molecular mechanism of bacterial lipoprotein modification- How, when and why? *FEMS Microbiol Rev* 2015;**39**:246–61. <https://doi.org/10.1093/femsre/fuu006>.
- 82 Tokuda H. Biogenesis of outer membranes in Gram-negative bacteria. *Biosci Biotechnol Biochem* 2009;**73**:465–73. <https://doi.org/10.1271/bbb.80778>.
- 83 Zückert WR. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochim Biophys Acta - Mol Cell Res* 2014;**1843**:1509–16. <https://doi.org/10.1016/j.bbamcr.2014.04.022>.
- 84 Mao G, Zhao Y, Kang X, Li Z, Zhang Y, Wang X, *et al*. Crystal structure of *E. coli* lipoprotein diacylglyceryl transferase. *Nat Commun* 2016;**7**:. <https://doi.org/10.1038/ncomms10198>.

- 85 Sankaran K, Wu HC. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J Biol Chem* 1994;**269**:19701–6.
- 86 Geukens N, De Buck E, Meyen E, Maes L, Vranckx L, Van Mellaert L, *et al.* The type II signal peptidase of *Legionella pneumophila*. *Res Microbiol* 2006;**157**:836–41. <https://doi.org/10.1016/j.resmic.2006.06.003>.
- 87 Gullón S, Arranz EIG, Mellado RP. Transcriptional characterisation of the negative effect exerted by a deficiency in type II signal peptidase on extracellular protein secretion in *Streptomyces lividans*. *Appl Microbiol Biotechnol* 2013;**97**:10069–80. <https://doi.org/10.1007/s00253-013-5219-9>.
- 88 Hussain M, Ichihara S, Mizushima S. Mechanism of Signal Peptide Cleavage in the Biosynthesis of the Major Lipoprotein of the *Escherichia coli* Outer Membrane. *J Biol Chem* 1982;**257**:5177–82.
- 89 Tokunaga M, Tokunaga H, Wu HC. Post-translational modification and processing of *Escherichia coli* prolipoprotein in vitro. *Proc Natl Acad Sci* 1982;**79**:2255–9. <https://doi.org/10.1073/pnas.79.7.2255>.
- 90 Tokunaga M, Loranger JM, Wu HC. Prolipoprotein modification and processing enzymes in *Escherichia coli*. *J Biol Chem* 1984;**259**:3825–30.
- 91 Pollitt S, Inouye S, Inouye M. Effect of amino acid substitutions at the signal peptide cleavage site of the *Escherichia coli* major outer membrane lipoprotein. *J Biol Chem* 1986;**261**:1835–7.
- 92 Wiktor M, Weichert D, Howe N, Huang C-Y, Olieric V, Boland C, *et al.* Structural insights into the mechanism of the membrane integral N-acyltransferase step in bacterial lipoprotein synthesis. *Nat Commun* 2017;**8**:15952.

- <https://doi.org/10.1038/ncomms15952>.
- 93 Jackowski S, Rock CO. Transfer of fatty acids from the 1-position of phosphatidylethanolamine to the major outer membrane lipoprotein of *Escherichia coli*. *J Biol Chem* 1986;**261**:11328–33.
- 94 Hillmann F, Argentini M, Buddelmeijer N. Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. *J Biol Chem* 2011;**286**:27936–46.
<https://doi.org/10.1074/jbc.M111.243519>.
- 95 Buddelmeijer N, Young R. The essential *Escherichia coli* apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. *Biochemistry* 2010;**49**:341–6. <https://doi.org/10.1021/bi9020346>.
- 96 Li D, Lyons JA, Pye VE, Vogeley L, Aragão D, Kenyon CP, *et al*. Crystal structure of the integral membrane diacylglycerol kinase. *Nature* 2013;**497**:521–4.
<https://doi.org/10.1038/nature12179>.
- 97 Mohiman N, Argentini M, Batt SM, Cornu D, Masi M, Eggeling L, *et al*. The ppm Operon Is Essential for Acylation and Glycosylation of Lipoproteins in *Corynebacterium glutamicum*. *PLoS One* 2012;**7**:.
<https://doi.org/10.1371/journal.pone.0046225>.
- 98 Gibson KJC, Eggeling L, Maughan WN, Krumbach K, Gurcha SS, Nigou J, *et al*. Disruption of Cg-Ppm1, a Polyprenyl Monophosphomannose Synthase, and the Generation of Lipoglycan-less Mutants in *Corynebacterium glutamicum*. *J Biol Chem* 2003;**278**:40842–50. <https://doi.org/10.1074/jbc.M307988200>.
- 99 Tschumi A, Nai C, Auchli Y, Hunziker P, Gehrig P, Keller P, *et al*. Identification of apolipoprotein N-acyltransferase (Lnt) in mycobacteria. *J Biol Chem*

- 2009;**284**:27146–56. <https://doi.org/10.1074/jbc.M109.022715>.
- 100 Konovalova A, Silhavy TJ. Outer membrane lipoprotein biogenesis: Lol is not the end. *Philos Trans R Soc B Biol Sci* 2015;**370**:. <https://doi.org/10.1098/rstb.2015.0030>.
- 101 Okuda S, Tokuda H. Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proc Natl Acad Sci U S A* 2009;**106**:5877–82. <https://doi.org/10.1073/pnas.0900896106>.
- 102 Taniguchi N, Matsuyama SI, Tokuda H. Mechanisms underlying energy-independent transfer of lipoproteins from LolA to LolB, which have similar unclosed β -barrel structures. *J Biol Chem* 2005;**280**:34481–8. <https://doi.org/10.1074/jbc.M507388200>.
- 103 Olatunji S, Yu X, Bailey J, Huang C, Zapotoczna M, Bowen K, *et al*. Structures of lipoprotein signal peptidase II from *Staphylococcus aureus* complexed with antibiotics globomycin and myxovirescin 2020. <https://doi.org/10.1038/s41467-019-13724-y>.
- 104 Sangith N, Sankaran K. Stringency of bacterial prolipoprotein signal peptidase (LspA) in recognition of signal peptides – Structure-function correlation. *Biochem Biophys Res Commun* 2017;**488**:413–7. <https://doi.org/10.1016/j.bbrc.2017.05.069>.
- 105 Rahman MS, Ceraul SM, Dreher-Lesnack SM, Beier MS, Azad AF. The *lspA* gene, encoding the type II signal peptidase of *Rickettsia typhi*: Transcriptional and functional analysis. *J Bacteriol* 2007;**189**:336–41. <https://doi.org/10.1128/JB.01397-06>.
- 106 Smith RA, Read AF. Antibiotic Resistance: A Primer and Call to Action. *Health*

- Commun* 2015;**30**:309–14.
<https://doi.org/10.1080/10410236.2014.943634>.Antibiotic.
- 107 Martinez JL. General principles of antibiotic resistance in bacteria. *Drug Discov Today Technol* 2014;**11**:33–9. <https://doi.org/10.1016/j.ddtec.2014.02.001>.
- 108 World Health Organisation. Overcoming antimicrobial resistance - WHO report on infectious diseases 2000 2000.
- 109 O’Neill J. *Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations*. 2014.
- 110 Venter H. Reversing Resistance to Counter Antimicrobial Resistance in the World Health Organisation’s Critical Priority of Most Dangerous Pathogens. *Biosci Rep* 2019:BSR20180474. <https://doi.org/10.1042/BSR20180474>.
- 111 The Pew Charitable Trusts. *Antibiotics Currently in Clinical Development*. 2019.
- 112 Mulcahy LR, Isabella VM, Lewis K. Pseudomonas aeruginosa Biofilms in Disease. *Microb Ecol* 2014;**68**:1–12. <https://doi.org/10.1007/s00248-013-0297-x>.
- 113 Dee Shortridge , Ana Gales , Jennifer Streit AT, Ronald N. Jones. Geographic and temporal patterns of antimicrobial resistance in Pseudomonas aeruginosa from the SENTRY Surveillance Program, 1997-2016. *Open Forum Infect Dis* 2019;**6**:S63–8. <https://doi.org/10.1093/ofid/ofy343>.
- 114 *Center for Disease Control and Prevention*. Antibiotic/Antrimicrobial Resistance Biggest Threats. 2013.
- 115 Willyard C. *The drug-resistant bacteria that pose the greatest health threats*. Nature. 2017.
- 116 Tacconelli E, Magrini N. *Global priority list of antibiotic-resistant bacteria to*

guide research, discovery, and development of new antibiotics. 2017.

- 117 López-García A, Rocha-Gracia RDC, Bello-López E, Juárez-Zelocualtecal C, Sáenz Y, Castañeda-Lucio M, *et al*. Characterization of antimicrobial resistance mechanisms in carbapenem-resistant *Pseudomonas aeruginosa* carrying IMP variants recovered from a Mexican hospital. *Infect Drug Resist* 2018;**11**:1523–36. <https://doi.org/10.2147/IDR.S173455>.
- 118 Boland C, Caffrey M. Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin 2016. <https://doi.org/10.1126/science.aad3747>.
- 119 Hu J, Xue Y, Lee S, Ha Y. The crystal structure of GXGD membrane protease FlaK. *Nature* 2014;**475**:528–31. <https://doi.org/10.1158/0008-5472.CAN-10-4002.BONE>.
- 120 Zhou R, Shi Y, Yang G. *Expression, Purification, and Enzymatic Characterization of Intramembrane Proteases*. vol. 584. 1st ed. Elsevier Inc.; 2017.
- 121 Sun L, Li X, Shi Y. Structural biology of intramembrane proteases: Mechanistic insights from rhomboid and S2P to γ -secretase. *Curr Opin Struct Biol* 2016;**37**:97–107. <https://doi.org/10.1016/j.sbi.2015.12.008>.
- 122 Bondar A, Val C, White SH. Rhomboid protease dynamics and lipid interactions. *Structure* 2009;**17**:395–405. <https://doi.org/10.1016/j.str.2008.12.017.Rhomboid>.
- 123 Sherratt AR, Braganza M V., Nguyen E, Ducat T, Goto NK. Insights into the effect of detergents on the full-length rhomboid protease from *Pseudomonas aeruginosa* and its cytosolic domain. *Biochim Biophys Acta - Biomembr* 2009;**1788**:2444–53. <https://doi.org/10.1016/j.bbamem.2009.09.003>.

- 124 Foo ACY, Harvey BGR, Metz JJ, Goto NK. Influence of hydrophobic mismatch on the catalytic activity of Escherichia coli GlpG rhomboid protease. *Protein Sci* 2015;**24**:464–73. <https://doi.org/10.1002/pro.2585>.
- 125 Frey L, Lakomek NA, Riek R, Bibow S. Micelles, Bicelles, and Nanodiscs: Comparing the Impact of Membrane Mimetics on Membrane Protein Backbone Dynamics. *Angew Chemie - Int Ed* 2017;**56**:380–3. <https://doi.org/10.1002/anie.201608246>.

CHAPTER 2. ELECTRON PARAMAGNETIC RESONANCE SPECTROSCOPY

2.1 Introduction

Electron paramagnetic resonance (EPR) is a widely used technique to study the structure and function of proteins. There are many excellent reviews which describe EPR in detail,¹⁻⁸ as well as the two most common EPR experiments: continuous-wave (CW-EPR)⁹⁻¹¹ and double electron electron resonance (DEER).¹²⁻¹⁸ There are also several reviews highlighting the advantages of EPR for studying membrane proteins specifically.¹⁹⁻²⁵ As such, here I will touch briefly on the theory of EPR (Section 2.3), but focus on the application of EPR to membrane protein systems (Section 2.4) and the practical considerations of EPR experimentation (Section 2.5). First, I will highlight biophysical approaches which are used to study membrane proteins, and why EPR is advantageous (Section 2.2).

2.2 Biophysical Techniques used to Study Membrane Proteins

As described extensively in Chapter 1, membrane proteins are difficult to study due to the requirement of a stabilizing membrane mimetic. Not only is it difficult to find a suitable mimetic for a protein of interest, these mimetics also increase the size and complexity of the system leading to experimental limitations. Here I will describe some common structural methods to study membrane proteins and their drawbacks (Figure 2.1). Then, I will explain why EPR is well suited for studies of membrane proteins.¹⁹⁻²⁵

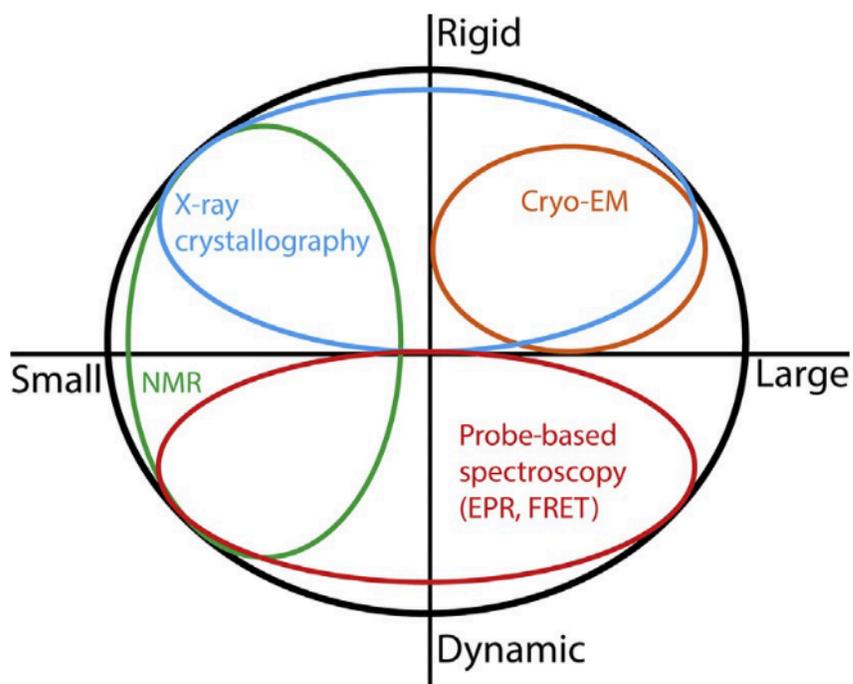


Figure 2.1. Biophysical techniques used to study membrane proteins. The size limitations of the protein system for each technique are shown on the x-axis (small or large) while the information probed (rigid or dynamic) is shown on the y-axis. Figure reprinted from Claxton *et. al* (2015), with permission.²²

2.2.1 Common Biophysical Techniques used to Study Membrane Proteins

2.2.1.1 X-ray Crystallography

The first membrane protein structure was determined by X-ray crystallography in 1985.²⁶ Crystallization of membrane proteins is difficult because the hydrophobic transmembrane helices do not form protein contacts in an ordered arrangement, which is necessary for crystallization.²⁰ While new techniques have been developed such as crystallizing in the lipid cubic phase,^{27–29} information about the protein's surroundings are not resolved.²³ While crystallography may be used to study proteins with a wide range of molecular weights, the crystal structure represents a snapshot of the protein and does not give much dynamic information (Figure 2.1).²² Different conformational states can be trapped and crystalized, but this may not represent the full range of conformations. However, crystal structures can be used to assist the choice of complementary experiments such as EPR. For example, the crystal structure may inform on the choice of spin label sites to probe the dynamics of certain regions, or to investigate changes in different membrane environments.²³

2.2.1.2 Cryo-Electron Microscopy

Cryo-electron microscopy (EM) is a more recently developed technique that can be used to determine the structure of membrane proteins. The first membrane protein structure determined by cryo-EM was in 2013,³⁰ and since, the use of the technique to study membrane proteins has grown drastically.³¹ Cryo-EM can be used to study large proteins with low concentrations comparable to EPR, but like crystallography it produces a rigid model and may not resolve flexible proteins or protein regions (Figure 2.1).²³ While cryo-

EM samples can be imaged in different conformations, a homogenous sample is still necessary to determine high-quality structures.³¹ Additionally, like crystallography it is difficult to resolve membrane components by cryo-EM. Membrane proteins studied via cryo-EM are solubilized in lipid nanoparticles, detergents, or amphipols. The density of the membrane mimetic can thus influence image alignment, which adds to the complexity of determining the structures of membrane proteins that do not have large soluble domains.³¹

2.2.1.3 Small Angle Scattering

Small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS) have traditionally been used with soluble proteins because free detergent or lipid, as well as the micelle or bicelle that surrounds the protein, makes it difficult to determine scattering specifically from the protein component.²³ However, new methods combining size exclusion chromatography with SAXS or SANS are able to separate the empty micelles or bicelles to increase the signal to noise ratio.^{32,33} Additionally, using deuterated components in SANS experiments allows for the separation of structural information from the protein and the membrane mimetic (see SANS Contrast Variation in section 5.2.2).³⁴ While SAXS and SANS techniques can shed light on large conformational changes of proteins, they are low-resolution and are usually used in combination with other biophysical techniques.²³

2.2.1.4 Nuclear Magnetic Resonance

Solution nuclear magnetic resonance, NMR, is a technique used to determine both the high resolution structure and dynamics of proteins.²² However, NMR has size limitations and can only be used with systems generally <30 kDa (Figure 2.1).²³ Thus,

NMR cannot be used to study all membrane proteins because of the increased size of the system due to the micelle, bicelle, or nanodisc membrane mimetic.^{20,23} Solid state NMR, on the other hand, can investigate membrane proteins in a variety of membranes including lipid vesicles, whole cell membranes, or bicelles.³⁵ NMR samples require a relatively large volume of concentrated sample which is not physiological, and may induce aggregation or protein destabilization.^{22,23} Additionally, expensive isotopes are needed to label the sample, which may not be practical for membrane proteins that have low yields or difficulty with purification.

2.2.1.5 Förster Resonance Energy Transfer

Förster resonance energy transfer (FRET) allows for the study of the dynamics of large protein systems similar to EPR (Figure 2.1). However, FRET requires the use of a fluorophore pair conjugated to the protein of interest. These probes are relatively large, and may interfere with protein folding, dynamics, and/or function.²³ Additionally, each FRET pair has a limited distance range that they can probe, so multiple sets of pairs are required to probe the full distance range that DEER can resolve. FRET is also limited because the fluorescent properties of fluorophores can be affected by their environment. The fluorescent properties may change in different membrane mimetics or even when contacting differing numbers of lipid molecules.²³ Thus, care must be taken when interpreting FRET results when conducted in membrane mimetics.

2.2.2 Advantages of using EPR to Study Membrane Proteins

Electron paramagnetic resonance (EPR) is well suited for studies of membrane proteins, as it can probe the structure and dynamics of systems with a wide range of molecular weights (Figure 2.1).^{19–25} Samples can be in a range of solvents and conditions to match physiological systems. As such, many different membrane mimetics can be used with EPR.²⁰ Not only does this allow researchers to use the membrane mimetic that demonstrates the best stability for their membrane protein of interest, it also allows for the comparison of multiple membrane systems to gain a better understanding of the protein's interactions with the surrounding membrane environment. EPR experiments only require the addition of a small probe containing a paramagnetic center (see Section 2.4.1), which is comparable in size to a tryptophan or phenylalanine side chain. Most protein systems do not contain a native paramagnetic center, so EPR can selectively probe regions of interest.²⁰ The small sample volumes and low protein concentrations that EPR requires are ideal for the study of membrane proteins as they are often hard to express and purify.^{19,20} While EPR is low-throughput and gives moderate structural resolution due to a limited number of spin labels, it is highly sensitive – up to three times that of NMR experiments.²⁰

The biggest advantage of using EPR to study membrane proteins is that it is one of few techniques (in addition to FRET, described in section 2.4.1.5) that is able to probe the dynamics of membrane protein systems. It is advantageous over FRET in that the probes are much smaller and thus have less of an effect, if any, on protein fold, function, and dynamics. Additionally, a wider selection of membrane mimetics can be used, including micelles, bicelles, liposomes, and nanodiscs because the EPR signal is not influenced by the membrane environment as FRET probes are.²² Recent work has shown EPR

measurements of membrane proteins in whole cells.^{17,36,37} Thus, EPR is an excellent technique to study membrane proteins in a physiological environment to obtain structural and dynamic information not obtainable by other biophysical methods (Table 2.1).

Structure	Dynamics
Distance information	Side chain internal motions
Secondary structure	Local backbone fluctuations
Local polarity	Conformational dynamics
Identification of trans-membrane segments	Motion of nitroxide on nanosecond timescale

Table 2.1. Structure and dynamics probed by EPR. Adapted from Hemminga *et. al* (2007).²¹

2.3 Primer on EPR Theory

An EPR instrument contains a magnet to generate a magnetic field, a microwave source, a resonator that holds the sample, and a computer to record and analyze the data.¹ Electron paramagnetic resonance (EPR) is the measurement of the absorption of electromagnetic radiation equaling the energy of the splitting of the energy states of an unpaired electron in a magnetic field.^{1,20} Electromagnetic radiation is coupled electric (E) and magnetic (B) fields which oscillate in a frequency range (ν) and have energy ($h\nu$) where h is the Planck constant.³⁸ Instruments are characterized by the frequency (ν) range in which they record: S-band (2-4 GHz), X-band (8-10 GHz), Q-band (~35 GHz), and W-band (~90 GHz).^{1,39,40} In this work CW spectra (see section 2.4.2) are collected at X-band, and DEER data (see section 2.4.3) are collected at Q-band.

2.3.1 Free Electron in a Magnetic Field

Electrons in a magnetic field (B_0) exist in two energy states (parallel or antiparallel to the field) and are denoted by the magnetic spin components, or Zeeman states, $m_s = +1/2$ and $m_s = -1/2$ (Figure 2.2).^{1,20,38} These spin states are separated by an energy (ΔE , also called Zeeman splitting), which increases as the magnetic field (B_0) is increased by the equation:

$$\Delta E = h\nu = g_e\beta B_0$$

Equation 2.1

where g_e is the electron's g-factor, which is a property of the electronic configuration of the radical, or free electron, in the spin label. The g-factor is a measure of the local field experienced by the electron, which is a combination of the external magnetic field and the

interaction between neighboring spins.³⁸ β is the electron's Bohr magneton, which is a property of an electron described by:

$$\beta = \frac{e\hbar}{2mc}$$

Equation 2.2

where m and e are the mass and charge of an electron, respectively, \hbar is Planck's constant divided by 2π , and c is the velocity of light.^{20,38}

Transitions of the unpaired electron spin between spin states occurs if electromagnetic radiation is absorbed that equals the Zeeman splitting (ΔE).^{20,38} If Zeeman splitting was the only factor, only one line would be observed in an EPR spectrum corresponding to the field strength required for the unpaired electron to flip states as described by Equation 2.1 (Figure 2.2).

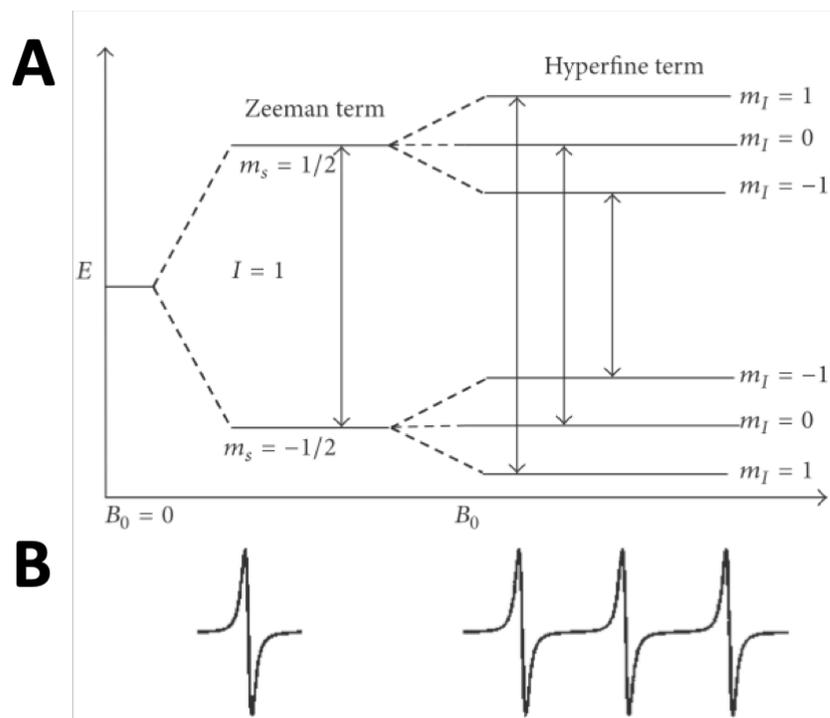


Figure 2.2. Energy diagram for nitroxide spin label in a magnetic field. A) Zeeman splitting results in two states and one energy transition, leading to a single line in the EPR spectrum (B, left). Hyperfine splitting further splits the high and low energy states into three distinct states, leading to three lines in the EPR spectrum (B, right). Figure reprinted from Sahu and Lorigan (2018), with permission.²⁰

2.3.2 Hyperfine Interaction

In a nitroxide paramagnetic spin label the unpaired electron interacts mainly with the nitrogen nucleus and this interaction is called the hyperfine interaction (A).²⁰ The hyperfine interaction causes slight changes in the allowed energy levels of the electron so each spin state (+1/2 and -1/2) is further split into 2I+1 states, where I is the spin quantum number of the atom.^{1,38} For a nitrogen (¹⁴N) nucleus, such as in a nitroxide spin label, I=1 so the hyperfine interaction results in three energy levels ($m_I = -1, 0, \text{ and } 1$). Thus, three lines are observed in EPR spectra (Figure 2.2) corresponding to the energy splitting.⁴¹

$$\Delta E = g_e \beta B_0 + A m_I$$

Equation 2.3

The other nearby nuclei, oxygen (¹⁶O) and carbon (¹²C), both have a spin quantum number of zero, and so do not contribute to further line splitting.¹

2.3.3 Anisotropy

The hyperfine interaction is usually anisotropic due to a non-uniform electron density of the nitrogen atom. The electron density of a nitrogen atom is concentrated in a p-orbital that is parallel to the z-axis of the nitroxide, and so the dipolar coupling dominates in the z direction.^{1,38} Thus, the g-factor (g_e) and hyperfine interaction (A) are dependent on the orientation of the spin label with respect to the magnetic field (B_0).⁴¹ Consequently, the position and splitting of the EPR lines are dependent on the anisotropic properties of the sample, as the energy of spin state transitions is altered following Equation 2.3. Magnetic anisotropy, therefore, leads to broadening of the CW lineshape. contrarily, if all

orientations with the field are equally sampled, the magnetic anisotropy is completely averaged and the CW spectra shows three sharp lines.³⁸

2.3.4 Relaxation

Another process that modulates the lineshape is relaxation. When the unpaired electron transitions from the low (n_B) to high (n_A) energy states (Figure 2.2), it must then relax or decay back to the low energy state by emitting a photon of energy equal to ΔE .^{39,40} Relaxation returns the net magnetization to equilibrium as defined by the Boltzmann distribution:

$$\frac{n_A}{n_B} = e^{\Delta E/kT} = e^{g\beta B_0/kT}$$

Equation 2.4

where $E=h\nu$, k is the Boltzmann constant, and T is temperature in Kelvin.^{38,41}

There are two relaxation processes which contribute to the overall relaxation of spins. These are spin-lattice relaxation (T_1), and spin-spin relaxation (T_2). T_1 relaxation is the transfer of thermal energy from the unpaired electron to the surrounding environment (lattice). Smaller T_1 values (higher amounts of exchange and faster relaxation) leads to a smearing or variation of the energy levels, and thus leads to a broad distribution in the lineshape.^{39,40} T_2 relaxation occurs due to the interaction of an individual spin with other spins in the system via dipolar and exchange interactions, as well as field inhomogeneity.^{39,40} T_2 relaxation reduces the net magnetization of the spins and leads to a reduction in signal intensity and broadening of the spectra.³⁸⁻⁴¹ For a nitroxide spin label, T_2 is much shorter than T_1 and so is the dominant contribution to the linewidth (ΔH):³⁸⁻⁴¹

$$\Delta H \propto \frac{1}{T_1} + \frac{1}{T_2}$$

Equation 2.5

2.3.5 Correlation Time

A CW EPR spectrum reflects the rotational correlation time (τ) of the nitroxide spin label on the pico- to nanosecond timescale. The rotational correlation time (τ) includes contributions from the correlation times of side chain internal motions (τ_i), fast backbone fluctuations (τ_{bb}), and overall protein rotational diffusion (τ_r).⁹ For small proteins (less than 15 kDa) τ_r will dominate and cause spectral broadening.¹ As τ_i and τ_{bb} reflect the protein's structure and dynamics, and are thus the desired components to investigate, it is advantageous for τ_r to be sufficiently long to have minimal effects on the lineshape ($\tau_r > 20$ ns).⁹ τ_r is influenced by both protein molecular weight (MW) and solvent viscosity (η) by:

$$\tau_r \approx (3 \times 10^{-13})\eta \times MW$$

Equation 2.6

The viscosity of the solvent can be increased by the addition of sucrose or Ficoll to reach $\tau_r > 20$ ns.⁹ Internal motions (τ_i) of the R1 side chain are largely dominated by χ_4 and χ_5 dihedral angles (Figure 2.3), and are thus anisotropic.⁴² Backbone fluctuations (τ_{bb}) on the nanosecond timescale increase the mobility of R1, while interactions of the spin label side chain with local protein structure decrease the mobility. Thus, CW spectra can be used to map backbone flexibility, classify secondary or tertiary structure, and identify conformational changes.⁹

2.4 Applications of EPR

2.4.1 Site-Directed Spin Labeling

EPR requires a paramagnetic center with an unpaired electron. As most proteins do not naturally contain an unpaired electron, they must be experimentally added via site-directed spin labeling (SDSL). Three methods have been developed to add paramagnetic spin labels to proteins of interest.⁸ The first is by reacting the spin label with the thiol group of a protein's cysteine residue. This is the most common method and will be described here.^{1,8,19,20,42} The second method is to spin label directly by peptide synthesis.⁸ In practice, this is only feasible when studying small proteins or peptides. The third method is to spin label via conjugation with unnatural amino acids.^{43,44} This method is useful when the protein has native cysteines that cannot be removed, or if multiple types of spin labels are desired in the same protein. Spin labeling via unnatural amino acids is described in further detail in Section 4.2.2.3.

For SDSL any reactive native cysteine residues must first be removed from the protein (generally these are replaced with alanine or serine residues) via mutagenesis. Then, the site(s) of interest to be spin labeled are mutated to cysteine residues. Considerations in the choice of spin labeling site are discussed in Section 2.5.1. When the thiol of the cysteine residue is reacted with the spin label, the spin label is covalently attached to the protein (Figure 2.3) and is denoted R1 for a Methanethiosulfonate spin label (i.e. residue C becomes R1). Methanethiosulfonate spin label (MTSL) is the most commonly used and most characterized spin label, and is used in this work (Figure 2.3).⁸ MTSL is comparable in size to a tryptophan or phenylalanine side chain and thus has

minimal impact on the fold or dynamics of the protein studied.³⁸ If necessary, spin labels can be removed from proteins by adding a denaturing agent.

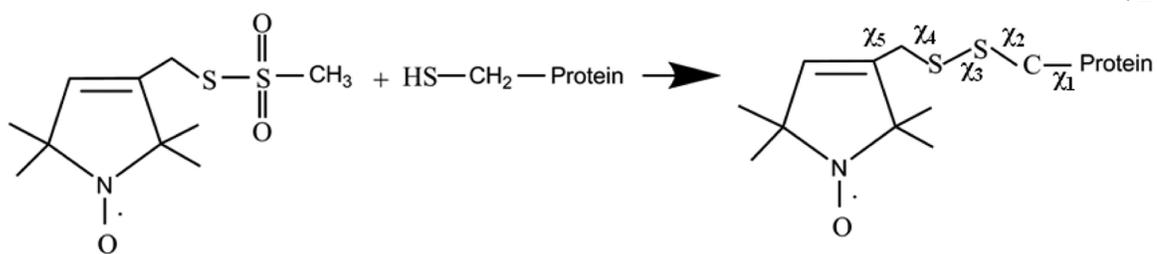


Figure 2.3. Nitroxide spin label reaction. Reaction of MTSL with the protein cysteine results in a spin labeled protein. The spin label has five sites of bond rotation labeled $\chi_1 - \chi_5$. Figure reprinted from Sahu, McCarrick, and Lorigan (2013), with permission.²

2.4.2 Continuous Wave EPR

Continuous wave (CW) EPR is an excellent technique to probe local protein effects such as dynamics, local environment, secondary structure, tertiary contacts, and solvent accessibility. Here I will describe how the spectral lineshape reflects the protein structure and dynamics. Other CW EPR applications, such as distance measurement in the 8-20 Å range, and power saturation experiments to determine membrane topology or binding, are described elsewhere^{2,20} and will not be addressed here.

CW spectra are recorded at a constant microwave frequency while sweeping the magnetic field. The absorption of microwave photons by the sample at specific resonant frequencies are detected and are recorded as the first derivative of the absorbance spectra.¹

In Section 2.3.1 and 2.3.2, I described the Zeeman and hyperfine splitting that results in three lines for a nitroxide spin label CW-EPR spectra (Figure 2.2). I will now touch on some of the other factors that contribute to the CW lineshape.

2.4.2.4 How to “Read” a CW Spectrum

To the untrained eye, CW spectra are just lines that are hard to interpret. However, CW lineshapes of nitroxide spin labels can give important information about the dynamics of the spin label. Here, I will explain the basic parts of the R1 lineshape and how to interpret the features of a CW spectra to gain information about the structure and dynamics of the protein.

In Section 2.3, I described the Zeeman and hyperfine splitting that results in three lines for a nitroxide spin label CW-EPR spectra (Figure 2.2). In the fast motional limit (~0.1 nsec, such as free spin label) the CW spectrum shows three sharp lines of equal

height. As motion slows, the lineshape broadens, the amplitudes decrease, and the lineshape becomes asymmetric (Figure 2.4).¹ Thus, the position of the spin label on the protein can be categorized as mobile, such as on a loop; semi-restricted, such as a surface exposed site containing secondary structure; or immobile, such as buried in the protein core (Figure 2.4).^{9,42} The broadening of spectra with decreasing mobility is the result of anisotropic and relaxation components becoming less averaged as described in Sections 2.3.3 and 2.3.4.

A typical semi-restricted spin label, such as on a surface exposed rigid α -helix, is shown in Figure 2.5.⁹ This spin label can be identified as semi-restricted due to the broadening and decreased amplitudes of the low and high field lines. The peak to peak (from peak max to min) separation of the first derivative lineshape, ΔH (Figure 2.5), is related to the rotational correlation time, and can be described by the approximation:¹

$$\tau = 6.5 \times 10^{-10} \Delta H \left[\left(\frac{A(0)}{A(-1)} \right)^{1/2} - 1 \right]$$

Equation 2.7

where $A(0)$ and $A(-1)$ are the peak to peak heights of the center and high field lines, respectively. Generally, the inverse of the central linewidth (ΔH^{-1}) is reported and is a report of nitroxide motion due to fluctuations of the backbone.^{9,45}

The incomplete averaging of the hyperfine tensor, which leads to broadening, can be separated into two components. A subset of nitroxide spins align parallel to the external magnetic field, and another perpendicular to it. These are defined as A_{par} and A_{perp} , respectively, and can be resolved in the broadened low field lineshape (Figure 2.5).⁹

Especially in the low mobility regime, the overall spectral splitting ($2A_{zz}'$) is a measurement of spin label mobility as the positioning of the high and low field lines are dependent on spin label movement.⁹ Smaller values of $2A_{zz}'$ indicate increased mobility while larger values indicate restricted motion.¹

Spectra may also be classified as “two-component” or “multi-component” if there are two resolved components corresponding to a “mobile” and “immobile” state of the spin label (Figure 2.6).⁹ Two-component spectra should not be confused with the presence of A_{par} and A_{perp} , as these are within the same lineshape, and two-component elements are distinct lineshapes (Figure 2.5). Multi-component spectra may be indicative of multiple side chain and/or protein conformational states indicating conformational dynamics.⁹

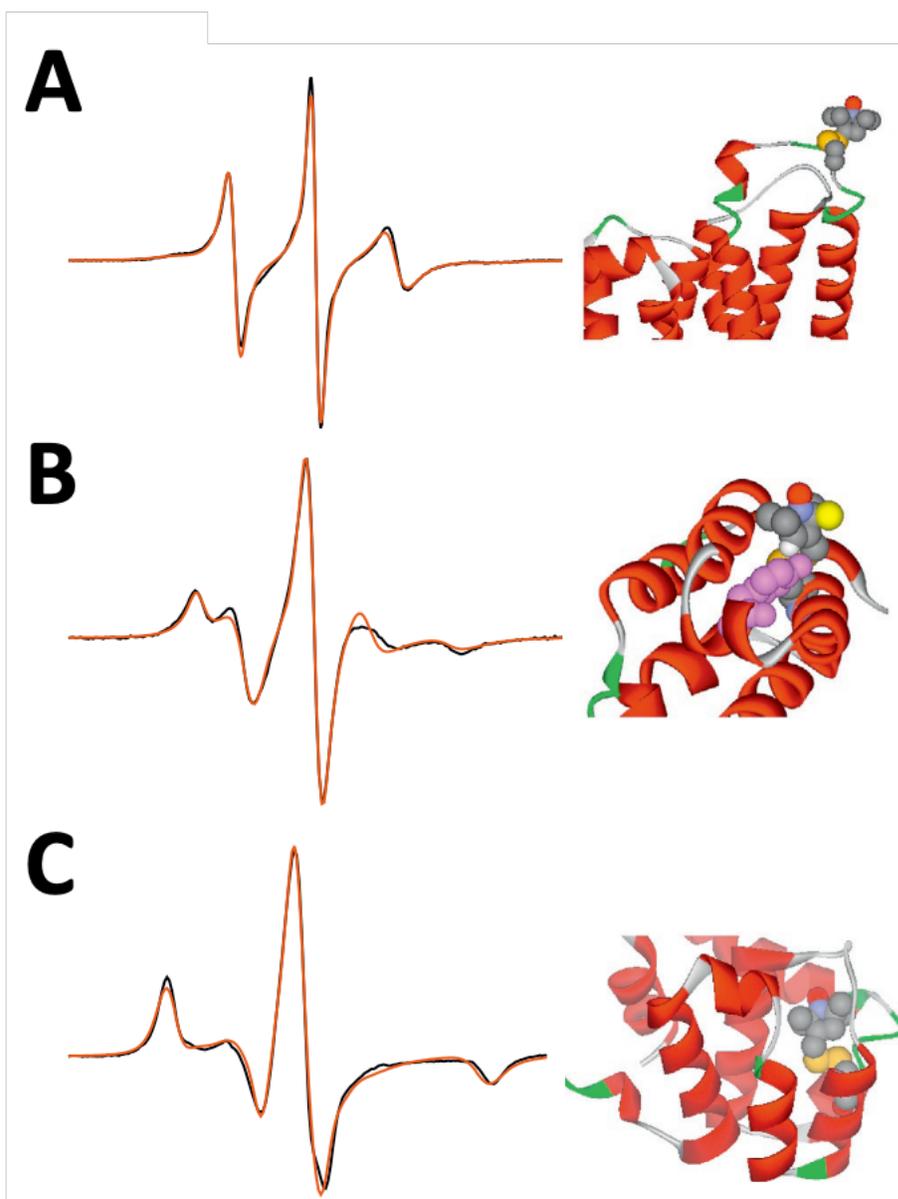


Figure 2.4. Mobile, semi-restricted, and immobile CW spectra. A) A mobile lineshape is seen when the spin label is attached to a loop residue. B) A semi-restricted lineshape is seen when the spin label is attached to a solvent exposed site that contains secondary structure, in this case an α -helix. C) An immobile lineshape is seen when the spin label is attached to a buried residue. Red trace shows experimental CW spectra while black shows simulated spectra. Figure adapted from Altenbach *et. al* (2015), with permission.⁹

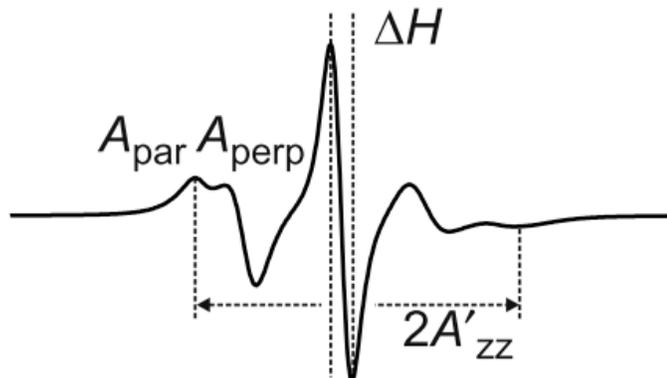


Figure 2.5. Characterization of CW lineshape. A single-component semi-restricted lineshape is shown of a spin label attached to a solvent exposed site that contains secondary structure (helix, Figure 2.4B), with labeled lineshape elements. Figure adapted from Altenbach *et. al* (2015), with permission.⁹

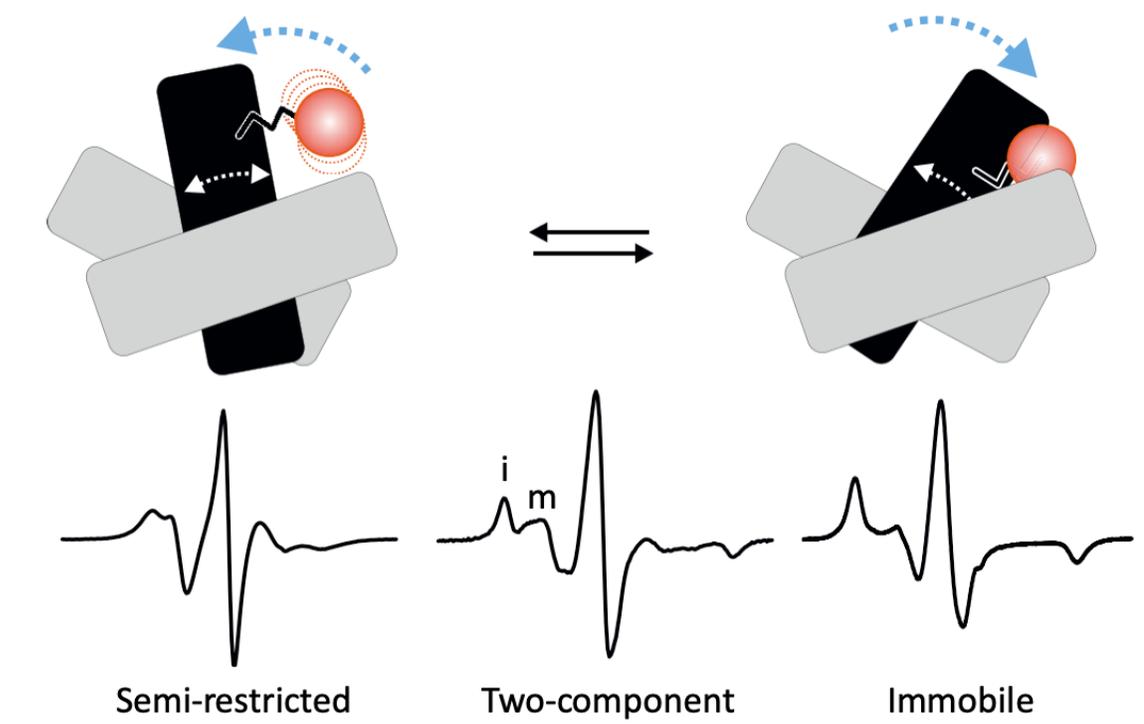


Figure 2.6. Two-component CW spectra resulting from conformational dynamics.

When a spin label samples both a semi-restricted (left) and an immobile (right) conformation due to conformational dynamics, the resulting lineshape is two-component (middle) with a mobile (m) and immobile (i) population. Figure adapted from Altenbach *et. al* (2015), with permission.⁹

2.4.3 Double Electron Electron Resonance

Double electron electron resonance (DEER) spectroscopy (also called pulsed-electron double resonance, PELDOR), can be employed to measure the distance distribution between two spin labels. These spin labels may be within the same protein, aiding in identification of protein structure and conformational dynamics; or, two different proteins can be singly labeled to investigate binding and the complex structure.²⁰ DEER extracts distance information from pairs of spin labels due to the distance dependence of the energy of the dipolar interaction between the unpaired electrons.²² Historically, this was measured by dipolar broadening of the CW spectrum which could only detect distances up to around 1.5 nm.²³ However, after the introduction of the 4-pulse DEER method in 2000, and an increase in the availability of commercial instrumentation to perform DEER, it became possible to measure longer spin label distances.⁴⁶ While 16 nm is the longest detected distance so far measured with DEER,⁴⁷ distances of membrane proteins generally fall within the 1.8 - 6 nm range.¹²

2.4.3.1 Dipolar Coupling

If two unpaired electrons are in close enough proximity, their magnetic dipoles will interact in a dipole-dipole interaction which is dependent on their distance (r) and angle (θ) with respect to the magnetic field (B) (Figure 2.7).¹⁴ The dipolar coupling (ω_{dd}) between the two electrons is defined as:

$$\omega_{dd} = \left(\frac{C}{r^3}\right) (1 - 3 \cos^2 \theta)$$

Equation 2.8

where C is proportional to the product of the g -values of the two spins and equals $52.2 \text{ MHz} \cdot \text{nm}^{-3}$ for nitroxide spin labels.¹² In order to determine the distance distribution between the spin labels, the dipolar coupling must be separated from other spin interactions such as g anisotropy and hyperfine couplings, which can be accomplished by using pulsed EPR experiments.^{14,16}

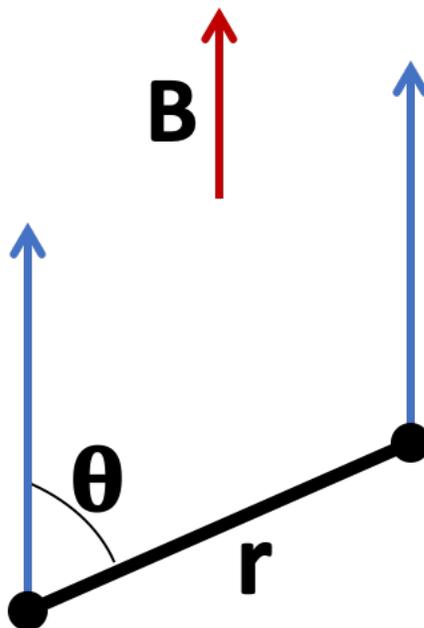


Figure 2.7. Orientation of spins. Two spins (black circles) are separated by a distance (r) and are oriented at an angle (θ) with respect to the magnetic field (B). Figure adapted from Jeschke (2012), with permission.¹²

2.4.3.2 Pulsed EPR

DEER experiments generally use a four-pulse sequence in order to extract the dipolar coupling, and thus the distance, between two electron spins (here named A and B). Simply, this is done by monitoring one set of spins while exciting another set of spins with a microwave frequency.²⁰

In the four-pulse DEER experiment (Figure 2.8), the first $\pi/2$ pulse occurs at a frequency ν_A which excites electron spins A. This pulse tips the magnetization of spins A into the x-y plane. Spins A precess, but lose coherence over time t_1 , due to T_2 relaxation of the spins (see Section 2.4.2.2). A portion of this relaxation is due to modulation of the angular rate of spin A by dipolar coupling from nearby spins B ($\pm\omega_{dd}$). After t_1 , a π pulse at frequency ν_A flips the spins 180° in the x-y plane which then refocus after an equal time (t_1), which gives a spin echo. This $\pi/2$ - π pulse sequence is called a Hahn echo.²¹

To determine the dipolar coupling, a π pump pulse is applied at frequency ν_B , flipping spins B (at an inversion efficiency fraction λ), at varied time intervals (t).³⁸ This pulse reverses the dipolar contribution experienced by spins A (from $+\omega_{dd}$ to $-\omega_{dd}$ and vice versa) causing spins A to accumulate a phase lag. This phase lag causes the spins to no longer be refocused, and the intensity of the echo decreases. Thus, the phase lag is a function of both the dipolar interaction (ω_{dd}) and the timing of the pump pulse (t), which determines how long spins A experience the dipolar field due to spins B, such that the phase difference is $\omega_{dd}t$.²¹ Lastly, after a time t_2 a second π pulse at ν_A refocuses spins A to generate the refocused echo which has been modulated by the phase lag which is finally collected after a second time t_2 (Figure 2.8).

The DEER signal between two spins A and B can thus be described by:

$$V(t)_{intra} = 1 - \lambda[1 - \cos(\omega_{ad}t)]$$

Equation 2.9

The entire DEER signal is the product of intra and intermolecular interactions:¹⁶

$$V(t) = V(t)_{intra}V(t)_{inter}$$

Equation 2.10

The intermolecular interactions have a background function:¹⁵

$$B(t) = e^{-Kt^{D/3}}$$

Equation 2.11

where D is the dimension and K is the background decay rate.^{12,15} The dimension for soluble proteins and membrane proteins in micelles or bicelles is D=3 as the spins are distributed homogeneously in three dimensions. For membrane proteins in liposomes D=2 as they are confined into a two dimensional bilayer.¹⁵ If there are multiple spin pair orientations which contribute to the signal, then the DEER signal takes the form:

$$V(t) = F(t)B(t)$$

Equation 2.12

where the form factor F(t) is the product of all possible spin pair contributions.¹² Thus, the background function [B(t)] can be fit and subtracted from the DEER signal, [V(t)] to isolate the signal from just the intramolecular components [V(t)_{intra} or F(t)] which contain the dipolar coupling and, therefore, the distance information.¹⁶

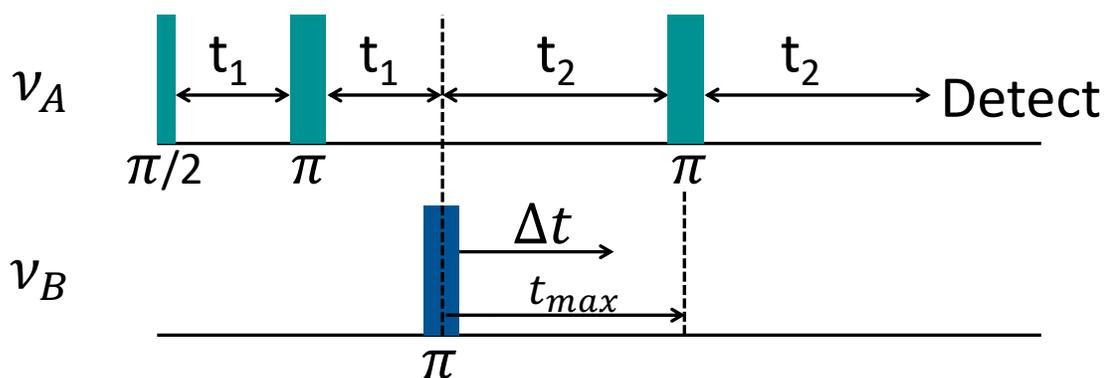


Figure 2.8. Four-pulse DEER experiment. A refocused echo is generated for spins A with a $\pi/2 - \pi - \pi$ sequence at an observer frequency ν_A separated by times t_1 and t_2 . The intensity of the echo is modulated by a π pump pulse at ν_B at various times t . The modulation of the echo gives information about the dipolar interaction and thus the distance between the spins A and B.

2.4.3.3 Distance Analysis

Distance information can be gleaned from simply looking at the background subtracted DEER signal. The form factor has an oscillation frequency that is inversely proportional to the cube of the mean distance between the spin labels (Figure 2.9). Thus, the longer the distance, the longer it takes for the signal to reach baseline.²¹ The decay of the form factor is related to the width of the distance distribution: a faster decay indicates increasing width (σ) of the distance distribution (Figure 2.9).^{15,21} Narrow distributions have multiple oscillations, but for wide distributions the DEER signal barely goes below baseline (Figure 2.9).²¹ If the spin pair has multiple conformations, the DEER signal will be additive of the two distances proportional to the percentage of each state (Figure 2.10).²²

Transforming the intramolecular DEER signal into a distance distribution, $P(r)$, is complex and there are several excellent reviews on this topic.^{12,21,46,48,49} Most commonly, DEER data is processed by Tikhonov regularization with software such as DeerAnalysis (written by Gunnar Jeschke)⁴⁹ or LongDistance (written by Christian Altenbach). As fitting the DEER signal time domain data is an “ill-posed” problem where small changes due to noise or error can induce large changes in $P(r)$, Tikhonov regularization is used to balance the smoothness and resolution of the DEER distribution.^{38,48} This is accomplished by plotting the log of the distribution smoothness (η) versus the log of the mean squared deviation (ρ) in an L-curve (Figure 2.11).¹⁵ The best fit of the regularization parameter (α) occurs at the corner of the L-curve. Large α values lead to over-smoothing of $P(r)$, while small α values lead to unrealistically sharp peaks (Figure 2.11).^{15,48}

The generated $P(r)$ plot yields the probability distribution of spin pair distances defined by the average distance (r_{av}) and the width of the peak (σ) (Figure 2.10). The

distance distribution informs on the collective dynamic processes of the spin labels including spin label side chain dynamics, backbone fluctuations, and protein conformations.²² Multimodal distance distributions may indicate a two state conformational change (Figure 2.10). Narrower distributions indicate a more rigid structure, while wider distributions indicate flexibility of one or both of the spin label sites. The fact that DEER yields a distance distribution, rather than a single distance, is advantageous in that it allows for the identification of both structural and conformational dynamic information about the protein of interest.

While it is impractical to collect DEER data for enough pairs to sufficiently determine the overall structure of a protein using distance restraints, DEER is an excellent way to investigate conformational changes in varying situations such as in different membrane systems or upon the addition of a ligand, inhibitor or binding partner. DEER data can be used in conjunction with molecular dynamics simulations to refine protein structures with experimentally derived constraints.^{20,50,51} Using the program Multiscale Modeling of Macromolecules (MMM), it is also possible to simulate DEER data from a crystal or NMR derived structure and compare it to experimental DEER data.⁵² This can aid in visualizing the distance distribution in the context of the protein structure, or identify discrepancies between experimental conditions.

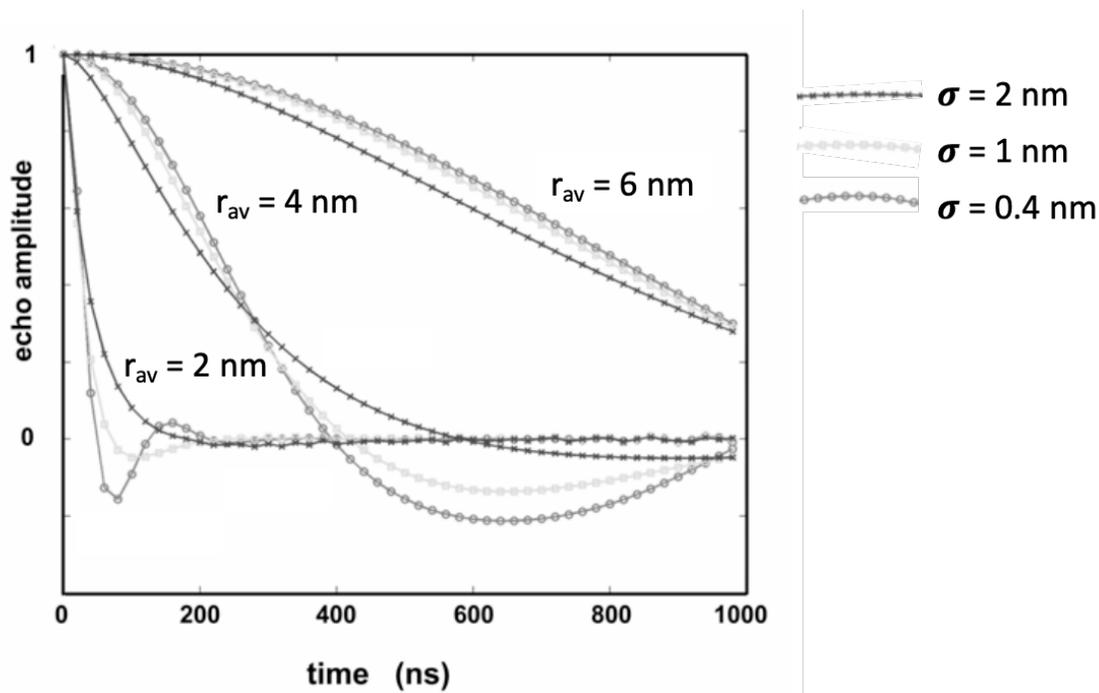


Figure 2.9. Effect of average radius and distribution width on DEER signal. A faster decay indicates a wider distribution (σ), while a higher oscillation frequency indicates a shorter average distance (r_{av}). Figure adapted from Hemminga *et. al* (2007), with permission.²¹

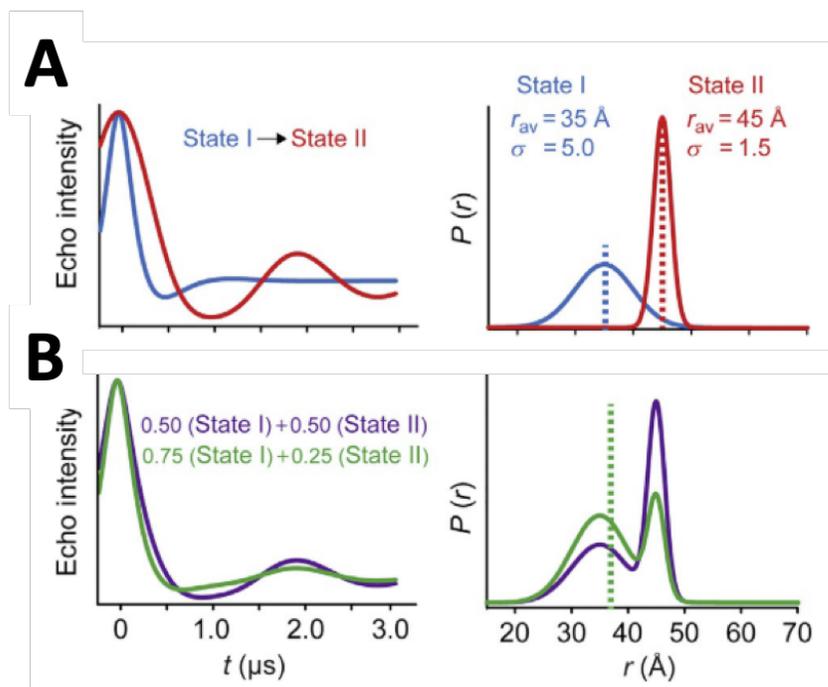


Figure 2.10. Multi-component DEER. A) Two states have unique dipolar evolution functions obtained using the DEER experiment (left, blue and red) resulting in unique distance distributions (right). B) If these states occur within the same DEER sample, the resulting DEER data will be additive of the two spectra with respect to the percentage of each state in the sample. The purple trace shows 50% of each state while the green trace shows 75% of state one and 25% of state two. Figure adapted from Claxton *et. al* (2015), with permission.²²

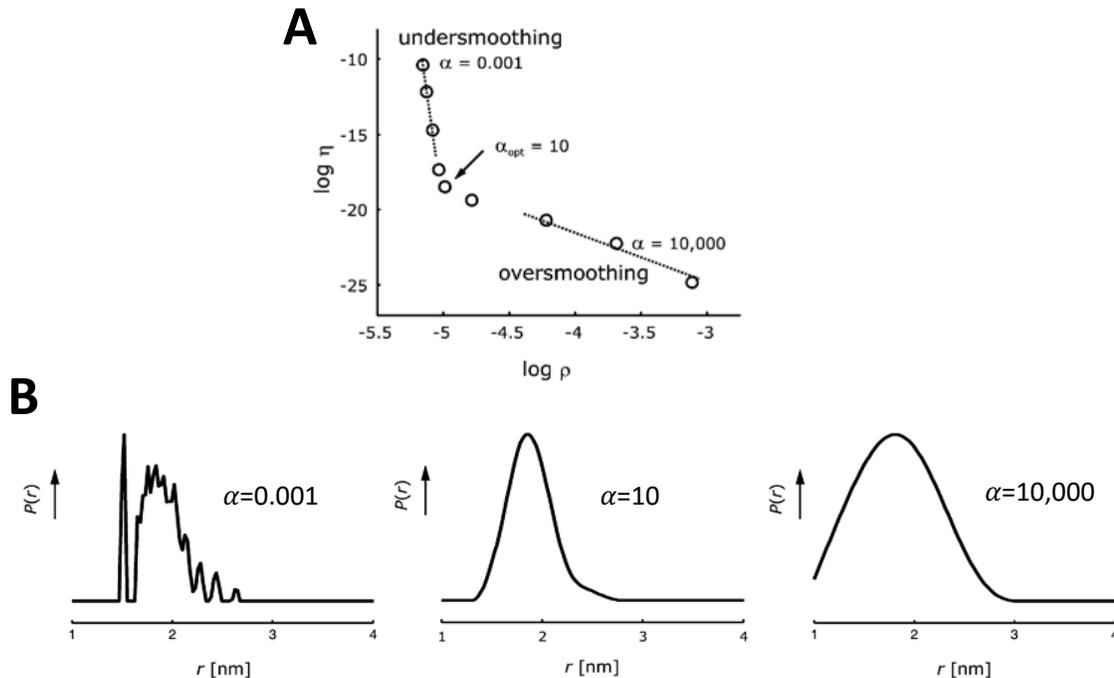


Figure 2.11. Tikhonov regularization L-curve. A) The correct choice of α is in the corner of the L-curve. B) Choosing an α that is too small yields unrealistically sharp peaks (left), while choosing an α that is too large results in over-smoothing (right). In this example $\alpha=10$ (middle) is the correct choice. Figure adapted from Jeschke *et. al* (2007), with permission.¹⁵

2.5 *Practical Considerations for EPR Experimentation*

2.5.1 Sample Preparation

The first consideration when preparing membrane proteins for EPR analysis is to determine if there are any native reactive cysteine residues. As the most frequently used spin labels, including MTSL (see Section 2.4.1), specifically react with the cysteine thiol to selectively label sites of interest, native cysteines must first be removed via mutagenesis and are replaced by another amino acid (most commonly alanine or serine).²⁰ If native reactive cysteines are important for the protein's fold and/or function (as in the lipoprotein substrate ICP, see Section 4.2.2.3), then they should not be removed. In this case, unnatural amino acids can be genetically incorporated at sites of interest which contain reactive groups that can be specifically spin labeled, generally via click-chemistry.^{43,44,53,54}

Second, spin label sites need to be identified. Mutations of conserved residues or residues buried in the core of the protein should be avoided to avoid interfering with the protein's fold and/or function. If available, a crystal structure can be used as a guide to choose sites that will not disrupt any important interactions or interfere with protein dynamics or function.^{13,22} To ensure that the spin label does not interfere with the protein fold, spin labeled protein can be analyzed by circular dichroism (CD)⁵⁵ or analyzed by NMR heteronuclear single quantum coherence spectroscopy (HSQC) to assess the secondary structure compared to unlabeled protein.⁵⁶ As practically only a limited number of sites are spin labeled, it is important to take time to critically choose sites that will yield the greatest amount of information, be that secondary structure, loop dynamics, ligand binding, distance restraints, etc. If probing conformational dynamics via DEER, choose spin label sites that you believe will have the largest changes. As DEER reports a distance

distribution, small changes can be contained within the same distribution peak and may be missed. To decrease the width of the distance distribution and increase resolution, choose spin label sites on exposed sites containing secondary structure. If the site is mobile, as on a loop, the flexibility will increase the width of the distance distribution and again conformational changes may be missed. To best understand conformational dynamics or distance restraints, make spin label pairs in a triangle, quadrangle, or pyramid shape.²² These complimentary distance restraints will better elucidate the conformation of the protein. Recently, computational methods have been developed to select the most informative pairs of spin labels to refine protein structures.^{50,51}

Third, sample quality must be considered. The sample should be pure, not have degradation products or aggregation, and not have unattached spin label contaminant.^{22,57} Not only will these issues affect the sample stability and concentration, but it will increase the background noise and can lead to errors in data analysis and, thus, impact interpretation of the data. Not only will destabilized protein show unrealistic spin label arrangement with contamination, but the spin label may have altered packing and show non-physiological tertiary contacts. Additionally, contamination and destabilization may alter the thermodynamics of conformational transitions, and effect the distribution of conformations.²² Along with sample quality, it is essential to know the oligomeric state the protein of interest. If the protein contains a pair of spin labels, protein oligomerization will cause DEER signal from both intra- and inter-protein spin labels. Thus, it is essential to first test single labeled protein to ensure that no DEER signal is seen from inter-protein contacts. A control without reactive cysteines should also be subject to the spin labeling process and analyzed to ensure that there are no other contributions to the DEER signal.^{1,57}

Fourth, the membrane mimetic must be considered. While EPR is suitable for the study of membrane proteins in a wide range of membrane mimetics including micelles, bicelles, nanodiscs, amphipols, and liposomes, it is essential to be cognizant of the effects these mimetics might have on the protein structure and dynamics. For example, larger mimetics may allow for increased conformational dynamics while smaller ones may constrict the protein. Additionally, mimetics may influence the quality of data recorded.⁵⁷ For example, the heterogeneous distribution of multiple spin labeled proteins within liposomes creates high background and poor DEER modulation leading to a reduction in the measurable distance range.^{12,20} This reduction in distance range is due to the spatial distribution of spins being reduced from three dimensions to two.²² Comparatively, proteins are generally distributed in a one to one ratio in nanodiscs and bicelles so the spatial distribution in three dimensions is maintained.²² Having the proper ratios of detergents or lipids in the system is also important. Varying lipid ratios impairs nanodisc formation, while a high detergent to protein ratio decreases spin label efficiency which decreases the EPR signal.²²

Lastly, the effect of solvent and additives needs to be considered. While EPR measurements can be recorded in a wide range of solvents and conditions, the researcher must be aware of the effects that changing these conditions from experiment to experiment may bring. For example, in this work inhibitors dissolved in DMSO were added to protein samples and studied by CW-EPR. However, DMSO alone caused significant lineshape changes (Figure 2.12). Thus, the inhibitor had to be aliquoted, dried down, and resuspended in the protein solution so that no DMSO was added to the protein sample. Consideration of all components of the protein sample – the placement of spin labels, sample quality,

membrane mimetic, and solution conditions – is essential when conducting EPR experiments.

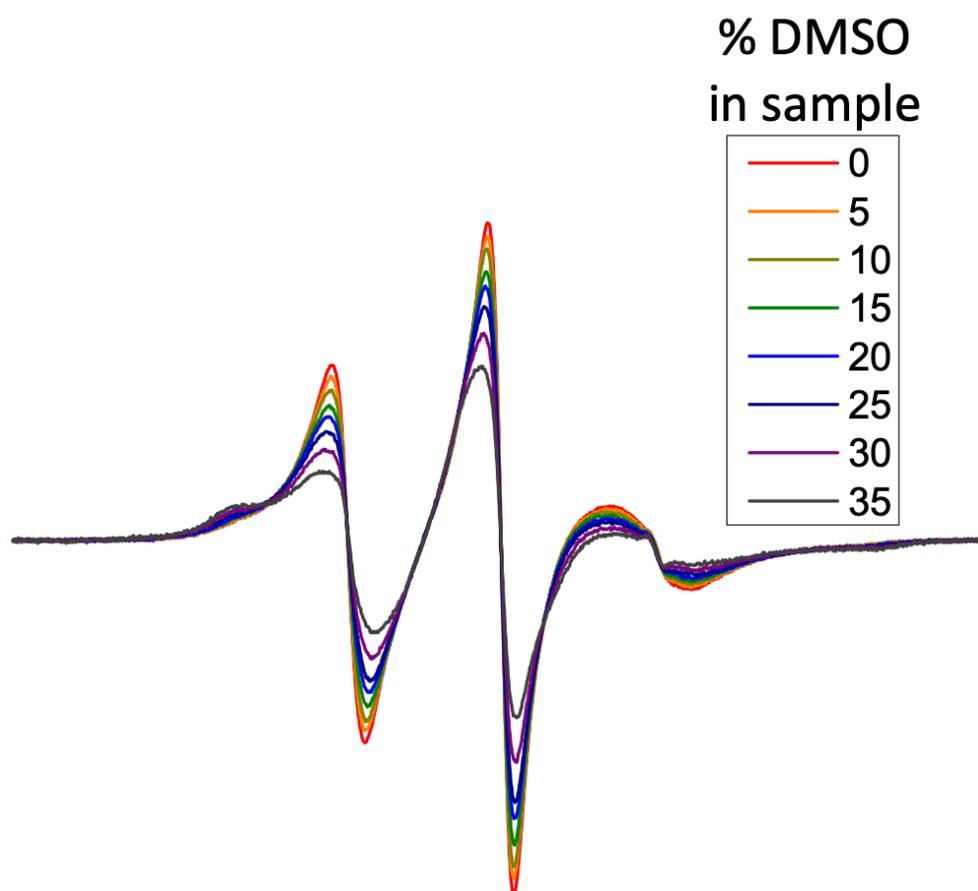


Figure 2.12. DMSO addition causes significant CW lineshape changes. CW spectra are shown for WT LspA spin labeled at A63R1 in 0-35% DMSO.

2.5.2 CW Spectral Analysis

As mentioned in Section 2.5.1, one must be aware of the effects of the solvent used. When assessing the effect of additives (drugs, inhibitors, lipids, etc.), controls must be completed to make sure that the solvent that these additives are in do not affect the lineshape (Figure 2.12). For CW spectra, it is especially important to not change the polarity of the solution, as polar solvents increase polarization of the N-O bond of the nitroxide which increases the isotropic hyperfine coupling.¹

When analyzing two-component CW spectra, one must be aware of the different situations that may give rise to this lineshape. Two components may be derived from two conformational states of the protein and, thus, of the spin label. However, the two components may also arise from two rotamers of the nitroxide side chain, one which places the nitroxide in an orientation that interacts with a nearby side chain giving the “immobile” lineshape, and one that faces outward giving the “mobile” lineshape. The origin of the two-component lineshape (either protein conformational dynamics or two spin label rotamers) can be resolved by osmolyte perturbation.⁹ The addition of an osmolyte solution, such as sucrose or polyethylene glycol (PEG), has little effect on the rotameric state of the side chain, but will shift the conformational equilibria of the protein toward the least solvent-exposed or most compact state. Thus, if the two-component spectrum remains in the presence of osmolytes, then the components arise from two side chain rotamers. If the two-component spectrum disappears, then the two components were indeed from multiple conformational states.⁹ While modeling can be completed to determine the energy difference between the two components, the CW spectra gives no information about the exchange rate between states.⁹

Correctly processing CW spectra is essential in order to directly compare multiple CW spectra. First, the spectrum is baseline corrected such that the baseline is flat. Then, the spectrum is phased to maximize the second derivative of the sample.⁴¹ Finally, the spectrum is normalized. As the number of scans increases the signal intensity, normalizing allows spectra collected with various numbers of scan to be directly compared.

2.5.3 DEER Spectral Analysis

As mentioned in Section 2.5.1, while EPR can be conducted in multiple membrane mimetics, the effects of these mimetics must be considered. DEER is a good technique to investigate structure and/or dynamic changes in different membrane environments. However, membranes, particularly liposomes pose additional challenges for DEER analysis as the heterogeneous distribution of spin labeled proteins within the liposome alter relaxation times due to spin diffusion.^{20,23} These can lead to poor DEER modulation and thus raise difficulties in accurate and precise distance measurements.^{20,23} Deuteration of solvents, proteins, and membrane mimetics can help with relaxation times to improve the quality of the data; however, this can be costly and difficult.²³

The DEER distribution measures the distance between the unpaired electrons. Thus, when analyzing distance distribution data, it is important to take into consideration the added distance of the spin label.¹² For a nitroxide spin label the unpaired electron resides between the N – O bond (Figure 2.3), which is about 0.8 nm from the protein backbone.²³ The spin label may sample many rotamer orientations as demonstrated in Figure 2.13. The conformational space that the spin label samples is contained within the distance distribution. Software such as MMM can help identify contributions of the spin

label to the distribution by modeling the spin labels onto the protein at the spin labeling sites and simulating the DEER distribution between them.⁵² MMM can “fit” the simulated spectra to the experimental DEER data by selecting the rotamer population that is sampled in the DEER distribution.⁵² However, it is critical to consider if the fit distribution makes physiological sense. For example, it does not make sense if only a small portion of the rotamers are sampled in the fit distribution, as this would not occur in solution. Acquiring CW spectra of each of the spin labels individually can aid in this assessment by determining how mobile the spin label is at each site.

Understanding the limits of the distance information that can be extracted from DEER data is essential for proper interpretation. The lower distance limit that can be measured with DEER depends on the frequency of the instrument, but generally falls within 1.7-2.0 nm.^{12,15} One full oscillation period is required for accurate distance determination, which requires that t_{max} is long enough to do so.¹⁶ Thus, accurate measurement of the maximal average distance (r_{av}) and the width of the distribution (σ) is dependent on t_{max} :¹²

$$r_{av,max} \approx 5 \sqrt[3]{\frac{t_{max}}{2 \mu s}} \text{ nm}$$

Equation 2.13

$$\sigma_{max} \approx 4 \sqrt[3]{\frac{t_{max}}{2 \mu s}} \text{ nm}$$

Equation 2.14

This means that for a sample with $t_{max} = 3.5 \mu s$, which is common for membrane protein samples, $r_{av,max} \approx 6$ nm and $\sigma_{max} \approx 5$ nm.¹² For membrane proteins, the upper distance limit for interpretation of peak asymmetries or shoulders is only 3.6 nm.¹² Above this distance, peaks should only be thought of as Gaussian and peak asymmetries cannot be interpreted.

Accurate distance information, both average and width of distribution, is additionally influenced by the signal to noise ratio and error associated with background subtraction, and Tikhonov regularization. Figure 2.11 shows the drastic effects that the choice of α has on the resulting distance distribution.¹⁵ Background signal is dependent on protein concentration due to crowding. If spin label concentration is high (200-500 μM), the background signal will dominate and is difficult to remove. A short t_{max} will also cause an ill-defined background, and can result in artifacts in the distance distribution.²² On the other hand, longer t_{max} can decrease the signal to noise ratio leading to error in the analysis.²² Additionally, aggregation or dimerization can yield non-physiological longer distances in the distance distribution.¹² Thus, it is essential to always test a singly labeled protein with DEER to ensure that there are no background contributions from aggregation, contaminant, or dimer.

These considerations are all taken into account in this work when the LspA CW and DEER spectra were analyzed in Chapter 3.

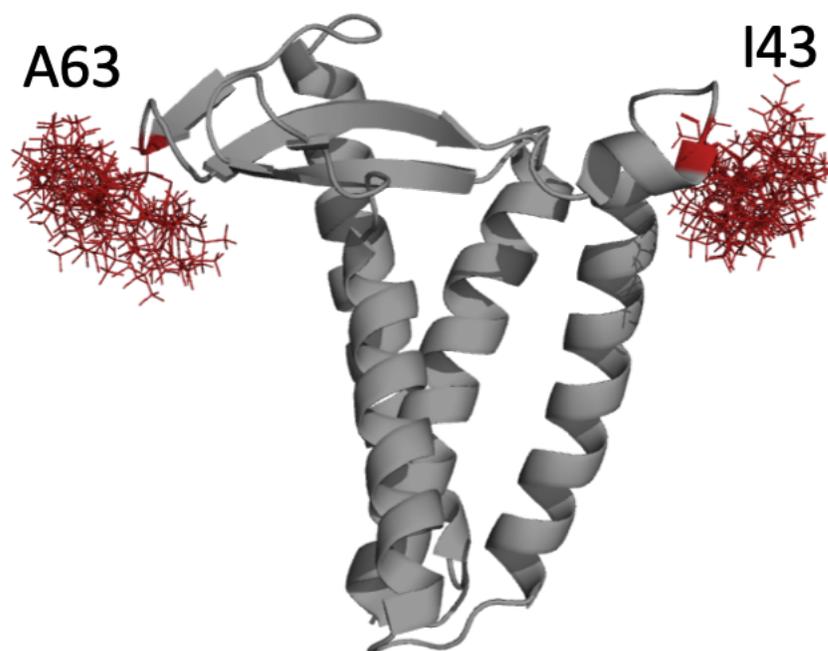


Figure 2.13. Spin label rotamers. MTSL spin label can occupy many rotamers, as shown by the twenty lowest energy conformations at two spin labeling sites. Protein shown is LspA (PDB ID: 5dir) and is spin labeled at I43 and A63 as described in Chapter 3.

2.6 References

- 1 Klug CS, Feix JB. Methods and Applications of Site-Directed Spin Labeling EPR Spectroscopy. *Methods Cell Biol* 2008;**84**:617–58. [https://doi.org/10.1016/S0091-679X\(07\)84020-9](https://doi.org/10.1016/S0091-679X(07)84020-9).
- 2 Sahu ID, McCarrick RM, Lorigan GA. Use of electron paramagnetic resonance to solve biochemical problems. *Biochemistry* 2013;**52**:5967–84. <https://doi.org/10.1021/bi400834a>.
- 3 Eaton SR, Eaton GR, Berliner LJ. *Biomedical EPR - Part B: Methodology, Instrumentation and Dynamics*. 2005.
- 4 Fanucci GE, Cafiso DS. Recent advances and applications of site-directed spin labeling. *Curr Opin Struct Biol* 2006;**16**:644–53. <https://doi.org/10.1016/j.sbi.2006.08.008>.
- 5 Hubbell WL, Gross A, Langen R, Lietzow MA. Recent advances in site-directed spin labeling of proteins. *Curr Opin Struct Biol* 1998;**8**:649–56.
- 6 Hubbell WL, Lopez CJ, Altenbach C, Yang Z. Technological advances in site-directed spin labeling of proteins. *Curr Opin Struct Biol* 2013;**23**:725–33. <https://doi.org/10.1016/j.pain.2013.06.005>. Re-Thinking.
- 7 Hubbell WL, Cafiso DS, Altenbach C. Identifying conformational changes with site-directed spin labeling. *Nat Struct Biol* 2000;**7**:735–9. <https://doi.org/10.1038/78956>.
- 8 Klare JP, Steinhoff HJ. Spin labeling EPR. *Photosynth Res* 2009;**102**:377–90. <https://doi.org/10.1007/s11120-009-9490-7>.
- 9 Altenbach C, López CJ, Hideg K, Hubbell WL. Exploring Structure, Dynamics,

- and Topology of Nitroxide Spin-Labeled Proteins Using Continuous-Wave Electron Paramagnetic Resonance Spectroscopy. *Methods Enzymol* 2015;**564**:59–100. <https://doi.org/10.1016/bs.mie.2015.08.006>.
- 10 Tilegenova C, Elberson BW, Cortes DM, Cuello LG. CW-EPR Spectroscopy and Site-Directed Spin Labeling to Study the Structural Dynamics of Ion Channels. *Potassium Channels Methods Protoc.*, vol. 1684. 2018. p. 279–88.
- 11 Columbus L, Hubbell WL. A new spin on protein dynamics. *Trends Biochem Sci* 2002;**27**:288–95. [https://doi.org/10.1016/S0968-0004\(02\)02095-9](https://doi.org/10.1016/S0968-0004(02)02095-9).
- 12 Jeschke G. DEER Distance Measurements on Proteins. *Annu Rev Phys Chem* 2012;**63**:419–46. <https://doi.org/10.1146/annurev-physchem-032511-143716>.
- 13 Jeschke G. Conformational dynamics and distribution of nitroxide spin labels. *Prog Nucl Magn Reson Spectrosc* 2013;**72**:42–60. <https://doi.org/10.1016/j.pnmrs.2013.03.001>.
- 14 Jeschke G. Distance measurements in the nanometer range by pulse EPR. *ChemPhysChem* 2002;**3**:927–32. [https://doi.org/10.1002/1439-7641\(20021115\)3:11<927::AID-CPHC927>3.0.CO;2-Q](https://doi.org/10.1002/1439-7641(20021115)3:11<927::AID-CPHC927>3.0.CO;2-Q).
- 15 Jeschke G, Polyhach Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys Chem Chem Phys* 2007;**9**:1895–910. <https://doi.org/10.1039/b614920k>.
- 16 Reginsson GW, Schiemann O. Pulsed electron-electron double resonance: beyond nanometre distance measurements on biomacromolecules. *Biochem J* 2011;**434**:353–63. <https://doi.org/10.1042/BJ20101871>.
- 17 Roser P, Schmidt MJ, Drescher M, Summerer D. Site-directed spin labeling of

- proteins for distance measurements: In vitro and in cells. *Org Biomol Chem* 2016;**14**:5468–76. <https://doi.org/10.1039/c6ob00473c>.
- 18 Schiemann O, Prisner TF. Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Q Rev Biophys* 2007;**40**:1–53. <https://doi.org/10.1017/S003358350700460X>.
- 19 Sahu ID, Lorigan GA. Biophysical EPR Studies Applied to Membrane Proteins. *J Phys Chem Biophys* 2015;**5**:doi:10.4172/2161-0398.1000188. <https://doi.org/10.4172/2161-0398.1000188>.
- 20 Sahu ID, Lorigan GA. Site-Directed Spin Labeling EPR for Studying Membrane Proteins. *Biomed Res Int* 2018. <https://doi.org/10.1155/2018/3248289>.
- 21 Hemminga MA, Berliner LJ. *ESR Spectroscopy in Membrane Biophysics*. 2007.
- 22 Claxton DP, Kazmier K, Mishra S, McHaourab HS. *Navigating Membrane Protein Structure, Dynamics, and Energy Landscapes Using Spin Labeling and EPR Spectroscopy*. vol. 564. 1st ed. Elsevier Inc.; 2015.
- 23 Bordignon E, Bleicken S. New limits of sensitivity of site-directed spin labeling electron paramagnetic resonance for membrane proteins. *Biochim Biophys Acta - Biomembr* 2018;**1860**:841–53. <https://doi.org/10.1016/j.bbamem.2017.12.009>.
- 24 Cafiso DS. Identifying and quantitating conformational exchange in membrane proteins using site-directed spin labeling. *Acc Chem Res* 2014;**47**:3102–9. <https://doi.org/10.1021/ar500228s>.
- 25 McHaourab HS, Steed PR, Kazmier K. Toward the fourth dimension of membrane protein structure: Insight into dynamics from spin-labeling EPR spectroscopy. *Structure* 2011;**19**:1549–61. <https://doi.org/10.1016/j.str.2011.10.009>.

- 26 Deisenhofer J, Epp O, Huber R, Michel H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3Å resolution. *Nature* 1985;**318**:618–24.
- 27 Caffrey M, Cherezov V. Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* 2009;**4**:706–31. <https://doi.org/10.1007/978-94-007-6232-9-4>.
- 28 Ujwal R, Abramson J. High-throughput Crystallization of Membrane Proteins Using the Lipidic Bicelle Method. *J Vis Exp* 2012:1–7. <https://doi.org/10.3791/3383>.
- 29 Faham S, Ujwal R, Abramson J, Bowie JU. Chapter 5 Practical Aspects of Membrane Proteins Crystallization in Bicelles. *Curr Top Membr* 2009;**63**:109–25. [https://doi.org/10.1016/S1063-5823\(09\)63005-2](https://doi.org/10.1016/S1063-5823(09)63005-2).
- 30 Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 2013;**504**:107–12. <https://doi.org/10.1038/nature12822>.
- 31 Cheng Y. Membrane protein structural biology in the era of single particle cryo-EM. *Curr Opin Struct Biol* 2018;**52**:58–63. <https://doi.org/10.1016/j.sbi.2018.08.008>.
- 32 Johansen NT, Pedersen MC, Porcar L, Martel A, Arleth L. Introducing SEC-SANS for studies of complex self-organized biological systems. *Acta Crystallogr Sect D* 2018;**74**:1178–91. <https://doi.org/10.1107/S2059798318007180>.
- 33 Pérez J, Vachette P. A Successful Combination: Coupling SE-HPLC with SAXS. In: Chaudhuri B, Muñoz IG, Qian S, Urban VS, editors. *Biol. Small Angle Scatt*.

- Tech. Strateg. Tips*. Singapore: Springer Singapore; 2017. p. 183–99.
- 34 Gabel F. Applications of SANS to Study Membrane Protein Systems. In: Chaudhuri B, Muñoz IG, Qian S, Urban VS, editors. *Biol. Small Angle Scatt. Tech. Strateg. Tips*. Singapore: Springer Singapore; 2017. p. 201–14.
- 35 Mandala VS, Williams JK, Hong M. Structure and Dynamics of Membrane Proteins from Solid-State NMR. *Annu Rev Biophys* 2018;**47**:201–22.
<https://doi.org/10.1016/j.physbeh.2017.03.040>.
- 36 Bonucci A, Ouari O, Guigliarelli B, Belle V, Mileo E. In-Cell EPR: Progress towards Structural Studies Inside Cells. *ChemBioChem* 2019;**20**:.
<https://doi.org/10.1002/cbic.201900291>.
- 37 Joseph B, Jaumann EA, Sikora A, Barth K, Prisner TF, Cafiso DS. In situ observation of conformational dynamics and protein ligand–substrate interactions in outer-membrane proteins with DEER/PELDOR spectroscopy. *Nat Protoc* 2019;**14**:2344–69. <https://doi.org/10.1038/s41596-019-0182-2>.
- 38 Kieber MK. *Neisserial Opa protein dynamics and interactions with human CEACAM1*. University of Virginia; 2018; 2018.
- 39 Schosseler PM. *Electron paramagnetic resonance study of the copper(II) complexation with carbonate ligands in aqueous solution and at calcium carbonate surfaces*. 1998; 1998.
- 40 *Basic Concepts of EPR*. ETHzurich. 2020.
- 41 Nyenhuis SB. *Elucidating the Specificity of Membrane Interactions of Synaptotagmin-1 in Synchronous Neurotransmitter Release*. University of Virginia; 2018; 2018.

- 42 Columbus L, Kálai T, Jekő J, Hideg K, Hubbell WL. Molecular motion of spin labeled side chains in α -helices: Analysis by variation of side chain structure. *Biochemistry* 2001;**40**:3828–46. <https://doi.org/10.1021/bi002645h>.
- 43 Evans EGB, Millhauser GL. *Genetic Incorporation of the Unnatural Amino Acid p-Acetyl Phenylalanine into Proteins for Site-Directed Spin Labeling*. vol. 563. 1st ed. Elsevier Inc.; 2015.
- 44 Young TS, Ahmad I, Yin JA, Schultz PG. An Enhanced System for Unnatural Amino Acid Mutagenesis in *E. coli*. *J Mol Biol* 2010;**395**:361–74. <https://doi.org/10.1016/j.jmb.2009.10.030>.
- 45 Columbus L, Hubbell WL. Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* 2004;**43**:7273–87. <https://doi.org/10.1021/bi0497906>.
- 46 Pannier M, Veit S, Godt A, Jeschke G, Spiess HW. Dead-Time Free Measurement of Dipole-Dipole Interactions between Electron Spins. *J Magn Reson* 2000;**142**:331–40. <https://doi.org/10.1023/A:1011323413057>.
- 47 Schmidt T, Wälti MA, Baber JL, Hustedt EJ, Clore GM. Long Distance Measurements up to 160 Å in the GroEL Tetradecamer Using Q-Band DEER EPR Spectroscopy. *Angew Chemie - Int Ed* 2016;**55**:15905–9. <https://doi.org/10.1002/anie.201609617>.
- 48 Edwards TH, Stoll S. Optimal Tikhonov regularization for DEER spectroscopy. *J Magn Reson* 2018;**288**:58–68. <https://doi.org/10.1016/j.jmr.2018.01.021>.
- 49 Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham JE, *et al*. DeerAnalysis 2006- a Comprehensive Software Package for Analyzing Pulsed

- ELDOR Data. *Appl Magn Reson* 2006;**30**:473–98.
- 50 Hays JM, Cafiso DS, Kasson PM. Hybrid Refinement of Heterogeneous Conformational Ensembles Using Spectroscopic Data. *J Phys Chem Lett* 2019;**10**:3410–4. <https://doi.org/10.1021/acs.jpcllett.9b01407>.
- 51 Hays JM, Kieber MK, Li JZ, Han JI, Columbus L, Kasson PM. Refinement of Highly Flexible Protein Structures using Simulation-Guided Spectroscopy. *Angew Chemie - Int Ed* 2018;**57**:17110–4. <https://doi.org/10.1002/anie.201810462>.
- 52 Jeschke G. MMM: A toolbox for integrative structure modeling. *Protein Sci* 2018;**27**:76–85. <https://doi.org/10.1002/pro.3269>.
- 53 Xie J, Schultz PG. An expanding genetic code. *Methods* 2005;**36**:227–38. <https://doi.org/10.1016/j.ymeth.2005.04.010>.
- 54 Fleissner MR, Brustad EM, Kalai T, Altenbach C, Cascio D, Peters FB, *et al.* Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc Natl Acad Sci* 2009;**106**:21637–42. <https://doi.org/10.1073/pnas.0912009106>.
- 55 Mchaourab HS, Lietzow MA, Hideg K, Hubbell WL. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 1996;**35**:7692–704. <https://doi.org/10.1021/bi960482k>.
- 56 Lo RH, Kroncke BM, Solomon TL, Columbus L. Mapping membrane protein backbone dynamics: A comparison of site-directed spin labeling with NMR 15N-relaxation measurements. *Biophys J* 2014;**107**:1697–702. <https://doi.org/10.1016/j.bpj.2014.08.018>.
- 57 Kroncke BM, Columbus L. Identification and removal of nitroxide spin label contaminant: Impact on PRE studies of α -helical membrane proteins in detergent.

Protein Sci 2012;**21**:589–95. <https://doi.org/10.1002/pro.2038>.

CHAPTER 3. CONFORMATIONAL DYNAMICS OF LspA

3.1 Introduction

As described in Chapter 1, with the rapid growth of antibiotic resistance, it is imperative that new drugs are developed which target novel pathways. The lipoprotein processing pathway is a novel pathway for antibiotic drug targeting as the enzymes involved, Lgt, LspA, and Lnt (Figure 1.9), are essential in some organisms including *E. coli*, *S. enterica*, *M. tuberculosis*, and *S. coelicolor*, and have no mammalian homologs.^{1,2} Bacterial lipoproteins are characterized by an N-terminal lipid moiety that serves as a membrane anchor, and serve a wide range of functions including in signal transduction, stress sensing, virulence, cell division, sporulation, nutrient uptake, antibiotic resistance, adhesion, and trigger the activation of host innate immune responses.^{3,4} If lipoproteins are not processed correctly they cannot serve these vital functions and the bacteria will be compromised. Lipoprotein signal peptidase (LspA) is an aspartyl protease that carries out the second step in the lipoprotein processing pathway - cleaving the transmembrane helical signal peptide of lipoproteins after lipidation by Lgt.

3.1.1 Project Motivation

The structures of LspA from *P. aeruginosa* with the antibiotic globomycin bound, and *S. aureus* with globomycin or myxovirescin bound have been determined by crystallization in the lipid cubic phase.^{2,5} As described in Section 1.2.3, LspA consists of four transmembrane helices, a periplasmic helix (PH), and a four-stranded β -sheet, called the β -cradle, which rests on the membrane. In the *S. aureus* structures, that authors termed the PH the extracellular loop (EL) as it is from a gram-positive bacteria which does not

contain a periplasm.⁵ The catalytic residues (D124 and D143 in *P. aeruginosa* LspA, or D118 and D136 in *S. aureus* LspA) and 14 additional highly conserved residues that surround the active site were identified.^{2,5} The extensive conservation indicates that mutations arising to block antibiotic binding in the active site, as occurs in antibiotic resistance, would also likely interfere with the binding and cleavage of substrate.² Thus, LspA is a powerful target to combat the development of antibiotic resistance.

The apo and lipoprotein substrate bound structures of LspA have remained elusive and thus, the conformational dynamics associated with substrate binding and signal peptide cleavage is not understood. Even less is known of the role of the lipid environment in substrate binding and LspA cleavage, as the active site lies within the inner membrane. Based on identified lipid (from the crystallization conditions) and globomycin interactions in the crystal structure, Vogeley *et. al* proposed an orientation of the lipoprotein substrate – LspA complex in which the β -cradle and highly conserved periplasmic helix (PH) “clamp” the substrate in place (Figure 3.1 IV).^{2,5} However, in order for the substrate to enter the active site, we propose that there must be a conformational change in which the β -cradle and PH expands to a “open” state to allow the substrate to enter the active site (Figure 3.1 II and III). We also hypothesize that in the apo state it is possible for the PH and β -cradle to come closer together to sample a “closed” conformation (Figure 3.1 I). This more compact structure would be energetically favorable as it would bury the charged residues of the active site.

Based on this model, the hypothesis that the β -cradle and PH must undergo conformational dynamics in order to allow the lipoprotein substrate to enter the constricted active site (Figure 3.1) is investigated using electron paramagnetic resonance (EPR) studies

(Section 3.2) and molecular dynamics (MD) simulations (Section 3.3). This hybrid approach allows visualization of structures consistent with experimental EPR restraints (Section 3.4).

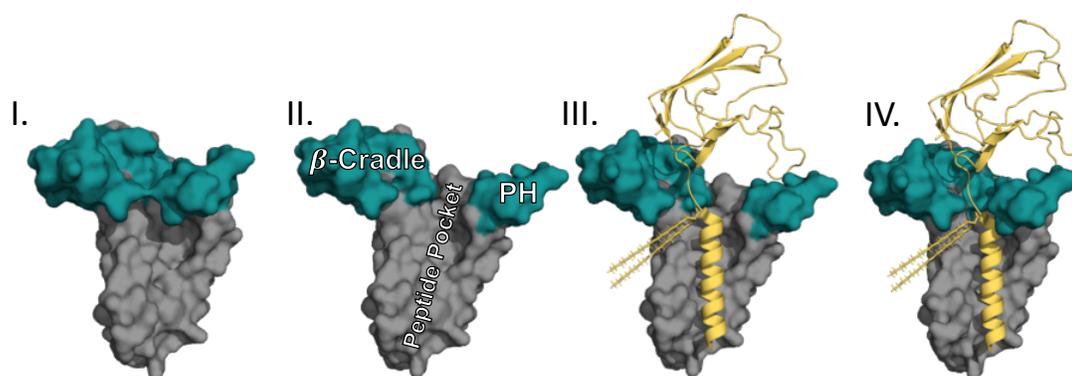


Figure 3.1. Hypothesized LspA conformational dynamics. The *P. aeruginosa* LspA crystal structure with the antibiotic globomycin represents the bound state (IV, “bound”). In order for the substrate to enter the active site we hypothesize that there are conformational dynamics of the PH and β -cradle (II & III, “open”). Without substrate the domains may also come closer together (I, “closed”) so that the PH and β -cradle hide the active site.

3.1.2 Support of Proposed Conformational Dynamics

Evolutionary couplings (ECs) are correlated mutations in protein sequences which can identify close or interacting pairs of residues, or residues that both interact with a mutual third moiety such as another residue or a substrate.⁶⁻⁸ Interestingly for LspA, along with ECs between helical residues, there are several EC pairs between residues of the β -cradle and PH (Figure 3.2). Of note, four of these pairs have very high probabilities (0.658-0.995) (Figure 3.3). These ECs between the β -cradle and PH suggest that there is a conformation where the EC residues are either interacting with each other, or both interacting with substrate or lipid, supporting the hypothesized closed conformation of LspA (Figure 3.1 I).

More recently (after the start of this project), the *S. aureus* LspA structure was solved in the presence of globomycin (PDB ID 6RYO) and myxovirescin (PDB ID 6RYP).⁵ While the *P. aeruginosa* (PDB ID 5DIR) and *S. aureus* globomycin bound structures are very similar, the myxovirescin bound LspA structure is much more compact or “closed” as the EL folds down and blocks the active site (Figure 3.4). Additionally, while globomycin binds toward the PH side of the active site, the myxovirescin inhibitor samples a second binding site more toward the β -cradle side of the active site.⁵

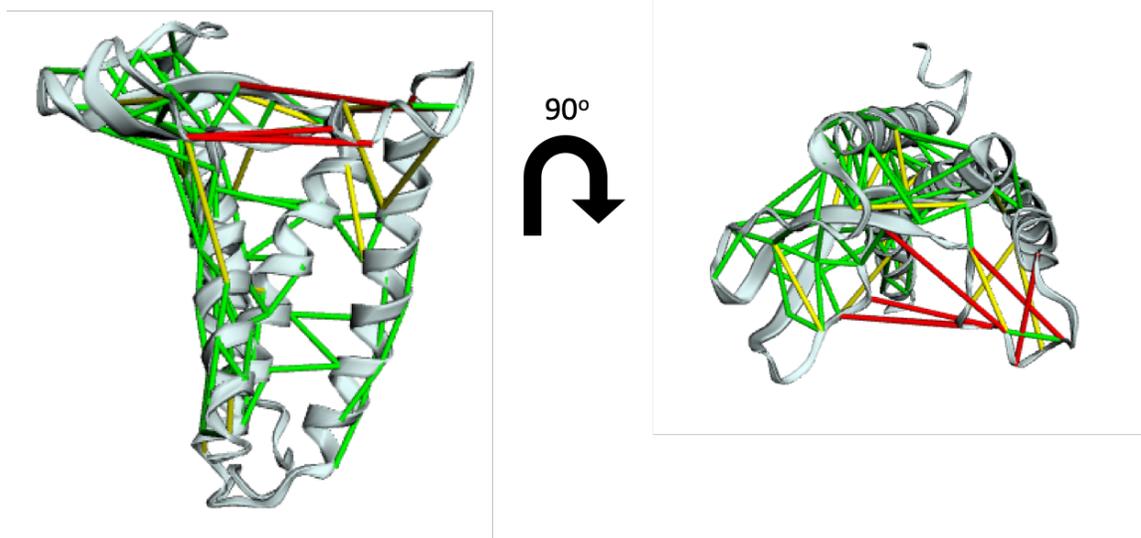


Figure 3.2. LspA evolutionary couplings. Evolutionary couplings were generated and visualized with the GREMLIN coevolution analysis server as lines between EC pair residues.⁹ Pairs are colored based on the distance between coupled residues: green $< 5 \text{ \AA}$ apart, yellow 5- 10 \AA , and red $> 10 \text{ \AA}$.¹⁰

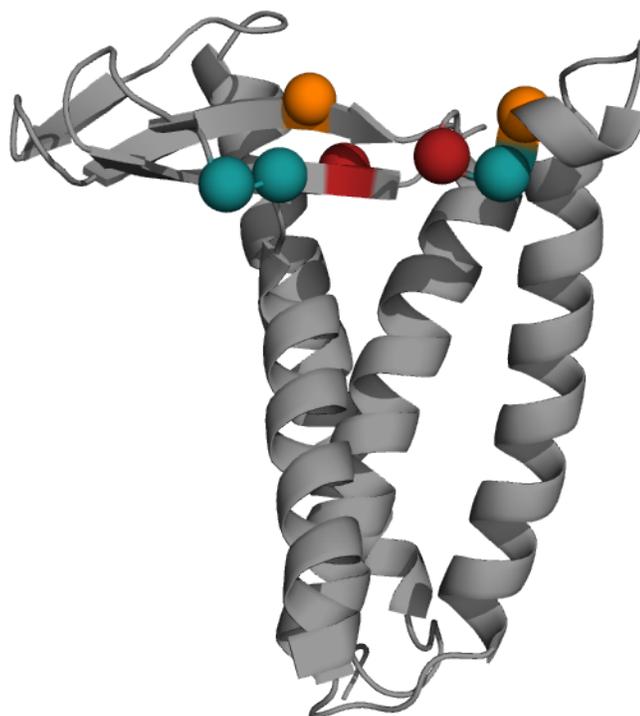


Figure 3.3. Evolutionary couplings between the β -cradle and PH. ECs determined by the GREMLIN server suggest contacts between the β -cradle and PH. Probability between teal residues (59 to 136/137): 0.883/0.995, orange residues (52 to 60): 0.934, and red residues (58 to 123): 0.653.

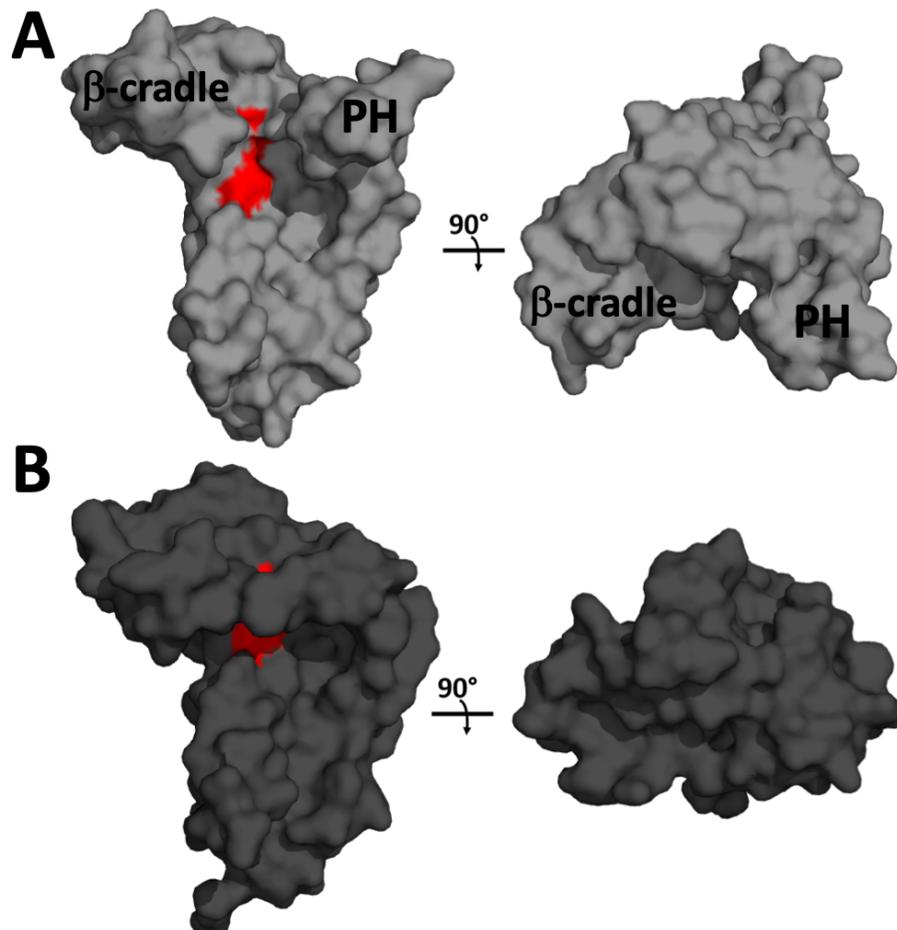


Figure 3.4. LspA conformational dynamics is supported by crystal structures. The *P. aeruginosa* LspA crystal structure with the antibiotic globomycin (A, PDB ID 5DIR), and *S. aureus* LspA with the antibiotic myxovirescin (B, PDB ID 6RYP) provide two structures that support the hypothesis that there are conformational dynamics of the PH and β -cradle. The catalytic dyad is colored red.

3.2 *Electron Paramagnetic Resonance Studies*

In order to investigate the conformational dynamics between the PH and β -cradle, CW and DEER EPR experiments were conducted with spin labels placed on these domains. As described in Chapter 2, EPR is an excellent technique to study the conformational dynamics of membrane proteins as it uses small labels to probe the environment of specific sites, and does not have protein size limitations.

3.2.1 Continuous Wave EPR Studies

3.2.1.1 CW in FC12 Micelles

P. aeruginosa LspA solubilized in FC12 micelles was spin labeled (see methods) at two sites in the PH (A57R1 and A63R1) and three sites in the β -cradle (M36R1, I43R1, and V41R1) (Figure 3.5A). Continuous wave (CW) EPR spectra indicate that each of these sites in apo LspA are in regions with backbone fluctuations in the nanosecond timescale (Figure 3.5B). As expected from the globomycin bound crystal structure, the CW spectra are broadened upon addition of globomycin for sites in the PH, but not in the β -cradle, with the exception of M36 which is on a loop site (Figure 3.5B). The spectral broadening indicates that sites in the PH have a decrease in mobility of the spin label upon globomycin addition, while sites in the β -cradle do not show a change in mobility. Thus, these data suggest that the PH becomes less dynamic when LspA is bound to globomycin. Broadening was most dramatic for site A63R1, and a wider range of globomycin concentrations was tested for this site (Figure 3.6). With increasing concentrations of globomycin, an additional spectral component appears in the A63R1 lineshape (Figure 3.6, arrow), indicating a second more immobile conformation of the spin label. It is interesting to note

that changes in the CW spectra were seen with up to around 100 molar excess of globomycin while the reported IC_{50} of globomycin is reported to be in the nM range.¹¹⁻¹⁴ Assuming that LspA at 100 molar excess is fully bound, fitting of the A63R1 CW spectral changes indicates a K_d of 6.8 mM (Figure 3.7). While this is not a reliable measurement due to the error in the fit, it is 1000 fold greater than the reported K_m which is in the μ M range (Table 4.1).^{5,13} This fit uncertainty is due to the appearance of the second spectral component (Figure 3.6), so the broadening in the lineshape is not a linear combination of bound and apo conformations.

The appearance of the second lineshape component in the CW spectra indicates that there are two binding modes of globomycin: one in which A63R1 remains mobile (but less mobile than the apo state) and another in which A63R1 is in a more rigid conformation. Interestingly, a comparison of the location of globomycin and myxovirescin in the *S. aureus* LspA structures indicates that although only one conformation is observed for each antibiotic in these crystal structures, and both antibiotics interact with the same catalytic residues, their conformations in the active site are different. Myxovirescin occupies more space toward the PH side of the active site and thus the PH has more tertiary contacts with the inhibitor, while globomycin occupies more space toward the β -cradle side of the active site. The residue corresponding to *P. aeruginosa* A63 is S61 in *S. aureus* LspA. S61 is seen to interact with N53 in the myxovirescin structure, but in the globomycin bound structure is on a solvent accessible loop (Figure 3.8). S61 is therefore more restricted in the myxovirescin structure than in the globomycin structure. It is possible that with increasing concentrations, globomycin also samples this second binding site leading to the second more immobile conformation of the spin label.

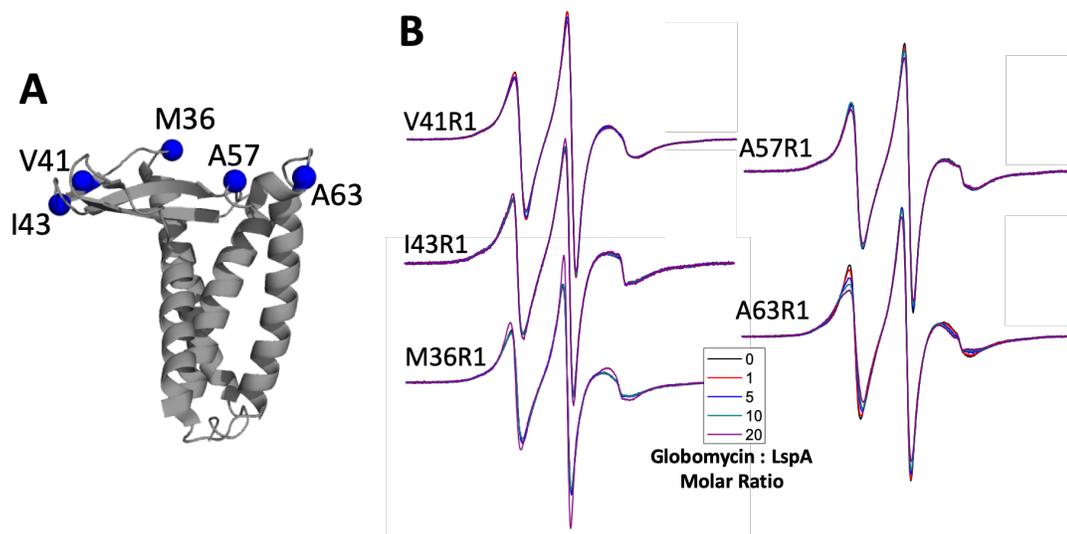


Figure 3.5. CW spectra. A) Spin label sites of *P. aeruginosa* LspA. B) CW spectra of sites shown in A at 0 (black), 1 (red), 5 (blue), 10 (green), and 20 (purple) globomycin to LspA molar ratio.

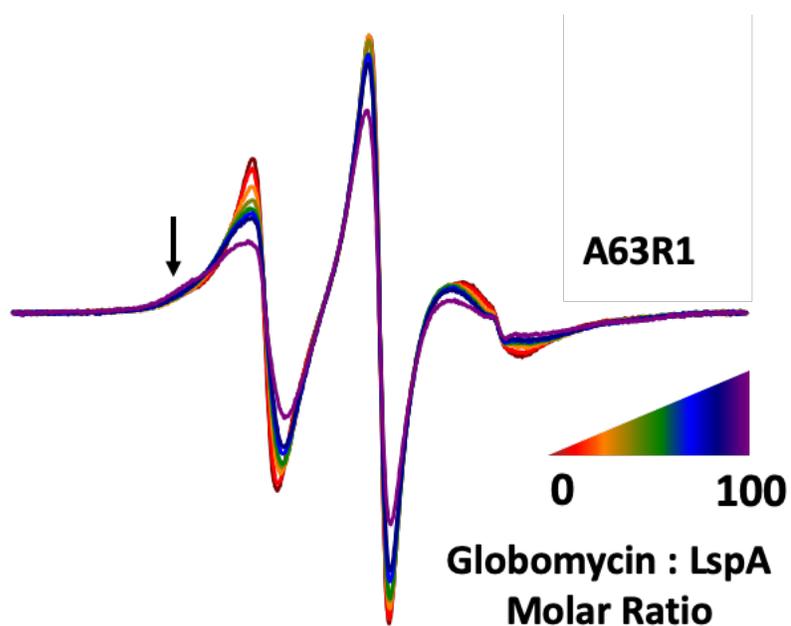


Figure 3.6. A63R1 CW spectra. Increasing concentrations of globomycin (from 0-100 globomycin to LspA molar ratio, red to purple) broaden the CW lineshape and show a second structural element (arrow) indicating a more rigid conformation.

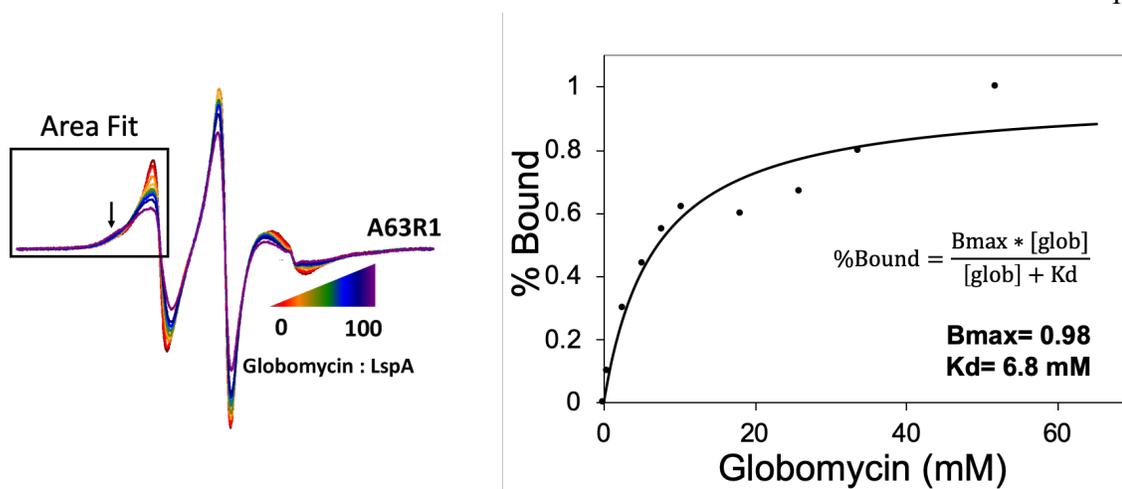


Figure 3.7. Fit of A63R1 CW spectra. Computed spectra were generated by the equation: $x(\text{bound spectrum}) + (1-x)(\text{free spectrum})$ with 0 molar ratio being the “free spectrum” and 100 molar ratio being the “bound spectrum”. Each experimental spectrum was compared to the computed spectra and the percent bound (x) was determined for the best fit between computed and experimental spectra. The percent bound vs. globomycin concentration plot was used then to fit the B_{max} and K_d .

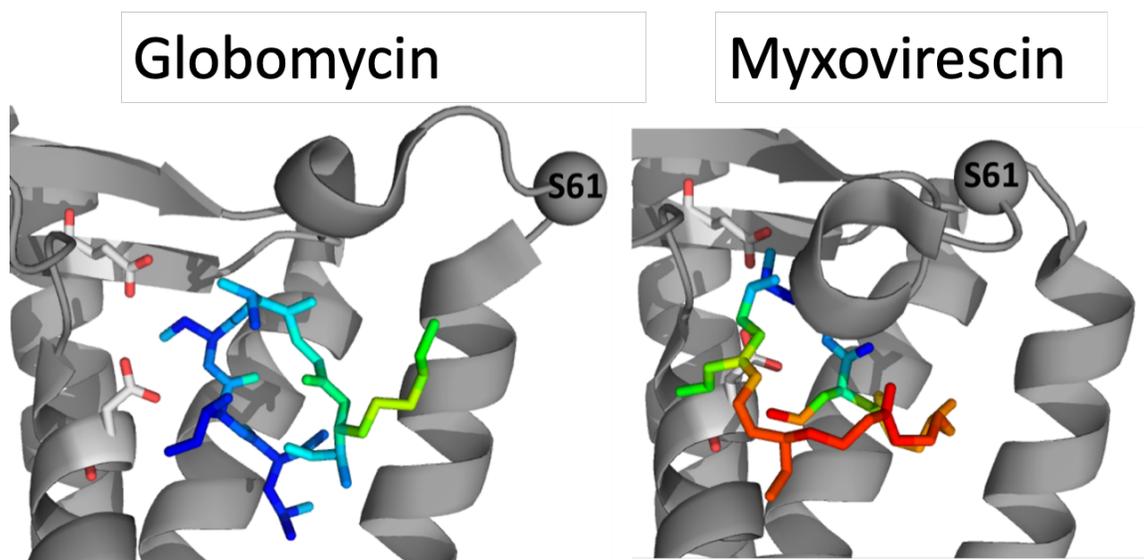


Figure 3.8. Location of spin label site in *S. aureus* LspA. S61, the analogous residue to A63 in *P. aeruginosa* LspA, is on a solvent exposed loop in the globomycin bound structure, but has tertiary contacts in the myxovirescin bound structure. The antibiotics are colored by B-factor from blue (low) to red (high). PDB IDs 6RYO (globomycin bound, left) and 6RYP (myxovirescin bound, right).

3.2.1.2 CW with the Addition of Lipids

In order to investigate the role of the membrane environment, CW spectra were acquired at various q values (lipid to detergent ratio) for A63R1 LspA in DMPC/FC12 and POPG/FC12 bicelles. The addition of both DMPC and POPG (Figure 3.9) caused broadening of the CW spectra, although to a lesser extent than the addition of globomycin (Figure 3.6). The DMPC/FC12 spectra did not broaden until $q=0.75$, while POPG/FC12 spectra showed broadening starting at $q=0.3$.

A63R1 LspA was tested in DMPC/FC12 at higher q -values to determine if further broadening would occur (Figure 3.10). Again, spectra from $q=0$ to $q=0.5$ mostly overlaid. Above $q=0.5$ the lineshape broadened and the spectra from $q=0.75$ to $q=2$ also overlaid. This suggests that the lipid and detergent molecules of the bicelles are mixed until $q=0.5$, and then then become segregated with lipids in the core and detergents in the rim as in an ideal bicelle (see Chapter 5 for further characterization of bicelles). This agrees with SAXS data completed by a lab member, Nicole Swope. As described in Section 5.1.2, SAXS is able to measure the headgroup to headgroup distance (L) between amphiphiles in a bicelle. Nicole's data show that for DMPC/FC12 bicelles, L increases from $q=0$ to $q=0.4$ and then levels off around 48 Å for $q \geq 0.5$ (Figure 3.11). Again, this suggests mixing up to $q=0.4$ and then segregation at higher q values. The broadening observed upon lipid addition indicates that the addition of lipid decreases the mobility of the spin label.

DMPC/FC12 bicelles were chosen for all further experiments as the segregation has been characterized (Figure 3.11), and they are well behaved in solution. Twenty molar excess globomycin was added to A63R1 LspA DMPC/FC12 bicelle samples at $q = 0.3$, 0.5, 0.75, and 1. Surprisingly, there were no lineshape changes upon the addition of

globomycin (Figure 3.12) as were seen in FC12 micelles ($q=0$). Since there are no lineshape changes, it is impossible to determine from these data if globomycin is in fact binding to LspA in bicelles. The altered membrane environment may impact the partitioning of globomycin into the bicelle or LspA active site, or alter the conformational dynamics of LspA such that globomycin cannot enter the active site.

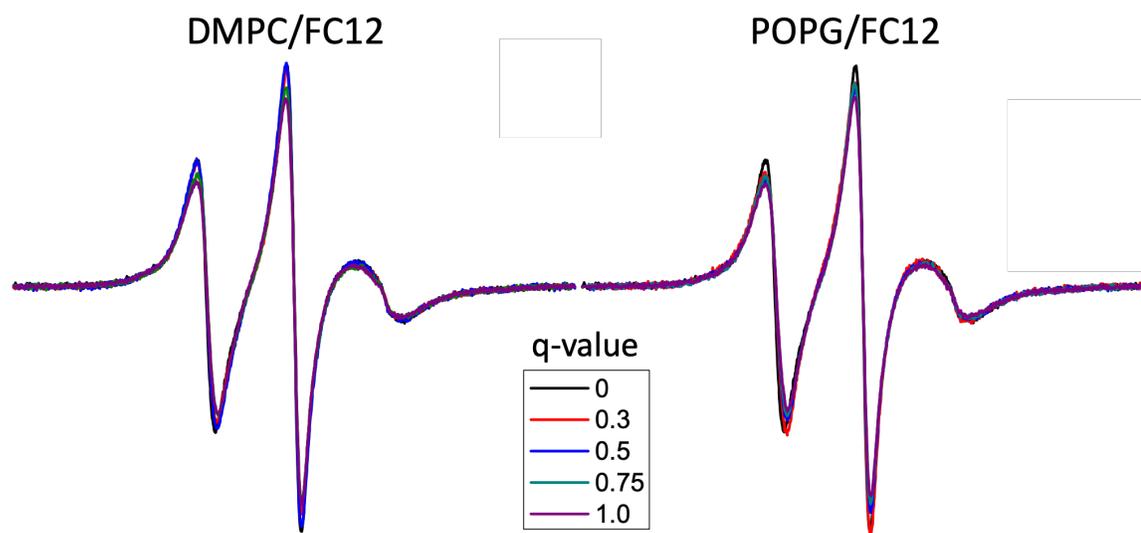


Figure 3.9. CW spectra with the addition of lipid. CW of A63R1 LspA in DMPC/FC12 (left) and POPG/FC12 (right) bicelles at q values of 0 (black), 0.3 (red), 0.5 (blue), 0.75 (green), and 1 (purple) show broadening upon the addition of lipid. 6% amphiphile concentration was used.

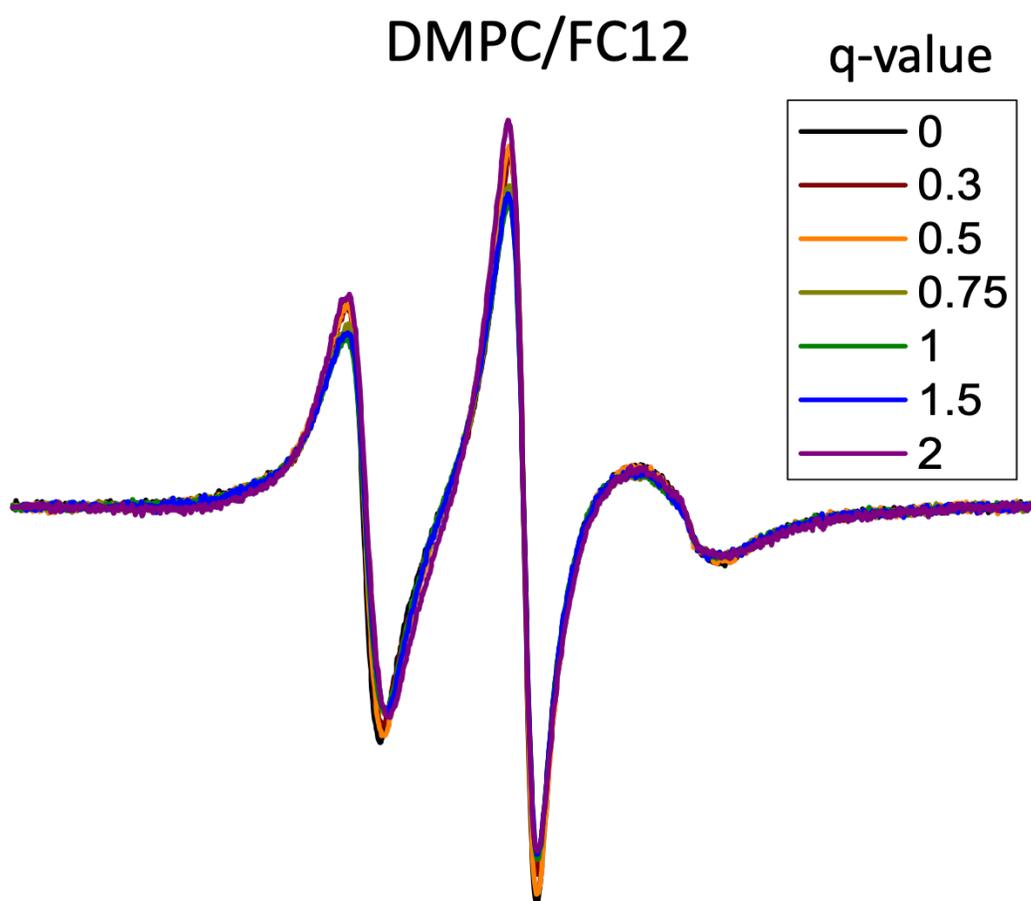


Figure 3.10. DMPC/FC12 CW titration. CW for A63R1 LspA in DMPC/FC12 bicelles were acquired from $q=0$ to $q=2$ at 6% amphiphile concentration.

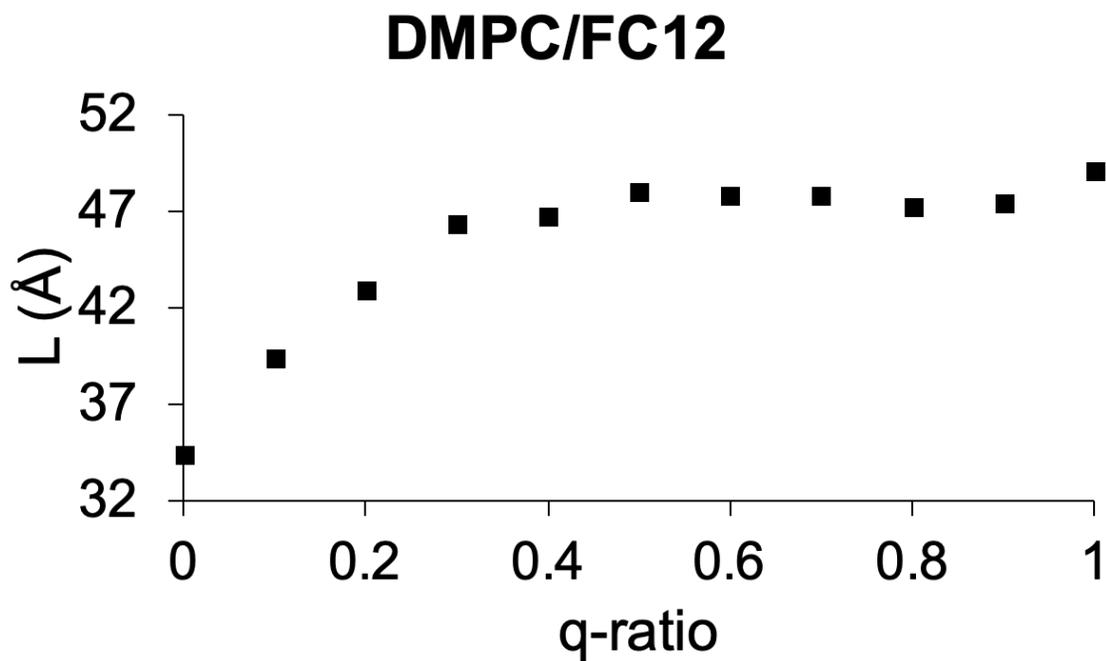


Figure 3.11. DMPC/FC12 SAXS data. Headgroup to headgroup distance L was determined from SAXS data of DMPC/FC12 bicelles from $q=0$ to 1. This work was performed by Nicole Swope (unpublished).

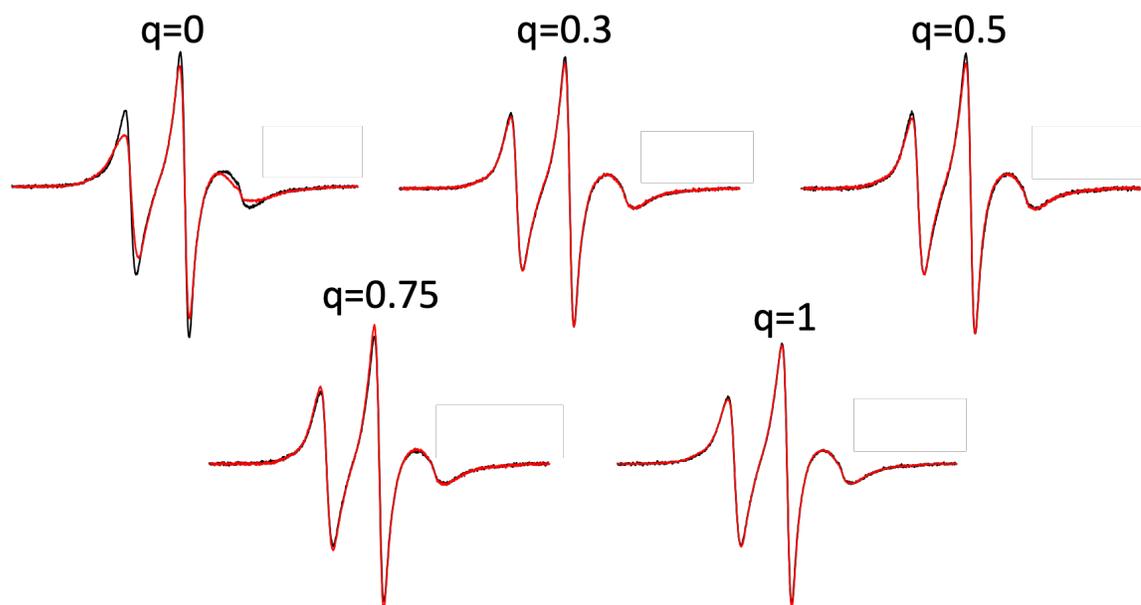


Figure 3.12. CW of DMPC/FC12 LspA with globomycin. CW spectra without (black) and with (red) globomycin for A36R1 LspA in DMPC/FC12 bicelles from $q=0$ to $q=1$.

3.2.2 Double Electron Electron Resonance Studies

3.2.2.1 I43R1/A63R1 in FC12 Micelles

To better understand the conformation changes involving the β -cradle and PH between the apo and globomycin bound states, double electron-electron resonance (DEER) EPR was conducted with *P. aeruginosa* LspA doubly spin labeled at I43R1 (on the β -cradle) and A63R1 (on the PH) (Figure 3.5A). As discussed in Chapter 2, the DEER data for this sample can be realistically interpreted in the 1.8-6.0 nm range.¹⁵ In the apo state, the distance between I43R1 and A63R1 is broad, from 1.8-4.5 nm, and is centered around 2.9 nm (Figure 3.13). Upon addition of globomycin, three populations are observed at 2.7, 3.6, and 4.5 nm with the longer distance conformation being the most populated (Figure 3.13). The two shorter globomycin distance peaks fit within the apo distribution, and there is a large degree of uncertainty in this area (Figure 3.13). These data suggest that the β -cradle and PH are closer together in the apo state than in the globomycin bound state as in Figure 3.1 I. As stated previously, the apo “closed” state may be important to occlude the charged catalytic residues from the hydrophobic membrane.

Multiscale Modeling of Macromolecules (MMM) is a program which allows for the addition of spin labels at sites of interest on protein structures and the subsequent simulation of the DEER distribution between the two spin labeled sites.¹⁶ Experimental DEER data can then be uploaded, and the simulated DEER distribution can be fit to the experimental data by selecting spin label rotamers which fit within the experimental distribution. In this way the apo and globomycin DEER distributions were analyzed against the simulated DEER distributions for each of the three LspA structures (Figure 3.14), with I41 and S61 being the homologous residues to I43 and A63 in *S. aureus* LspA. The

globomycin bound DEER distribution matches well with the MMM simulated spectra for both globomycin structures, while the apo DEER distribution matches best with the *S. aureus* myxovirescin bound structure. The fit MMM distributions only cover a portion of the apo DEER distribution for the globomycin bound structures, while it covers the entire distribution in the myxovirescin bound structure (Figure 3.14 top). The coverage of the MMM fit is exemplified in the percentage of total rotamers that contribute to the fit MMM distribution (Table 3.1). While only 12.3 and 20.2% of rotamers are contained within the MMM distribution fit to the apo DEER data for *P. aeruginosa* and *S. aureus* globomycin bound structures, respectively, 99.6% are contained within the fit MMM distribution for *S. aureus* LspA bound to myxovirescin (Table 3.1). For the globomycin bound DEER distributions, 99-100% of rotamers are contained within the fit MMM distributions for all three structures, because each is entirely contained within the experimental DEER distribution. However, the fit MMM distributions only cover a portion of the globomycin experimental distribution. The globomycin bound fit MMM distributions both fit well to the more prominent peak at 4.5 nm, while the myxovirescin bound structure fit MMM distribution only covers the shorter peaks in the distribution (Figure 3.14 bottom). Thus, the MMM analysis demonstrates that the globomycin bound structures (PDB IDs 5DIR and 6RYO) represent the globomycin bound experimental DEER peak at 4.5 nm, while the myxovirescin bound structure (PDB ID 6RYP) is a representation of the apo DEER results.

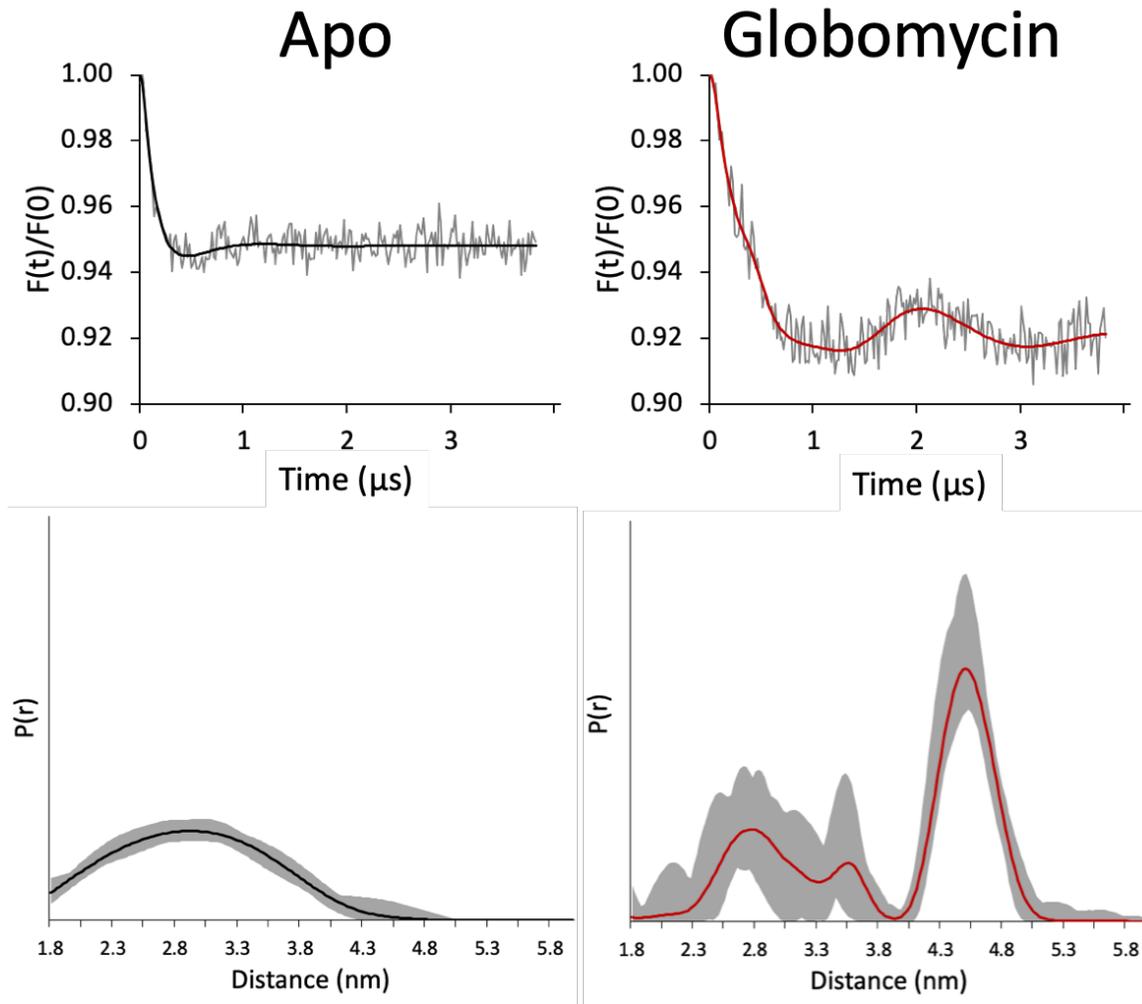


Figure 3.13. I43R1/A63R1 DEER analysis. Top) Best fit of the DEER echo decay for apo (black, left) and globomycin bound (red, right) LspA. Bottom) DEER distributions for apo (black, left) and globomycin bound (red, right) LspA. Error is shown in gray. The addition of globomycin at a 20:1 molar ratio shows the appearance of a longer distribution around 4.5 nm.

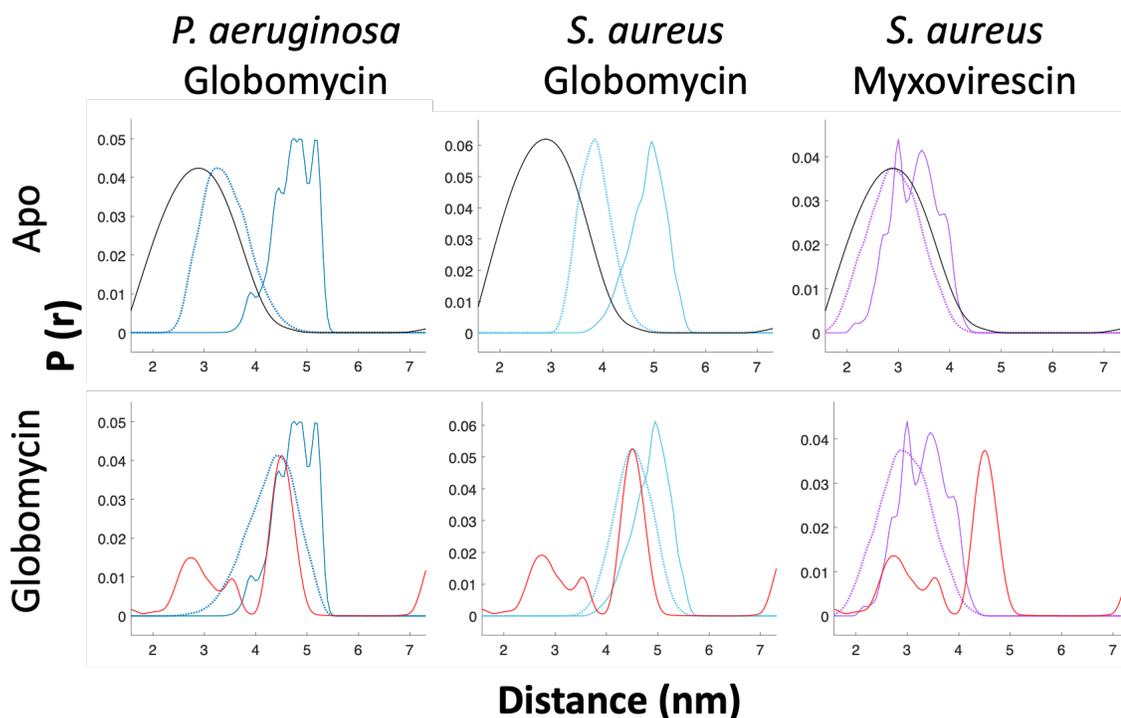


Figure 3.14. I43R1/A63R1 LspA MMM simulated spectra. MMM simulated spectra for *P. aeruginosa* globomycin bound structure (teal, left, PDB ID 5DIR), *S. aureus* globomycin bound structure (blue, middle, PDB ID 6RYO), and *S. aureus* myxovirescin bound structure (purple, right, PDB ID 6RYP), were compared and fit to apo (top, black) and globomycin bound (bottom, red) I43R1/A63R1 DEER data in FC12 micelles. Simulated spectra with no rotamer bias are shown as solid lines while fit simulated spectra selecting for rotamers which fit the experimental distribution are shown as dashed lines. I43R1/A63R1 corresponds to I41R1/S61R1 in *S. aureus* LspA.

Table 3.1 Percent of rotamers in fit
MMM distributions (Fig. 3.14)

				DEER Dataset	
Crystal Structure	PDB ID	Species	Antibiotic	Apo	Glob
	5dir	<i>P. aeruginosa</i>	Globomycin	12.3	99
	6ryo	<i>P. aeruginosa</i>	Globomycin	20.2	100
	6ryp	<i>S. aureus</i>	Myxovirescin	99.6	100

3.2.2.2 M36R1/A63R1 in FC12 Micelles

Another spin label pair, M36R1 (on the β -cradle) and A63R1 (on the PH) was additionally tested via DEER spectroscopy in FC12 micelles (Figure 3.5A). The DEER distribution for this spin label pair was not as informative as the I43R1/A63R1 pair, as it yielded broader distributions for both apo and globomycin bound states from 1.8-4.8 nm (Figure 3.15). The apo distribution appears Gaussian and is centered at 2.9 nm, while the globomycin bound distribution has a major peak centered at 2.7 nm and shoulder to the right. The uncertainty of the fit in this region make the validity of these shoulder peaks questionable. The broadness of the spectra may be due to M36R1 residing on a loop allowing it to sample a greater volume of conformational space (Figure 3.5A). However, the CW spectrum for M36R1 shows a similar lineshape to I43R1 indicating that the two sites have a similar amount of flexibility (Figure 3.5B).

MMM analysis was similarly uninformative as the apo and globomycin bound DEER distributions are similar and they both fit the simulated MMM distributions for the globomycin bound structures (M36 corresponds to I34 of *S. aureus* LspA). However, the simulated distribution for myxovirescin-bound LspA showed shorter distances than observed in either DEER distribution. (Figure 3.16). The myxovirescin structure MMM fit distribution was closer for the globomycin DEER distribution than for the apo DEER distribution because this peak has a slightly shorter average distance (Figure 3.16). The simulated distribution contains distances that are below the limit of DEER observation (<1.8 nm), and so it cannot be determined from this DEER analysis if those shorter distances are sampled between M36 and A63 in the LspA structure. However, broadening due to interactions of the magnetic dipoles of the spin labels was not observed in the

M36R1/A63R1 CW spectra, which would indicate a pairwise distance of $\sim 8-20 \text{ \AA}$.¹⁷

Overall, these data suggest that the distances sampled between M36 and A63 do not change upon the addition of globomycin, and agree with the globomycin bound crystal structures.

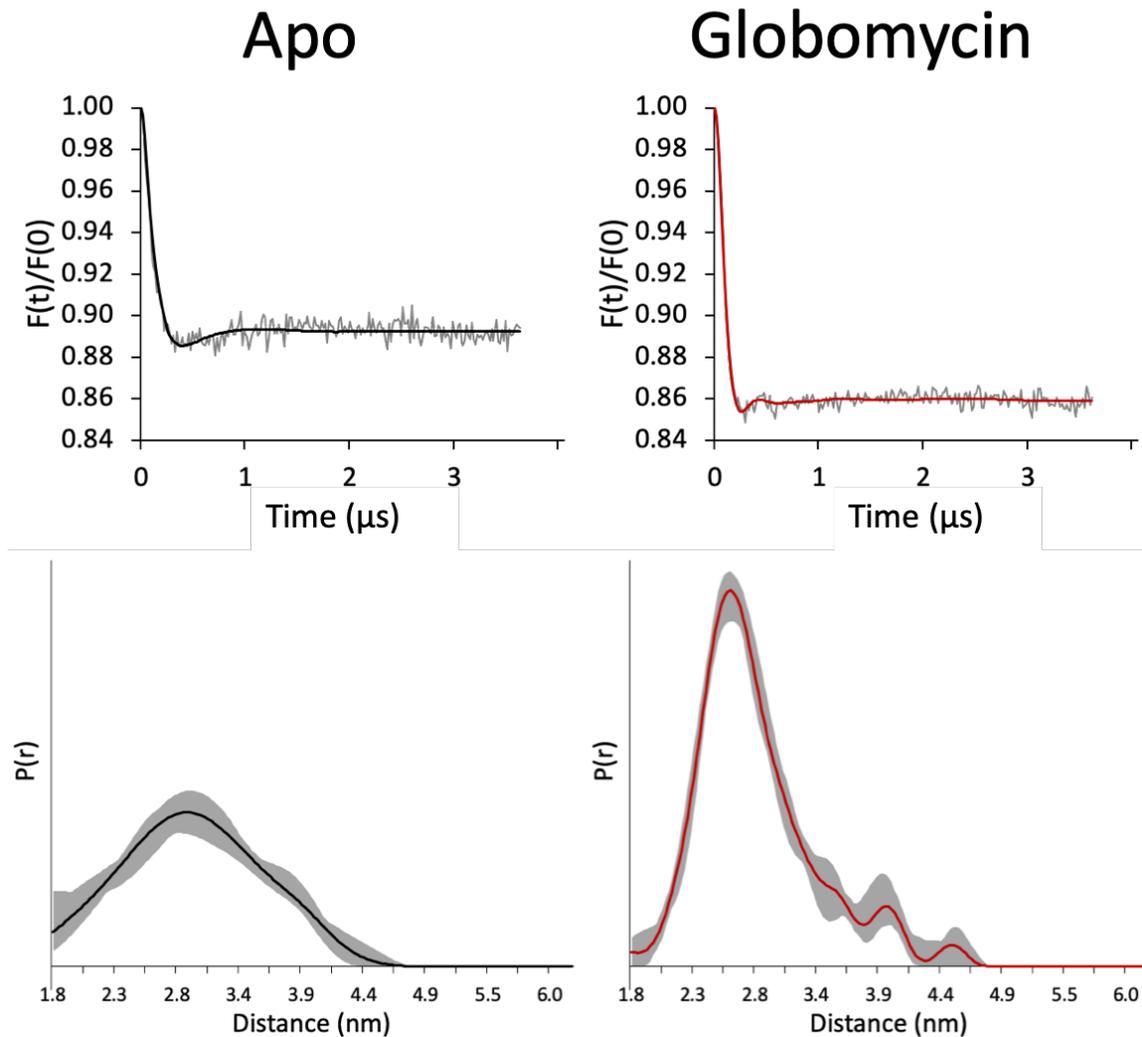


Figure 3.15. M36R1/A63R1 DEER analysis. Top) Best fit of the DEER echo decay for apo (black, left) and globomycin bound (red, right) LspA. Bottom) DEER distributions for apo (black, left) and globomycin bound (red, right) LspA in FC12 micelles. Error is shown in gray. Globomycin was added at a 20:1 molar ratio.

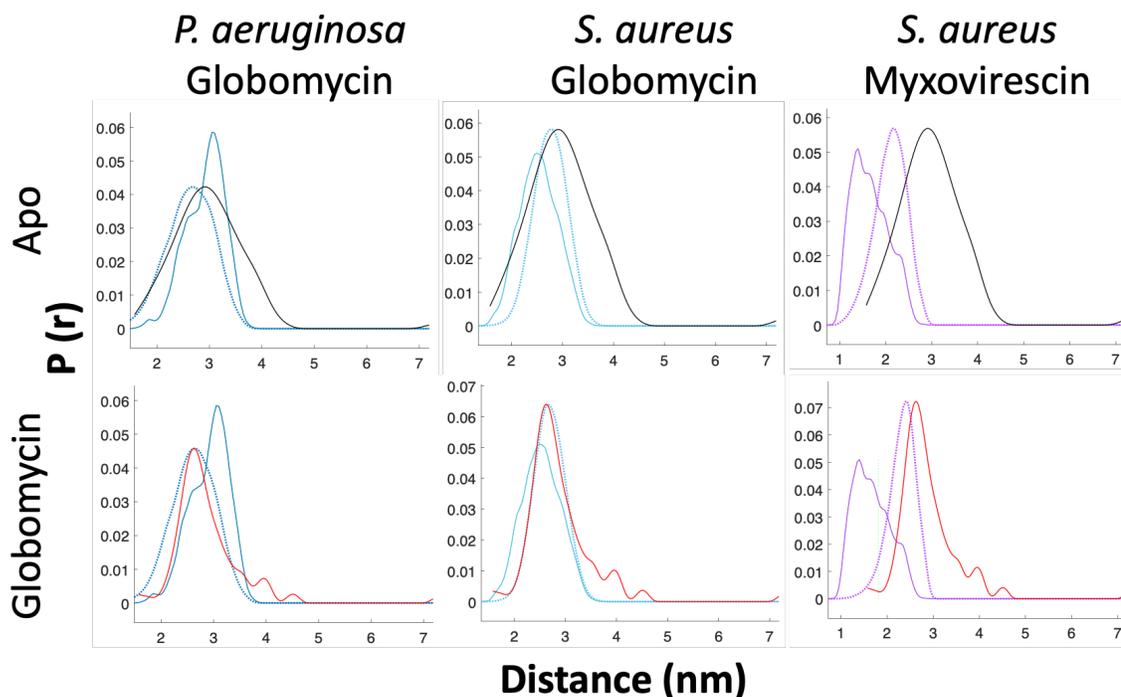


Figure 3.16. M36R1/A63R1 LspA MMM simulated spectra. MMM simulated spectra for *P. aeruginosa* globomycin bound structure (teal, left, PDB ID 5DIR), *S. aureus* globomycin bound structure (blue, middle, PDB ID 6RYO), and *S. aureus* myxovirescin bound structure (purple, right, PDB ID 6RYP), were compared and fit to apo (top, black) and globomycin bound (bottom, red) M36R1/A63R1 DEER data in FC12 micelles. Simulated spectra with no rotamer bias are shown as solid lines while fit simulated spectra selecting for rotamers which fit the experimental distribution are shown as dashed lines. M36R1/A63R1 corresponds to I34R1/S61R1 in *S. aureus* LspA.

3.2.2.3 I43R1/A63R1 in DMPC/FC12 Bicelles

To probe the effects of the membrane mimetic, the I43R1/A63R1 spin label pair was again tested with DEER spectroscopy with LspA solubilized in DMPC/FC12 bicelles with $q=1$ (50% of each amphiphile component) at an amphiphile concentration of 18%. A q value of 1 was chosen as bicelles with this lipid to detergent ratio are segregated, as demonstrated by the CW spectra of A63R1 in various DMPC/FC12 q -value bicelles (Figure 3.10) and SAXS analysis (Figure 3.11).

Strikingly, the echo decays of apo and globomycin bound LspA in bicelles are very similar, resulting in very similar distance distributions (Figure 3.17). Both show a very broad peak from 1.8 to 4.8 nm and a second, less populated peak at 5.4 nm. While the broad peaks seem to have slightly different populations (with the apo distribution having a peak at 2.7 nm and a shoulder peak at 3.7 nm and the globomycin distribution having more of a single Gaussian distribution), the uncertainties of these peaks are large (Figure 3.17). The error seen in the distance distributions, along with the similarity of the echo decays, do not allow for the interpretation of distinct distributions in the 1.8 to 4.7 nm range. Thus, here, globomycin does not change the distance distribution. Comparing the I43R1/A63R1 micelle and bicelle DEER distributions, the shorter broad distributions are aligned while the longer distance peak has increased in distance from 4.5 nm in the micelle distribution to 5.4 nm in the bicelle distribution (Figure 3.18). This increase in pairwise distance may be due to the bicelles being larger than the micelles, allowing the protein to sample a wider conformational state.

Simulated DEER distributions of I43R1/A63R1 by MMM for the globomycin bound structures do not fit well to either the apo or the globomycin DEER data in

DMPC/FC12 bicelles (Figure 3.19). The simulated distribution lies between the two experimental DEER peaks, and the fit distributions shifts left to fit a portion of the broad shorter distance distribution. The shorter, broad distance peak of the DEER distributions align better with the myxovirescin structure (Figure 3.19).

As described in Section 3.2.1.2, the CW spectra did not show any lineshape changes upon the addition of globomycin when LspA is solubilized in DMPC/FC12 bicelles (Figure 3.12). The lack of change between apo and globomycin bound LspA in the CW spectra as well as the DEER distributions suggests that globomycin is not binding to LspA in the bicelle membrane mimetic. The alternative could be that globomycin is binding but does not change the conformational dynamics of the β -cradle and PH. In this case, more research would need to be conducted to determine why LspA would have changes in conformational dynamics in a micelle but not a bicelle upon the addition of globomycin.

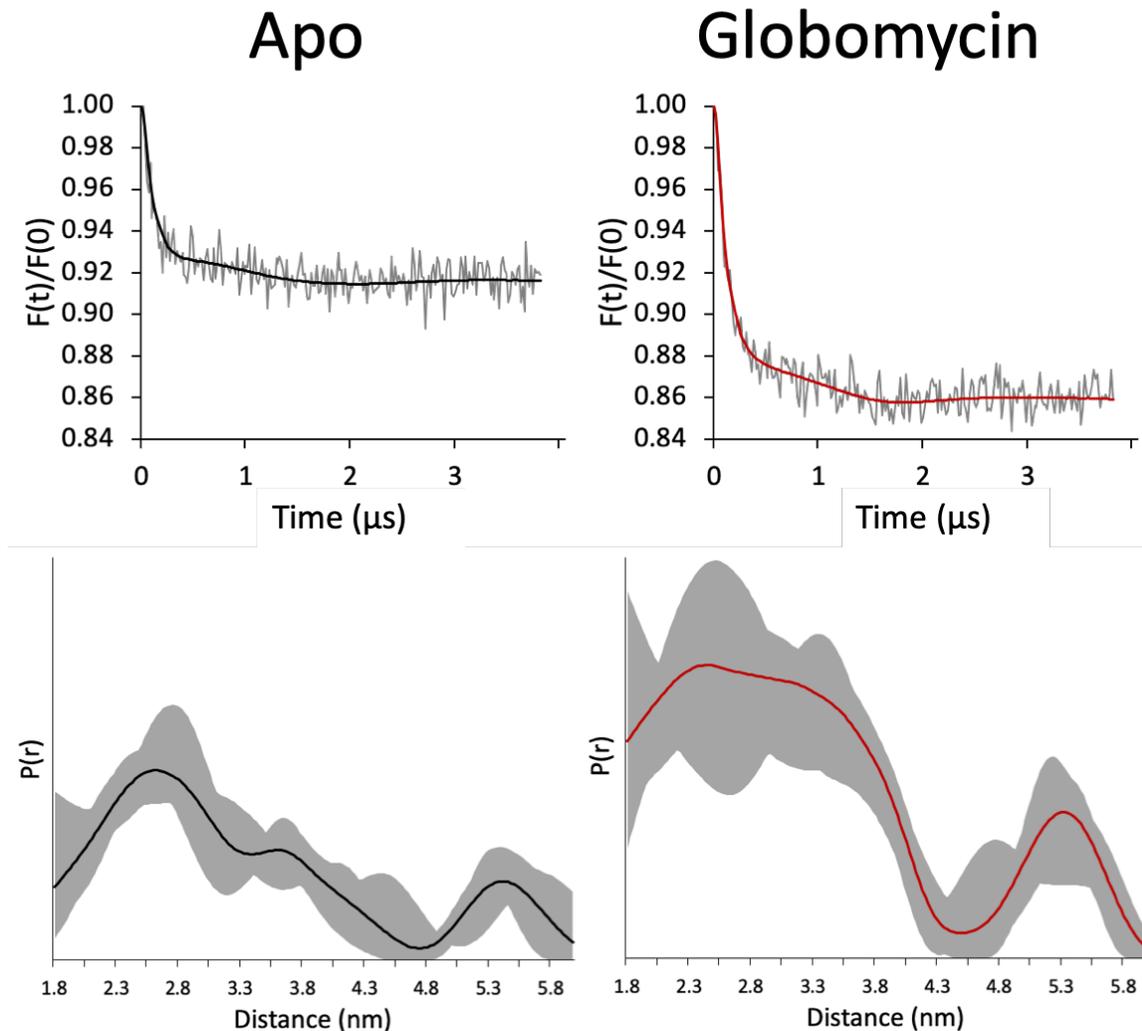


Figure 3.17. I43R1/A63R1 DEER analysis in DMPC/FC12 bicelles. Top) Best fit of the DEER echo decay for apo (black, left) and globomycin bound (red, right) LspA. Bottom) DEER distributions for apo (black, left) and globomycin bound (red, right) LspA. Error is shown in gray. Globomycin was added at a 20:1 molar ratio and DMPC/FC12 was at a q-ratio of 1.

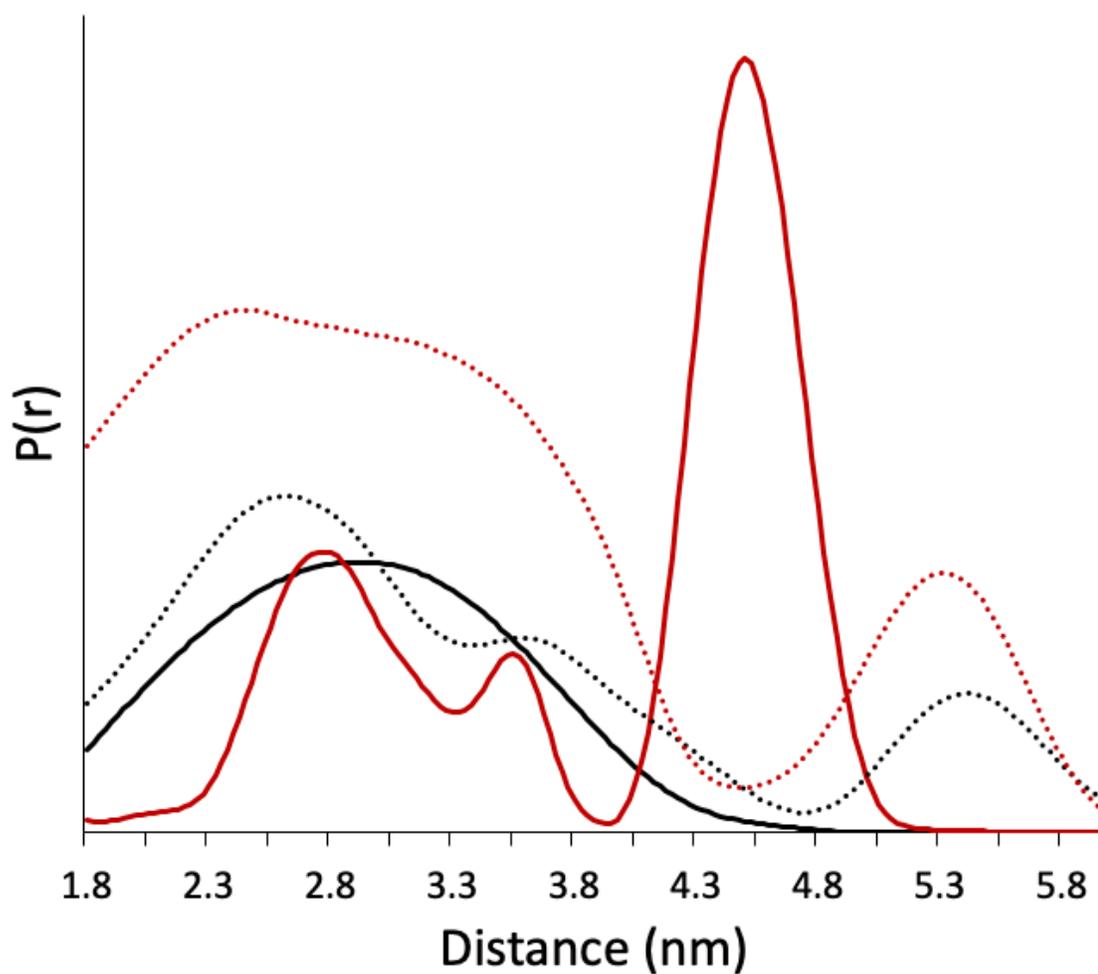


Figure 3.18. I43R1/A63R1 DEER in micelles compared to bicelles. Apo (black) and globomycin bound (red) I43R1/A63R1 LspA DEER distributions are shown for protein in FC12 micelles (solid lines) and DMPC/FC12 bicelles (dotted lines) at $q=1$. Error is not shown for clarity.

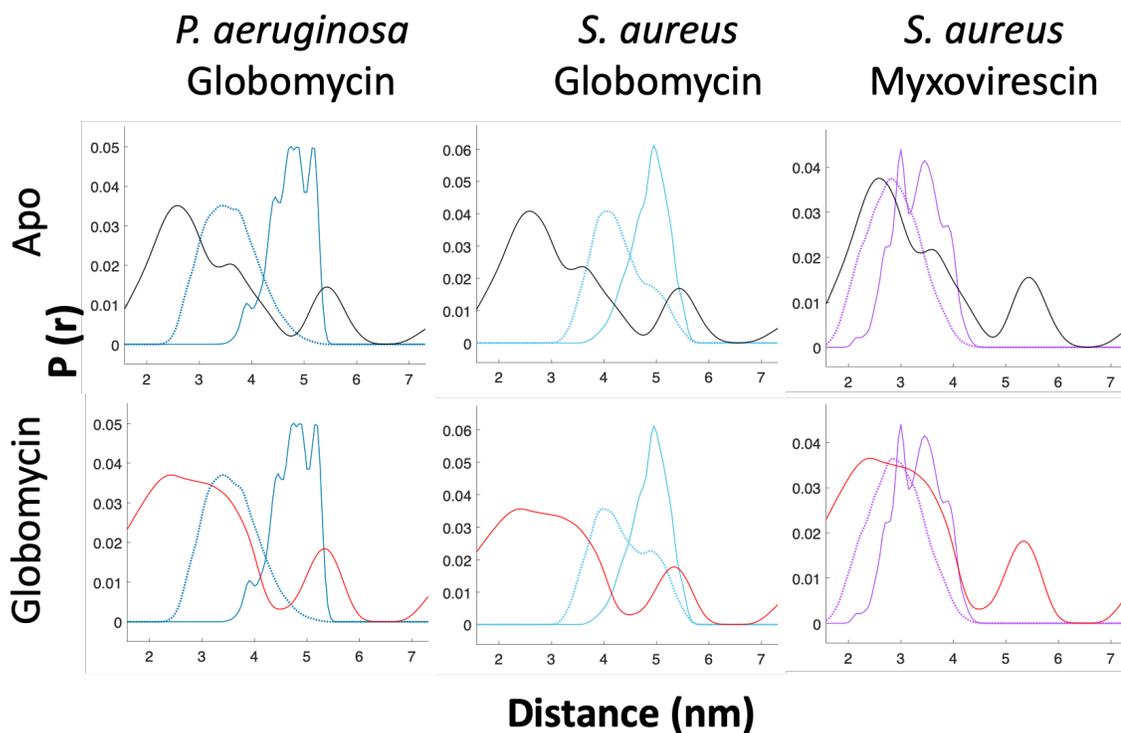


Figure 3.19. I43R1/A63R1 LspA MMM simulated spectra compared to DEER in DMPC/FC12 bicelles. MMM simulated spectra for *P. aeruginosa* globomycin bound structure (teal, left, PDB ID 5DIR), *S. aureus* globomycin bound structure (blue, middle, PDB ID 6RYO), and *S. aureus* myxovirescin bound structure (purple, right, PDB ID 6RYP), were compared and fit to apo (top, black) and globomycin bound (bottom, red) I43R1/A63R1 DEER data in DMPC/FC12 bicelles at $q=1$. Simulated spectra with no rotamer bias are shown as solid lines while fit simulated spectra selecting for rotamers which fit the experimental distribution are shown as dashed lines. I43R1/A63R1 corresponds to I41R1/S61R1 in *S. aureus* LspA.

3.3 Molecular Dynamics Simulations

In order to visualize the conformational states of LspA represented in the DEER data, molecular dynamics (MD) simulations were performed in collaboration with Dr. Owen Vickery and Dr. Phill Standfeld of the University of Warwick. The *P. aeruginosa* LspA crystal structure (PDB ID 5DIR) was first equilibrated in a palmitoyloleoylphosphatidylglycerol (POPG) palmitoyloleoylphosphatidylethanolamine (POPE) (1:3 mole ratio) bilayer, and subsequently three 500 ns simulations were performed for both apo and globomycin bound LspA (see methods Section 3.6.4 for details).

To compare MD and DEER data, the distances between residues spin labeled in EPR experiments (I43-A63 and M36-A63) were measured for each simulation (Figure 3.20). Then, histograms of these distances were made to produce a distance distribution similar to the DEER data (Figure 3.21). However, these distance measurements are between the α -carbons of the two residues while the DEER distance measurements are between the unpaired electron of the spin label. Thus, care needs to be taken when comparing these results. Combining the histograms from all three simulations for each condition (apo and globomycin) and each distance pair (I43-A63 and M36-A63), allowed for a more global interpretation of each data set (Figure 3.22).

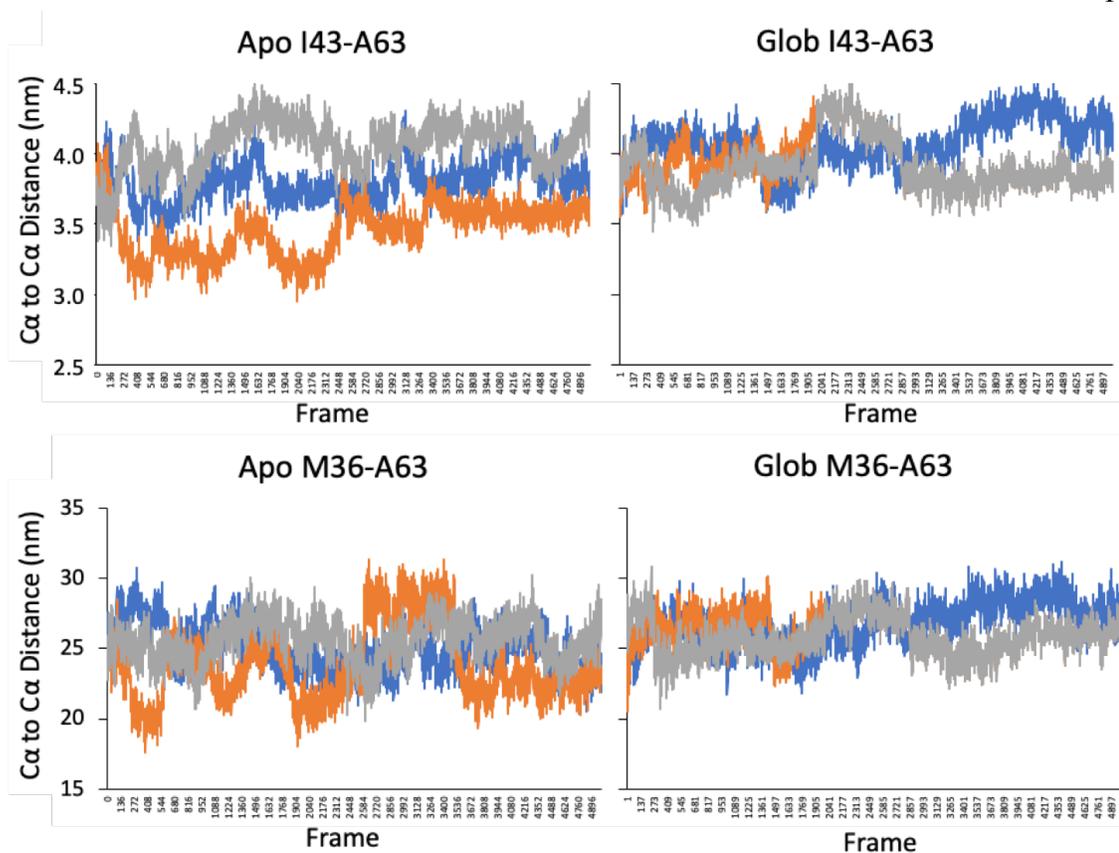


Figure 3.20. MD distance measurements. The distance between the α -carbons of I43-A63 (top) and M36-A63 (bottom) were measured for apo (left) and globomycin bound (right) simulations. The distances derived from the three simulations for each condition are represented by gray, blue, or orange traces.

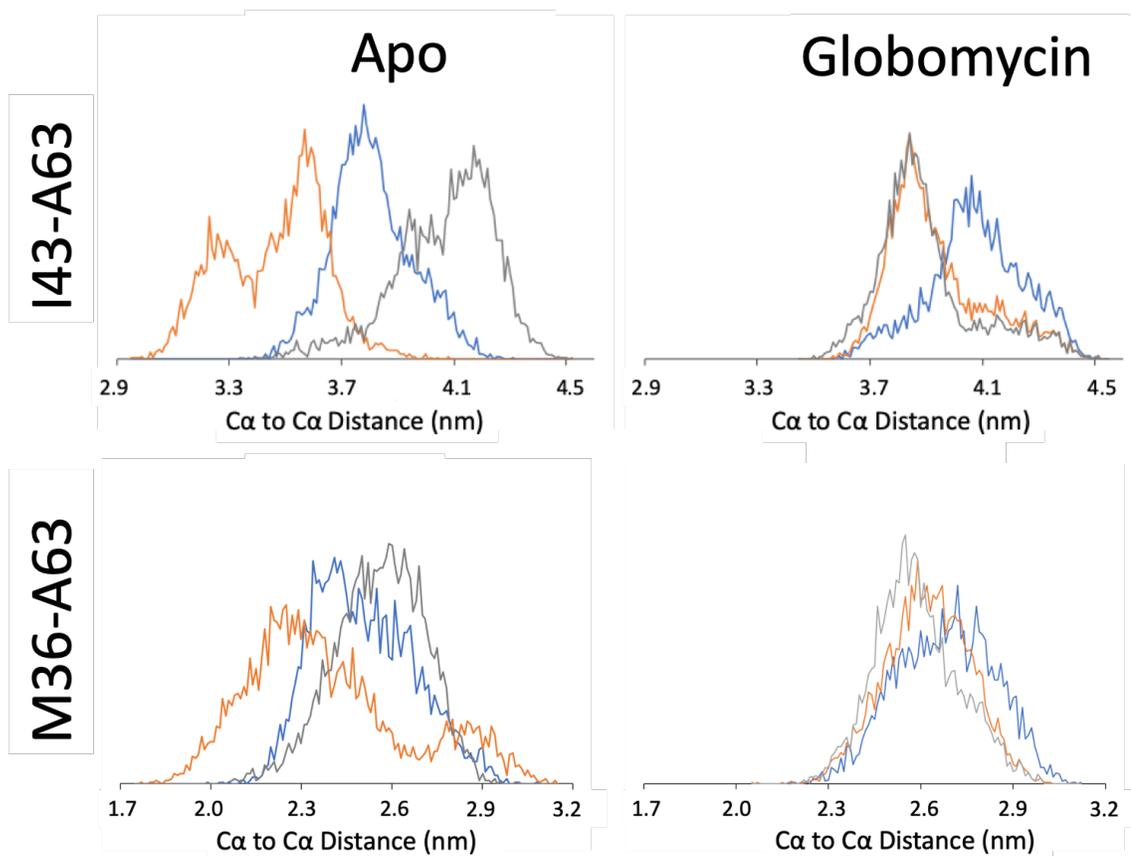


Figure 3.21. MD distance histograms. The measured simulation distances seen in Figure 3.20 were compiled into a histogram to better visualize the populations of the distance distribution for each simulation trajectory.

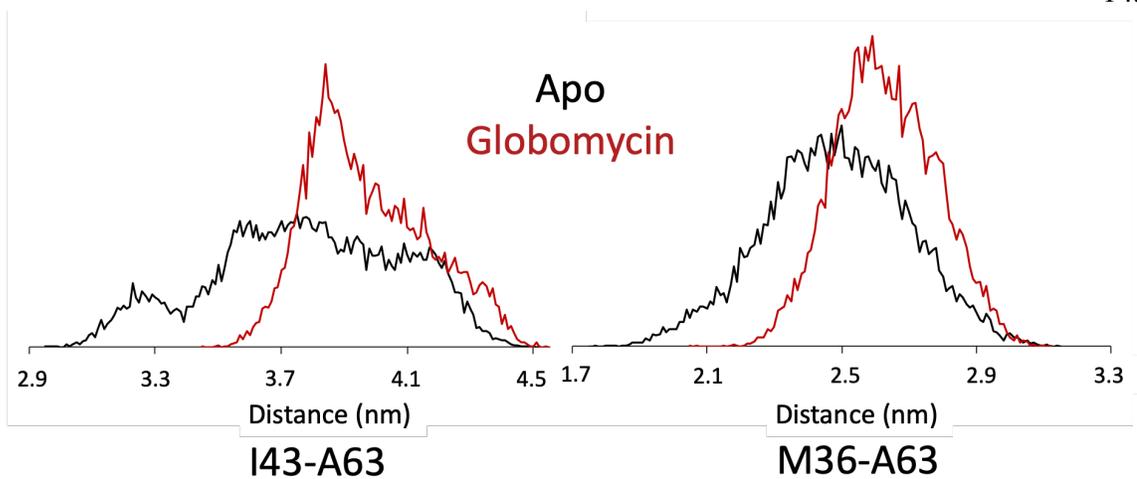


Figure 3.22. Combined MD histograms. Making a histogram combining all three simulation sets for each condition resulted in a more global analysis of each condition. The distance distribution for apo simulations is shown in black and globomycin bound simulations is shown in red for the I43-A63 pair (left) and the M36-A63 pair (right).

3.4 Comparison of EPR and MD Studies

3.4.1 Comparing distance distributions

MD derived distances are measured from $C\alpha$ to $C\alpha$ while DEER distances are measured between the spin labels. Thus, the DEER derived distances are around 8-10 Å longer than the MD derived distances. Additionally, the MD simulations were each run for 500 ns and so may not capture all conformations sampled in solution.

For the M36-A63 pair, the apo and globomycin MD distributions are broad and cover a similar distance range with apo sampling 1.8-3.1 nm centered at 2.5 nm and globomycin sampling 2.1-3.1 nm centered at 2.6 nm. These results are consistent with the DEER data in that the distributions are similar, except that in the DEER results the apo distribution was centered at a slightly longer distance than the globomycin distribution, and in the MD data the globomycin distribution is centered at a slightly longer distance than the apo distribution (Figures 3.15 and 3.22). Additionally, some population of longer distances (up to 4.9 nm) are seen in the DEER distribution (Figure 3.15). These inconsistencies may have arisen due to the short timescale of the MD simulations or the varying membrane environment, but overall the apo and globomycin distributions are both broad and similar.

On the other hand, the I43-A63 MD analysis did not yield results similar to the DEER analysis. The globomycin bound I43-A63 MD distribution was from 3.5-4.6 nm but peaked at 3.8 nm. The apo MD distribution was broad from 3.0-4.5 nm but had a shorter peak centered at 3.2 nm and then a second broad peak from 3.6-4.2 nm (Figure 3.22). Thus, the apo structure samples shorter distances than the globomycin bound structure in the MD simulations, but they both sample up to the same maximum distance (Figure 3.22).

Contrarily, the experimental DEER results in micelles showed that apo and globomycin bound structures both had a similar minimum distance, but the globomycin bound structure sampled a population of longer distances at 4.5 nm. The MD distributions also do not agree with the experimental DEER results in bicelles where a similar distance range was observed for both states (Figure 3.15).

3.4.2 MMM analysis of MD conformations

In order to directly compare MD and DEER results, structure files were extracted from the MD simulation, spin labels were added via MMM, and the DEER distribution was simulated (Figure 3.23). The simulated MMM distribution was then fit to the experimental DEER results by selecting the spin label rotamers that agree with the DEER distribution (Figure 3.23).

An MD structure was extracted corresponding to the most populated I43/A63 apo conformation at 3.8 nm. This structure yielded a simulated DEER distribution with a peak around 5 nm, far from the 2-4 nm peak observed in the experimental DEER distribution in micelles (Figure 3.24A), and was between the two experimental distribution peaks seen in the bicelle DEER data (Figure 3.25A). Thus, another structure was extracted corresponding to the shorter distance peak at 3.2 nm. This structure also yielded a simulated DEER distribution longer than the observed experimental distribution for micelles (Figure 3.24B), and was again between the two distance distribution peaks seen in the bicelle DEER data (3.25B). Fitting the simulated distribution to the micelle DEER data allowed for partial fitting of the experimental DEER distribution, however this distribution only contains 66.2% of spin label rotamers (Figure 3.26B). Thus, the fit MMM distribution is overfitting

the data. Finally, the structure with the shortest I43-A63 distance observed in the MD simulations (2.9 nm) was extracted (Figure 3.24C and 3.25C). This structure yielded a MMM simulated distance distribution around 4.1 nm. The fit distribution accounts for 91.1% of spin label rotamers for the micelle sample and fits the shorter distance distribution in the bicelle sample better. Thus, the apo LspA structure in both micelles and bicelles samples shorter distances experimentally than are observed in the MD simulations.

For the globomycin I43/A63 distance pair, a structure was extracted at the most populated distance of 3.8 nm. MMM analysis of this LspA structure showed a peak similar to the longer distance peak seen in the micelle DEER distribution, and the MMM fit distribution covered this peak well (Figure 3.26A). The simulated MMM distribution for this structure covered the shorter half of the longer distance seen in the globomycin bicelle DEER distribution, but when fit covers the shorter peak in the distribution (Figure 3.27A). As this peak is more populated, the MMM fit distribution is biased toward the shorter population. Another structure corresponding to a median pairwise distance of 4.1 nm was also extracted and analyzed. This structure has a simulated distribution at about 5.5 nm, which covers the micelle globomycin experimental peak at 4.5 nm when fit (Figure 3.26B). However, the shorter distances from 2-3 nm were not sampled by this MD-derived structure. The MMM simulated distribution for this extracted MD structure fits perfectly to the longer distance peak seen in the bicelle globomycin DEER distribution (Figures 3.27B and 3.28). It also fits to the same peak in the apo bicelle DEER distribution (Figure 3.28A). However, when these simulated distributions are fit to the experimental distributions, the program again biases toward the more populated shorter distances and no

longer covers these longer peaks (Figure 3.28B). This exemplifies the caution that needs to be taken when fitting MMM distributions.

The simulated DEER distributions for structures corresponding to the most populated M36-A63 distances for apo (2.5 nm) and globomycin (2.6 nm) both fit well to the experimental DEER data (Figure 3.29). Both simulated distributions were slightly longer than the DEER data, but when fit by MMM cover the experimental distributions precisely.

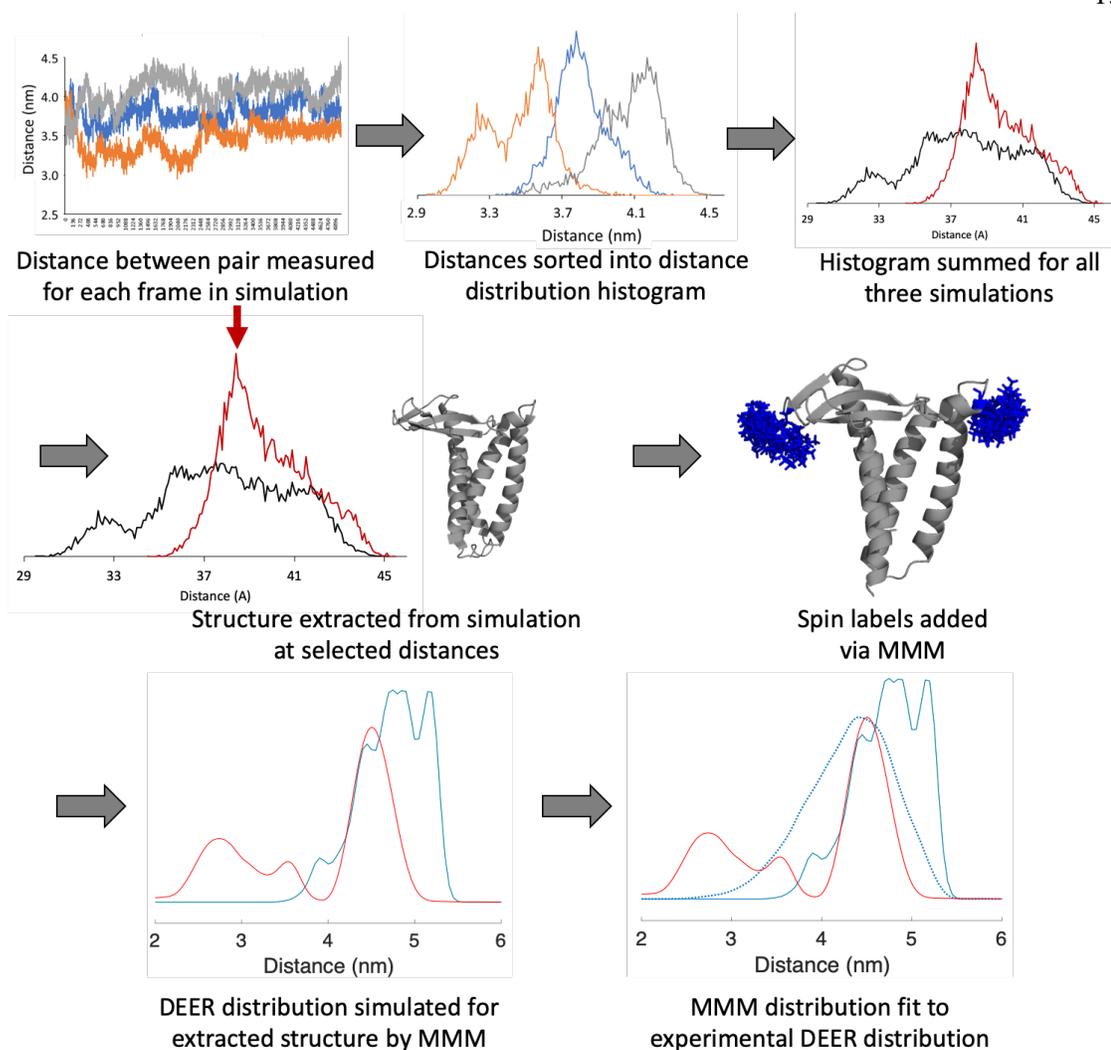


Figure 3.23. MD analysis workflow. Distances between pairs are measured for each MD simulation, formatted into histograms, and used to choose which structures to analyze. Spin labels are added to this structure and the DEER distribution is simulated by MMM. The simulated MMM distribution is then fit to the experimental DEER results.

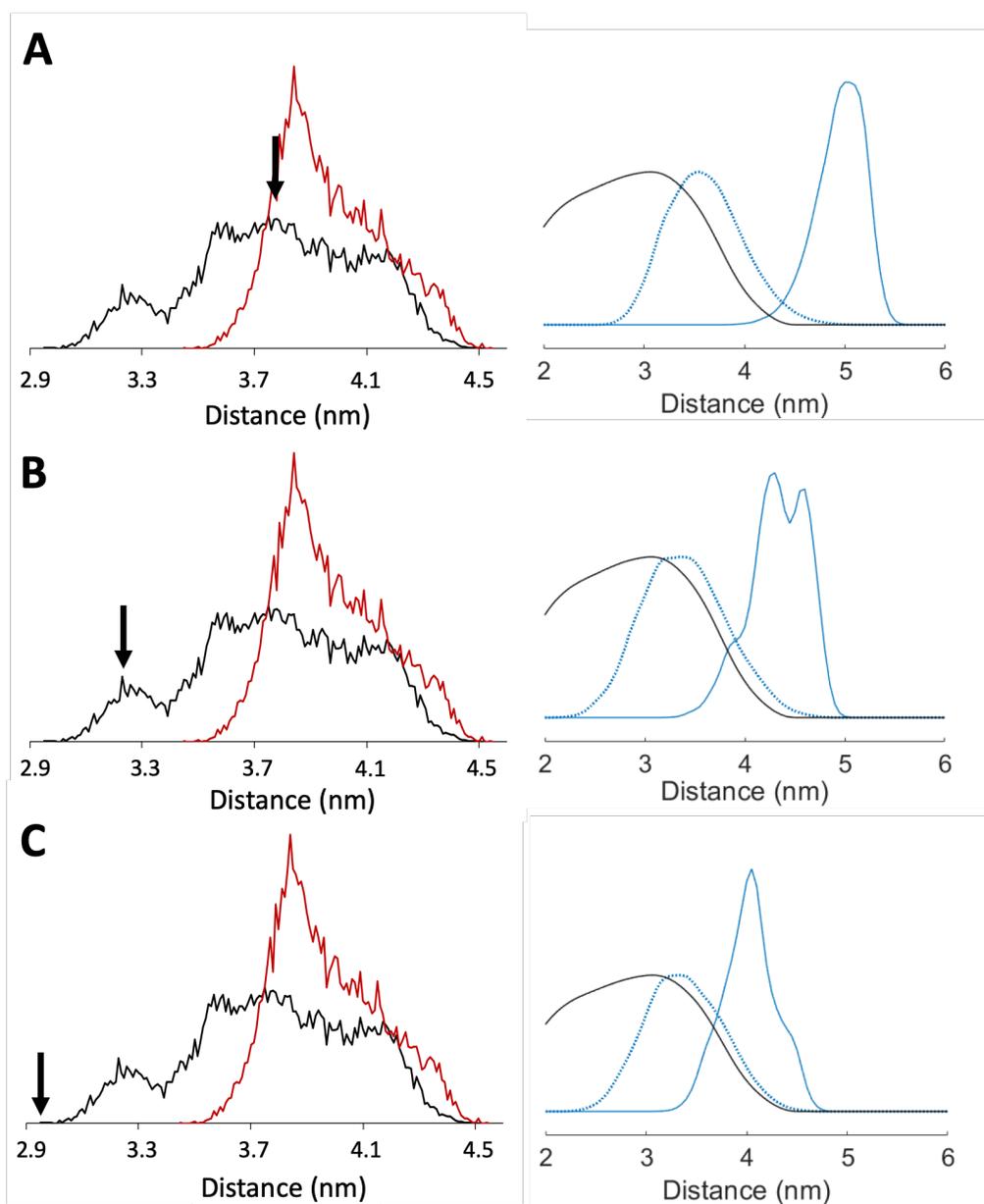


Figure 3.24. I43R1/A63R1 apo MD MMM analysis compared to micelle DEER data.

Structures were chosen at the most populated I43-A63 distance (A) the maximum of the shorter distance peak (B) and the shortest observed distance (C) for the apo LspA simulations. These were used to simulate the DEER distributions (solid blue line) which were then fit (dotted blue line) by selecting spin label rotamers that fit to the experimental apo micelle DEER distribution (black).

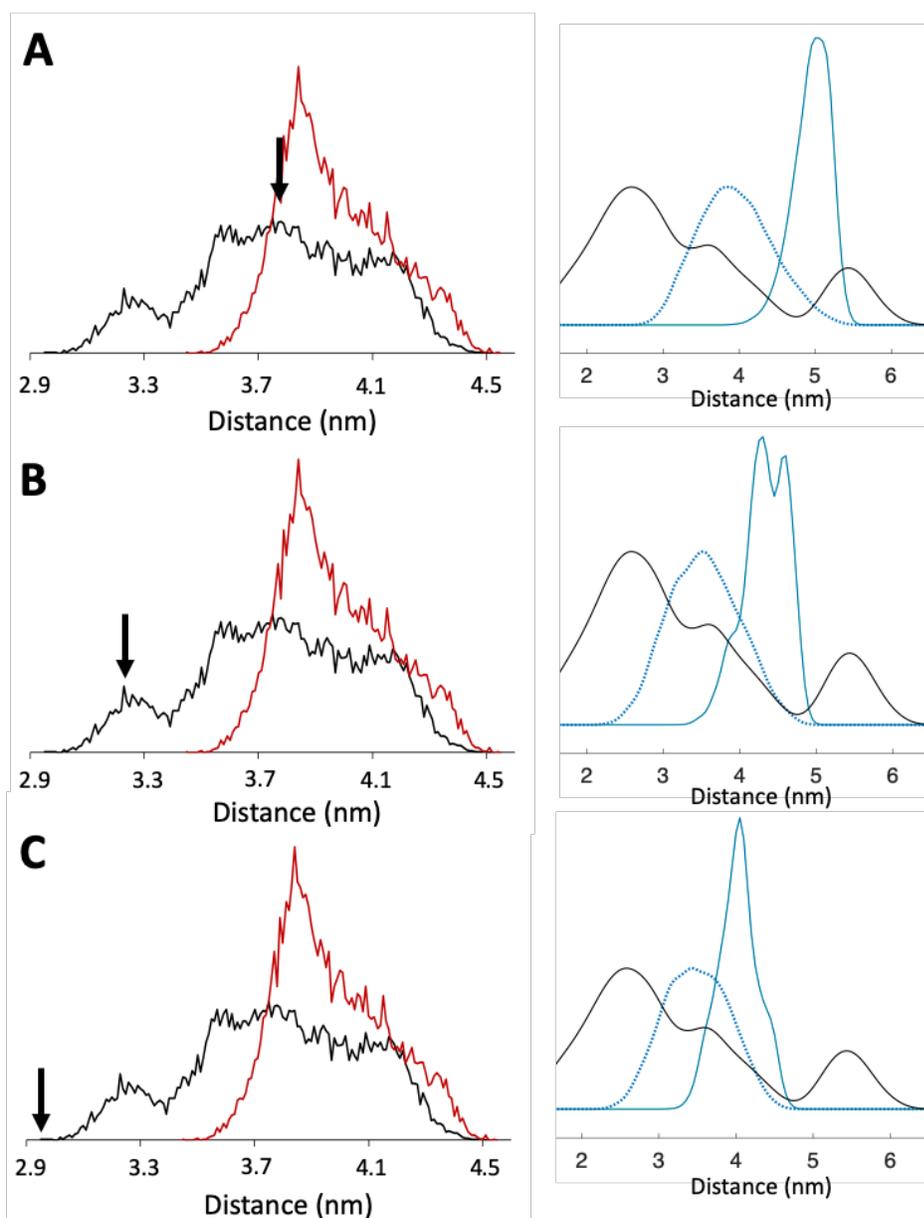


Figure 3.25. I43R1/A63R1 apo MD MMM analysis compared to bicelle DEER data.

Structures were chosen at the most populated I43-A63 distance (A) the maximum of the shorter distance peak (B) and the shortest observed distance (C) for the apo LspA simulations. These were used to simulate the DEER distributions (solid blue line) which were then fit (dotted blue line) by selecting spin label rotamers that fit to the experimental apo bicelle DEER distribution (black).

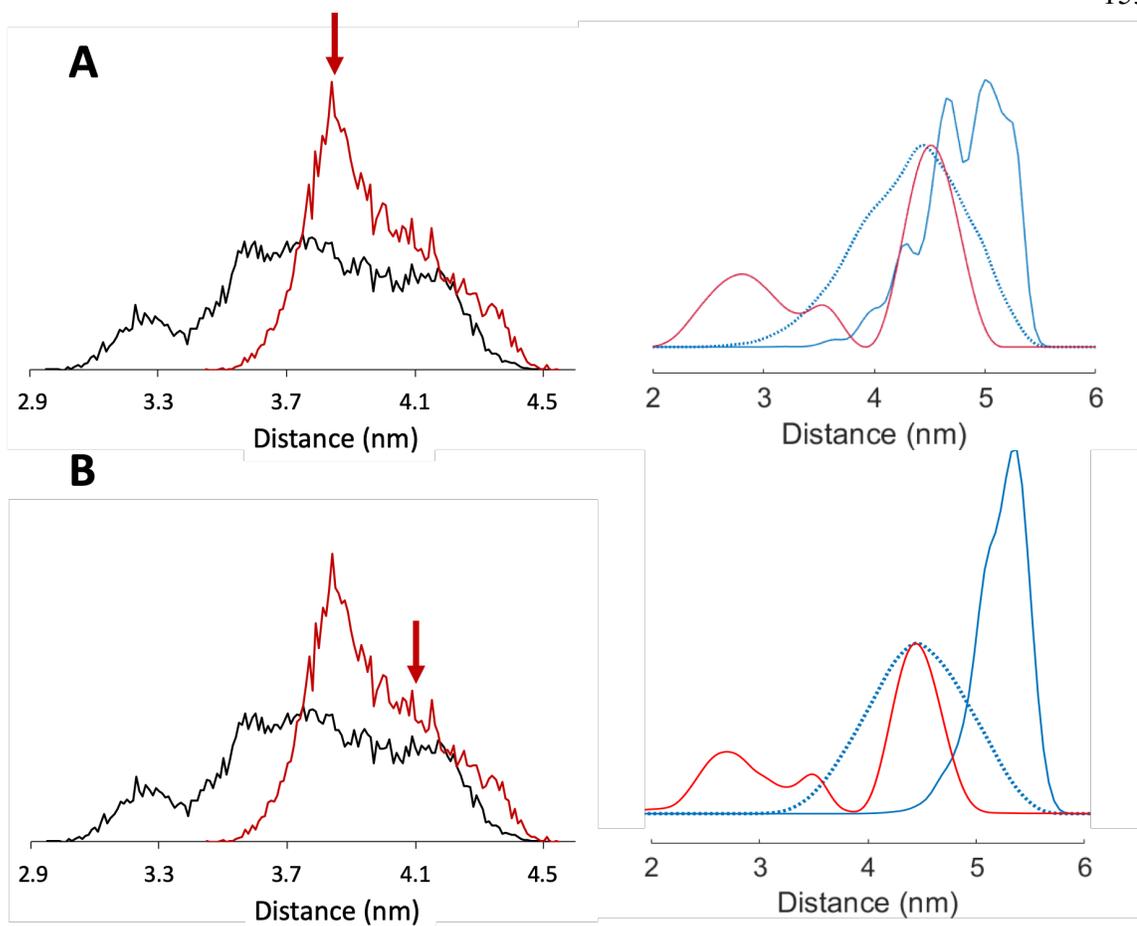


Figure 3.26. I43R1/A63R1 globomycin MD MMM analysis compared to micelle DEER data. Structures were chosen at the most populated I43-A63 distance for globomycin bound LspA (A), and at a median distance of 4.1 nm (B) (left). These structures were used to simulate the DEER distributions (solid blue line) which were then fit (dotted blue line) by selecting spin label rotamers that fit to the experimental micelle DEER distribution (red).

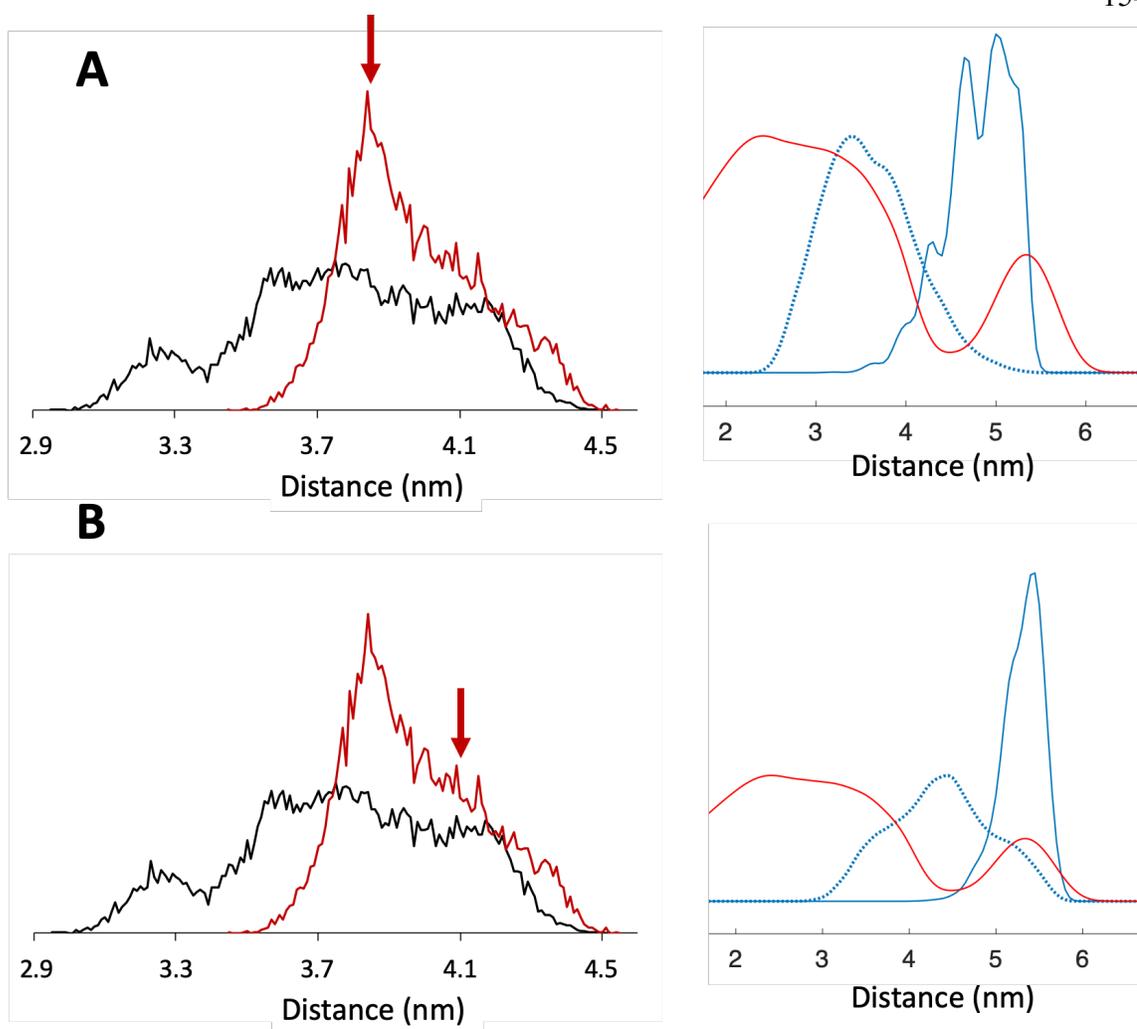


Figure 3.27. I43R1/A63R1 globomycin MD MMM analysis compared to bicelle DEER data. Structures were chosen at the most populated I43-A63 distance for globomycin bound LspA (A), and at a median distance of 4.1 nm (B) (left). These structures were used to simulate the DEER distributions (solid blue line) which were then fit (dotted blue line) by selecting spin label rotamers that fit to the experimental bicelle DEER distribution (red).

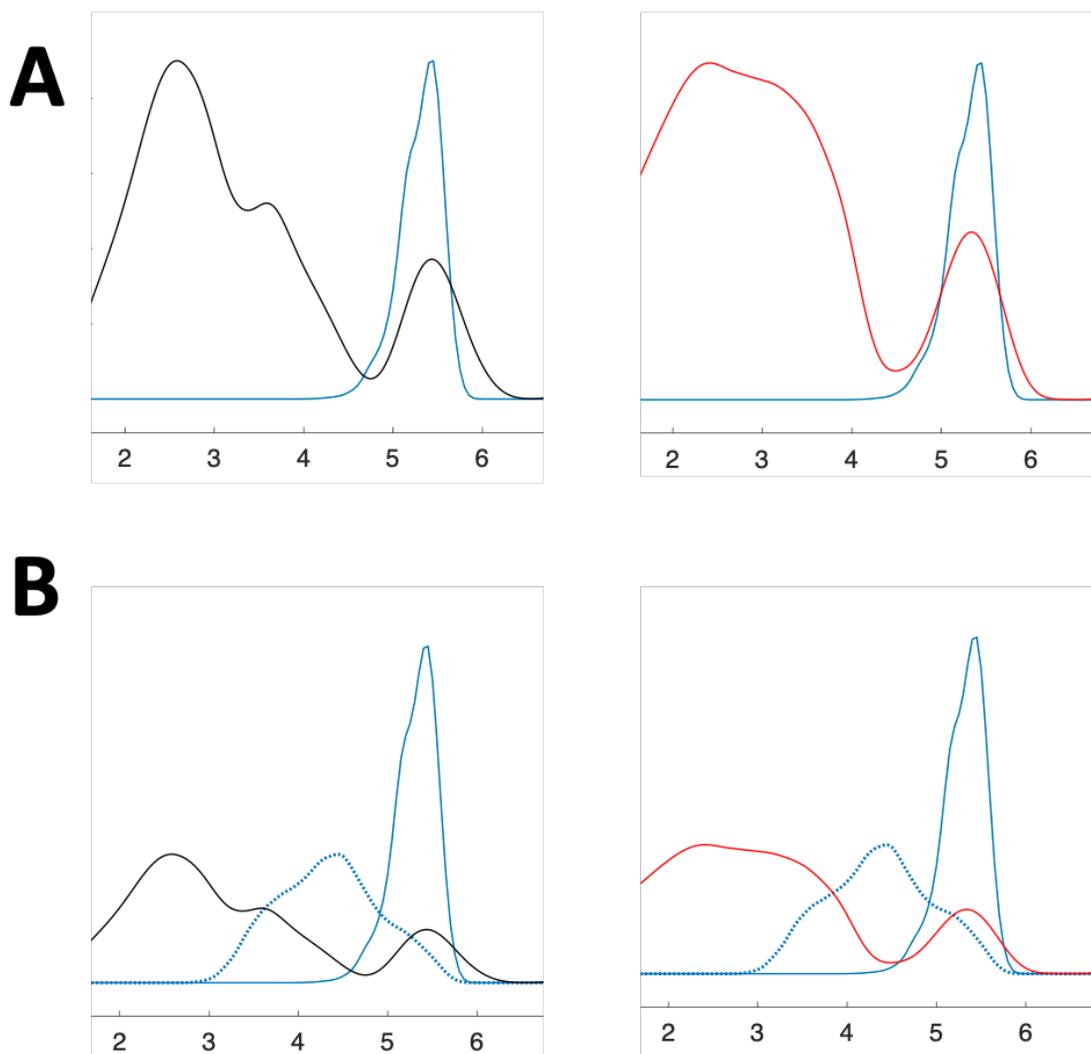


Figure 3.28. MD MMM distributions fit to bicelle DEER data. The simulated DEER distribution for the MD structure with a I43-A63 distance of 4.1 nm (Figure 3.27B) aligns with the longer distance peak seen in both apo (left) and globomycin bound (right) LspA in bicelles (A). However, when the MMM distribution is fit (B) the more populated shorter distances bias the fit distribution so that it does not fit either of the peaks. Simulated DEER distribution shown as solid blue line, fit simulated distribution shown by dotted blue line, and experimental DEER distribution shown by black (apo) or red (globomycin) lines.

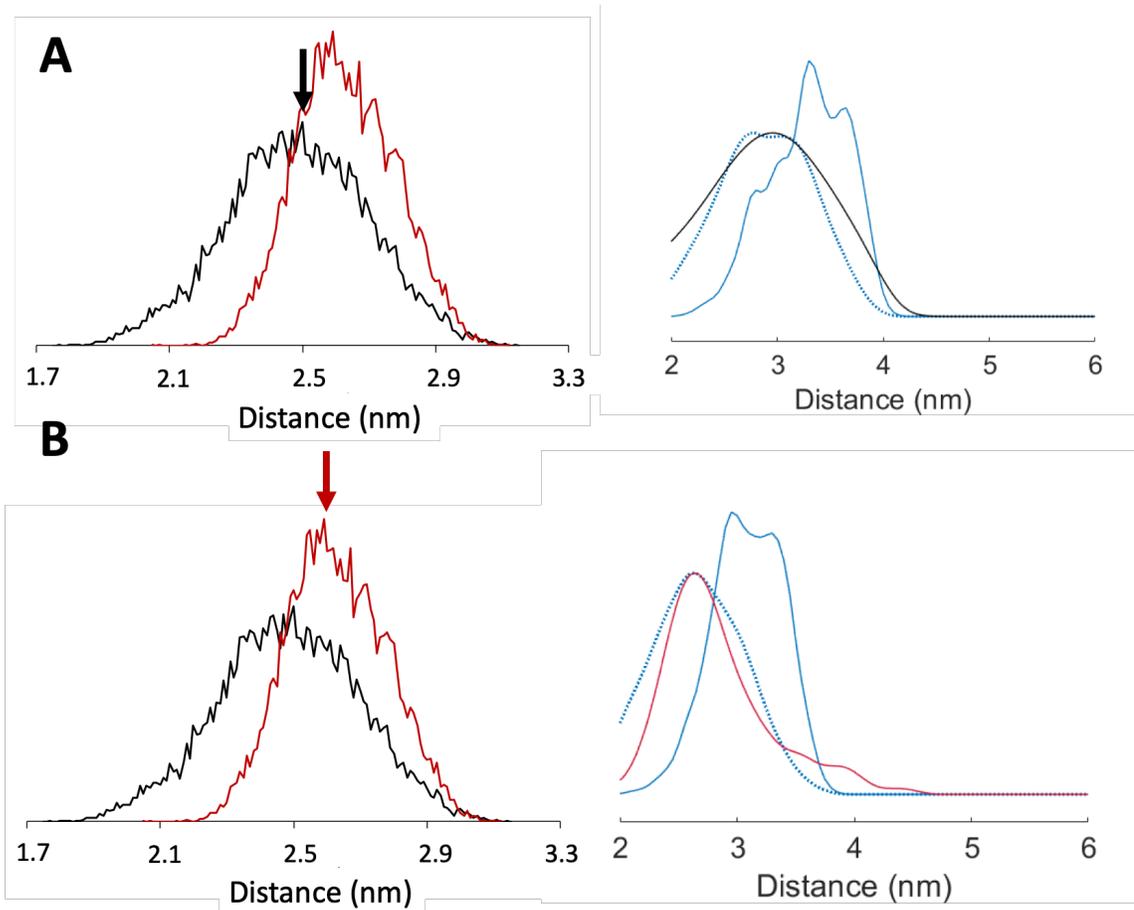


Figure 3.29. M36R1/A63R1 MD MMM analysis. Structures were chosen at the most populated M36-A63 distance for apo (A) and globomycin bound (B) LspA (left). These were used to simulate the DEER distributions (solid blue line) which were then fit (dotted blue line) by selecting spin label rotamers that fit to the experimental micelle DEER distribution (black, apo and red, globomycin).

3.4.3 Generating model structures

Three structures were extracted from the molecular dynamics simulations which represent closed, intermediate, and open states which are sampled in the experimental DEER distributions (Figures 3.30-3.33).

The closed conformation has an I43-A63 distance of 3.2 nm and a M36-A63 distance of 2.5 nm (Figure 3.30, gray). As described above, this structure contributes to a portion of the apo DEER distributions in both micelles and bicelles, but shorter I43-A63 distances are observed in the experimental DEER distribution than are observed in the MD simulated distribution (Figure 3.24B, 3.25B). This suggests that the closed structure may not be as constricted as observed experimentally. However, this closed structure is in close alignment with the myxovirescin crystal structure. Interestingly, while the orientations of the PH and β -cradle in the closed MD structure and myxovirescin crystal structure align well and create the same constricted structure, the A63 (S61 in *S. aureus* LspA) spin label orientation is influenced by the curvature of the PH (Figure 3.34). Thus, the distance distribution for the myxovirescin crystal structure is shorter and agrees better with the apo DEER distribution in micelles (Figure 3.14), while the closed MD derived structure results in a slightly longer distance distribution (Figure 3.24B). This suggests that the same constricted structure is observed in the DEER data and MD simulation, and that the differences seen in the distance distributions are due to differing orientations of the spin labels and not movement of the PH. This difference exemplifies the complexity of interpreting DEER distance distributions due to the flexibility of the spin label, and shows that small changes in side chain orientation can greatly influence the distance distribution. While the MD globomycin structure does not sample the closed structure distance pairs, it

is possible that over a longer timescale the globomycin bound LspA structure could sample this structure, as indicated by a second more immobile component in the CW spectra upon globomycin addition, and the overlap of the globomycin and apo DEER distributions at shorter distances.

The intermediate conformation state has an I43-A63 distance of 3.8 nm and a M36-A63 distance of 2.6 nm corresponding to the most populated distance states for each of these pairs in the globomycin trajectories (Figure 3.30, red). This structure aligns well with the globomycin crystal structures. The simulated distribution for this structure agrees well with the globomycin induced population in the micelle DEER distribution at 4.5 nm (Figure 3.26A), and resides between the two distance distributions of the bicelle globomycin bound DEER distribution (Figure 3.27A). Thus, this structure best represents the globomycin bound state in a micelle membrane mimetic. It is possible that this state could also represent a myxovirescin bound structure if myxovirescin can occupy both binding sites as hypothesized for globomycin. Preliminary MD studies suggest that myxovirescin can occupy the two binding sites (personal communication with Dr. Denis Wolan), but future work would need to be completed to collect DEER data for myxovirescin bound LspA to determine if this is observed in solution.

The open state has an I43-A63 distance of 4.1 nm and a M36-A63 distance of 2.7 nm (Figure 3.30, blue). This structure agrees very well with the longer distance peaks seen in the apo and globomycin bound bicelle distance distributions around 5.5 nm (Figure 3.28A), and also covers distances seen in the globomycin bound micelle DEER distribution (Figure 3.26B). Thus, this structure represents a more open conformation that is seen in both the micelle and bicelle DEER data, but is more prominent in the bicelle DEER data.

The PH - β -cradle orientation in this structure is wider than is seen in any crystal structure. This open structure may occur in either the apo or substrate bound states, as the structure was sampled in both the apo and globomycin bicelle DEER distributions and is seen in both the apo and globomycin bound MD simulations. However, this conformation was not sampled in the micelle apo DEER distribution. Perhaps the size of the micelle constricted the conformational dynamics of the protein, or perhaps the open conformation is more stabilized by the bicelle environment.

Comparison of these three structures yields a constricted conformation between the PH and β -cradle for the closed structure, an intermediary conformation for the intermediate structure, and the widest conformation for the open structure (Figure 3.30 and 3.31). The differences between these conformations is due to a movement of the PH domain. All other parts of the protein do not show conformational change. The intermediate structure agrees well with the globomycin bound crystal structures, while the closed MD derived structure agrees well with the myxovirescin crystal structure. The open MD derived structure is not represented by a crystal structure.

The PH is oriented much closer to the β -cradle in the closed structure. At the most constricted point, the β -cradle and PH backbone are only 6.2 Å apart which completely occludes access to the active site, as can be visualized by a surface representation (Figure 3.31). This closed conformation keeps the negatively charged residues in the active site from interacting with the hydrophobic lipid environment, which are more exposed in the globomycin structure, and most exposed in the open structure (Figure 3.32).

While the MD derived intermediate and open structures show a trigonal cavity allowing for the lipoprotein, signal peptide, and DAG moiety of the lipoprotein substrate

to fit in the active site, the closed structure does not show this trigonal feature. Notably, the groove where DAG is hypothesized to reside is blocked, and there is a much thinner space for the linker from the signal peptide to the lipoprotein soluble domain (Figure 3.33). While the intermediate structure shows the trigonal cavity, there is obstruction where the signal peptide helix is hypothesized to reside. This makes sense, as while globomycin binds in the active site and mimics the lipoprotein lipobox,² it does not contain the signal peptide. Although not completely clear, the open conformation has the most space in the signal peptide groove. Thus, this open structure of LspA may represent a conformation that allows the prelipoprotein to enter into and bind in the active site of LspA in the correct orientation for signal peptide cleavage.

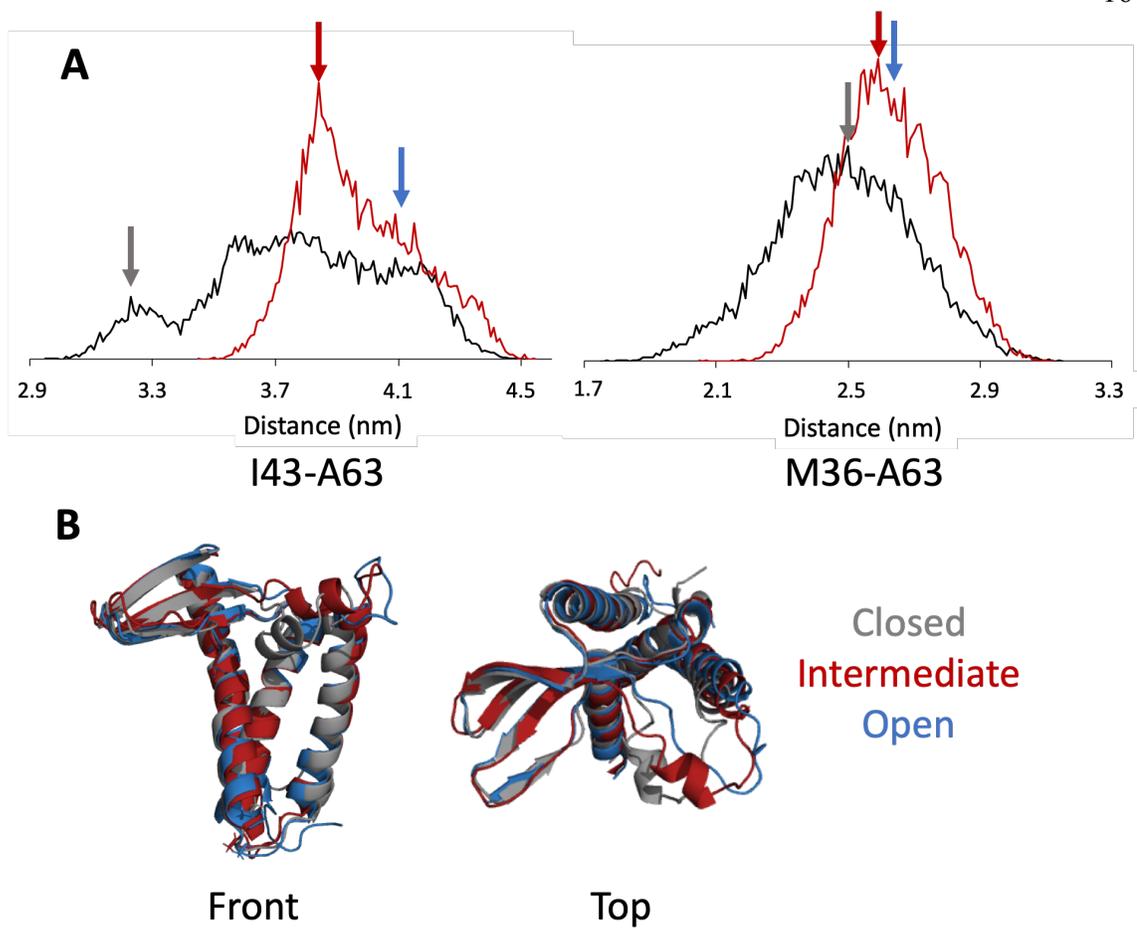


Figure 3.30. MD derived model structures. A) Distances of I43-A63 (left) and M36-A63 (right) for each structure are shown. B) Closed (gray), intermediate (red), and open (blue) structures show movement of the PH domain.

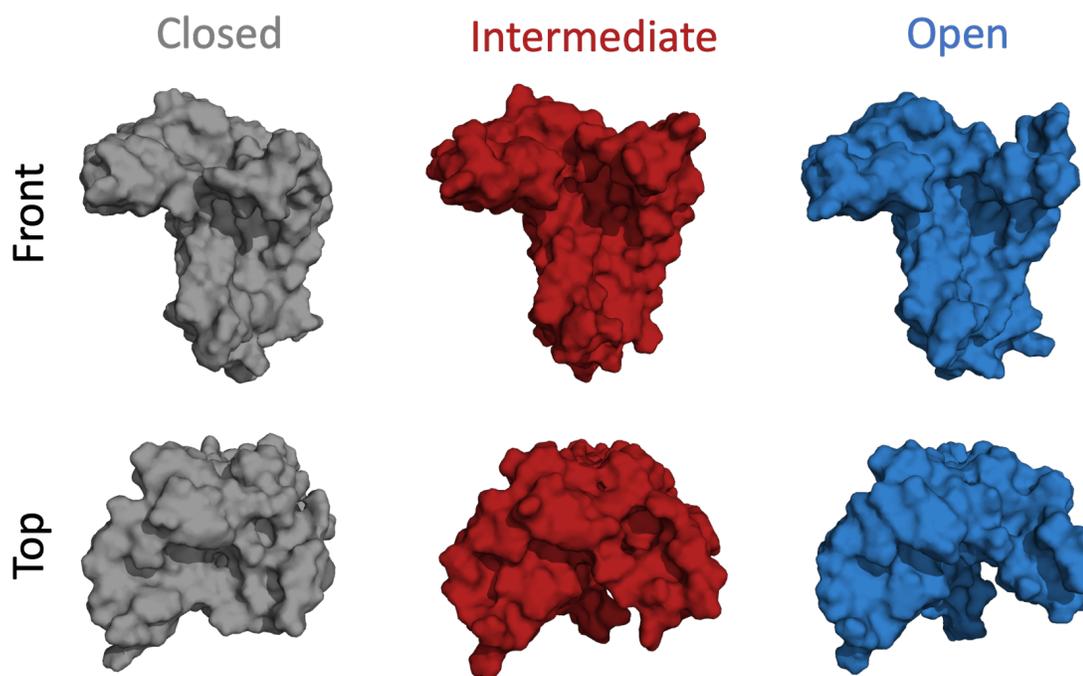


Figure 3.31. Surface representation of structures. From the front view (top) and top view (bottom), the surface representations clearly show the opening of the active site from the closed (gray) to intermediate (red) to open (blue) conformations through the movement of the PH domain relative to the β -cradle.

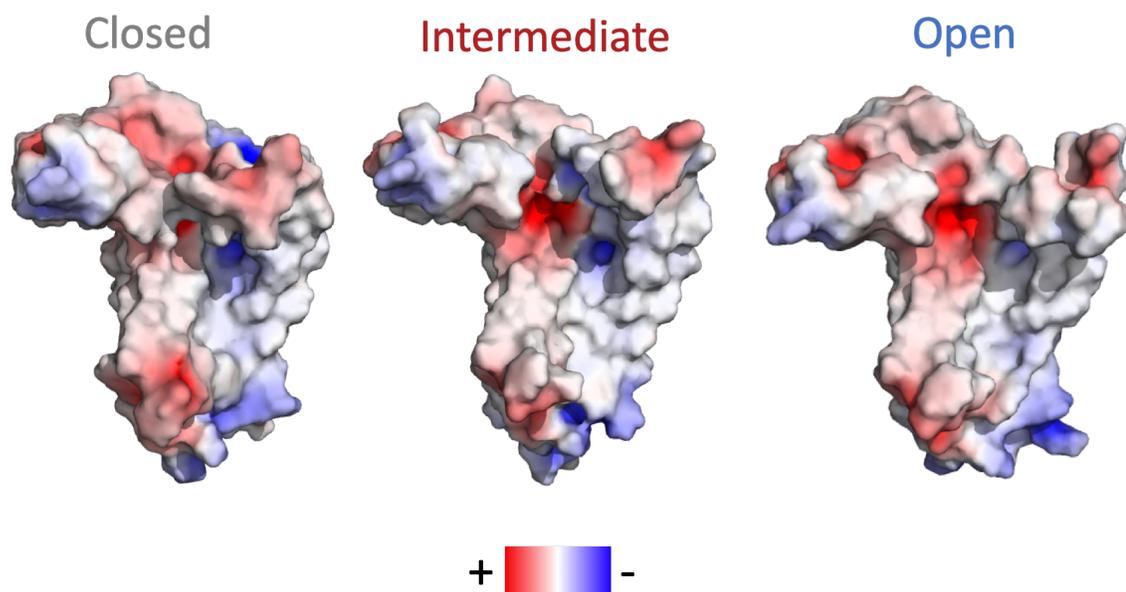


Figure 3.32. Protein electrostatics. Charged residues in the active site are hidden in the closed conformation. Positive charged residues are shown in red and negative charged residues in blue.

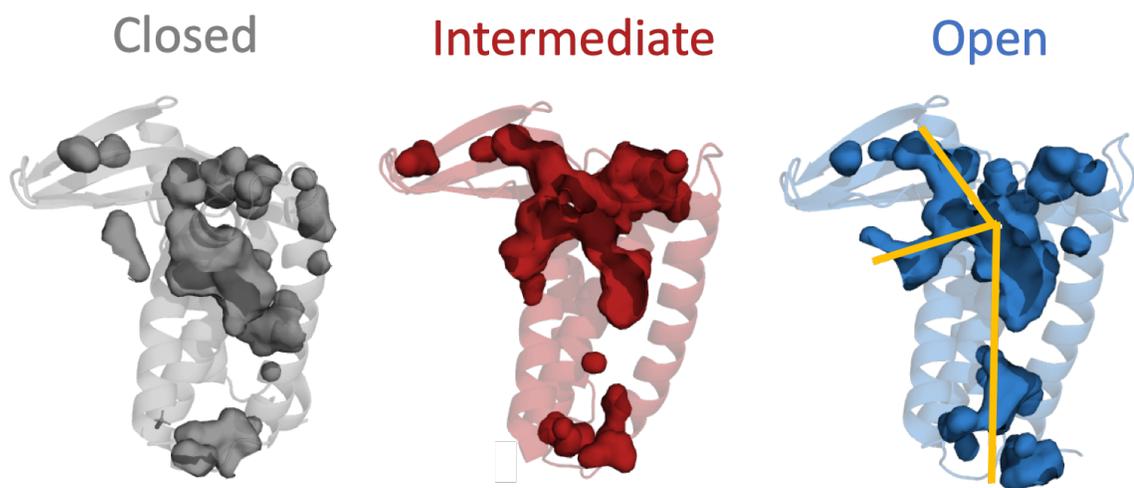


Figure 3.33. Cavity fill models. Cavity fill models generated in pyMOL show the space available in the active site and binding grooves. The trigonal cavity where the prelipoprotein is hypothesized to bind is shown as yellow lines on the open structure.

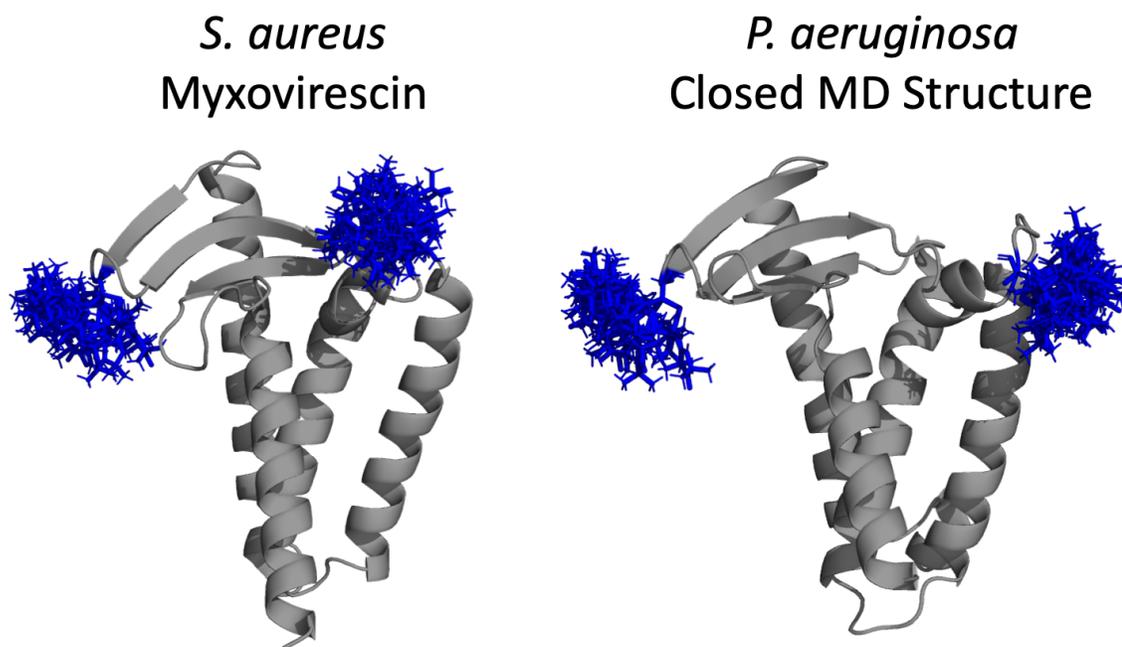


Figure 3.34. Spin labeled LspA. Although the *S. aureus* myxovirescin crystal structure (PDB ID 6RYP) and the *P. aeruginosa* closed MD derived structure have a similar constricted structure, the curvature of the PH causes the spin labels to be facing in different directions leading to differing distance distributions.

3.5 Conclusions

LspA is an excellent target for the development of antibiotic drugs as it is essential in several bacteria and may be impervious to the development of antibiotic resistance.^{1,2} Here we report the first study of LspA conformational dynamics, which allows for a better understanding of the mechanism of LspA. EPR data show changes in LspA conformation upon binding of the antibiotic globomycin and suggest a more closed PH - β -cradle conformation in the apo state. MD simulations agree with experimental DEER data. Slight discrepancies in EPR and MD data, such as the shorter distances observed in the apo DEER distribution, may arise due to differences in the membrane mimetics used and the timescale sampled by MD.

Complementation of EPR and MD data allowed for the generation of three model structures which represent the conformational states of LspA. These structures suggest that in the apo state the β -cradle and PH adopt a closed conformation that occludes the binding site and hides the charged residues; in the globomycin bound state the active site is opened to allow for the binding of the antibiotic; and in the open state the β -cradle and PH are wider apart allowing for the entrance and binding of prelipoprotein substrate in the active site.

3.6 Methods

3.6.1 Expression, Purification, and Spin Labeling of LspA

The *Pseudomonas aeruginosa* (strain PAO1) LspA gene was purchased in a pET28b vector with an N-terminal 6xHis tag and thrombin cleavage sequence (General Biosystems Inc.). Cysteine residues were introduced via PIPE Mutagenesis¹⁸ or

QuikChange¹⁹ (Agilent, Santa Clara, CA), and gene sequencing confirmed the correct mutations (Genewiz Inc.). The plasmids were transformed into C41(DE3) *E. coli* cells (Lucigen) and cultures were grown in Luria-Burtani (LB) media supplemented with 50 mg/L kanamycin to and OD600 around 0.8. Expression was induced with 1 mM isopropyl- β -thio-D-galactoside (IPTG) for around 18 hours at 25°C. Cells were harvested by centrifugation at 5,000g for 10 min at 4 °C and the pellets were frozen at -20 °C.

Cells were resuspended in Buffer A (20 mM phosphate, pH 7.2, 150 mM NaCl) with one Complete Protease Inhibitor Cocktail tablet (Roche) and passed twice through a high-presser Nano DeBEE homogenizer (BEE International). Cell debris were removed via centrifugation at 18,000g for 30 min. The membrane fraction was subsequently separated by ultra-centrifugation at 150,000g for 1 hour at 4 °C. The membrane pellet was resuspended in 30 mL Buffer A and frozen at -80 °C. The membrane resuspension was thawed and 10 mL Buffer A containing fos choline-12 (FC12) was added such that the final concentration of FC12 was 1.8% (w/v) and was allowed to rock at 4 °C for at least one hour. Unsolubilized material was then removed by ultracentrifugation at 100,000g for 45 min. Recombinant LspA protein was bound to a Ni²⁺ immobilized metal affinity chromatography column, washed with Buffer A containing 40 mM imidazole.

For spin labeling, 10 mL Buffer A containing 40 mM imidazole and 0.7 mM MTSL/R1 spin label (S-(2, 2, 5, 5-tetramethyl-2,5- dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate, Toronto Research Chemicals Inc.) was added and the column was allowed to nutate at 4 °C overnight. The next day, the column was further washed with 20 mL Buffer A with 40 mM imidazole to remove free R1 and spin labeled protein was eluted in 10 mL Buffer A containing 300 mM imidazole. Imidazole was removed over a PD-10

column (GE Healthcare Biosciences) and the protein was concentrated to around 300 mM using a 10kDa molecular weight cutoff concentrator (Millipore). Protein purity was assessed by SDS-PAGE and the protein was confirmed as LspA with MALDI-TOF mass spectrometry.

3.6.2 Continuous-Wave EPR

CW EPR experiments on singly labeled LspA proteins in FC12 detergent micelles were performed on a CW X-Band EMX spectrometer (Bruker Biospin) at room temperature. The samples were loaded into 0.6 mm glass capillary tubes (VitroCom) with a sample volume around 7 μ L. Spectra were processed with Bruker software (WinEPR), and analyzed with LabView programs (Base2 and ADJ) provided by Dr. Christian Altenbach and Dr. Wayne Hubbell (UCLA).

For globomycin studies, globomycin (Sigma) was resuspended in DMSO at a concentration of 10 mg/mL. The appropriate amount was aliquoted into tubes and dried in a lyophilizer. The dried globomycin was resuspended with spin labeled LspA sample for EPR studies. This procedure was done because DMSO was shown to have a large impact on CW spectra (Figure S6).

3.6.3 Pulsed EPR

Double-labeled LspA proteins in FC12 detergent micelles were measured using pulsed EPR with a Q-band Bruker ELEXSYS E580 Spectrometer at Q-band and 80 K. All samples were prepared to a final protein concentration around 300 μ M with 20% deuterated glycerol with a sample volume of 15 μ L. The samples were loaded into quartz capillaries

with a 1.6 mm od x 1.1 mm id x 100 mm length (VitroCom) and flash frozen in liquid nitrogen. A four pulse DEER sequence was used with 16-step phase cycling (20 ns $\pi/2$ and two 40 ns π observed pulses, and a π pump pulse).²⁰ The pump frequency was set at the maximum of the nitroxide spectrum and the observed frequency was set to 75 MHz lower. Increasing inter-pulse delays at 16 ns increments were used with a 16-step phase cycle during data collection. Accumulation times were typically between 18 and 24 hours, with a dipolar evolution time around 4 μ s. Dipolar evolution data were processed using DEERAnalysis2018 software using Tikhonov regularization to generate the distance distributions.^{15,21} Simulated DEER distributions were generated using MMM.¹⁶

3.6.4 Molecular Dynamics Simulations

All MDS were performed using GROMACS v5.0.5.²² The Martini 2.2 force field²³ was used to run an initial 1 μ s Coarse Grained (CG) MD simulation to permit the assembly and equilibration of a palmitoyloleoylphosphatidylglycerol (POPG) palmitoyloleoylphosphatidylethanolamine (POPE) (1:3 mole ratio) bilayer around *P. aeruginosa* LspA.²⁴ The end-snapshot of the CG simulations was then converted to atomic detail with the crystal structure aligned with the CG protein within the assembled lipid bilayer.²⁵ The systems were equilibrated for 1 ns with the protein restrained before 500 ns of unrestrained atomistic MD for apo and globomycin bound simulations.²⁶ Three 500 ns simulations were performed for both apo and globomycin bound LspA. The parameters for globomycin were manually created, modifying the standard amino acid force field parameters to generate the correct stereochemistry and to complete the cyclic structure of the antibiotic. The dagylated cysteine parameters were created from lipid parameters for

diacylglycerol and appended to the thiol side chain of cysteine. Systems were neutralized with 150 mM NaCl. *In silico* mutagenesis was performed and figures were prepared using PyMol (Schrödinger). VMD was used to analyze molecular dynamics trajectories.²⁷

3.6.5 Evolutionary Couplings

Evolutionary contacts for *P. aeruginosa* LspA were generated using the GREMLIN server⁹ and were analyzed in PyMol (Schrödinger).

3.7 *References*

- 1 Buddelmeijer N. The molecular mechanism of bacterial lipoprotein modification- How, when and why? *FEMS Microbiol Rev* 2015;**39**:246–61.
<https://doi.org/10.1093/femsre/fuu006>.
- 2 Vogeley L, El Arnaout T, Bailey J, Stansfeld PJ, Boland C, Caffrey M. Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. *Science (80-)* 2016;**351**:876–80.
<https://doi.org/10.1126/science.aad3747>.
- 3 Nakayama H, Kurokawa K, Lee BL. Lipoproteins in bacteria: Structures and biosynthetic pathways. *FEBS J* 2012;**279**:4247–68.
<https://doi.org/10.1111/febs.12041>.
- 4 Szewczyk J, Collet JF. *The Journey of Lipoproteins Through the Cell: One Birthplace, Multiple Destinations*. vol. 69. 1st ed. Elsevier Ltd.; 2016.
- 5 Olatunji S, Yu X, Bailey J, Huang C, Zapotoczna M, Bowen K, *et al*. Structures of lipoprotein signal peptidase II from *Staphylococcus aureus* complexed with

- antibiotics globomycin and myxovirescin 2020. <https://doi.org/10.1038/s41467-019-13724-y>.
- 6 Marks DS, Hopf TA, Sander C. Protein structure prediction from sequence variation. *Nat Biotechnol* 2012;**30**:1072–80. <https://doi.org/10.1038/nbt.2419>.
 - 7 Marks DS, Colwell LJ, Sheridan R, Hopf TA, Pagnani A, Zecchina R, *et al*. Protein 3D structure computed from evolutionary sequence variation. *PLoS One* 2011;**6**:. <https://doi.org/10.1371/journal.pone.0028766>.
 - 8 Tang Y, Huang YJ, Hopf TA, Sander C, Marks DS, Montelione GT. Protein structure determination by combining sparse NMR data with evolutionary couplings. *Nat Methods* 2015;**12**:751–4. <https://doi.org/10.1038/nmeth.3455>.
 - 9 Balakrishnan S, Kamisetty H, Carbonell JG, Lee S-I, Langmead CJ. Learning generative models for protein fold families. *Proteins Struct Funct Bioinforma* 2011;**79**:1061–78. <https://doi.org/10.1002/prot.22934>.
 - 10 Kamisetty H, Ovchinnikov S, Baker D. Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *Proc Natl Acad Sci* 2013;**110**:15674–9. <https://doi.org/10.1073/pnas.1314045110>.
 - 11 Dev IK, Ray PH. Rapid assay and purification of a unique signal peptidase that processes the prolipoprotein from *Escherichia coli* B. *J Biol Chem* 1984;**259**:11114–20.
 - 12 Tokunaga M, Loranger JM, Wu HC. Prolipoprotein modification and processing enzymes in *Escherichia coli*. *J Biol Chem* 1984;**259**:3825–30.
 - 13 Kitamura S, Owensby A, Wall D, Wolan DW. Lipoprotein Signal Peptidase Inhibitors with Antibiotic Properties Identified through Design of a Robust In

- Vitro HT Platform. *Cell Chem Biol* 2018.
<https://doi.org/10.1016/j.chembiol.2017.12.011>.
- 14 Kitamura S, Wolan DW. Probing substrate recognition of bacterial lipoprotein signal peptidase using FRET reporters. *FEBS Lett* 2018.
<https://doi.org/10.1002/1873-3468.13155>.
- 15 Jeschke G. DEER Distance Measurements on Proteins. *Annu Rev Phys Chem* 2012;**63**:419–46. <https://doi.org/10.1146/annurev-physchem-032511-143716>.
- 16 Jeschke G. MMM: A toolbox for integrative structure modeling. *Protein Sci* 2018;**27**:76–85. <https://doi.org/10.1002/pro.3269>.
- 17 Klug CS, Feix JB. Methods and Applications of Site-Directed Spin Labeling EPR Spectroscopy. *Methods Cell Biol* 2008;**84**:617–58. [https://doi.org/10.1016/S0091-679X\(07\)84020-9](https://doi.org/10.1016/S0091-679X(07)84020-9).
- 18 Klock HE, Lesley SA. The Polymerase Incomplete Primer Extension (PIPE) Method Applied to High-Throughput Cloning and Site-Directed Mutagenesis Heath. *Methods Mol. Biol. High Throughput Protein Expr. Purif.*, vol. 498. 2009. p. 91–103.
- 19 Xia Y, Xun L. Revised Mechanism and Improved Efficiency of the QuikChange Site-Directed Mutagenesis Method. In: Reeves A, editor. *Vitr. Mutagen. Methods Protoc.* New York, NY: Springer New York; 2017. p. 367–74.
- 20 Pannier M, Veit S, Godt A, Jeschke G, Spiess HW. Dead-Time Free Measurement of Dipole-Dipole Interactions between Electron Spins. *J Magn Reson* 2000;**142**:331–40. <https://doi.org/10.1023/A:1011323413057>.
- 21 Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham JE, *et al.*

- DeerAnalysis 2006- a Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl Magn Reson* 2006;**30**:473–98.
- 22 Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013;**29**:845–54.
<https://doi.org/10.1093/bioinformatics/btt055>.
- 23 De Jong DH, Liguori N, Van Den Berg T, Arnarez C, Periole X, Marrink SJ. Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex. *J Phys Chem B* 2015;**119**:7791–803.
<https://doi.org/10.1021/acs.jpcc.5b00809>.
- 24 Stansfeld PJ, Goose JE, Caffrey M, Carpenter EP, Parker JL, Newstead S, *et al.* MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* 2015;**23**:1350–61.
<https://doi.org/10.1016/j.str.2015.05.006>.
- 25 Stansfeld PJ, Sansom MSP. From coarse grained to atomistic: A serial multiscale approach to membrane protein simulations. *J Chem Theory Comput* 2011;**7**:1157–66. <https://doi.org/10.1021/ct100569y>.
- 26 Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 2004;**25**:1656–76.
<https://doi.org/10.1002/jcc.20090>.
- 27 Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *J Mol Graph* 1996. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).

CHAPTER 4. LspA ACTIVITY ASSAY

4.1 *Introduction*

4.1.1 Project Motivation

As antibiotic resistance is a growing medical concern, it is imperative that drugs are developed to target novel pathways. Lipoprotein signal peptidase (LspA) has been identified as an antibiotic drug target, as it targets a novel pathway, is essential in some bacteria, and does not have a mammalian homolog (see Chapter 1). The crystal structure of LspA has been determined (Chapter 1),^{1,2} and the conformational dynamics of the protein has been studied (Chapter 3). However, in order to efficiently develop drugs to target novel proteins, it is essential to have a biological, quantitative, reproducible, and high-throughput activity assay to test the effectiveness of the developed drugs. An activity assay is currently lacking for LspA. Here, a LspA activity assay is sought that will be used to gain a deeper understanding of the protein's mechanism, test requirements for LspA activity, and ultimately be used to test the efficiency of inhibitors in future antibiotic development.

4.1.2 Previous Work

4.1.2.1 Initial LspA Studies

Early biochemical studies of LspA were completed in the 1980s using radioactively-labeled signal peptide.³⁻¹¹ While studying a newly developed antibiotic, globomycin, two groups separately discovered that the signal peptide conjugated form of lipoprotein (later named prolipoprotein) accumulated in the cytoplasmic membrane. They thus determined that a signal peptidase (later named LspA) was blocked from carrying out

cleavage of the signal peptide.^{6,9,12} Radioactively-labeled signal peptide assays were also used to determine that lipid modification was necessary for LspA activity, and that cleavage occurred in the cytoplasmic membrane.^{4,5,10} The requirements of the residue before the cleavage site (directly before the lipid-modified cysteine) were also tested. While glycine is strongly conserved, serine is the upper size limit for the side chain at this position, as valine, leucine, and threonine substitutions were not processed.⁷

Subsequent studies identified changes in cell growth, colony morphology, protein transcription, and lipoprotein biogenesis when LspA was mutated, knocked out, or inhibited. These studies were completed on gram-negative bacteria: *Escherichia coli*,^{5,9,13–15} *Enterobacter aerogenes*,¹⁶ *Myxococcus xanthus*,^{17,18} *Legionella pneumophila*,¹⁹ *Rickettsia typhi*,²⁰ and *Pseudomonas aeruginosa*¹; as well as gram-positive bacteria: *Bacillus subtilis*,^{15,21,22} *Staphylococcus aureus*,^{2,23} *Staphylococcus carnosus*,²⁴ *Streptococcus suis*,²⁵ *Streptococcus pneumoniae*,²⁶ *Streptomyces scabies*,²⁷ *Streptomyces lividans*,²⁸ *Streptomyces coelicolor*,^{29,30} and *Streptococcus pyogenes*.¹³ These studies showed that LspA is essential in most gram-negative bacteria including *E. coli*,^{8,11} *S. coelicolor*,³⁰ and *R. typhi*.²⁰ While LspA is not essential in gram-positive bacteria, loss of LspA function impairs production of secretory proteins, alters cellular metabolism impairing growth and development, reduces bacterial replication, and triggers stress responses.^{31,32} Interestingly, LspA of *Mycobacterium tuberculosis* is not essential for growth but is essential for virulence and infection.³³ Mutations of LspA in *S. aureus* were also shown to reduce the bacteria's ability to survive in human blood, demonstrating the importance of LspA in both gram-negative and gram-positive bacteria.²

4.1.2.2 Gel Based Assay

More recently, a qualitative gel-based assay was used to measure the activity of LspA. The first published gel-based assay used a maltose-binding protein (MBP) lipoprotein fusion construct and was used to test *E. coli* LspA produced by cell-free expression.³⁴ The MBP sequence was N-terminal to the lipoprotein signal peptide, and thus after cleavage the free MBP-signal peptide was visualized at 40 kDa and the mature lipoprotein was visualized at 10 kDa on a protein gel.³⁴ However, the lipoprotein sequence, including the signal peptide sequence and lipobox sequence, were not specified, and the lipoprotein was not dagylated. As all other studies report that the lipoprotein must be dagylated for LspA cleavage,^{1,5,10} this raises concern about the validity of the assay. Additionally, the cleaved MBP was not visualized on an anti-His western blot, even though the construct had both N-terminal and C-terminal His-tags. Lastly, as it is still unknown how the lipoprotein partitions into the active site, having a tag as large as MBP on the cytoplasmic side of the signal peptide may be problematic.

Another gel-based assay was published which was used to study both *Pseudomonas aeruginosa*¹ and *Staphylococcus aureus*² LspA.^{1,2} This assay utilizes a *P. aeruginosa* lipoprotein, inhibitor of cysteine peptidase (ICP), as the substrate. *P. aeruginosa* Lgt is used in order to lipidate the signal peptide of ICP. LspA activity is assessed by the presence of the cleaved signal peptide on a protein gel (Figure 4.1).¹

For this coupled assay ICP, Lgt, and LspA must first be individually recombinantly expressed and purified. Then, preproICP, Lgt, and the lipid substrate dioleoylphosphatidylglycerol (DOPG) are incubated for one hour at 37°C in order for ICP to be dagylated resulting in proICP. Finally, LspA is added to start the reaction and the

reaction mixture is incubated at 37°C. Various protein concentrations and incubation times may be used. Lastly, the reaction is stopped by the addition of sodium dodecyl sulfate (SDS) buffer and the reactions are analyzed by polyacrylamide gel electrophoresis (SDS-PAGE). LspA activity is assessed by the presence of the cleaved signal peptide (3.3 kDa) and by a gel shift in the molecular weight of ICP as it is converted from proICP (15.9 kDa) to mature ICP (12.6 kDa) (Figure 1). The gel-based assay was used to test LspA mutants, including to identify the catalytic dyad residues of D124 and D143 for *P. aeruginosa* and D118 and D136 for *S. aureus*, as asparagine mutations were not active. The assay was also used to identify other important or essential residues for each construct.^{1,2}

One problem with this assay is that the cleaved signal peptide is small, and thus it can be challenging to image the peptide on a gel effectively. The authors addressed this problem by using InstantBlue™ gel stain which minimizes stain time and does not require destain, lowering the chance of peptide diffusion out of the gel. The gel shift from proICP to ICP is quite small as well. From the first paper the group published to the second they added a thrombin cleavage sequence (SSGLVPRGSH) N-terminal to the signal peptide sequence, increasing the molecular weight from 3.3 to 4.2 kDa.^{1,2} Furthermore, proICP, ICP, and LspA are all similar molecular weights making analysis difficult.¹

Additionally, the gel-based assay raises some other questions. First, the assay is complex using three separate recombinant membrane proteins. Each of these proteins is purified in a different buffer and is solubilized in a different detergent as protein purifications were individually optimized. ICP uses a HEPES buffer at pH 7.1 and lauryl maltose neopentyl glycol (LMNG) detergent, Lgt uses a tris buffer at pH 7.5 with LMNG detergent, *P. aeruginosa* LspA uses MES buffer at pH 6.2 with *n*-dodecylphosphocoline

(FC12) detergent, and *S. aureus* LspA uses MES buffer at pH 6.2 with *n*-dodecyl- β -D-maltoside (DDM) detergent. Thus, when the proteins are mixed during the assay the detergents and buffers also mix, changing the pH and forming FC12/LMNG or DDM/LMNG mixed micelles with unknown effects. The changing conditions could have negative effects on the protein's solubility, activity, or rates of reaction. Additionally, as all of the components of the assay are membrane proteins, it is unknown how the micelles combine to allow for preproICP/Lgt or proICP/LspA binding and release of products, further complicating the ability to determine reaction rates.²

Second, as preproICP must first be processed by Lgt to proICP it is impossible to accurately know the concentration of proICP at the start of the LspA reaction. Additionally, as the Lgt reaction is not stopped before the addition of LspA, Lgt may continue processing preproICP to proICP as the LspA reaction is occurring, making accurate measurement of proICP concentration impossible.²

Third, ICP is a lipoprotein from *P. aeruginosa* and may not be an optimal substrate for all LspA orthologs. *P. aeruginosa* LspA was shown to have three times the activity of *S. aureus* LspA.² While lipoproteins have a conserved signal and lipobox sequence across bacterial species, this lipoprotein may not be an optimal substrate for both LspA proteins and could have contributed to the differences in reaction rates.² However, lipoproteins may simply be processed at different rates due to variability in the signal sequence and lipobox, even within the same species. As only one lipoprotein was tested with the gel-based assay, additional *P. aeruginosa* lipoproteins would need to be tested with *P. aeruginosa* LspA in order to test this hypothesis.

Fourth, while the authors of this study wrote that they could not determine kinetic parameters from this assay, they did publish them in the supplement.² The kinetic parameters were derived from quantification of the gel bands using Image Lab, and these values allowed them to conclude that comparatively, *P. aeruginosa* LspA is faster (larger V) and more efficient (larger k_{cat}) than *S. aureus* LspA (Table 4.1). Using the quantitated gel bands, they also determined IC_{50} values for globomycin and myxovirescin inhibition (Table 4.2). As this is not truly a quantitative assay, and there are issues with assessing the gel bands as described above, we believe kinetics parameters cannot be determined from it, and a more quantitative assay needs to be developed. The gel-based assay is purely qualitative and can give a ‘yes or no’ answer to whether LspA is active. Thus, it can serve some utility when testing mutants or different conditions, taking into consideration the drawbacks described above, but should not be used to determine kinetic parameters.

Compared to the FRET assay described in the next section (4.1.2.3), this gel-based assay is advantageous in that it uses a native lipoprotein and is thus more comparable to physiological conditions. However, in order for an assay to be used in drug development, it must be simple, quantitative, and high throughput. This assay is none of the above, and thus a more simple, quantitative, and high-throughput assay is required.

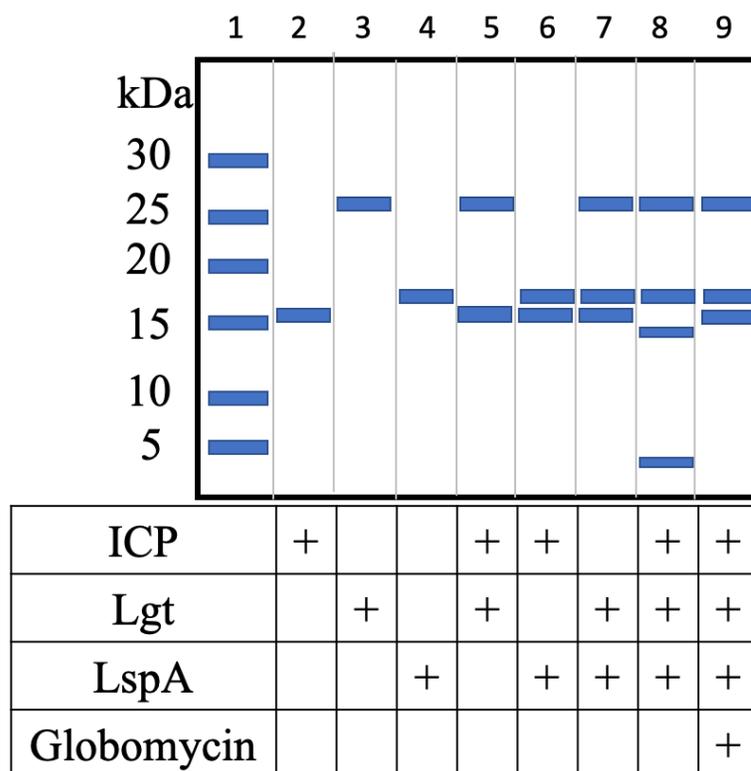


Figure 4.1. Schematic of gel-based LspA assay. After assay incubation, the reaction mixture is run on an SDS-PAGE gel and ICP, Lgt, LspA, and signal peptide bands are visualized. LspA cleavage of ICP is seen by the presence of the signal peptide at ~4 kDa and a reduction in the molecular weight of ICP (lane 8). Globomycin inhibits cleavage (lane 9). Lgt dagylation is required for LspA activity (lane 6). Gray lines between wells are shown for clarity. Molecular weight marker shown in lane 1.

4.1.2.3 Förster Resonance Energy Transfer Assay

A more quantitative Förster resonance energy transfer (FRET) assay for LspA was developed by the Wolan laboratory. This assay utilizes a peptide substrate containing an N-terminal dabsyl quencher and a C-terminal EDANS fluorophore (Figure 4.2).^{13,14} The peptide sequence between the fluorophores is VTGCAK with a dagylated cysteine, mimicking the lipoprotein lipobox which contains the conserved residue sequence [LVI][ASTVI][GAS][C].^{14,35} Before cleavage there is no fluorescence as the dabsyl quenches the fluorescence from the fluorophore. LspA activity is observed by an increase in fluorescence as the peptide is cleaved and the dabsyl quencher diffuses away from the EDANS fluorophore (Figure 4.2). This assay was used to study LspA from *Escherichia coli* and *Streptococcus pyogenes*.^{13,14} The FRET assay concept was repeated with an aminobenzoic acid fluorophore, nitro-tyrosine quencher, and LALAGCSS linker sequence with dagylated cysteine to study LspA from *Pseudomonas aeruginosa* and *Staphylococcus aureus*.²

This FRET-based LspA assay was used to determine activity for *E. coli*, *P. aeruginosa*, and *S. aureus* LspA (Table 4.1). By measuring the fluorescence at various concentrations of FRET substrate, kinetic parameters were determined. K_m values are similar for *E. coli* and *P. aeruginosa* LspA, while the k_{cat} of *P. aeruginosa* LspA was about three fold higher than *E. coli* LspA (Table 4.1).^{2,14} *S. aureus* LspA appears to be much slower and have a larger K_m than the other two orthologs (Table 4.1).²

IC_{50} values, or the concentration of inhibitor that results in half-maximal enzyme activity, were also determined for globomycin and myxovirescin inhibition of *E. coli*, *P. aeruginosa*, and *S. aureus* LspA using the FRET assay (Table 4.2). IC_{50} values were

similar for globomycin and myxovirescin for each LspA ortholog. By analyzing the ratio of IC_{50} to enzyme concentration, these results suggest that the inhibitors are less potent for *E. coli* LspA (1.2 for globomycin, 1.2 for myxovirescin), than *P. aeruginosa* LspA (0.57, 0.53) or *S. aureus* LspA (0.55, 0.50) (Table 4.2). An inhibitor library was also tested against *E. coli* LspA using the FRET assay. Out of this library the inhibitor with the most potency had an IC_{50} value of 99 nM, which is almost 10 fold less than globomycin (1.2 nM).¹⁴

While the FRET-based LspA assay is quantitative and high-throughput (can be performed in 1,536 well plates¹⁴), it also has some drawbacks. Two different FRET probes were used, one for *E. coli* LspA and the other for *P. aeruginosa* and *S. aureus* LspA. These FRET probes have different lipobox sequences which may not be processed by LspA in the same manner. Similar to the gel-based assay, the conditions for the experiment were optimized for each LspA ortholog separately. Thus, different enzyme and substrate concentrations were used for each experiment (see footnotes for Tables 4.1 and 4.2) and different buffers and solubilizing detergents were used for each ortholog. K_m and k_{cat} do not vary with enzyme concentration, but the differing buffers and detergents could play an unidentified role in the enzyme reaction rate by changing protein solubility, dynamics, and/or substrate partitioning. IC_{50} values do change with enzyme concentration, as they measure the inhibitor concentration required for 50% inhibition at that specified enzyme concentration. As different enzyme concentrations were used in these experiments, care should be taken when comparing the IC_{50} results.

The major drawback of the FRET LspA assay is that it does not correctly mimic the lipoprotein substrate. While residues of the lipobox are present, this substrate does not contain a complete signal peptide. Signal peptides are highly conserved across lipoproteins

(see Chapter 1), and prolipoprotein recognition by LspA is believed to occur due to the trigonal orientation of the signal peptide, DAG, and lipoprotein.¹ Therefore, the FRET substrate may not be recognized by LspA in the same manner as lipoproteins, or oriented properly in the active site. The FRET substrate will also partition in and out of the active site differently than physiological lipoproteins. Thus, FRET substrates may not give accurate kinetic results and may give false hits when screening for inhibitors.²

Lastly, different activity measurements and IC₅₀ values were obtained from the FRET and gel-based assays (Table 4.1 and 4.2).^{2,14} While activity measurements of k_{cat} and V were similar for *P. aeruginosa* LspA, the observed specific activity of *S. aureus* LspA was 20-fold lower by the FRET assay than the gel-based assay. Additionally, IC₅₀ values recorded by the two assays differed by three orders of magnitude.² Even taking into consideration the differing enzyme concentrations used in the two assays, this is a large difference. These discrepancies highlight the concerns of quantitating the gel-based assay, having a non-physiological substrate for the FRET assay, and having varying, uncharacterized reaction conditions for each enzyme studied. Due to these concerns we aimed to develop an assay that was both biological and quantitative.

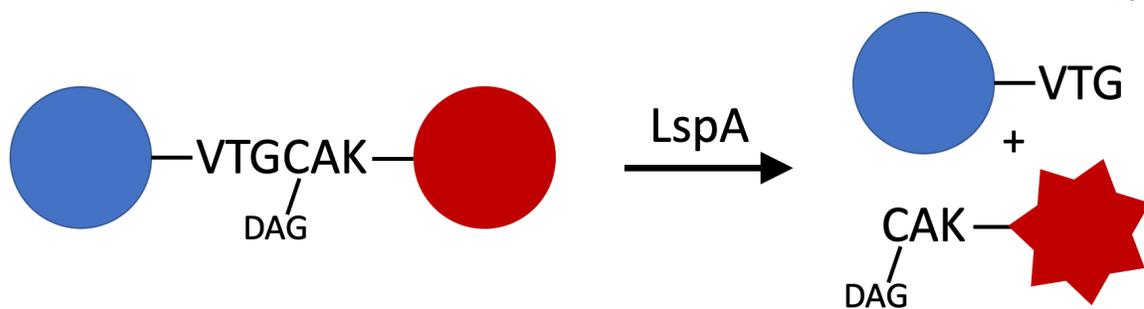


Figure 4.2. Schematic of FRET-based assay. The FRET substrate construct contains a quencher (blue) and fluorophore (red) connected by lipobox residues including the essential dagylated cysteine. Upon cleavage by LspA, the quencher separates from the fluorophore and fluorescence is observed (right). Figure adapted from Kitamura *et al.* (2018), with permission.¹⁴

Table 4.1: LspA activity values determined from assays					
	FRET Assay ^a			Gel-Based Assay ^b	
Organism	K_m (μM)	k_{cat} (10^{-3} s^{-1})	V (nmol/mg*min)	k_{cat} (10^{-3} s^{-1})	V (nmol/mg*min)
<i>E.coli</i>	14.2±4.6	10±0.1	-	-	-
<i>P. aeruginosa</i>	10	30.86	87.52	31.96	90.6
<i>S. aureus</i>	47	0.50	1.62	9.55	31.2

^aLspA and FRET substrate concentration for *E. coli* were 1 nM and 50 μM , for *P. aeruginosa* was 100 nM and 30 μM , and for *S. aureus* was 300 nM and 80 μM .^{2,14} ^bLspA concentrations for gel-shift assay were 500 nM and preproICP substrate were 12 μM .²

Table 4.2: IC ₅₀ values determined from assays								
Organism	FRET Assay ^a				Gel-Based Assay ^b			
	Globomycin		Myxovirescin		Globomycin		Myxovirescin	
	nM	Ratio ^c	nM	Ratio ^c	nM	Ratio ^c	nM	Ratio ^c
<i>E. coli</i>	1.2	1.2	1.2	1.2	-	-	-	-
<i>P. aeruginosa</i>	57	0.57	53	0.53	640	1.28	1090	2.18
<i>S. aureus</i>	167	0.55	151	0.50	170700	341	160	0.32

^aLspA and FRET substrate concentration for *E. coli* were 1 nM and 50 μ M, for *P. aeruginosa* were 100 nM and 30 μ M, and for *S. aureus* were 300 nM and 80 μ M.^{2,14} ^bLspA concentrations for gel-shift assay were 500 nM and preproICP substrate were 12 μ M.²

^cRatio of IC₅₀ to LspA enzyme concentration used in the assay.

4.2 Attempts to Optimize Activity Assay

4.2.1 Repeating Published Assays

While a quantitative biological assay is sought, the first step of this project was to repeat the published assays to compare my preparations of *P. aeruginosa* LspA to published results. However, I could never successfully repeat either the published gel-based or FRET assay.

4.2.1.1 Gel and Western Assay

I have successfully expressed and purified ICP, Lgt, and LspA from *P. aeruginosa*. The yields for each of these proteins are very low: generally ~0.75 mg/L for wild-type LspA, and ~0.2 mg/L for ICP and Lgt. Interestingly, adding an N-terminal tag to ICP such as TST-ICP (described in section 4.2.2.3) or FP-ICP (described in section 4.2.2.1) increased yields up to 1.3 mg/L. Mutations of LspA for EPR spin labeling or mutant studies (Chapter 3) had varying effects on yields with A63C increasing yields up to 1.8 mg/L and A63C/D124N/D143N almost abolishing expression. LspA preparations are pure as assessed by SDS-PAGE gels (Figure 4.3), however, the purifications of Lgt and ICP are generally less pure as other contaminant bands could be seen. Size exclusion chromatography was not performed as the yields were already low and the loss of protein would be too great. To decrease the variability and unknowns of detergent mixing and pH changes in the reaction mixture, all proteins were purified in FC12 detergent and in phosphate buffer at pH 7.2.

The LspA gel-based assay was completed by first incubating ICP, Lgt, and DOPG at 37°C for one hour. LspA was then added to the reaction mixture and the reaction was

further incubated for one hour at 37°C. The reaction was stopped by the addition of an equal volume of SDS loading buffer and the reactions were run on an SDS-PAGE gel. However, cleaved signal peptide was never observed on the gel (Figure 4.3).

Several gel types were used including 4-20% and 8-16% Tris-Glycine gels, and 16.5% Tris-Tricine gels. 8-16% Tris-Glycine and Tris-Tricine gels are better for visualizing small molecular weight bands such as the signal peptide. InstantBlue™ gel stain was used to minimize stain time and does not require destain, preventing the diffusion of peptide out of the gel.¹ Even with these precautions cleaved signal peptide was never observed.

Anti-His western blots were also completed. The anti-His antibody showed bands for all three proteins (Figure 4.4), as all have an N-terminal 6x His-tag. LspA showed the darkest band, suggesting that the antibody has a stronger affinity for this construct over the others. As the ICP His-tag is N-terminal (Figure 4.5A), after cleavage the His-tag would remain conjugated to the signal peptide and a band would show around 4 kDa while the proICP band would disappear. Thus, if the signal peptide was diffusing out of the gel and the difference in molecular weight between proICP and ICP was too small to resolve, the western blot would show the disappearance of the ICP band and indicate cleavage. However, again, signal peptide cleavage was not observed as the ICP band remained constant (Figure 4.4).

A new ICP construct was developed (hereafter called FP-ICP, Figure 4.5B), which allowed for fluorescent labeling of the signal peptide. This construct was mainly developed for a fluorescence polarization assay described in section 4.2.2.1; however, it was also used with the gel-based assay as visualization of signal peptide cleavage proved elusive with the

above described methods. FP-ICP contains a 6x His-tag, followed by a tetracysteine tag, followed by a thrombin sequence, and finally the ICP lipoprotein (Figure 4.5B). Fluorescein arsenical hairpin binder ethanedithiol (FlAsH-EDT₂) specifically binds to the tetracysteine tag (CCPGCC) and serves as the fluorescent probe.³⁶ The gel based assay was completed as described above using labeled FP-ICP as the substrate, and the gel was imaged with a fluorescence imager. Similar to analysis of the western blots, as the fluorescent probe is N-terminal to the lipoprotein, signal peptide cleavage could be assessed by the presence of the cleaved signal peptide and/or a disappearance of the FP-ICP band. As there is the addition of the tetracysteine tag and the thrombin sequence, the cleaved signal peptide molecular weight was increased to 4.7 kDa. However, neither the cleaved signal peptide, nor a decrease in molecular weight of FP-ICP, nor a decrease in fluorescence of the FP-ICP band were seen, and thus activity again was not confirmed. Unexpectedly, the fluorescence of the probe seemed to depend on dacylation of the protein. FP-ICP alone, or FP-ICP with only LspA showed less fluorescence intensity than when Lgt and DOPG were added to dacylate the substrate (Figure 4.6). Perhaps this is due to differing orientations of the signal peptide in the micelle with and without dacylation. However, as fluorescence changed throughout the reaction time, we did not feel that the assay could be reliably quantitated, even if activity was observed.

In order to have better control of the concentration of ICP substrate added to the reaction, another construct was made with a C-terminal Strep-tag® II tag (hereafter STT-ICP, Figure 4.5C). Strep-tag® II is a short 8 amino acid purification tag with the sequence WSHHPQFEK. Strep-tag® II binds strongly to Strep-Tactin® resin. Strep-Tactin® is a derivative of streptavidin and Strep-tag® II binds in the biotin binding pocket. However,

the binding affinity of Strep-tag® II to Strep-Tactin® is 100 times greater than to streptavidin. This strong binding allows for a one-step purification with low wash volumes, and yields highly purified protein with a desthiobiotin elution step.³⁷ STT-ICP was expressed and purified and has much better yields and purity than the original ICP construct. The STT-ICP protein will be used so that ICP and Lgt can be separated after the dagylation reaction (previously both were purified by His tags). Purification of dagylated ICP would allow for a known concentration of substrate to be added to the LspA assay. Quantifying the purified, dagylated ICP does assume, however, that Lgt is dagylating all of the STT-ICP that is put into the reaction mixture.

As the gel-based assay is a coupled assay, a lack of signal peptide cleavage could mean that one (or both) of the two reactions are not occurring. Either Lgt is not properly adding DAG to the preprolipoprotein, or LspA is not cleaving the substrate (or both). Lgt activity can be visualized by a gel shift upon dagylation of preprolipoprotein.³⁸ However, a gel shift of Lgt was not observed in the published gel based assay,^{1,2} nor was it observed in my assay attempts. Therefore, it is possible that the Lgt dagylation reaction is not occurring, which is not allowing the lipoprotein to be recognized by LspA. DOPG, the substrate used in the published LspA gel-based assay,^{1,2} and DPPG, the substrate seen to have most activity in the Lgt assay,³⁸ were both tried with no success (Figure 4.3).

It is concerning that activity was never observed with the gel-based assay. However, as the gel-based assay has many limitations and complications as described (section 4.1.2.2), I continued on to try the FRET assay. The FRET assay was also beneficial as it can be quantified.

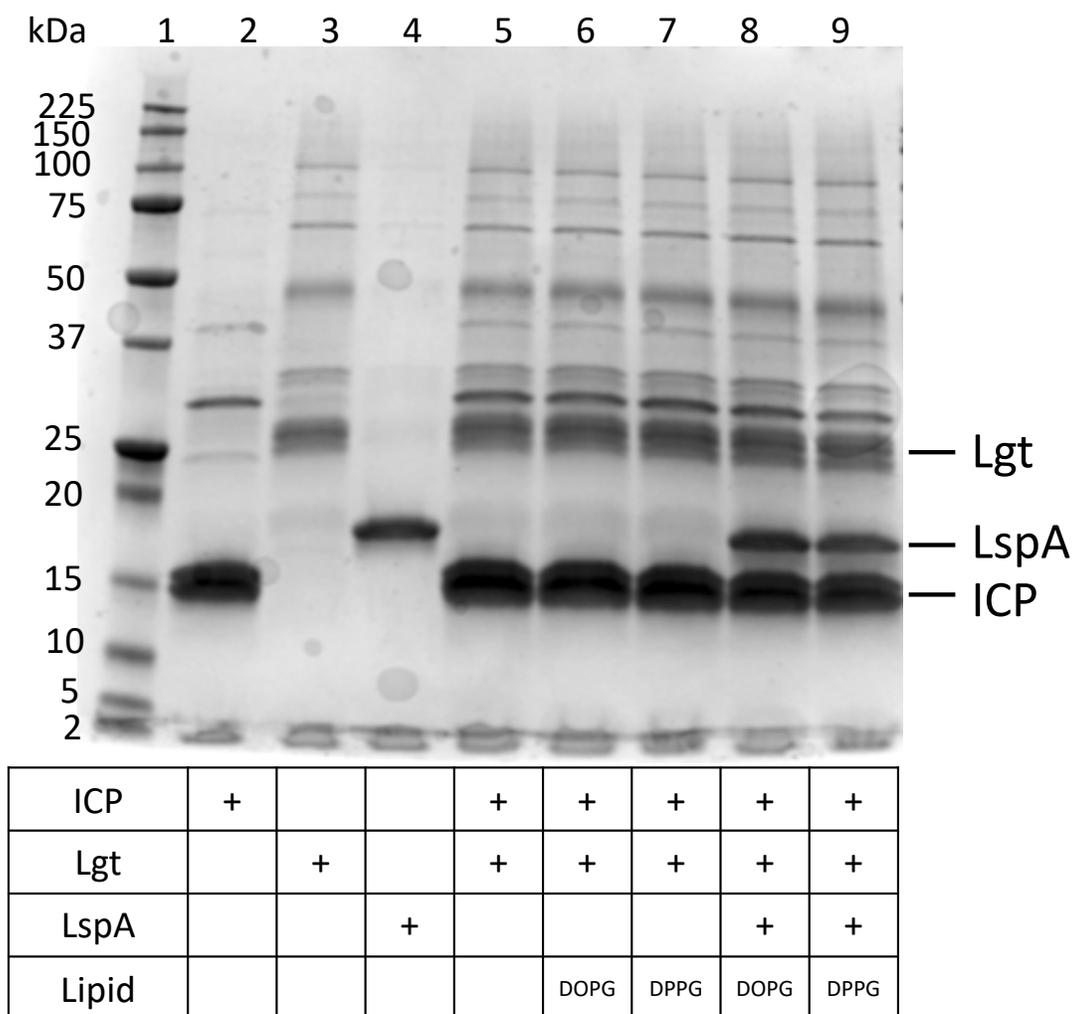


Figure 4.3. Gel-based LspA assay. LspA activity assay was completed, reactions run on an 8-16% Tris-Glycine gel, and stained with InstantBlue™ stain. Both DOPG and DPPG lipid substrates were used. Cleaved signal peptide was not observed as expected in complete reaction mixtures (lanes 8 and 9). Molecular weight marker shown in lane 1. ICP construct is STT-ICP.

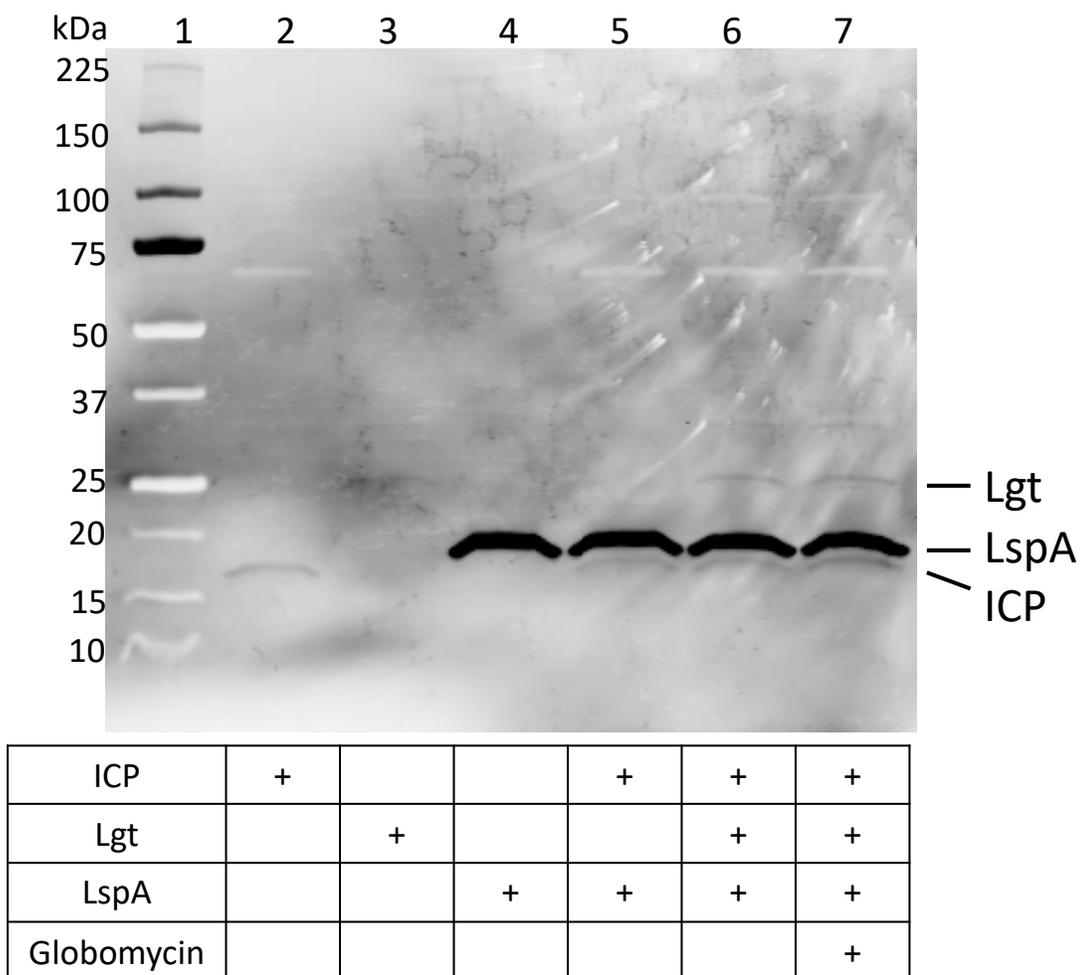


Figure 4.4. Gel-based LspA assay western blot. LspA activity assay was completed, reactions run on a 4-20% Tris-Glycine gel, transferred to a nitrocellulose membrane, and blotted with an anti-His tag antibody. ICP, Lgt, and LspA all have 6x His-tags. The cleaved signal peptide would carry the His-tag, however cleavage was not observed in complete reaction (lane 7). Molecular weight marker shown in lane 1.

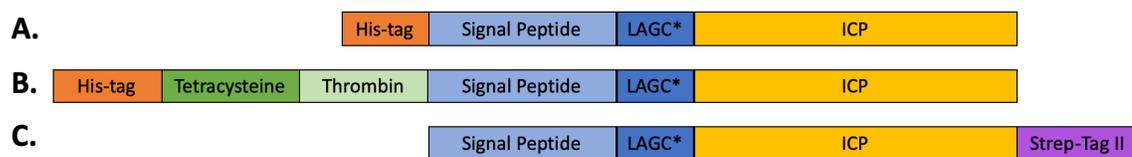


Figure 4.5. Schematic of ICP constructs. A) Original ICP construct containing N-terminal 6x-His tag. B) FP-ICP construct containing 6x-His tag, tetracysteine tag, and thrombin sequence. FLAsH fluorophore attaches to the tetracysteine tag. C) STT-ICP construct containing C-terminal Strep-Tag® II.

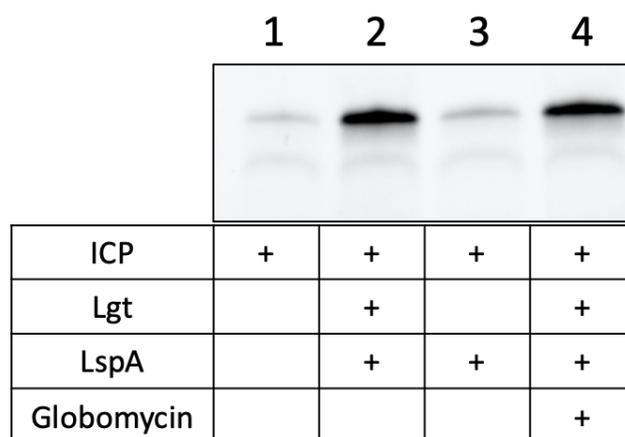


Figure 4.6. Fluorescent gel-based assay with FP-ICP construct. Only FP-ICP bands are shown. Unexpectedly, fluorescence increases when Lgt dagylates the substrate (lanes 2 and 4).

4.2.1.2 Förster Resonance Energy Transfer Assay

Förster resonance energy transfer (FRET) substrate with an N-terminal dabsyl quencher and a C-terminal EDANS fluorophore linked by the lipobox sequence VTGCAK with a dacylated cysteine^{13,14} (Figure 4.2) was generously provided by Dr. Dennis Wolan.

The published FRET assay used n-dodecyl β -D-maltoside (DDM) detergent for protein solubilization and thus DDM was contained in the reaction buffer. Since LspA, ICP, and Lgt were solubilized in dodecylphosphocoline (FC12) detergent, the first step was determining if the detergent itself had any effect on the FRET probe excitation and emission spectra. Excitation wavelengths ranged from 300 to 400 nm, and emission wavelengths ranged from 400 to 600 nm in 1 nm increments. Excitation and emission spectra were taken of the FRET probe in buffer containing DDM or FC12 before adding LspA, and after incubating with LspA (solubilized in DDM or FC12, respectively) for one hour (Figure 4.7). FRET probe in FC12 buffer showed excitation maxima at 350 nm and emission maxima at 485 nm, while FRET probe in DDM buffer showed excitation maxima at 355 nm and emission maxima at 468 nm. These values are similar to the reported emission maxima of 355 nm, but lower than the reported excitation maxima of 516 nm for this construct.¹³ Interestingly, LspA incubation with the FRET probe increased fluorescence intensity in DDM buffer as published, although to a lesser extent, while the intensity in FC12 buffer slightly decreased with LspA incubation (Figure 4.7).

Next, the fluorescence intensity of the FRET probe was measured over time at the maximal excitation and emission wavelengths after addition of LspA. As described in section 4.1.2.3, LspA activity would be observed as an increase in fluorescence as the FRET probe is cleaved and the quencher diffuses away from the fluorophore (Figure 4.2).

As an average of two measurements, the fluorescence intensity of the DDM reaction set increased over time while the fluorescence intensity of the FC12 reactions increased for about five minutes and then decreased over time (Figure 4.8A). These effects, however were negligible as seen when the blank control is subtracted (Figure 4.8B). DDM LspA showed increasing fluorescence intensity at the 5 min timepoint (Figure 4.9A). However, fitting the first 20 min did not give a linear increase in velocity as expected (Figure 4.9B).

As a quantitative, biological, universal LspA assay is desired, the FRET assay does not check all of the boxes, as described in section 4.2.1.2. Most importantly, the FRET assay is not biological as it does not contain the lipoprotein signal peptide. While in theory the FRET assay is quantitative and has the potential to be high throughput, it could not be repeated in my hands even in the same buffer conditions. Having different fluorescent intensities in different detergents is also an issue, as this does not make the assay universal for the study of different LspA orthologs that may be purified in different detergents. Differing fluorescence intensities also raises questions about how the FRET probe behaves in different detergent environments. The probe must partition into the micelle that LspA is in to be cleaved. Thus, if the probe behaves differently in different detergents, the assessment of LspA activity will be impacted. These concerns not only make the FRET assay undesirable, but it shows that the kinetic parameters derived from it are not solely due to the cleavage of the probe by LspA, and that the partitioning rate into and out of the LspA active site and micelle must also be considered. Of note, the Wolan group who developed this assay also saw no activity as measured by the FRET assay when they used FC12 detergents, either with the *E. coli* LspA that they expressed and purified (as was used in their publications^{13,14}) or with *P. aeruginosa* LspA that I sent to them (personal

communication with Dr. Dennis Wolan and Dr. Seiya Kitamura). As this raised major concerns, we set out to develop our own LspA assay that would be biological, quantitative, reproducible, and universal.

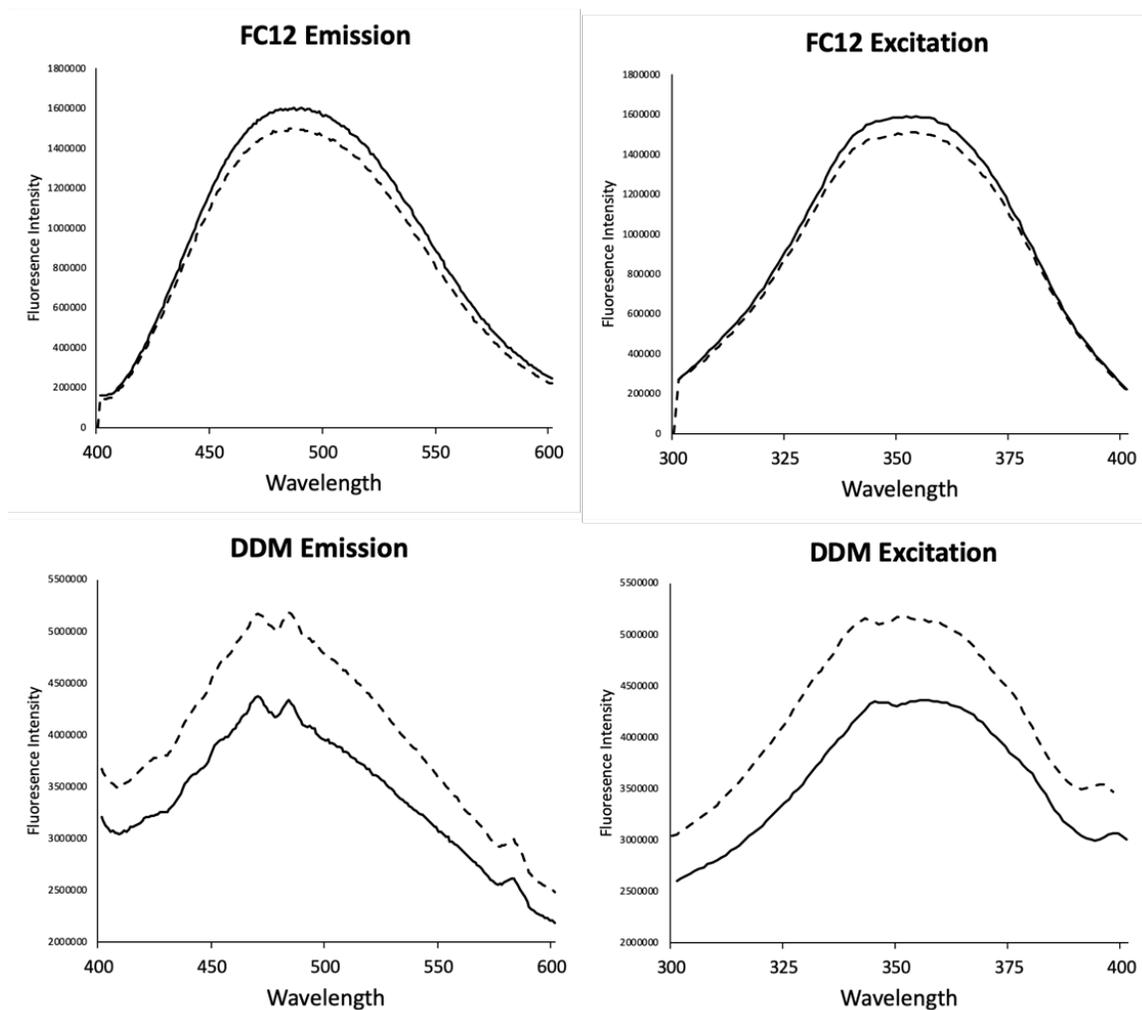


Figure 4.7. Emission and excitation spectra of FRET probe. Excitation wavelengths (right) ranged from 300 to 400 nm, and emission wavelengths (left) ranged from 400 to 600 nm in 1 nm increments. Measurements were taken in FC12 buffer (top) and DDM buffer (bottom), and were recorded prior to LspA addition (solid line), and after incubation with LspA for one hour (dotted line).

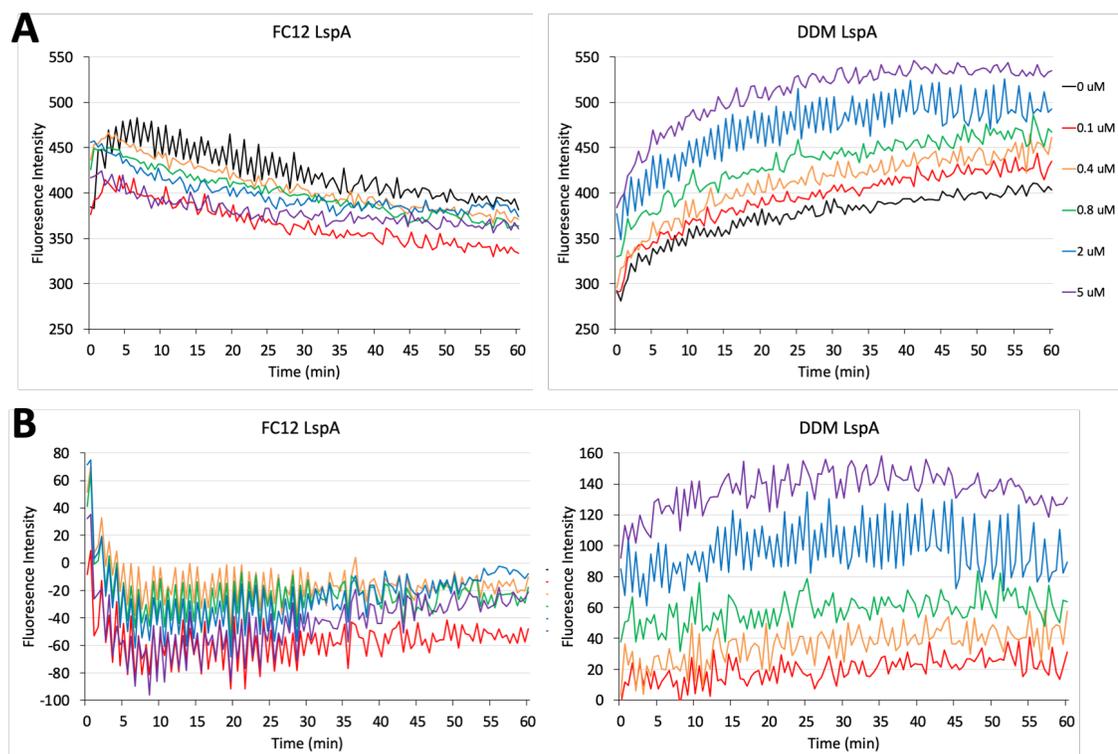


Figure 4.8. FRET assay in FC12 and DDM. FRET fluorescence intensity was measured every 30 seconds for one hour after LspA addition. LspA concentrations were 0 (black), 0.1 μM (red), 0.4 μM (yellow), 0.8 μM (green), 2 μM (blue), and 5 μM (purple) and reactions were carried out for FC12 (left) and DDM (right) solubilized protein. An average of two measurements is shown for raw data (A) and blank control subtracted (B).

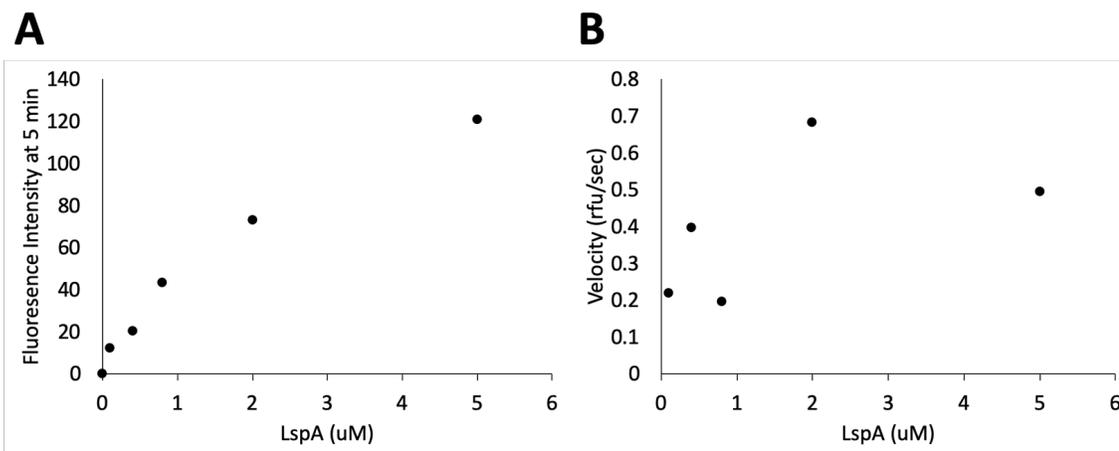


Figure 4.9. Analyzing DDM FRET assay. A) Fluorescence intensity increases at the 5 min time point with increasing LspA concentration in DDM. B) Velocity as determined by fitting first 20 min of control subtracted spectra do not show a linear increasing trend with LspA concentration.

4.2.2 Developing a New LspA Assay

As the gel-based assay and FRET assay are not biological, quantitative, or reproducible, we sought to develop a novel LspA assay that would satisfy those criteria.

4.2.2.1 Fluorescence Polarization Assay

Development of a fluorescence polarization (FP) assay was attempted to quantitatively assess LspA cleavage of a lipoprotein signal peptide as has been done for other proteases.³⁹⁻⁴¹ An FP-ICP construct was developed in which a tetracysteine motif (CCPGCC) was added to the N-terminus of the lipoprotein ICP (Figure 4.5B). This motif binds the small fluorogenic biarsenical FIAsh-EDT₂ with high affinity, thus creating a fluorescent probe at the cytoplasmic side of the signal peptide (Figure 4.10).^{36,42-44} FIAsh-EDT₂ is added in equimolar ratio to FP-ICP and nutated overnight at room temperature. The next day, Lgt and DOPG are added for one hour at 37°C so that the essential DAG is added to the cysteine of the FP-ICP lipobox. The reaction is started by adding LspA and is incubated at 37°C for one hour (Figure 4.11). FP is monitored at each step of the reaction.

Upon signal peptide cleavage by LspA it is expected that the anisotropy of the probe will decrease due to a decrease in the molecular weight and thus an increase in the rate of rotation of the probe.⁴⁰⁻⁴² The presence of the inhibitor globomycin will ablate cleavage such that the anisotropy of the probe will not change. The FP assay developed here is expected to better assess LspA activity than the FRET-based assay because the FP-ICP substrate is more biological than the synthetically made peptide FRET probe. In particular, the FP probe contains the signal peptide of the lipoprotein. As signal peptides are highly conserved, it is likely that they are important for LspA recognition.^{7,8,35} It is important for

the assay to be biologically relevant as it will be used to probe the effectiveness of antibiotic drug candidates.

Characterization of the preproICP bound FAsH probe emission and excitation spectra shows a maximum excitation at 288 nm and a maximum emission at 348 nm (Figure 4.12). The fluorescence intensity increases when Lgt and DOPG are added to produce proICP, and again when LspA is added to produce mature ICP (Figure 4.13A), but the maximal excitation and emission wavelengths do not change (Figure 4.12). An increase in the FAsH fluorescence between labeled preproICP and proICP was also observed when this construct was used with the gel-based assay (Figure 4.6). The changes in fluorescence intensity are a problem because in order to directly compare reaction anisotropies, as in this assay, it is assumed that the fluorescence intensity does not change throughout the reaction.

It was expected that the anisotropy of the FAsH probe would not change upon conversion of preproICP to proICP by dagylation, but would decrease upon signal peptide cleavage by LspA. Contrarily, anisotropy decreased when Lgt and DOPG were added, and again when LspA was added (Figure 4.13B). These unexpected results may be due to the varying fluorescence intensity throughout the reaction. Additionally, the anisotropy and fluorescence values varied greatly between reaction trails (data not shown), and thus a different assay approach was pursued.

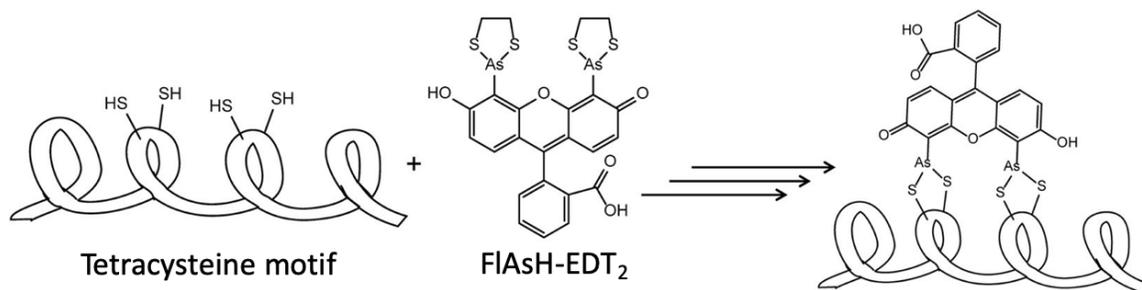


Figure 4.10 FIAsh-EDT₂ coordination. FIAsh (middle) binds with high affinity to the tetracysteine motif CCPGCC (left) to create a fluorescent probe on the protein of interest (right). Figure adapted from Fernandes *et al.* (2017), with permission.³⁶

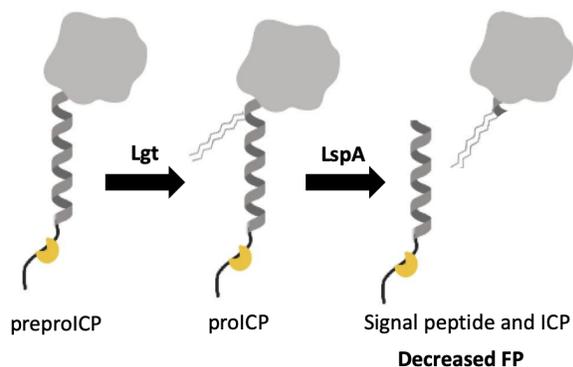


Figure 4.11. Schematic of FP assay. FAsH (yellow) coordinates to the tetracysteine motif on the N-terminus of FP-ICP (gray). In the first step, Lgt dagylates preproICP to proICP. In the second step LspA cleaves the signal peptide. Activity is assessed by a decrease in fluorescence polarization (FP) of the FAsH fluorophore.

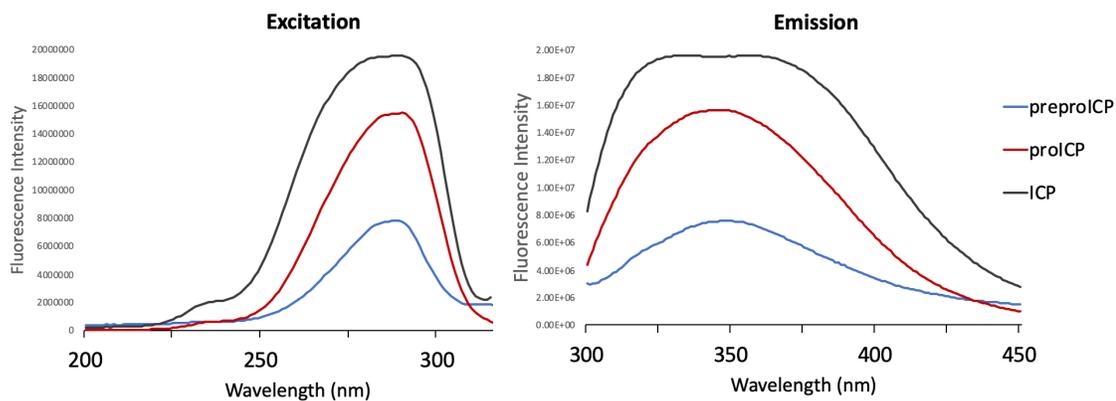


Figure 4.12. Emission and excitation spectra of FP probe. Excitation wavelengths (right) ranged from 200 to 320 nm, and emission wavelengths (left) ranged from 300 to 450 nm in 1 nm increments. Spectra were taken at each step of the reaction: FIAsh labeled preproICP (blue), dagylated proICP (red), and after LspA incubation (black).

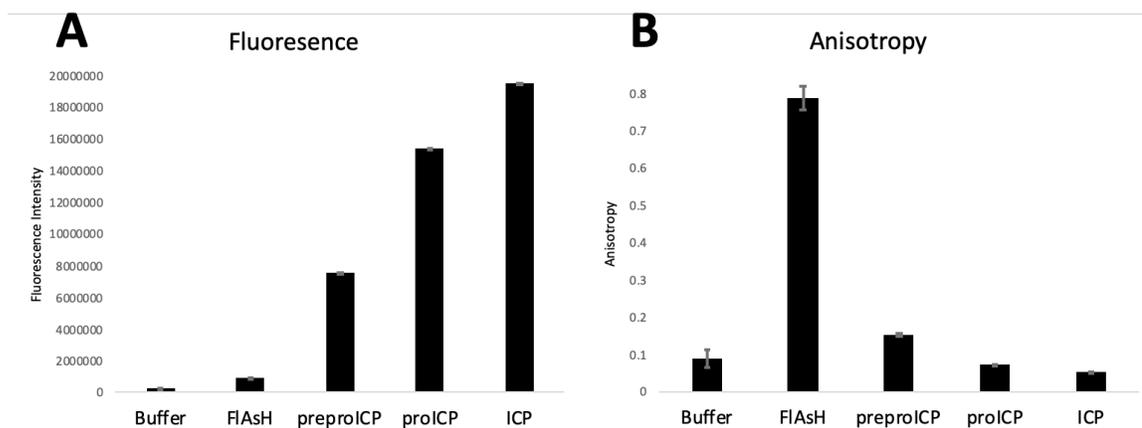


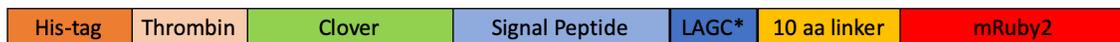
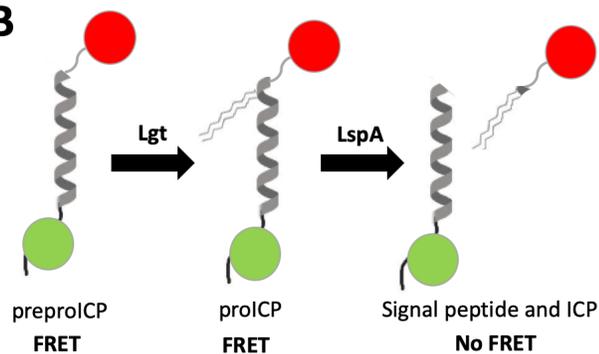
Figure 4.13. FP assay results. A) Fluorescence intensity was seen to increase throughout the reaction, while B) anisotropy was seen to decrease.

4.2.2.2 Förster Resonance Energy Transfer Assay with Signal Peptide

Concurrent to the development of the fluorescence polarization assay described above, I also worked to develop a Förster resonance energy transfer (FRET) assay. Contrary to the published FRET assay described in section 4.1.2.3 and 4.2.1.2, this FRET assay would be more biological as it would contain the signal peptide. The substrate consists of the ICP signal peptide with fluorophores on both termini (Figure 4.14A). In the first step, Lgt would be used to ligate the preprolipoprotein to proICP. Upon LspA cleavage, the fluorophores would diffuse away from one another and the measured FRET would decrease (Figure 4.14B).

Fluorophores mRuby2 and Clover were chosen due to their high degree of donor emission and acceptor absorbance overlap, leading to a high Förster radius (r_0) of 6.3 nm.⁴⁵ The r_0 is the distance between fluorophores at which energy transfer is 50% efficient. It was important to choose a fluorophore pair with a long r_0 as they reside on opposite sides of the ICP signal peptide which is around 4 nm, and span the micelle. The effect, if any, of the micelle detergents on the fluorophores is unknown. Thus, Clover and mRuby2 were also chosen because they are relatively bright and photostable compared to other commonly used fluorophores.⁴⁵

The ICP-FRET construct and the Clover construct were doubly digested with EcoRI and Bam HI. ICP-FRET and mRuby2 were doubly digested with HincII and HindIII. The digested DNA was purified using a DNA cleanup kit (New England Biolabs). Subsequently, the 5' end of the vector (ICP-FRET) was dephosphorylated, and the insert (Clover or mRuby2) phosphorylated. Unfortunately, ligation via an instant sticky-end ligase or quick ligation (New England Biolabs) was never successful.

A**B**

4.14. Schematic of new FRET assay. A) Planned FRET construct with Clover (green) and mRuby2 (red) fluorophores. B) Fluorophore conjugated ICP construct (Clover, green; mRuby2, red; ICP signal peptide, gray) is diglylated by Lgt. LspA cleavage activity is assessed by a decrease in FRET.

4.2.2.3 Electron Paramagnetic Resonance Assay

As fluorescent probes, in either FP or FRET forms, did not yield good results for a LspA activity assay, an electron paramagnetic resonance (EPR) assay was pursued. This assay would involve spin labeling the ICP construct on either end of the signal peptide and dagylating with Lgt. LspA activity would be observed by a decrease in double electron electron resonance (DEER) signal after incubation with LspA. As described in Chapters 2 and 3, DEER allows for the determination of a distance distribution between two spin labels. Prior to completing the assay, DEER analysis will show the distance between the spin labeled sites. As the signal peptide is around 4 nm, this falls well within the distance range probed by DEER (generally 1.8 nm to around 6 nm depending on quality of data, see Chapter 2). When the ICP construct is cleaved, the spin labels will diffuse away from one another and a DEER signal will not be observed (Figure 4.15).

This assay is biological as it contains the dagylated signal peptide. It is also semi-quantitative as the amount of DEER signal can be determined and would correlate with the percentage of ICP that has not been cleaved. As the DEER samples are flash frozen, time points can be taken during the reaction. A major complication, however, is that spin labels are normally conjugated to cysteine residues in the protein. For this protocol, any native cysteine residues must first be removed. As a cysteine residue is the essential amino acid in the lipoprotein lipobox which gets dagylated, all native cysteine residues cannot be removed. Therefore, incorporation of unnatural amino acids (UAA) and subsequent spin labeling will be used.⁴⁶

During protein translation, transfer RNAs (tRNAs) transport amino acids to the ribosome. The anticodon of the tRNA matches with the complementary codon of the

mRNA and allows for the correct amino acid to be added to the growing peptide. For this to occur, the tRNA must be “charged” or bound to its correct amino acid. Aminoacyl-tRNA synthetase (aaRS) catalyzes the binding between a specific tRNA and the correct amino acid.⁴⁷ UAAs can be recombinantly incorporated into *E. coli* expressed proteins by introducing an orthogonal tRNA and aaRS pair that delivers the desired unnatural amino acid in response to a codon which does not encode any of the 20 natural amino acids.⁴⁸

As the tRNA must not recognize any codons which code for the 20 natural amino acids, a stop codon is used to code for the UAA. The amber stop codon (TAG) is most commonly used for UAA incorporation as it is the least frequent stop codon in *E. coli* (~9% of genes, none of which are essential), and therefore should not affect bacterial growth. It is thus essential that the protein of interest not use this stop codon. The aaRS must also not aminoacylate any endogenous *E. coli* tRNAs. Therefore, a tRNA/aaRS pair is generally used from a different organism.⁴⁹ Furthermore, studies have shown that inducible control of the aaRS allows for an increased yield of recombinant proteins containing the UAA.⁵⁰ Hence, to express proteins with UAA, two plasmids must be co-expressed. The first encodes the protein of interest with TAG codon(s) at the desired site(s), and the second encodes for the orthogonal tRNA/aaRS pair. Both plasmids are under inducible expression and are selected for by different antibiotics (Figure 4.16).⁴⁶ In this way more than 40 UAAs have been genetically incorporated into recombinant proteins.⁵¹

For this EPR based assay, the incorporation of the UAA p-acetylphenylalanine (pAcF) is desired. pAcF contains a keto functional group that can be reacted with a hydroxylamine reagent to generate a nitroxide spin label termed K1.⁵¹ The pEVOL plasmid will be used which has been developed to contain the tRNA/aaRS pair for pAcF

incorporation (Figure 4.16).⁵⁰ TAG mutations have been made at specific sites on either side of the ICP signal peptide. Importantly, the TST-ICP construct was used in this case because it contains a C-terminal purification tag. C-terminal purification is important to ensure that the full protein has been synthesized and that it is not truncated at the TAG site(s).⁴⁶ The constructs for this assay have been made, however the assay is still in development.

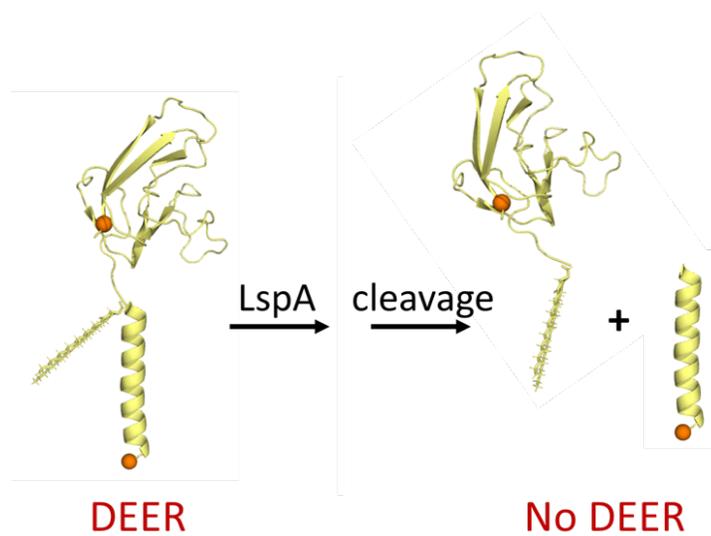


Figure 4.15. Schematic of DEER assay. Spin labels (orange) will be placed on either side of the ICP signal peptide (yellow). After LspA cleavage the cleaved substrate will diffuse away from the signal peptide and a loss of DEER signal will be observed.

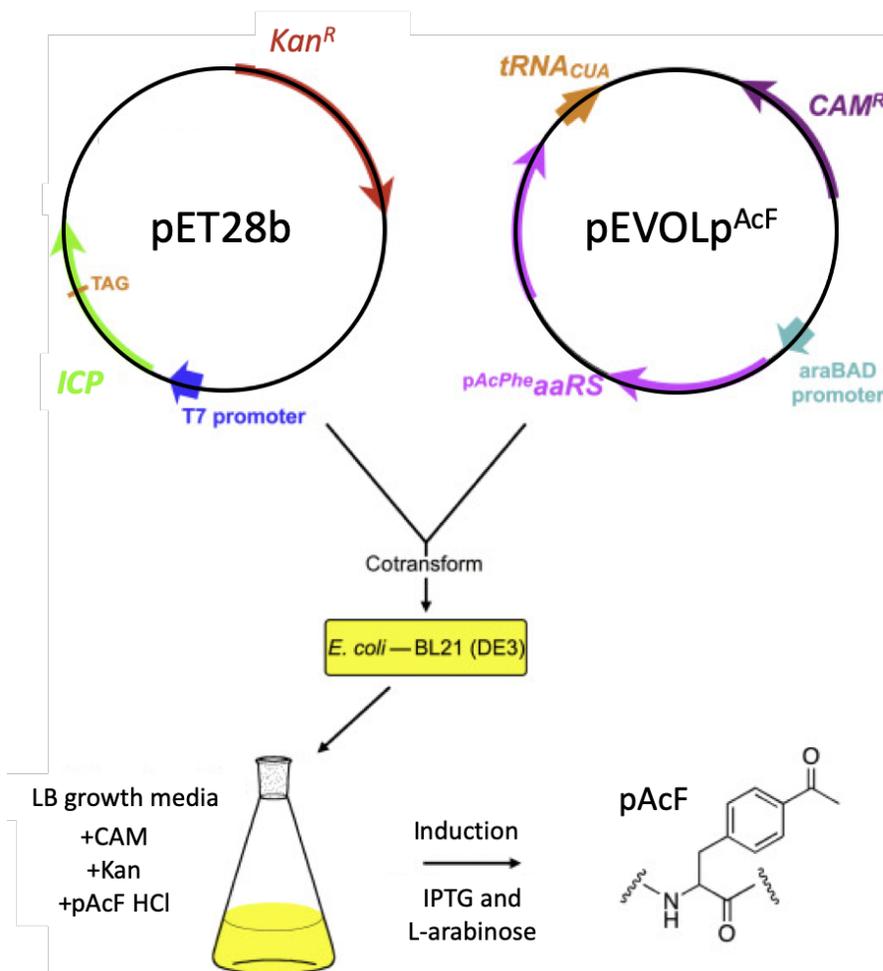


Figure 4.16. Unnatural amino acid incorporation. To incorporate the unnatural amino acid pAcF, a pET28b plasmid containing ICP and the pEVOL plasmid containing the tRNA/aaRS pair are coexpressed in *E. coli*, selected for by kanamycin (Kan) and chloramphenicol (CAM) antibiotics, and induced with IPTG and L-arabinose. Adapted from Evans and Millhauser (2015), with permission.⁴⁶

4.3 Conclusions

LspA is a promising antibiotic drug target. In order for therapeutics to be developed to target LspA, an activity assay must be developed to test the effectiveness of the compounds. This assay must be biological, quantitative, reproducible, and high throughput. So far, a gel-based assay and a FRET assay have been published for LspA. The gel-based assay is biological, but it is not quantitative. While the FRET assay is potentially quantitative and high-throughput, it is not biological. Unfortunately, neither published assay was able to be reproduced, raising concerns about this system as a whole.

Here, new FP, FRET, and EPR based assays were introduced. The FP assay is not promising as the fluorescence intensity is seen to change throughout the reaction. Both the FRET and EPR assays are still in the development stages. Both are biological and quantitative, but they are not particularly high throughput. Both the FRET and EPR assays are complex, requiring three recombinantly expressed proteins, and efficient dacylation of LspA by Lgt, before the LspA reaction can be assessed. The EPR assay is even more challenging due to the need for unnatural amino acid incorporation, spin labeling, and DEER spectroscopy. Thus, while these assays will continue to be developed, a more high throughput assay is desired and will continue to be investigated.

4.4 Methods

4.4.1 Protein Expression and Purification

The *Pseudomonas aeruginosa* (strain PAO1) *lspa* and *lgt* genes were purchased in a pET28b vector with an N-terminal 6x His-tag and thrombin cleavage sequence (General Biosystems). Genes coding for *P. aeruginosa* ICP were purchased in a pET28b vector with

an N-terminal 6x His-tag (ICP); N-terminal 6x His-tag, tetracysteine motif, and thrombin cleavage sequence (FP-ICP); and C-terminal Strep-Tag II (STT-ICP) (General Biosystems Inc.) (Figure 4.5).

The plasmids were transformed into C41(DE3) *E. coli* cells (Lucigen) and cultures were grown in Luria-Burtani (LB) media supplemented with 50 mg/L kanamycin to an OD₆₀₀ around 0.8. Expression was induced with 1 mM isopropyl- β -thio-D-galactoside (IPTG) for around 18 hours at 25°C. Cells were harvested by centrifugation at 5,000g for 10 min at 4 °C and the pellets were frozen at -20 °C.

Cells were resuspended in Buffer A (20 mM phosphate, pH 7.2, 150 mM NaCl) with one Complete Protease Inhibitor Cocktail tablet (Roche) and passed twice through a high-presser Nano DeBEE homogenizer (BEE International). Cell debris were removed via centrifugation at 18,000g for 30 min. The membrane fraction was subsequently separated by ultra-centrifugation at 150,000g for 1 hour at 4 °C. The membrane pellet was resuspended in 30 mL Buffer A and frozen at -80 °C. The membrane resuspension was thawed and 10 mL Buffer A containing fos choline-12 (FC12) was added such that the final concentration of FC12 was 1.8% (w/v), and was allowed to rock at 4 °C for at least one hour. Unsolubilized material was then removed by ultracentrifugation at 100,000g for 45 min. Recombinant protein was bound to a Ni²⁺ immobilized metal affinity chromatography column, and washed with 40 mL Buffer A containing 40 mM imidazole and 0.14% (w/v) FC12. The protein was eluted in 20 mL Buffer A containing 300 mM imidazole and 0.14% (w/v) FC12. Imidazole was removed by dialysis in 4 L Buffer A three times for one hour each, and the protein was concentrated using a 10 kDa molecular weight

cutoff concentrator (Millipore). Protein purity was assessed by SDS-PAGE and the protein identity was confirmed with MALDI-TOF mass spectrometry.

4.4.2 Gel-Based Assay

The gel based assay was completed as previously described.^{1,2} In the first step 1.5-10 μM Lgt, 10-50 μM ICP (ICP or FP-ICP), and 200 lipid (DOPG or DPPG) were incubated for one hour at 37 °C. Then, 2.5-10 μM LspA was added and allowed to incubate for another hour at 37 °C. The reaction was stopped by the addition of an equal volume of SDS buffer (final reaction volumes were 15-100 μL). The reactions were then run on a gel according to the recommended settings. 4-20% and 8-16% Tris-Glycine gels, and 16.5% Tris-Tricine gels were used (Bio Rad). The gels were stained with InstantBlue™ (Sigma) or Coomassie blue stain and imaged with a Gel Doc EZ system (Bio Rad).

For western blots, a protein gel was run in the same manner and the protein was transferred to a nitrocellulose membrane via a Trans-Blot® Turbo™ system (Bio Rad). The membrane was blocked with a 5% (w/v) dry milk TBS-T blocking buffer for two hours, and then the primary anti-His antibody (ThermoFisher) was allowed to incubate overnight at a 1:1000 dilution. The next day, the membrane was washed three times in TBS-T buffer for 15 minutes each and the secondary goat anti-mouse IRDye 800CW (LI-COR) antibody was added and incubated for one hour. The membrane was then imaged with an Odyssey® imaging system (LI-COR).

4.4.3 FRET Assay

The FRET assay was completed as previously described with FRET substrate generously provided by Dr. Denis Wolan.^{13,14} FRET probe excitation and emission spectra were first acquired. Excitation wavelengths ranged from 300 to 400 nm, and emission wavelengths ranged from 400 to 600 nm in 1 nm increments. These spectra were taken of the FRET probe in both DDM and FC12 buffers before adding LspA, and after incubating with LspA for one hour (Figure 4.7). FRET probe in FC12 buffer showed excitation maxima at 350 nm and emission maxima at 485 nm, while FRET probe in DDM buffer showed excitation maxima at 355 nm and emission maxima at 468 nm. These maximal excitation and emission values were used when completing the assay.

For the assay, 0-5 μM LspA (solubilized in 0.14% FC12 or DDM) and 50 μM FRET substrate were added in either FC12 or DDM buffers to a black 96-well plate with a final reaction volume of 200 μL . The assay fluorescence was then recorded at the maximal excitation and emission values on a SpectraMax M5 microplate reader (Molecular Devices).

FP assay. For the newly developed FP assay, FP-ICP, Lgt, and LspA were expressed and purified as described above. Equimolar FIAsh fluorophore (Cayman Chemical) was added to the purified FP-ICP and allowed to incubate in the dark at 4 °C overnight.

FP measurements were taken on a FluoroMax3 fluorimeter (Horiba). FP probe excitation and emission spectra were first acquired. Excitation wavelengths ranged from 200 to 320 nm, and emission wavelengths ranged from 300 to 450 nm in 1 nm increments. For the assay, combinations of 10 μM FIAsh labeled FP-ICP, 5 μM Lgt, 150 μM DOPG, and 6 μM LspA were mixed and incubated for 1 hr at 37 °C. Fluorescence and anisotropy

measurements were taken in triplicate at the maximum excitation (288 nm) and maximum emission (348 nm) wavelengths with a slit of 2 nm.

4.4.4 New FRET Assay

Fluorophore constructs (Clover and mRuby2) were purchased in pBAD vectors (Addgene). An ICP-FRET construct with an N-terminal 6x His-tag and thrombin sequence and a C-terminal extension containing the first 20 amino acids of mRuby2 ending at the HincII cut site was purchased (General Biosystems). This extension was required for cloning purposes. Clover, mRuby2, and the ICP construct were amplified and sequencing was confirmed (Genewiz). Double digestion of Clover and the ICP-FRET or mRuby2 and ICP-FRET was completed with EcoRI-HF and BamHI-HF, or HindIII and HincII, respectively, using the standard protocol (New England BioLabs). The cleaved DNA was purified via gel extraction (QIAquick Gel Extraction Kit, Qiagen) or DNA cleanup (Monarch® PCR & DNA Cleanup Kit, New England BioLabs). The vector was dephosphorylated with a Quick Desphosphylation Kit (New England BioLabs), and the insert was phosphorylated via T4 polynucleotide kinase (New England BioLabs). Ligation was carried out using an Instant Sticky-End Ligase Master Mix for the Clover/FRET-ICP pair (New England BioLabs) or a Quick Ligation Kit for the mRuby2/FRET-ICP pair (New England BioLabs) and transformed. Unfortunately, ligation was never successful.

4.4.5 EPR Assay

ICP was purchased in a pET28b vector with a C-terminal Strep-tag® II tag (Figure 4.5C) (General Biosystems). TAG mutations were made via PIPE mutagenesis⁵² or a

QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent). Sites before (after amino acid 11, X11) and after (X38) the signal peptide were chosen as spin labeling sites as well as one further in the ICP sequence (X143).

The plasmids were transformed into C41(DE3) *E. coli* cells (Lucigen) and cultures were grown in Luria-Burtani (LB) media supplemented with 50 mg/L kanamycin to an OD₆₀₀ around 0.8. Expression was induced with 1 mM isopropyl- β -thio-D-galactoside (IPTG) for around 18 hours at 25°C. Cells were harvested by centrifugation at 5,000g for 10 min at 4 °C and the pellets were frozen at -20 °C.

For purification, cells were resuspended in Buffer A (20 mM phosphate, pH 7.2, 150 mM NaCl) with one Complete Protease Inhibitor Cocktail tablet (Roche) and passed twice through a high-presser Nano DeBEE homogenizer (BEE International). Cell debris were removed via centrifugation at 18,000g for 30 min. The membrane fraction was subsequently separated by ultra-centrifugation at 150,000g for 1 hour at 4 °C. The membrane pellet was resuspended in 30 mL Buffer A and frozen at -80 °C. The membrane resuspension was thawed and 10 mL Buffer A containing fos choline-12 (FC12) was added such that the final concentration of FC12 was 1.8% (w/v), and was allowed to rock at 4 °C for at least one hour. Unsolubilized material was then removed by ultracentrifugation at 100,000g for 45 min.

Recombinant protein was bound to a Strep-Tactin® chromatography column (iba Lifesciences), and washed with 6 x 2 mL Buffer A containing 0.14% (w/v) FC12. The protein was eluted in 6 x 1 mL Buffer A containing 2.5 mM desthiobiotin (iba Lifesciences) and 0.14% (w/v) FC12. Desthiobiotin was removed by dialysis in 4 L Buffer A three times

for one hour each, and the protein was concentrated using a 10 kDa molecular weight cutoff concentrator (Millipore). Protein purity was assessed by SDS-PAGE.

The pEVOL pAcF plasmid was generously provided by the laboratory of Dr. Wayne Hubbell. The next step will be to co-transform the pEVOL pAcF and pET28b ICP (with TAG mutations) plasmids to be able to express ICP with pAcF unnatural amino acids at the mutation sites.

4.5 References

- 1 Vogeley L, El Arnaout T, Bailey J, Stansfeld PJ, Boland C, Caffrey M. Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. *Science (80-)* 2016;**351**:876–80.
<https://doi.org/10.1126/science.aad3747>.
- 2 Olatunji S, Yu X, Bailey J, Huang C, Zapotoczna M, Bowen K, *et al.* Structures of lipoprotein signal peptidase II from *Staphylococcus aureus* complexed with antibiotics globomycin and myxovirescin 2020. <https://doi.org/10.1038/s41467-019-13724-y>.
- 3 Novak P, Dev IK. Degradation of a signal peptide by protease IV and oligopeptidase A. *J Bacteriol* 1988;**170**:5067–75.
<https://doi.org/10.1128/jb.170.11.5067-5075.1988>.
- 4 Tokunaga M, Tokunaga H, Wu HC. Post-translational modification and processing of *Escherichia coli* prolipoprotein in vitro. *Proc Natl Acad Sci* 1982;**79**:2255–9.
<https://doi.org/10.1073/pnas.79.7.2255>.
- 5 Tokunaga M, Loranger JM, Wu HC. Prolipoprotein modification and processing

- enzymes in *Escherichia coli*. *J Biol Chem* 1984;**259**:3825–30.
- 6 Hussain M, Ichihara S, Mizushima S. Mechanism of Signal Peptide Cleavage in the Biosynthesis of the Major Lipoprotein of the *Escherichia coli* Outer Membrane. *J Biol Chem* 1982;**257**:5177–82.
- 7 Pollitt S, Inouye S, Inouye M. Effect of amino acid substitutions at the signal peptide cleavage site of the *Escherichia coli* major outer membrane lipoprotein. *J Biol Chem* 1986;**261**:1835–7.
- 8 Wu HC, Tokunaga M, Tokunaga H, Hayashi S, Giam C-Z. Posttranslational modification and processing of membrane lipoproteins in bacteria. *J Cell Biochem* 1983;**22**:161–71. <https://doi.org/10.1002/jcb.240220305>.
- 9 Hussain M, Ichihara S, Mizushima S. Accumulation of glyceride-containing precursor of the outer membrane lipoprotein in the cytoplasmic membrane of *Escherichia coli* treated with globomycin. *J Biol Chem* 1980;**255**:3707–12.
- 10 Inouye S, Franceschini T, Sato M, Itakura K, Inouye M. Prolipoprotein signal peptidase of *Escherichia coli* requires a cysteine residue at the cleavage site. *EMBO J* 1983;**2**:87–91. <https://doi.org/10.1002/j.1460-2075.1983.tb01386.x>.
- 11 Yamagata H. Temperature-sensitive prolipoprotein signal peptidase in an *Escherichia coli* mutant: use of the mutant for an efficient and convenient assay system. *J Biochem* 1983;**93**:1509–15.
- 12 Inukai M, Takeuchi M, Shimizu K, Arai M. Mechanism of action of globomycin. *J Antibiot (Tokyo)* 1978;**31**:1203–5. <https://doi.org/10.7164/antibiotics.31.1203>.
- 13 Kitamura S, Wolan DW. Probing substrate recognition of bacterial lipoprotein signal peptidase using FRET reporters. *FEBS Lett* 2018.

<https://doi.org/10.1002/1873-3468.13155>.

- 14 Kitamura S, Owensby A, Wall D, Wolan DW. Lipoprotein Signal Peptidase Inhibitors with Antibiotic Properties Identified through Design of a Robust In Vitro HT Platform. *Cell Chem Biol* 2018.
<https://doi.org/10.1016/j.chembiol.2017.12.011>.
- 15 Hayashi S, Chang SY, Chang S, Giam CZ, Wu HC. Modification and processing of internalized signal sequences of prolipoprotein in *Escherichia coli* and in *Bacillus subtilis*. *J Biol Chem* 1985;**260**:5753–9.
- 16 Isaki L, Kawakami M, Beers R, Hom R, Wu HC. Cloning and nucleotide sequence of the *Enterobacter aerogenes* signal peptidase II (*lsp*) gene. *J Bacteriol* 1990;**172**:469–72. <https://doi.org/10.1128/jb.172.1.469-472.1990>.
- 17 Paitan Y, Orr E, Ron EZ, Rosenberg E. A Nonessential Signal Peptidase II (*Lsp*) of *Myxococcus xanthus* Might be Involved in Biosynthesis of the Polyketide Antibiotic TA. *J Bacteriol* 1999;**181**:5644–51. [https://doi.org/10.1016/S0378-1097\(98\)00542-4](https://doi.org/10.1016/S0378-1097(98)00542-4).
- 18 Xiao Y, Wall D. Genetic redundancy, proximity, and functionality of *lspA*, the target of antibiotic TA, in the *Myxococcus xanthus* producer strain. *J Bacteriol* 2014;**196**:1174–83. <https://doi.org/10.1128/JB.01361-13>.
- 19 Geukens N, De Buck E, Meyen E, Maes L, Vranckx L, Van Mellaert L, *et al*. The type II signal peptidase of *Legionella pneumophila*. *Res Microbiol* 2006;**157**:836–41. <https://doi.org/10.1016/j.resmic.2006.06.003>.
- 20 Rahman MS, Ceraul SM, Dreher-Lesnick SM, Beier MS, Azad AF. The *lspA* gene, encoding the type II signal peptidase of *Rickettsia typhi*: Transcriptional and

- functional analysis. *J Bacteriol* 2007;**189**:336–41.
<https://doi.org/10.1128/JB.01397-06>.
- 21 Tjalsma H, Kontinen VP, Prágai Z, Wu H, Meima R, Venema G, *et al*. The role of lipoprotein processing by signal peptidase II in the Gram-positive eubacterium *Bacillus subtilis*. *J Biol Chem* 1999;**274**:1698–707.
<https://doi.org/10.1074/jbc.274.3.1698>.
- 22 Kurokawa K, Kim MS, Ichikawa R, Ryu KH, Dohmae N, Nakayama H, *et al*. Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in *Staphylococcus aureus*. *J Bacteriol* 2012;**194**:3299–306.
<https://doi.org/10.1128/JB.00314-12>.
- 23 Zhao XJ, Wu HC. Nucleotide sequence of the *Staphylococcus aureus* signal peptidase II (*lsp*) gene. *FEBS Lett* 1992;**299**:80–4. [https://doi.org/10.1016/0014-5793\(92\)80105-P](https://doi.org/10.1016/0014-5793(92)80105-P).
- 24 Witke C, Götz F. Cloning and nucleotide sequence of the signal peptidase II (*lsp*)-gene from *Staphylococcus carnosus*. *FEMS Microbiol Lett* 1995;**126**:233–9.
[https://doi.org/10.1016/0378-1097\(95\)00015-W](https://doi.org/10.1016/0378-1097(95)00015-W).
- 25 de Greeff A, Hamilton A, Sutcliffe IC, Buys H, van Alphen L, Smith HE. Lipoprotein signal peptidase of *Streptococcus suis* serotype 2. *Microbiology* 2003;**149**:1399–407. <https://doi.org/10.1099/mic.0.26329-0>.
- 26 Khandavilli S, Homer KA, Yuste J, Basavanna S, Mitchell T, Brown JS. Maturation of *Streptococcus pneumoniae* lipoproteins by a type II signal peptidase is required for ABC transporter function and full virulence. *Mol Microbiol* 2008;**67**:541–57. <https://doi.org/10.1111/j.1365-2958.2007.06065.x>.

- 27 Widdick DA, Hicks MG, Thompson BJ, Tschumi A, Chandra G, Sutcliffe IC, *et al.* Dissecting the complete lipoprotein biogenesis pathway in *Streptomyces scabies*. *Mol Microbiol* 2011;**80**:1395–412. <https://doi.org/10.1111/j.1365-2958.2011.07656.x>.
- 28 Gullón S, Arranz EIG, Mellado RP. Transcriptional characterisation of the negative effect exerted by a deficiency in type II signal peptidase on extracellular protein secretion in *Streptomyces lividans*. *Appl Microbiol Biotechnol* 2013;**97**:10069–80. <https://doi.org/10.1007/s00253-013-5219-9>.
- 29 Munnoch JT, Widdick DA, Chandra G, Sutcliffe IC, Palmer T, Hutchings MI. Cosmid based mutagenesis causes genetic instability in *Streptomyces coelicolor*, as shown by targeting of the lipoprotein signal peptidase gene. *Sci Rep* 2016;**6**:1–10. <https://doi.org/10.1038/srep29495>.
- 30 Thompson BJ, Widdick DA, Hicks MG, Chandra G, Palmer T, Sutcliffe IC, *et al.* Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor*. *Mol Microbiol* 2010;**77**:no-no. <https://doi.org/10.1111/j.1365-2958.2010.07261.x>.
- 31 Buddelmeijer N. The molecular mechanism of bacterial lipoprotein modification—How, when and why? *FEMS Microbiol Rev* 2015;**39**:246–61. <https://doi.org/10.1093/femsre/fuu006>.
- 32 Paetzel M. Bacterial Signal Peptidases. *Bact. Cell Walls Membr.*, vol. 92. Springer International Publishing; 2019. p. 187–219.
- 33 Sander P, Rezwan M, Walker B, Rampini SK, Kroppenstedt RM, Ehlers S, *et al.* Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*.

- Mol Microbiol* 2004;**52**:1543–52. <https://doi.org/10.1111/j.1365-2958.2004.04041.x>.
- 34 Laguerre A, Löhr F, Henrich E, Hoffmann B, Abdul-Manan N, Connolly PJ, *et al.* From Nanodiscs to Isotropic Bicelles: A Procedure for Solution Nuclear Magnetic Resonance Studies of Detergent-Sensitive Integral Membrane Proteins. *Structure* 2016;1830–41. <https://doi.org/10.1016/j.str.2016.07.017>.
- 35 Szewczyk J, Collet JF. *The Journey of Lipoproteins Through the Cell: One Birthplace, Multiple Destinations*. vol. 69. 1st ed. Elsevier Ltd.; 2016.
- 36 Fernandes DD, Bamrah J, Kailasam S, Gomes GNW, Li Y, Wieden HJ, *et al.* Characterization of Fluorescein Arsenical Hairpin (FIAsH) as a Probe for Single-Molecule Fluorescence Spectroscopy. *Sci Rep* 2017;**7**:1–15. <https://doi.org/10.1038/s41598-017-13427-8>.
- 37 *Strep-tag®: The leading affinity tag in recombinant protein technology*. Iba: Solutions For Life Sciences. n.d. URL: <https://www.iba-lifesciences.com/strep-tactin-system-technology.html>.
- 38 Mao G, Zhao Y, Kang X, Li Z, Zhang Y, Wang X, *et al.* Crystal structure of E. coli lipoprotein diacylglyceryl transferase. *Nat Commun* 2016;**7**:. <https://doi.org/10.1038/ncomms10198>.
- 39 Liu X, Chen Y, Fierke CA. A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucleic Acids Res* 2014;**42**:e159. <https://doi.org/10.1093/nar/gku850>.
- 40 Blommel PG, Fox BG. Fluorescence anisotropy assay for proteolysis of specifically labeled fusion proteins. *Anal Biochem* 2005;**336**:75–86.

<https://doi.org/10.1016/j.ab.2004.09.023>.

- 41 Hall MD, Yasgar A, Peryea T, Braisted JC, Jadhav A, Simeonov A, *et al*. Fluorescence polarization assays in high-throughput screening and drug discovery: A review. *Methods Appl Fluoresc* 2016;**4**:. <https://doi.org/10.1088/2050-6120/4/2/022001>.
- 42 Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, *et al*. New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: Synthesis and biological applications. *J Am Chem Soc* 2002;**124**:6063–76. <https://doi.org/10.1021/ja017687n>.
- 43 Adams SR, Tsien RY. Preparation of the membrane-permeant biarsenicals, FLAsH-EDT2 and ReAsH-EDT2 tagged proteins. *Nat Protoc* 2008;**3**:1527–34. <https://doi.org/10.1038/nprot.2008.144.Preparation>.
- 44 Griffin BA, Adams SR, Tsien RY. Specific covalent labeling of recombinant protein molecules inside live cells. *Science (80-)* 1998;**281**:269–72. <https://doi.org/10.1126/science.281.5374.269>.
- 45 Lam AJ, St-Pierre F, Gong Y, Marshall JD, Cranfill PJ, Baird MA, *et al*. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat Methods* 2012;**9**:1005–12. <https://doi.org/10.1038/nmeth.2171>.
- 46 Evans EGB, Millhauser GL. *Genetic Incorporation of the Unnatural Amino Acid p-Acetyl Phenylalanine into Proteins for Site-Directed Spin Labeling*. vol. 563. 1st ed. Elsevier Inc.; 2015.
- 47 Rajendran V, Kalita P, Shukla H, Kumar A, Tripathi T. Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. *Int J Biol Macromol*

- 2018;**111**:400–14. <https://doi.org/10.1016/j.ijbiomac.2017.12.157>.
- 48 Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of *Escherichia coli*. *Science (80-)* 2001;**292**:498–500. <https://doi.org/10.1002/1873-3468.12333>.
- 49 Xie J, Schultz PG. An expanding genetic code. *Methods* 2005;**36**:227–38. <https://doi.org/10.1016/j.ymeth.2005.04.010>.
- 50 Young TS, Ahmad I, Yin JA, Schultz PG. An Enhanced System for Unnatural Amino Acid Mutagenesis in *E. coli*. *J Mol Biol* 2010;**395**:361–74. <https://doi.org/10.1016/j.jmb.2009.10.030>.
- 51 Fleissner MR, Brustad EM, Kalai T, Altenbach C, Cascio D, Peters FB, *et al.* Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc Natl Acad Sci* 2009;**106**:21637–42. <https://doi.org/10.1073/pnas.0912009106>.
- 52 Klock HE, Lesley SA. The Polymerase Incomplete Primer Extension (PIPE) Method Applied to High-Throughput Cloning and Site-Directed Mutagenesis Heath. *Methods Mol. Biol. High Throughput Protein Expr. Purif.*, vol. 498. 2009. p. 91–103.

CHAPTER 5. BICELLE PROPERTIES

5.1 Introduction

5.1.1 Project Motivation

As described in chapter 1, membrane proteins hold a wide variety of essential functions and comprise a large percentage of drug targets. However, in order to study membrane proteins *in vitro*, a membrane mimetic must be used.

For nearly two decades, bicelles have had a wide variety of applications, most commonly as bilayer mimetics for structural^{1–14} and functional^{15–20} investigations of membrane-associated proteins. A bicelle is a bilayer micelle; a disc-shape aggregate typically formed by a mixture of detergents (Figure 5.1A) and lipids (Figure 5.1B). Bicelle self-assembly was first determined in 1984^{21,22} and since, bicelles have been characterized using many methods such as small angle neutron scattering (SANS),^{23,24} and NMR.^{25–31} The classically described bicelle contains a central disk-shaped lipid bilayer encircled by a rim of detergents which screen the hydrophobic lipid tails from water (Figure 5.1C).^{27,28,32} Thus, in the “ideal” bicelle (term coined by Vold and Prosser in 1996³²) the lipid and detergent molecules are segregated spatially. Bicelles vary in size and shape depending on the ratio of lipid to detergent (known as the *q*-value)³², the structure of the lipid and detergent monomers²⁵, total concentration of amphiphiles^{33,34}, and temperature.^{29,30} For solution NMR structural studies, bicelles with low *q*-values (< 0.7 ; also known as fast-tumbling “isotropic” bicelles) have demonstrated some utility for polytopic integral membrane proteins.^{24,31,35,36} Several of these studies suggest that the stabilization of membrane protein fold is due to the more “bilayer” nature of bicelles compared to micelles.

That is, the segregated lipid core in bicelles is more similar in structure to the native membrane.

Recent investigations of binary mixtures of detergents with different alkyl chain lengths and head groups indicated that these compositions are fully mixed.^{37,38} This observation led to a hypothesis that bicelles with q -values below 1, for which the detergent concentration is higher than the lipid concentration, may not have segregated lipid cores, as previously suggested.³⁴ Here, we investigate the structure and segregation of bicelles with q -values less than 1 formed by dihexanoylphosphatidylcholine (DHPC; Figure 5.1A) and dimyristoylphosphatidylcholine (DMPC; Figure 5.1B), which have been studied for almost 30 years.^{26,39–41}

Several measurable structural and physical properties allow the mixing of lipids and detergents to be tested. As with mixed micelles, the average head group – head group distance (L) is expected to vary with concentration in mixed bicelles and can be determined via small angle X-ray scattering (SAXS), a model-free measurement.^{37,38} Since the two components of the bicelle have different scattering length densities, small angle neutron scattering (SANS) can determine their degree of mixing. The gel to liquid phase transition temperature (T_m), measured using the fluorescence anisotropy of diphenylhexatriene (DPH), is an independent measurement of the extent of bilayer formation. Thus, fluorescence anisotropy can be used to determine the extent of mixing. Finally, the shape, size, and lipid-detergent mixing of bicelles can be quantified directly using molecular dynamics (MD) simulations.

Here I will describe the structural investigation of DMPC/DHPC bicelles. SAXS had been previously completed in the laboratory by Ryan Oliver and set the stage for this

investigation. Ashton Brock conducted the SANS experiments, and I fit the data, interpreted the results, and wrote the manuscript. These were complemented by collaborations with the groups of Peter Tieleman for molecular dynamics simulations and Jebrell Glover for fluorescence anisotropy experiments.⁴²

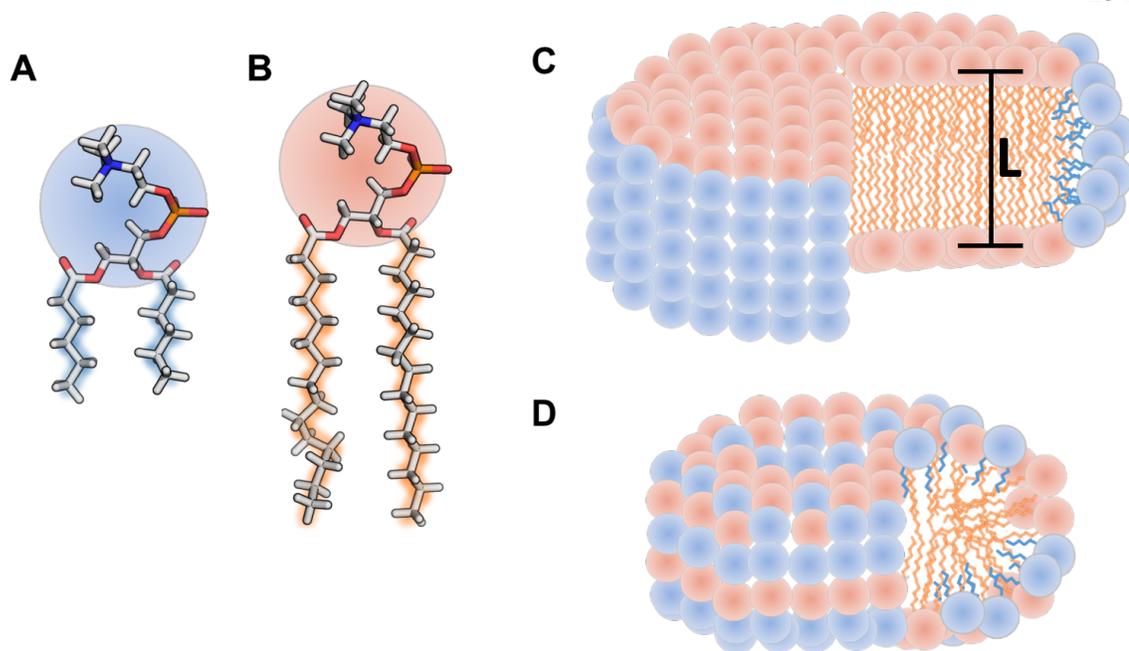


Figure 5.1. Detergent, lipid, and bicelle structures. Structures of detergent DHPC (A), lipid DMPC (B), and cartoons of idealized bicelles (C) and mixed micelles (D). Head group to head group distance (L) is shown in (C).

5.1.2 Previous Work: SAXS

Small-angle X-ray scattering (SAXS) experiments provide low resolution structural information about a protein or complex.^{43,44} Here, SAXS is used to investigate the structure of DMPC/DHPC bicelles at various q -values.

For SAXS, mixtures of DMPC/DHPC were prepared from $q = 0.1$ to 1.0 , in increments of 0.1 , at 6% total amphiphile weight per volume in 10 mM phosphate buffer, pH 6.6 with 7% D₂O. A freeze-thaw cycle was performed on mixtures which were not optically transparent after vortex mixing. Small-angle scattering measurements were collected at the Advanced Photon Source (Argonne, IL) and processed as described previously.^{45,46}

In the SAXS scattering profile, the second maximum (Q_{\max}) corresponds to the distance between opposing electron rich head groups, L .⁴⁵ For an ideal bicelle with a fully segregated core, the average head group to head group distance (L) equals to twice the length of DMPC tails plus a head group (one half on each side) (~ 43 Å; Figure 5.1C). But, if the detergent and lipid components mix, then the parameter L will be less and decrease linearly with the concentration of DHPC in the core.^{37,38}

Our data show that at q -values from 0.5 to 1 , the model-free dimension L remains constant at 42 Å (Figure 5.2), suggesting a segregated bicelle (core). However, below $q=0.5$, L varies linearly with decreasing q -values (Figure 5.2), indicative of lipid and detergent mixing in the core. A linear fit of the data below $q=0.5$ produces a y -intercept of 22 Å, the approximate L of pure DHPC, while a linear fit of the data above $q=0.5$ produces a y -intercept of 42 Å, the approximate L of pure DMPC (Figure 5.2B).^{38,45} These data suggest that DHPC and DMPC mix between $q=0.1$ and $q=0.5$ and are segregated for $q>0.5$.

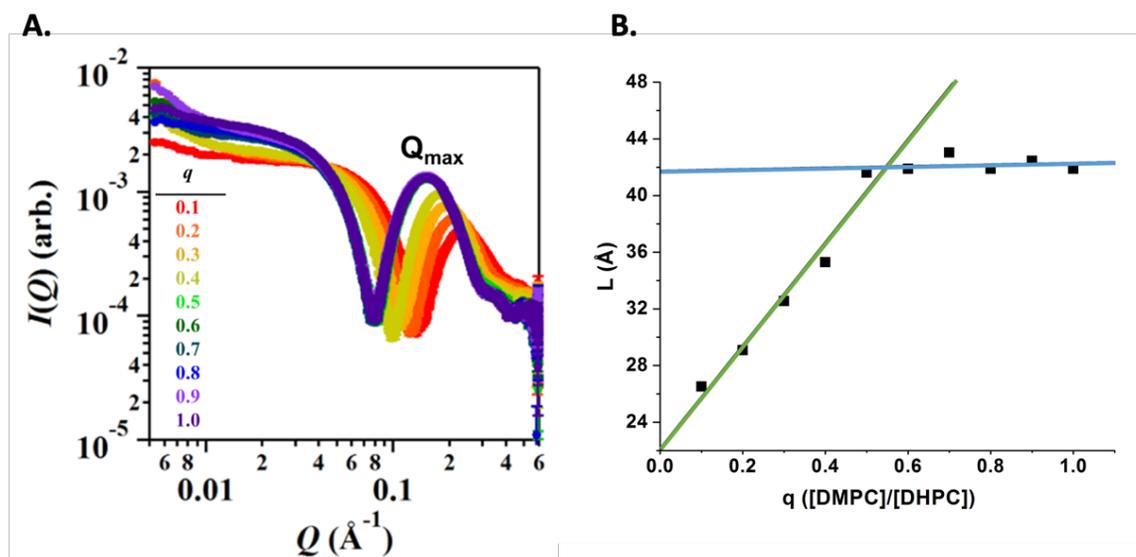


Figure 5.2. SAXS measurements of $q=0.1$ to 1 bicelles. Bicelle dimensions vary with q -values between 0 and 0.5. The L values (B) are measured directly from the SAXS scattering profiles (A) for bicelles with varying q -values according to $L=2\pi/Q_{max} \approx 2(1.5 + 1.265n_c) + t$, where n_c is the number of carbons in the alkyl chain and t is the head group thickness. 6% (w/w) amphiphile concentration was used. Linear fits to the data points for $q \leq 0.5$ (green) and data points for $q \geq 0.5$ (blue) are shown in B.

5.2 SANS

In order to expand upon the global structural information obtained by SAXS, small-angle neutron scattering (SANS) was used to determine structural information about the components of the bicelle as well as lipid – detergent mixing.

5.2.1 SANS Methods

For SANS experiments, DMPC/DHPC bicelles were prepared in 10 mM phosphate buffer, pH 6.6 at $q = 0.3$ and 0.7 at 6% (w/v) total amphiphile concentrations. DMPC-d54, an analog of DMPC with deuterated acyl chains, was used as the lipid component to increase the difference in scattering length density (SLD). A freeze-thaw cycle was performed on mixtures which were not optically transparent after vortex mixing.

Solvent deuteration was modified for contrast variation experiments by changing the D₂O content in the buffer from 0% to 99% via dialysis. Neutron scattering data were obtained at the High-Flux Isotope Reactor (Oak Ridge, TN). The experimental match point for each data set was calculated from the square root of the total signal intensity as a function of the percentage of D₂O in the solvent.⁴⁷

5.2.2 SANS Contrast Variation

While SAXS gives information about the shape of the overall complex, small-angle neutron scattering (SANS) is able to differentiate between components of a multicomponent system.⁴³ The intensity (I) of the SANS signal is related to the neutron scattering length density (SLD, ρ), the particle volume (V), and properties of the neutron wavelength used (q) for N_i particle species:⁴⁴

$$I(q) = \sum_i N_i \langle |\Delta\rho_i e^{iq \cdot r} dV_i|^2 \rangle$$

Equation 5.1

The SLD and volume can be calculated for each component of the system, and the wavelength properties are known. If the SLD of a component matches the SLD of the surrounding solvent, the component will be “matched out” and will not contribute to the SANS scattering profiles (Figure 5.3). In this way, SANS experiments can be conducted at various H₂O:D₂O ratios to focus on the structural information from specific parts of the complex system.^{43,47-50} As multiple variables can be solved using an equal number of equations, systems measuring multiple parameters (as in this case) can be fit by using an equal or higher number of SANS scattering profiles at different D₂O concentrations.

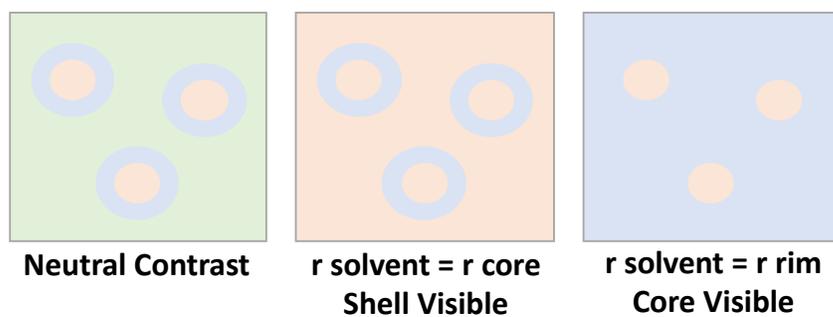


Figure 5.3. Schematic of SANS contrast variation. When the scattering length density (SLD) of the solvent is equal to the SLD of one of the components, it becomes matched out and does not contribute to the scattering intensity.

5.2.3 SANS Fitting Approach

The SasView cylinder-based core-shell bicelle model was used for all data fits (Figure 5.4).⁵¹ The scattering length density (SLD) of the solvent was calculated using known SLD values of $-5.5 \times 10^{-7} \text{ \AA}^{-2}$ and $6.3 \times 10^{-6} \text{ \AA}^{-2}$ for H₂O and D₂O, respectively. The chemical composition of the phosphocholine headgroup along with reported molecular volume was used to calculate the face SLD to be $1.44 \times 10^{-6} \text{ \AA}^{-2}$.⁵² This procedure was repeated to determine SLDs for whole DHPC ($6.71 \times 10^{-7} \text{ \AA}^{-2}$), whole d54-DMPC ($5.39 \times 10^{-6} \text{ \AA}^{-2}$), DHPC tail ($-7.49 \times 10^{-8} \text{ \AA}^{-2}$), and d54-DMPC tail ($7.04 \times 10^{-6} \text{ \AA}^{-2}$) (Table 5.1).^{30,53-55} Ratios of DHPC and d54-DMPC whole lipid SLDs were calculated for the rim SLD, while ratios of DHPC and d54-DMPC tails were calculated for the core SLD. These SLD values were confirmed using the online resource MULCh: modules for the analysis of contrast variation data.⁵⁶ The face thickness range was set to 7-12 Å based on known length of the phosphocholine headgroup while the rim thickness range was set between 10-25 Å based on the lengths of DHPC and d54-DMPC.^{30,52,55,57} The original length range was 15-36 Å based on the lengths of two DMPC or DHPC tails as well as the SAXS data. The radius started at 20 Å but had no set range since there was no prior evidence to the radius length. The $q=0.7$ 6% (w/w) total amphiphile concentration (C_L) and $q=0.3$ 6% C_L Q ranges were set as 0.02-0.20 and 0.035-0.3 respectively to remove noise. The following solvent D₂O percentages were fit for each set so as to not fit within 20% of the contrast match point (CMP): $q=0.7$ 6% C_L : 0, 10, 20, 30, 80, 90, 100; and $q=0.3$ 6% C_L : 0, 10, 20, 70, 80, 90, 100. All scattering profiles in a set were fit starting with the above parameter ranges and SLD values corresponding to the amount of mixing observed in the MD simulations ($q=0.3$: 76% DHPC and 24% DMPC corresponding to a core SLD of $1.64 \times 10^{-6} \text{ \AA}^{-2}$ and a

rim SLD of $1.80 \times 10^{-6} \text{ \AA}^{-2}$; $q=0.7$: 49% DHPC and 51% DMPC corresponding to a core SLD of $3.56 \times 10^{-6} \text{ \AA}^{-2}$ and a rim SLD of $3.08 \times 10^{-6} \text{ \AA}^{-2}$. Values for radius, rim thickness, face thickness, and length as well as SLDs for core and rim were converged upon after several rounds of fitting (Figure 5.5 and 5.6). The percent of DHPC and DMPC in the core and rim were calculated by the ratios of pure DHPC and DMPC SLDs [SLD=(%DMPC)(DMPC SLD)+(DHPC SLD)] where the percentages of DMPC and DHPC add to 100%. The radii were calculated as radius+rim and $\frac{1}{2}$ length+face thickness.

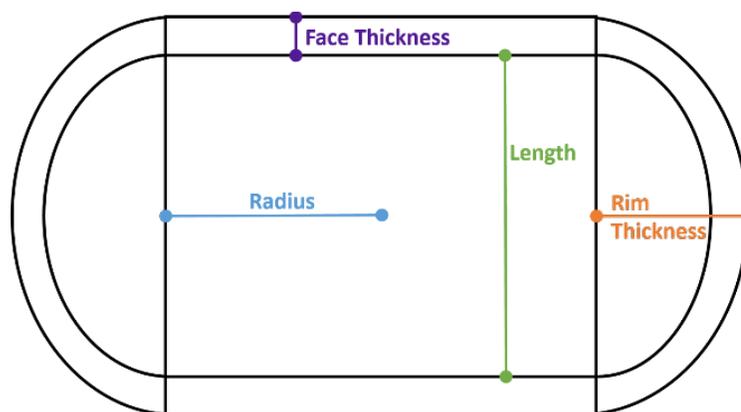


Figure 5.4. Schematic of core-shell bicelle model. Core-shell bicelle model used to fit the experimental SANS scattering profiles for radius, rim thickness, length, and face thickness values.

Component	Scattering length density ($\times 10^{-6} \text{ \AA}^{-2}$)
DMPC with deuterated alkyl chains	5.39
DHPC (rim in ideal bicelle)	0.671
PC head group (face)	1.44
Deuterated dimyristoyl chains (core in ideal bicelle)	7.04
Dihexanoyl chains	-0.075

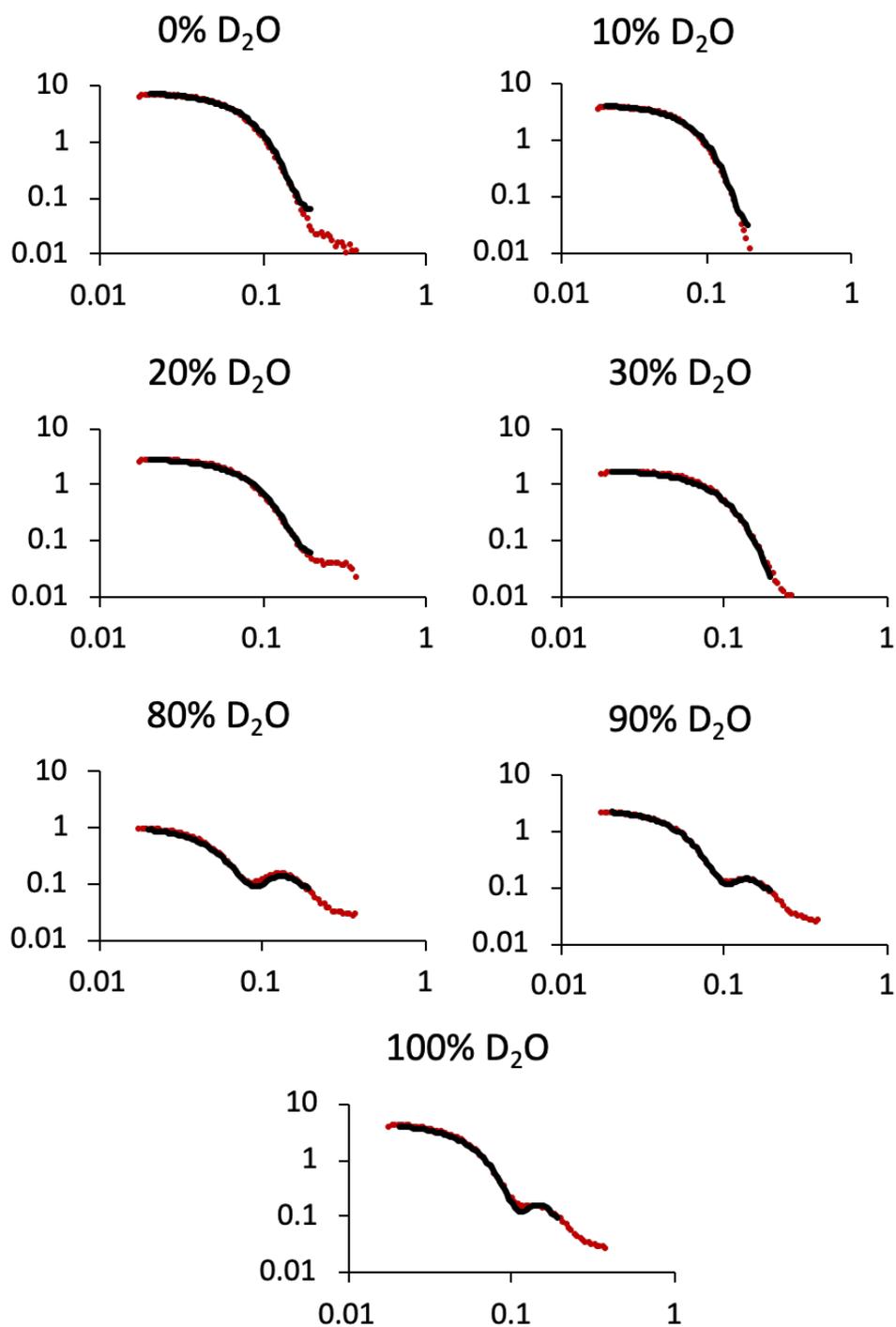


Figure 5.5. SANS core-shell bicelle model fits for $q=0.7$. Core-shell bicelle model fits (black) to the experimental SANS scattering profiles (red) for $q = 0.7$ bicelles in different solvent D_2O concentrations.

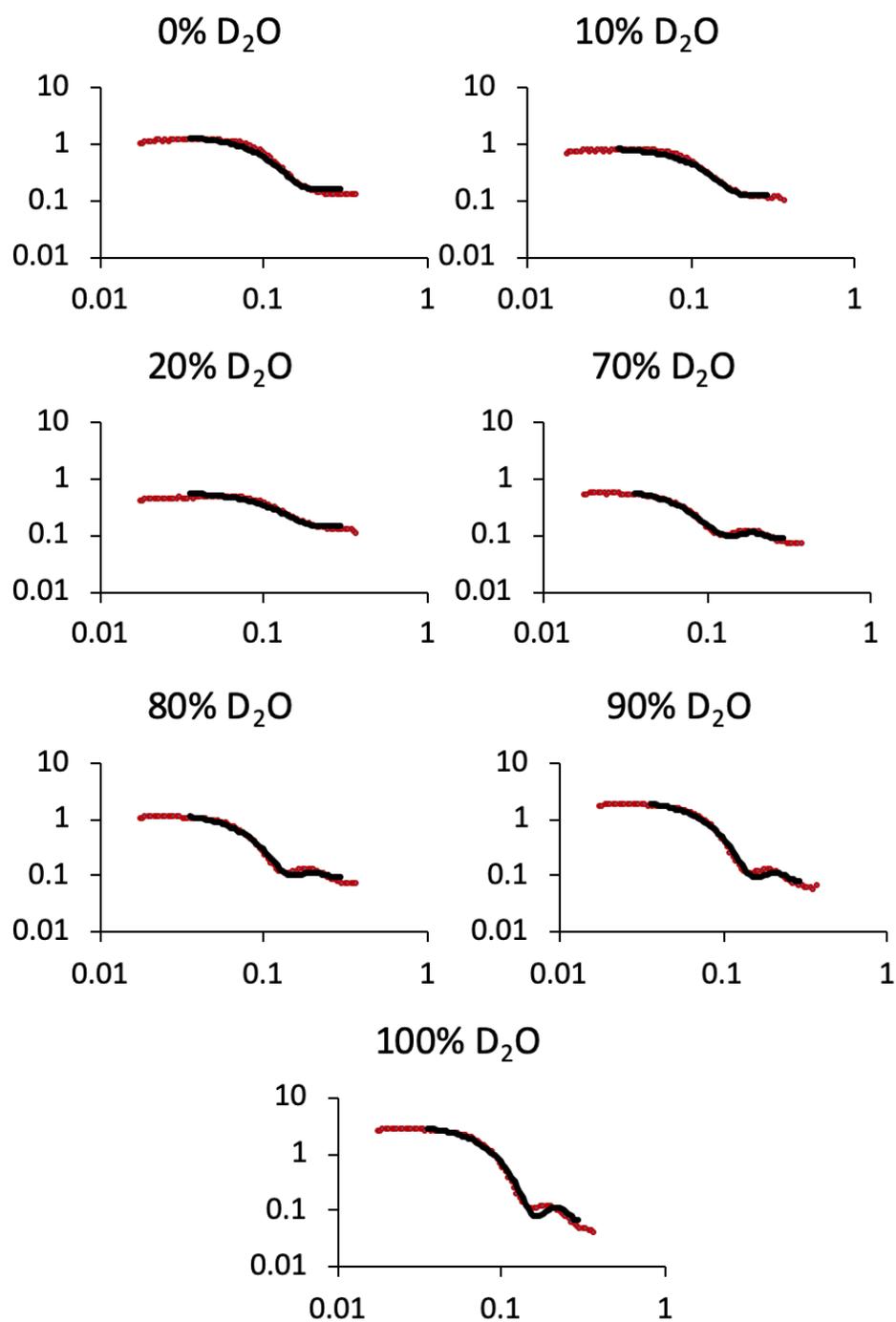


Figure 5.6. SANS core-shell bicelle model fits for $q=0.3$. Core-shell bicelle model fits (black) to the experimental SANS scattering profiles (red) for $q = 0.3$ bicelles in different solvent D_2O concentrations.

5.2.4 SANS Results

To investigate the structure of bicelles, SANS experiments were conducted on bicelles with q -values of 0.3 and 0.7 with different solvent scattering length densities (varied percentages of D_2O in H_2O). Each scattering profile (Figure 5.5 and 5.6) was fit to the core-shell bicelle model (Figure 5.4). The obtained dimensions (Tables 5.2 and 5.3) provide information about the shape of each bicelle. These results show that higher detergent concentrations reduce the size and result in a more spherical shaped bicelle (Table 5.4).

The classical bicelle model predicts that the concentration of lipids and detergents in the core and rim will deviate from their bulk concentration: the bilayer forming lipid DMPC is expected to preferentially partition to the core, while the detergent DHPC to the rim. SANS contrast variation was used to investigate the extent of mixing in bicelles with q -values of 0.3 and 0.7.

In SANS experiments, bicelles formed by DHPC with protonated alkyl chains and DMPC with deuterated alkyl chains were used to distinguish a segregated versus a mixed bicelle.⁵¹ Deviations from the DHPC or DMPC alkyl chain scattering length density (SLD; Table 5.1) in the “rim” and “core”, respectively, indicates lipid/detergent mixing because of the SLD contrast between DHPC and DMPC. To verify the SLD values and the effective q -values of the bicelle the theoretical match points and the experimental match points were compared and are in good agreement (Table 5.5). The percent of DHPC and DMPC in the core and rim were calculated by the ratios of pure DHPC and DMPC SLDs (Table 5.1) [$SLD=(\%DMPC)(DMPCSLD)+(\%DHPC)(DHPC\ SLD)$] where the percentages of DMPC and DHPC add to 100%. The SLD values from the core-shell bicelle fits to the SANS data

(Figure 5.5 and 5.6) indicate that the core composition is 38 – 77% DHPC and 23 – 62% DMPC in $q=0.3$ bicelles and the core composition is 37 – 49% DHPC and 51 – 63% DMPC in $q=0.7$ bicelles. Although a broad range of DHPC is observed for the $q=0.3$ bicelles, fully mixed values (76% for bulk value) are observed while they are not for $q=0.7$ bicelles (56% for bulk value) (Tables 5.2-5.4).

Table 5.2: Fit parameters for the SANS profiles of $q=0.7$ bicelles using the core-shell bicelle model						
D ₂ O	Radius	Rim	Face	Length	Core DMPC	Rim DHPC
%	(Å)				%	
0	17	12	7	19	63	59
10	15	11	7	23	52	49
20	14	11	7	18	52	49
30	11	11	7	25	51	49
80	27	19	12	36	52	49
90	19	19	12	36	55	52
100	16	13	7	32	59	51
AVG	17	15	8	27	55	51
Range	10-27	11-23	7-12	18-36	51-63	49-59

Table 5.3: Fit parameters for the SANS profiles of $q=0.3$ bicelles using the core-shell bicelle model						
D ₂ O	Radius	Rim	Face	Length	Core DMPC	Rim DHPC
%	(Å)				%	
0	6	12	7	16	24	77
10	4	13	7	15	23	76
20	5	10	7	15	32	77
30	21	10	12	23	24	78
80	15	11	7	21	60	100
90	14	11	7	19	62	100
100	15	11	7	18	56	100
AVG	11	11	8	18	40	87
Range	4-21	10-13	7-12	15-23	23-62	76-100

Table 5.4: Comparison of q=0.3 and q=0.7 bicelles				
	q=0.7		q=0.3	
	Average	Range	Average	Range
Radius	17	10-27	11	4-21
Rim Thickness	15	11-23	11	10-13
Face Thickness	8	7 - 12	8	7 - 12
Length	27	18-36	18	15-23
Core % DMPC	55	51-63	40	23-62
Rim % DHPC	51	49-59	87	76-100

Table 5.5: Match point comparison for tested bicelles				
Bulk q value	Effective q value*	Match point (% D₂O)		
		Theoretical for bulk q value	Theoretical for effective q value	Experimental
0.70	0.78	54	56	57
0.30	0.32	40	41	41

*Effective q-value is calculated by subtracting the critical bicelle concentration (cbc) of DHPC monomer from the overall concentration of amphiphiles ($q_{\text{eff}} = [\text{DMPC}] / ([\text{DHPC}] - \text{CBC})$; $\text{CBC}_{\text{DHPC}} = 7\text{mM}$)

5.3 Collaborations

5.3.1 Molecular Dynamics

In order to gain a better understanding of the size, shape, and lipid – detergent mixing in DMPC/DHPC bicelles, molecular dynamics simulations at various q-ratios were performed in collaboration with the group of Peter Tieleman.⁴²

5.3.1.1 MD Methods

MD simulations were performed for mixtures of DMPC (lipid) and DHPC (detergent) in water. The Gromacs v. 4.6.5 software⁵⁸ with the Stockholm lipid (Slipid) force field^{59–61} and TIP3p water model⁶² were used. The starting configuration constituted a loose spherical aggregate of DMPC and DHPC in a water cube. The ratio of DMPC/DHPC was varied to produce a range of q-values (0.3 and 0.7), resulting in 6 different system setups with varying numbers of lipids and detergents in the box. The hydration level was selected at 1/400 lipid (and detergent) to water ratio, which corresponds to ~100 mM and ~9% (w/v) amphiphile. Aggregation of lipids and detergents in the course of simulation resulted in one or more bicelles/micelles in the simulation box. We considered the largest and most stable aggregate for each setup; the actual number of molecules in the considered aggregate was therefore lower than the total number of molecules in the box. Note that the actual q-value of the aggregate differed from the total value in the box.

The standard simulation parameters for Slipids force field were used. The temperature was maintained at 303 K with the v-rescale thermostat⁶³ with a relaxation time of 1 ps. The pressure was maintained at 1 bar with the isotropic coupling scheme and

Parinello-Rahman barostat⁶⁴ with the time constant of 3 ps. A cut-off of 1 nm was used for van der Waals and electrostatic interactions; the energy-pressure dispersion correction and the PME method^{65,66} were used for long-range interactions. The integration time step was 2 fs, the neighbor list was updated every 10 steps. The simulation time was 1 microsecond for each setup.

To characterize the mixing of lipids and detergents, we calculated the enrichment values of DHPC around DMPC. The enrichment value is given by the ratio of the local concentration of DHPC around DMPC to the bulk concentration (defined as the molar ratio of DHPC and corresponding to $1/(q+1)$ at the given q -value). The calculations were performed using custom python scripts employing MDAnalysis library.⁶⁷ To characterize the bicelle shape, the bicelle headgroup layer was fitted to a 3D ellipsoid using custom Matlab scripts, and the principal radii of the ellipsoid were calculated. We also calculated the SAXS scattering profiles of the bicelle structures obtained in MD using FoXS software.⁶⁸

5.3.1.2 MD Results

Molecular dynamics simulations were performed of $q=0.3$ and $q=0.7$ bicelles (Figure 5.7). At higher q , bicelles become less spherical compared to lower q , as evident from the principal radii of an ellipsoid fitted to the aggregate shape (Table 5.6).

In MD simulations, the segregation of lipids and detergents can be quantified by comparing the local concentration of DHPC around DMPC. There is on average 76% DHPC around DMPC in $q=0.3$ bicelles and 49% in 0.7 bicelles. It is interesting to note that full segregation was not observed in either case indicating a certain degree of mixing even

in isotropic bicelles with $q > 0.5$, similar to previously reported simulations.³⁴ Thus, mixed nearly-spherical micelles were observed for $q \approx 0.3$ and partially segregated ellipsoid bicelles for $q \approx 0.7$ (representative structures are shown in Figure 5.7).

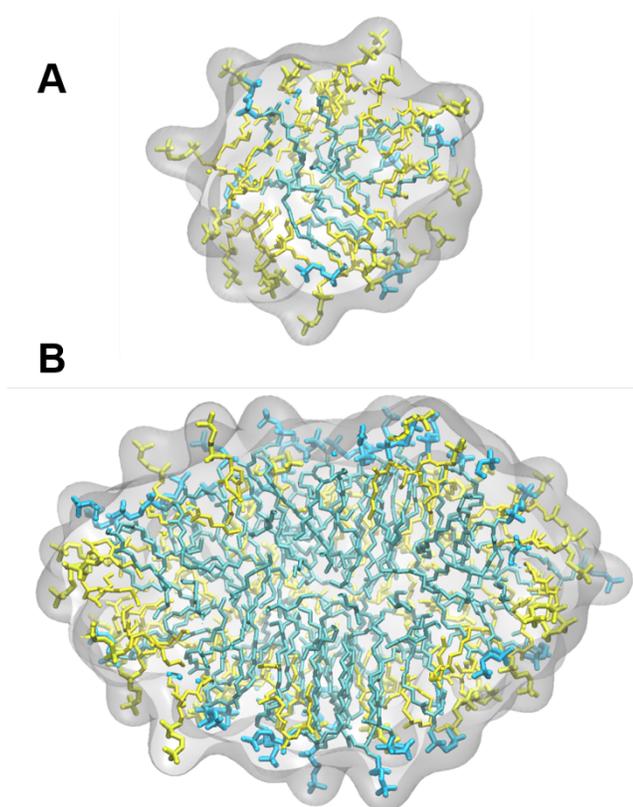


Figure 5.7. Snapshots from MD bicelle simulations. A) $q=0.3$ and B) $q=0.7$ bicelles (9% (w/w) amphiphile) visualized in VMD.⁶⁹ DHPC and DMPC are rendered as sticks and colored yellow and blue, respectively. The surface is shown as transparent gray and a portion of the bicelle is removed to view the interior distribution of the DMPC and DHPC tails.

Table 5.6: Bicelle properties from MD simulations								
#	n (n*)	q (q*)	Rg,nm	L,nm	a,nm	b,nm	c,nm	comment
1	34 (33)	0.3 (0.3)	1.6	2.6	2.4	2	1.8	mixed micelle
2	75 (41)	0.3(0.32)	1.7	2.8	2.4	2.1	1.8	mixed micelle
3	75 (74)	0.5(0.51)	2.1	3.5	3.0	2.5	2.3	ellipsoidal bicelle
4	150(110)	0.5(0.72)	2.5	4.2	4.2	2.8	2.3	segregated bicelle
5	120(116)	0.7(0.76)	2.6	4.2	4.0	3.1	2.2	segregated bicelle
6	150 (96)	0.7(0.92)	2.4	4.2	3.5	3.1	2.2	segregated bicelle

Here n is total number of lipids and detergents in the simulation box, n* is the actual number of lipids and detergents in the aggregate, q is q-value (lipid-to-detergent ratio) in the simulation box, q* is the actual q-value in the bicelle (or micelle), Rg is radius of gyration, L parameter is determined from the second peak of the SAXS profile using FoXS software, a, b, and c are the principal radii of an ellipsoid, fitted to the bicelle surface (headgroups layer).

5.3.2 Fluorescence Anisotropy

To gain a better understanding of DMPC/ DHPC mixing in various q-ratio bicelles, fluorescence anisotropy was employed in collaboration with the group of Jebrell Glover.⁴²

5.3.2.1 Anisotropy Methods

For fluorescence anisotropy experiments, 2.3% (w/w) DMPC – DHPC mixtures with q-values ranging from 0 to 1.50 (0.05, 0.30, 0.50, 0.75, 1.00, 1.50) were prepared on a 3-gram scale using the following methodology. First, a DMPC stock was made by dissolving the lipid to 50 mg/mL in chloroform. To a 2 mL microcentrifuge tube, DMPC in chloroform was added (83, 370, 510, 630, 720, and 827 μ L respectively). After this, DPH dissolved in methanol was added to each tube to give a final concentration of 6 μ M, and the samples were then dried under vacuum overnight. The samples were then hydrated by the addition of 2.78 mL of water, and 75 μ L buffer (400 mM HEPES, 4.0 M NaCl pH 7.4). Finally, DHPC was added as a 25% (w/w) stock to achieve clear homogeneous solutions (250, 196, 169, 146, 129, and 107 μ L respectively). Vesicles were prepared by dissolving 16 mg of DMPC in 600 μ L of chloroform. To this DPH dissolved in methanol was added to give a final concentration of 30 μ M (1:500 DPH to lipid molar ratio). The sample was then dried under vacuum overnight. The sample was further rehydrated using 2.78 mL of water and 75 μ L buffer (400 mM HEPES, 4.0 M NaCl pH 7.4). This solution was sonicated for 5 minutes using a microtip to generate small unilamellar vesicles. The solution became clear and was then centrifuged at 20,000 x g for 5 minutes to remove titanium and large lipid aggregates. This solution was then allowed to sit 12 hours to allow

fusion of small unilamellar vesicles. The solution was then centrifuged at 20,000 x g for 5 minutes. The supernatant was diluted 4 fold and to be used in fluorescence experiments.

Fluorescence emission spectra of lipid and detergent mixtures were acquired with magnetic stirring using a 1 × 1 cm quartz cuvette on an Agilent Eclipse fluorometer (Santa Clara, CA). The excitation and emission slit widths were both set to 5 nm. The fluorescence emission intensity was measured (excitation 355 nm, emission 430 nm) with polarizers parallel to each other (both oriented at 0° from vertical) and repeated in the perpendicular configuration (excitation 0° and emission 90°). The correction factor for emission monochromator transmission efficiency was obtained from the ratio of emission intensity at 0° and 90° with the excitation polarizer oriented at 90°. Melting curves were generated for both pure DMPC vesicles and DMPC-DHPC bicelles by examining the change in diphenylhexatriene (DPH) anisotropy as a function of temperature over the range of 2 - 36 °C. Each melting temperature was determined from the inflection point of the melting curve using a sigmoidal fit in Igor Pro 6.22A (WaveMetrics, Inc., Lake Oswego, OR). Melting temperatures obtained represent the average of two experiments.

5.3.2.2 Anisotropy Results

The fluorescence anisotropy of diphenylhexatriene (DPH) detects changes in the fluidity of lipid bilayers as a function of temperature, from which the main phase transition temperature (T_m) of a lipid bilayer can be determined.⁷⁰⁻⁷² To benchmark this technique, the T_m of pure DMPC vesicles was measured to be 23.1 ± 0.4 °C, consistent with other methods (Figure 5.8).³⁴ The T_m of bicelles is expected to be identical to that of DMPC vesicles if DPH partitions into a region comprised purely of DMPC. However, if significant

mixing between DMPC and DHPC occurs, a decrease of T_m compared to DMPC vesicles will be observed, as DHPC disrupts acyl chain packing between DMPC molecules.

Analysis of the melting curves (Figure 5.8A) yielded the T_m for each q -value. Comparison of the T_m values obtained from the anisotropy measurements to T_m values for ideally mixed DHPC/DMPC vesicles indicates significant differences at all q -values less than 1.0, suggesting that these bicelles do not fit a fully mixed bicelle model (Figure 5.8B).³⁴ This data agrees with previously reported T_m values derived from FTIR spectroscopy of various q -value bicelles;³⁴ however, it does not support recent NMR data indicating similar lipid/detergent mixing in low and high q bicelles.⁷³ As the q -value increased, the T_m asymptotically approached the melting temperature of a pure DMPC bilayer. Only for $q \geq 1.0$ a T_m close to that of a pure DMPC bilayer is obtained (± 1 °C) in agreement with FTIR measurements (Figure 5.8B).³⁴ This suggests a variation in the lipid/detergent mixing at q -values below 1.0.

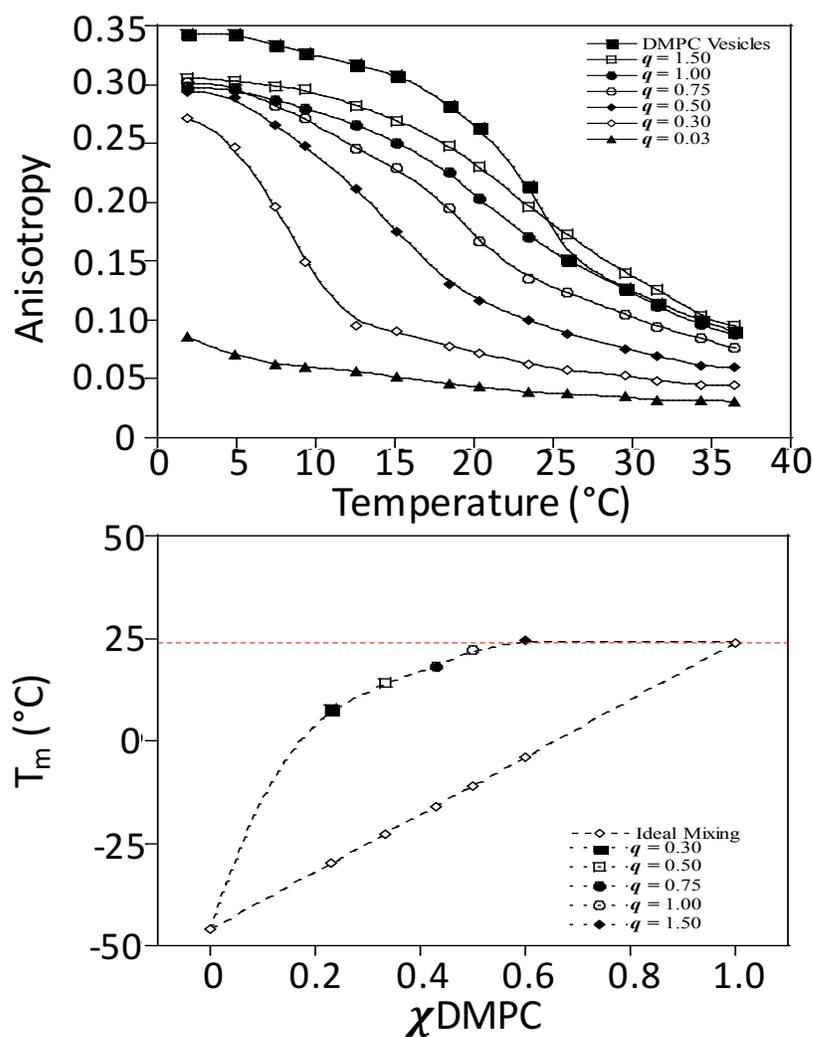


Figure 5.8. Fluorescence anisotropy of bicelles. A) Temperature dependence of the anisotropy value for DPH fluorescence reconstituted into bicelles with varying q -values (2.3% (w/w) amphiphile). The inflection point of each melting curve was taken as the T_m . B) Experimentally determined and calculated T_m values for bicelle solutions as a function of mole fraction DMPC. The linear black dashed line represents T_m values for ideal mixing ($T_m = \chi_{\text{DHPC}} \times T_m(\text{DHPC}) + \chi_{\text{DMPC}} \times T_m(\text{DMPC})$). The red dashed line shows the T_m of pure DMPC bilayers (23.1 $^{\circ}\text{C}$). Errors in each T_m measurement were approximately $\pm 0.2 - 0.4$ $^{\circ}\text{C}$.

5.4 Conclusions

5.4.1 Bicelle Shape

In order to gain a better understanding of the bicelle membrane mimetic environment, bicelle structural features were determined with SAXS, SANS, and MD simulations. SANS obtained dimensions (Tables 5.2-5.4) agree with SAXS data (Table 5.7) and provide additional information about the shape of the bicelle. Higher detergent concentrations reduce the size and result in a more spherical shaped bicelle. The simulation results support the SANS and SAXS analysis. At higher q , bicelles become less spherical compared to lower q , as evident from the principal radii of an ellipsoid fitted to the aggregate shape (Tables 5.6-5.7). This trend is observed in the SANS models; however, the average radii from the SANS models are slightly shorter than the MD models. Some discrepancies are expected due to differences in the methods related to ensemble properties (multiple bicelles in experiments with a certain degree of polydispersity vs. a single bicelle in the simulation box). However, the MD dimensions are within the ranges obtained from the SANS fits. The small radius is comparable to half the SAXS-derived L dimension (the SAXS value is smaller by half of a head group since L is the distance measured from the middle of each head group). Furthermore, L values derived from the simulated SAXS data (from the MD obtained bicelle structures) are equal to the SAXS values for $q=0.7$ (42 Å), but are somewhat larger than those for $q=0.3$ bicelles (Figure 5.2, Table 5.6). Altogether, the difference in the radii between $q=0.3$ and $q=0.7$ bicelles is indicative of a structural change in isotropic bicelles above and below $q\approx 0.5$. The linear changes in L observed in SAXS experiments, and the overall geometry determined by all three methods suggest that the bicelles with $q < 0.5$ do not have fully segregated lipid cores.

Table 5.7: Comparison of bicelle methods						
		Radius (Å)		DHPC		
q		1	2	Expected [#] (%)	Observed (%)	Ratio [^]
0.7	SAXS	-	21	56	-	-
	SANS	32	22		45/51*	0.8/0.9
	MD	40	27		49	0.9
0.3	SAXS	-	16	76	-	-
	SANS	22	17		60/87*	0.8/1.1
	MD	24	19		76	1

Dash indicates the parameter is not determined. [#]If fully mixed, *Average values for the core/rim are given, [^]Ratio of expected to observed DHPC.

5.4.2 Lipid-Detergent Mixing

SANS, MD, and fluorescence anisotropy were used to investigate the extent of mixing in bicelles with q -values of 0.3 and 0.7. SANS and MD derived DHPC percentages are in good agreement (Table 5.7). Fluorescence anisotropy measurements suggest a variation in the lipid/detergent mixing at q -values below 1.0 (Figure 5.8), in agreement with the geometrical changes determined with SAXS (Figure 5.2), and the geometrical and lipid-detergent mixing observed in MD (Figure 5.7, Table 5.6) and SANS (Table 5.4) studies.

5.4.3 Overall Conclusions

We have shown using four independent methods – SAXS, SANS, MD, and fluorescence anisotropy - that bicelle properties vary with the lipid-to-detergent ratio. The data suggest that at q -values below 1 lipid and detergent molecules partially mix, and the bicelle structure deviates from the ideal bicelle model. With increasing q -values, the lipid-detergent aggregates transition from a spherical mixed micelle through an ellipsoidal micelle to a disc-like bicelle.

These results suggest that care should be taken in interpreting membrane protein structural changes in micelles and bicelles. Isotropic bicelles with q -values less than 0.5 likely present a micellar environment, and bicelles with q -values less than 1 may not fully capture bilayer properties. A recent NMR study inferred similar bicelle differences based on protein positioning using PRE experiments.³⁵ Changes in protein structure in a low- q micelle/bicelle may be related to the micelle shape, size, and fluidity, or specific interactions with the lipids rather than the claimed “more bilayer-like” feature. It is

interesting to note that segregation of lipids in low q bicelles may be protein mediated if the lipid interactions are pre-formed.^{74,75}

5.5 References

- 1 Dürr UHN, Soong R, Ramamoorthy A. When detergent meets bilayer: Birth and coming of age of lipid bicelles. *Prog Nucl Magn Reson Spectrosc* 2013;**69**:1–22. <https://doi.org/10.1016/j.pnmrs.2013.01.001>.
- 2 Marcotte I, Auger M. Bicelles as model membranes for solid-and solution-state NMR studies of membrane peptides and proteins. *Concepts Magn Reson Part A Bridg Educ Res* 2005;**24**:17–35. <https://doi.org/10.1002/cmr.a.20025>.
- 3 Faham S, Bowie JU. Bicelle crystallization: A new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J Mol Biol* 2002;**316**:1–6. <https://doi.org/10.1006/jmbi.2001.5295>.
- 4 McCaffrey JE, James ZM, Thomas DD. Optimization of bicelle lipid composition and temperature for EPR spectroscopy of aligned membranes. *J Magn Reson* 2015;**250**:71–5. <https://doi.org/10.1016/j.jmr.2014.09.026>.
- 5 Yamamoto K, Percy P, Lee DK, Yu C, Im SC, Waskell L, *et al*. Temperature-resistant bicelles for structural studies by solid-state NMR spectroscopy. *Langmuir* 2015;**31**:1496–504. <https://doi.org/10.1021/la5043876>.
- 6 Duc NM, Du Y, Zhang C, Lee SY, Thorsen TS, Kobilka BK, *et al*. Effective Application of Bicelles for Conformational Analysis of G Protein-Coupled Receptors by Hydrogen/deuterium Exchange Mass Spectrometry. *J Am Soc Mass Spectrom* 2015;**26**:808–17. <https://doi.org/10.1007/s10741-014-9462-7>. Natural.

- 7 Warschawski DE, Arnold AA, Beaugrand M, Gravel A, Chartrand É, Marcotte I. Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim Biophys Acta - Biomembr* 2011;**1808**:1957–74.
<https://doi.org/10.1016/j.bbamem.2011.03.016>.
- 8 Faham S, Ujwal R, Abramson J, Bowie JU. Chapter 5 Practical Aspects of Membrane Proteins Crystallization in Bicelles. *Curr Top Membr* 2009;**63**:109–25.
[https://doi.org/10.1016/S1063-5823\(09\)63005-2](https://doi.org/10.1016/S1063-5823(09)63005-2).
- 9 Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: Not just a soap opera. *Biochim Biophys Acta - Biomembr* 2004;**1666**:105–17.
<https://doi.org/10.1016/j.bbamem.2004.04.011>.
- 10 Poulos S, Morgan JLW, Zimmer J, Faham S. *Bicelles coming of age: An empirical approach to bicelle crystallization*. vol. 557. 1st ed. Elsevier Inc.; 2015.
- 11 Deshmukh M V., John M, Coles M, Peters J, Baumeister W, Kessler H. Inter-domain orientation and motions in VAT-N explored by residual dipolar couplings and ¹⁵N backbone relaxation. *Magn Reson Chem* 2006;**44**:89–100.
<https://doi.org/10.1002/mrc.1837>.
- 12 Gruss F, Hiller S, Maier T. Purification and Bicelle Crystallization for Structure Determination of the E. coli Outer Membrane Protein TamA. *BAM Complex*, vol. 1329. 2015. p. 259–70.
- 13 Ujwal R, Abramson J. High-throughput Crystallization of Membrane Proteins Using the Lipidic Bicelle Method. *J Vis Exp* 2012:1–7.
<https://doi.org/10.3791/3383>.
- 14 Faham S. Crystallization of bacteriorhodopsin from bicelle formulations at room

- temperature. *Protein Sci* 2005;**14**:836–40. <https://doi.org/10.1110/ps.041167605>.
- 15 Kaya AI, Iverson TM, Hamm HE. Functional Stability of Rhodopsin in a Bicelle System: Evaluating G Protein Activation by Rhodopsin in Bicelles. *Rhodopsin Methods Protoc.*, vol. 1271. 2010.
- 16 Mercredi PY, Bucca N, Loeliger B, Gaines CR, Mehta M, Bhargava P, *et al.* Structural and Molecular Determinants of Membrane Binding by the HIV-1 Matrix Protein. *J Mol Biol* 2016;**428**:1637–55. <https://doi.org/10.1016/j.jmb.2016.03.005>.
- 17 Musatov A, Siposova K, Kubovcikova M, Lysakova V, Varhac R. Functional and structural evaluation of bovine heart cytochrome c oxidase incorporated into bicelles. *Biochimie* 2016;**121**:21–8. <https://doi.org/10.1016/j.biochi.2015.11.018>.
- 18 Tian H, Naganathan S, Kazmi MA, Schwartz TW, Sakmar TP, Huber T. Bioorthogonal fluorescent labeling of functional G-protein-coupled receptors. *ChemBioChem* 2014;**15**:1820–9. <https://doi.org/10.1002/cbic.201402193>.
- 19 Whiles JA, Deems R, Vold RR, Dennis EA. Bicelles in structure-function studies of membrane-associated proteins. *Bioorg Chem* 2002;**30**:431–42. [https://doi.org/10.1016/S0045-2068\(02\)00527-8](https://doi.org/10.1016/S0045-2068(02)00527-8).
- 20 Lühns T, Zahn R, Wüthrich K. Amyloid formation by recombinant full-length prion proteins in phospholipid bicelle solutions. *J Mol Biol* 2006;**357**:833–41. <https://doi.org/10.1016/j.jmb.2006.01.016>.
- 21 Gabriel NE, Roberts MF. Spontaneous Formation of Stable Unilamellar Vesicles. *Biochemistry* 1984;**23**:4011–5. <https://doi.org/10.1021/bi00313a001>.
- 22 Gabriel NE, Roberts MF. Interaction of Short-Chain Lecithin with Long-Chain Phospholipids: Characterization of Vesicles That Form Spontaneously.

- Biochemistry* 1986;**25**:2812–21. <https://doi.org/10.1021/bi00358a012>.
- 23 Lin TL, Liu CC, Roberts MF, Chen SH. Structure of mixed short-chain lecithin/long-chain lecithin aggregates studied by small-angle neutron scattering. *J Phys Chem* 1991;**95**:6020–7. <https://doi.org/10.1021/j100168a057>.
- 24 Luchette PA, Vetman TN, Prosser RS, Hancock REW, Nieh MP, Glinka CJ, *et al.* Morphology of fast-tumbling bicelles: A small angle neutron scattering and NMR study. *Biochim Biophys Acta - Biomembr* 2001;**1513**:83–94. [https://doi.org/10.1016/S0005-2736\(01\)00358-3](https://doi.org/10.1016/S0005-2736(01)00358-3).
- 25 Mineev KS, Nadezhdin KD, Goncharuk SA, Arseniev AS. Characterization of Small Isotropic Bicelles with Various Compositions. *Langmuir* 2016;**32**:6624–37. <https://doi.org/10.1021/acs.langmuir.6b00867>.
- 26 Ram P, Prestegard JH. Magnetic field induced ordering of bile salt/phospholipid micelles: new media for NMR structural investigations. *BBA - Biomembr* 1988;**940**:289–94. [https://doi.org/10.1016/0005-2736\(88\)90203-9](https://doi.org/10.1016/0005-2736(88)90203-9).
- 27 Sanders CR, Prestegard JH. Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO. *Biophys J* 1990;**58**:447–60. [https://doi.org/10.1016/S0006-3495\(90\)82390-0](https://doi.org/10.1016/S0006-3495(90)82390-0).
- 28 Sanders CR, Schwonek JP. Characterization of Magnetically Orientable Bilayers in Mixtures of Dihexanoylphosphatidylcholine and Dimyristoylphosphatidylcholine by Solid-State NMR. *Biochemistry* 1992;**31**:8898–905. <https://doi.org/10.1021/bi00152a029>.
- 29 Shintani M, Matubayasi N. Morphology study of DMPC/DHPC mixtures by solution-state ¹H, ³¹P NMR, and NOE measurements. *J Mol Liq* 2015;**217**..

- 30 Triba MN, Warschawski DE, Devaux PF. Reinvestigation by phosphorus NMR of lipid distribution in bicelles. *Biophys J* 2005;**88**:1887–901. <https://doi.org/10.1529/biophysj.104.055061>.
- 31 Vold RR, Prosser RS, Deese AJ. Isotropic solutions of phospholipid bicelles: A new membrane mimetic for high-resolution NMR studies of polypeptides. *J Biomol NMR* 1997;**9**:329–35. <https://doi.org/10.1023/A:1018643312309>.
- 32 Vold RR, Prosser RS. Magnetically Oriented Phospholipid Bilayered Micelles for Structural Studies of Polypeptides. Does the Ideal Bicelle Exist? *J Magn Reson* 1996;**113**:267–71.
- 33 Glover KJ, Whiles JA, Wu G, Yu NJ, Deems R, Struppe JO, *et al.* Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophys J* 2001;**81**:2163–71. [https://doi.org/10.1016/S0006-3495\(01\)75864-X](https://doi.org/10.1016/S0006-3495(01)75864-X).
- 34 Beaugrand M, Arnold AA, Hénin J, Warschawski DE, Williamson PTF, Marcotte I. Lipid concentration and molar ratio boundaries for the use of isotropic bicelles. *Langmuir* 2014;**30**:6162–70. <https://doi.org/10.1021/la5004353>.
- 35 Piai A, Ru Q, Dev J, Chou JJ. Optimal Bicelle q for Solution NMR Studies of Protein Transmembrane Partition. *Chemistry (Easton)* 2017;**23**:1361–7. <https://doi.org/10.1183/09031936.00063810.The>.
- 36 Sanders CR, Landis GC. Facile Acquisition and Assignment of Oriented Sample NMR Spectra for Bilayer Surface-Associated Proteins. *J Am Chem Soc* 1994;**116**:6470–1. <https://doi.org/10.1021/ja00093a072>.
- 37 Oliver RC, Lipfert J, Fox DA, Lo RH, Kim JJ, Doniach S, *et al.* Tuning micelle

- dimensions and properties with binary surfactant mixtures. *Langmuir* 2014;**30**:13353–61. <https://doi.org/10.1021/la503458n>.
- 38 Columbus L, Lipfert J, Jambunathan K, Fox DA, Sim AYL, Doniach S, *et al.* Mixing and matching detergents for membrane protein NMR structure determination. *J Am Chem Soc* 2009;**131**:7320–6. <https://doi.org/10.1021/ja808776j>.
- 39 Sanders CR, Prosser RS. Bicelles: A model membrane system for all seasons? *Structure* 1998;**6**:1227–34. [https://doi.org/10.1016/S0969-2126\(98\)00123-3](https://doi.org/10.1016/S0969-2126(98)00123-3).
- 40 Diller A, Loudet C, Aussenac F, Raffard G, Fournier S, Laguerre M, *et al.* Bicelles: A natural ‘molecular goniometer’ for structural, dynamical and topological studies of molecules in membranes. *Biochimie* 2009;**91**:744–51. <https://doi.org/10.1016/j.biochi.2009.02.003>.
- 41 Sanders CR, Hare BJ, Howard KP, Prestegard JH. Magnetically-oriented phospholipid micelles as a tool for the study of membrane-associated molecules. *Prog Nucl Magn Reson Spectrosc* 1994;**26**:421–44. [https://doi.org/10.1016/0079-6565\(94\)80012-X](https://doi.org/10.1016/0079-6565(94)80012-X).
- 42 Caldwell TA, Baoukina S, Brock AT, Oliver RO, Root KT, Krueger JK, *et al.* Low-q Bicelles are Mixed Micelles. *J Phys Chem Lett* 2018;**9**:4469–73.
- 43 Chaudhuri BN. Emerging applications of small angle solution scattering in structural biology. *Protein Sci* 2015;**24**:267–76. <https://doi.org/10.1002/pro.2624>.
- 44 Chaudhuri B, Munoz IG, Qian S, Urban VS. *Biological Small Angle Scattering: Techniques, Strategies and Tips*. vol. 1009. 2017.
- 45 Lipfert J, Columbus L, Chu VB, Lesley SA, Doniach S. Size and shape of

- detergent micelles determined by small-angle X-ray scattering. *J Phys Chem B* 2007;**111**:12427–38. <https://doi.org/10.1021/jp0730161>.
- 46 Oliver RC, Lipfert J, Fox DA, Lo RH, Doniach S, Columbus L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS One* 2013;**8**:. <https://doi.org/10.1371/journal.pone.0062488>.
- 47 Research NC for N. *A SANS Experiment to Understand the Nature of Surfactant Stabilization in Suspensions of Spherical Colloidal Silica Particles*. 2004.
- 48 Gabel F. Biological Small Angle Scattering: Techniques, Strategies and Tips 2017;**1009**:201–14. <https://doi.org/10.1007/978-981-10-6038-0>.
- 49 Almgren M, Garamus VM, Asakawa T, Nan J. Contrast variation SANS investigation of composition distributions in mixed surfactant micelles. *J Phys Chem B* 2007;**111**:7133–41. <https://doi.org/10.1021/jp070271x>.
- 50 Brasher LL, Kaler EW. A Small-Angle Neutron Scattering (SANS) Contrast Variation Investigation of Aggregate Composition in Catanionic Surfactant Mixtures. *Langmuir* 2002;**12**:6270–6. <https://doi.org/10.1021/la960493e>.
- 51 *Sasview*. n.d.
- 52 Kucerka N, Kiselev M a, Balgavý P. Determination of bilayer thickness and lipid surface area in unilamellar dimyristoylphosphatidylcholine vesicles from small-angle neutron scattering curves: a comparison of evaluation methods. *Eur Biophys J* 2004;**33**:328–34. <https://doi.org/10.1007/s00249-003-0349-0>.
- 53 Kiselev MA, Zemlyanaya EV, Aswal VK. Sans study of the unilamellar DMPC vesicles. The fluctuation model of lipid bilayer. *Crystallogr Reports* 2004:S131–6.
- 54 Kučerka N, Tristram-Nagle S, Nagle JF. Structure of fully hydrated fluid phase

- lipid bilayers with monounsaturated chains. *J Membr Biol* 2006;**208**:193–202.
<https://doi.org/10.1007/s00232-005-7006-8>.
- 55 Tanford C. *The hydrophobic effect: formation of micelles and biological membranes*. 1973.
- 56 Whitten AE, Cai S, Trehwella J. MULCh : modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies . *J Appl Crystallogr* 2008;**41**:222–6. <https://doi.org/10.1107/s0021889807055136>.
- 57 Nagle JF, Tristram-Nagle S. Structure of lipid bilayers. *Biochim Biophys Acta - Rev Biomembr* 2000;**1469**:159–95. [https://doi.org/10.1016/S0304-4157\(00\)00016-2](https://doi.org/10.1016/S0304-4157(00)00016-2).
- 58 Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 2008;**4**:435–47. <https://doi.org/10.1021/ct700301q>.
- 59 Jämbeck JPM, Lyubartsev AP. Another piece of the membrane puzzle: Extending slipids further. *J Chem Theory Comput* 2013;**9**:774–84.
<https://doi.org/10.1021/ct300777p>.
- 60 Jämbeck JPM, Lyubartsev AP. An extension and further validation of an all-atomistic force field for biological membranes. *J Chem Theory Comput* 2012;**8**:2938–48. <https://doi.org/10.1021/ct300342n>.
- 61 Jämbeck JPM, Lyubartsev AP. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J Phys Chem B* 2012;**116**:3164–79. <https://doi.org/10.1021/jp212503e>.
- 62 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison

- of simple potential functions for simulating liquid water. *J Chem Phys* 1983;**79**:926–35. <https://doi.org/10.1063/1.445869>.
- 63 Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys* 2007;**126**:. <https://doi.org/10.1063/1.2408420>.
- 64 Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* 1981;**52**:7182–90. <https://doi.org/10.1063/1.328693>.
- 65 Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* 1993;**98**:10089–92. <https://doi.org/10.1063/1.464397>.
- 66 Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys* 1995;**103**:8577–93. <https://doi.org/10.1063/1.470117>.
- 67 Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. MDAAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J Comput Chem* 2011;**32**:. <https://doi.org/10.1002/jcc>.
- 68 Schneidman-Duhovny D, Hammel M, Sali A. FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* 2010;**38**:540–4. <https://doi.org/10.1093/nar/gkq461>.
- 69 Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *J Mol Graph* 1996. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- 70 Macdonald AG, Wahle KWJ, Cossins AR, Behan MK. Temperature, Pressure and Cholesterol Effects on Bilayer Fluidity; a Comparison of Pyrene

- Excimer/Monomer Ratios with the Steady-State Fluorescence Polarization of Diphenylhexatriene in Liposomes and Microsomes. *Biochim Biophys Acta* 1988;**938**:231–42.
- 71 Nelson SC, Neeley SK, Melonakos ED, Bell JD, Busath DD. Fluorescence anisotropy of diphenylhexatriene and its cationic Trimethylamino derivative in liquid dipalmitoylphosphatidylcholine liposomes: Opposing responses to isoflurane. *BMC Biophys* 2012;**5**:5. <https://doi.org/10.1186/2046-1682-5-5>.
- 72 Shinitzky M, Barenholz Y. Fluidity parameters of lipid regions determined by fluorescence polarization. *BBA - Rev Biomembr* 1978;**515**:367–94. [https://doi.org/10.1016/0304-4157\(78\)90010-2](https://doi.org/10.1016/0304-4157(78)90010-2).
- 73 Kot EF, Goncharuk SA, Arseniev AS, Mineev KS. Phase Transitions in Small Isotropic Bicelles. *Langmuir* 2018;**34**:3426–37. <https://doi.org/10.1021/acs.langmuir.7b03610>.
- 74 Morrison EA, Henzler-Wildman KA. Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile. *Biochim Biophys Acta - Biomembr* 2012;**1818**:814–20. <https://doi.org/10.1016/j.bbamem.2011.12.020>.
- 75 Laguerre A, Löhr F, Henrich E, Hoffmann B, Abdul-Manan N, Connolly PJ, *et al.* From Nanodiscs to Isotropic Bicelles: A Procedure for Solution Nuclear Magnetic Resonance Studies of Detergent-Sensitive Integral Membrane Proteins. *Structure* 2016;**24**:1830–41. <https://doi.org/10.1016/j.str.2016.07.017>.