

**Exploring BtuB in *Escherichia coli* via a new direct spin labeling approach  
and electron paramagnetic resonance spectroscopy**

Thushani Dilanka Nilaweera

Colombo, Sri Lanka

Master of Science in Chemistry, University of Iowa, 2013

Bachelor of Science in Biochemistry and Molecular Biology

University of Colombo, 2009

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

Department of Chemistry

University of Virginia

December, 2018



## Abstract

The TonB dependent transporters (TBDTs) are sophisticated machineries which enable active transport across the outer membrane (OM) of Gram-negative bacteria by utilizing the inner membrane's proton motive force through the TonB complex. Significant research has focused on understanding these transporters within isolated systems that do not truly mimic the physiological OM nor possess the required vital protein players for TBDTs to properly function. However, recently it has been possible to study a TBDT, the vitamin B<sub>12</sub> transporter, BtuB, directly in bacteria using electron paramagnetic resonance (EPR) spectroscopy.

Here an alternative spin labeling approach was developed in order to explore BtuB expressed in metabolically active *Escherichia coli* using continuous wave (CW) and pulsed EPR techniques. We have successfully developed the approach to spin labeled single and double cysteine residues that face the extracellular space of BtuB without significant background labeling. Labeling pairs of spin labels in BtuB proved to be particularly challenging because it required the use of mutants that alter the redox homeostasis in the periplasm. We demonstrated the utility of our approach in several ways. First, single spin labels on BtuB were shown to produce intermolecular double electron-electron resonance (DEER) data that cannot be represented by a purely random distribution of the protein, and that the data are most consistent with a crowding or clustering of the protein in the OM. Such clustering has been a topic of considerable current interest as OM protein (OMP) clustering has been implicated as a mechanism to drive protein turnover in the bacterial OM. There is little direct information on the molecular interactions that drive this association and EPR provides an approach to explore these interactions. In a second set of experiments, double spin-labeled BtuB were used to explore novel conformational states associated with the transport process. The DEER data for a spin pair where one label (T188R1) is in an extracellular loop and a second label is within the protein core

(V90R1) reveals two distance population where the short distance represents a previously unseen conformational state placing loop two of BtuB in close proximity to the core. Further investigations support that the core moves towards loop two and that this population increases in the presence of B<sub>12</sub>. Thus, this new conformational state is likely involved in the substrate transport and/or signal transduction processes. Third, we have shown that this system can be used for the study of the binding and position of colicin E3, a bacteriocidal protein produced by *E. coli* that uses BtuB as its primary receptor to target and kill other *E. coli*. Based on our preliminary work using a short fragment of the colicin E3 receptor domain (E3R), the spin labeled R domain is found to bind to BtuB, both in live *E. coli* cells and in the reconstituted bilayer systems. The second extracellular loop is unstructured in the crystal structure; however, based on the preliminary DEER data, this loop is in close proximity to colicin E3R.

A final topic addressed in this thesis involves the mechanism of signal transduction through BtuB, which involves a substrate-dependent unfolding of the Ton box, which couples BtuB to TonB. During the signal transduction processes, and possibly during transport, ion pairs may act as molecular switches. Based on site directed spin labeling (SDSL) coupled to EPR, the R14-D316 pair acts as a conformational switch, which is involved in the TonB box unfolding process. However, five conserved ion pair (R36-D515, R47-D548, D53-R526, R69-E419, and R111-E465) mutants inhibit the TonB box unfolding event even in the presence of B<sub>12</sub>. Interestingly, the R69, R111, E419, and E465 residues may form interchangeable salt bridges, where the signal can be propagated sequentially. Some of these alanine mutants altered the substrate binding, which affected the TonB box unfolding. This work demonstrates that charged side chains and ion pairs within the BtuB play a critical role in the function of BtuB.

## Table of contents

<b>Abstract</b>	<b>iii</b>
<b>Table of contents</b>	<b>v</b>
<b>List of figures</b>	<b>viii</b>
<b>List of tables</b>	<b>xv</b>
<b>Abbreviations</b>	<b>xvi</b>
<b>Acknowledgments</b>	<b>xviii</b>
<b>Chapter 1. Introduction</b>	<b>1</b>
1.1 Significance	1
1.2 Bacterial physiology	4
1.3 Gram-negative bacterial membranes	12
1.4 OM biogenesis	20
1.5 TonB dependent transporters	39
1.6 Bacteriocins	48
1.7 References	55
<b>Chapter 2. Materials and main techniques</b>	<b>61</b>
2.1 Materials	61
2.2 Molecular biology techniques	62
2.3 Microbiological techniques	73
2.4 Biochemical and biophysical techniques	80
2.5 EPR spectroscopy as a tool to study protein systems	87
2.6 References	101

<b>Chapter 3. Development of new sample preparation approach for EPR analysis using spin labeled live bacterial cells to investigate outer membrane transporters</b>	<b>104</b>
3.1 Abstract	104
3.2 Introduction	105
3.3 Methods	108
3.4 Results and discussion	118
3.5 Applications of finalized protocols	152
3.6 Factors to consider for future spin labeling of OMPs in live cells	153
3.7 References	157
<b>Chapter 4. Exploring live Gram-negative bacterial outer membrane proteins’ hierarchical organizations into domains</b>	<b>160</b>
4.1 Abstract	160
4.2 Introduction	161
4.3 Methods	165
4.4 Results and Discussion	168
4.5 Future directions	175
4.6 References	177
<b>Chapter 5. Exploring novel conformational state(s) of spin labeled BtuB in <i>Escherichia coli</i> cells.</b>	<b>179</b>
5.1 Abstract	179
5.2 Introduction	180
5.3 Methods	181

5.4 Results and Discussion	185
5.5 Future directions	199
5.6 References	201
<b>Chapter 6. Exploring the potential roles of BtuB conserved ion pairs during the signal transduction and/or substrate transport process</b>	<b>203</b>
6.1 Abstract	203
6.2 Introduction	204
6.3 Methods	207
6.4 Results and discussion	211
6.5 Future directions	227
6.6 References	229
<b>Appendix 1. Biophysical characterization of Colicin E3 receptor domain binding to the primary receptor, BtuB</b>	<b>230</b>
Abstract	230
Introduction	231
Methods	237
Results and Discussion	242
Future Directions	247
References	251
Appendix 2	253
Appendix 3	256
Appendix 4	258

## Table of figures

Figure 1.1. Bacterial cell cycle.	5
Figure 1.2. Chromosome replication in bacteria.	5
Figure 1.3. Bacterial growth curve.	9
Figure 1.4. The structure of typical LPS molecule consists of lipid A, a core unit, and an <i>O</i> -antigen.	13
Figure 1.5. Structures of different lipid A molecules.	14
Figure 1.6. Structure of rough oligosaccharide core of <i>E. coli</i> K-12 strain which consists of inner and outer core.	15
Figure 1.7. The structures of different lipids in the bacterial OM.	15
Figure 1.8. Crystal structures of a few OMPs.	19
Figure 1.9. The Lpt machinery and the LptD-E complex.	22
Figure 1.10. Transport of OMP and LP.	24
Figure 1.11. Crystal structure of Skp.	25
Figure 1.12. Crystal structure of DegP.	26
Figure 1.13. Crystal structures of SurA and FkpA.	27
Figure 1.14. The crystal structures of LolA and LolB.	28
Figure 1.15. The crystal structures of the BAM complex side views.	29
Figure 1.16. The crystal structures of BamA and accessory proteins.	31
Figure 1.17. Schematic representation of budding model.	32
Figure 1.18. Protein export into periplasm as post-translationally and co-translationally and undergo oxidative folding.	35
Figure 1.19. The conventional disulfide bond formation pathway in bacteria.	37
Figure 1.20. Structures of different siderophores synthesized by bacteria.	41
Figure 1.21. Structure of apo and ferrichrome-iron bound FhuA.	43
Figure 1.22. Structure of apo and ferric citrate bound FecA.	44
Figure 1.23. Structure of apo BtuB with calcium and CN-Cb bound BtuB.	45
Figure 1.24. Crystal structures of group A colicins.	51
Figure 1.25. Proposed mechanism for translocation of colicin E through OM.	52
Figure 1.26. Crystal structures of group B colicins.	53

Figure 2.1. DpnI recognition site.	63
Figure 2.2. Primer designing for V90C BtuB mutant.	64
Figure 2.3. Plasmid map of pET28.	69
Figure 2.4. Translated DNA sequence results of V90C-BtuB mutant using the Expsy bioinformatics resource portal.	72
Figure 2.5. NCBI protein blast is used for sequence alignment between the wild type and the translated plasmid sequence to identify the BtuB mutation.	72
Figure 2.6. A chromatogram of the DNA sequence visualized by Finch TV software.	73
Figure 2.7. Schematic representation of the heat-shock transformation process.	76
Figure 2.8. Different plating techniques.	78
Figure 2.9. Structures of n-octyl- $\beta$ -D-glucopyranoside and sodium dodecyl sulfate.	81
Figure 2.10. Schematic representation of affinity chromatography purification.	82
Figure 2.11. Elution profile of an IMAC protein purification using a gradient elution.	83
Figure 2.12. The net charge of protein changes depending on the pH of the surroundings.	84
Figure 2.13. Structures of 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine (DOPC) and 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine (POPC).	85
Figure 2.14. Different lipid systems.	86
Figure 2.15. Detergent mediated reconstitution of membrane proteins.	86
Figure 2.16. Structures of some nitroxide spin labels.	88
Figure 2.17. Spin labeling at the sulfhydryl group of a cysteine residue via disulfide bond formation with MTSSL generates an R1 side chain.	89
Figure 2.18. The rotating frames of on-resonance and off-resonance magnetization vectors.	91
Figure 2.19. The energy diagram for spin system with $m_s = \pm \frac{1}{2}$ and $m_I = 1$ .	93

Figure 2.20. Coordinates of nitroxide spin and the CW EPR spectra recorded along x-, y- and z-axis.	94
Figure 2.21. EPR spectra obtained at fast (10-400 ps), intermediate (1-100 ns) and rigid-limit ( $> 1 \mu\text{s}$ ).	95
Figure 2.22. EPR lineshape analysis.	96
Figure 2.23. Four-pulse sequence of the DEER experiment.	98
Figure 2.24. Dipole-dipole coupling between two electron spins (S1 and S2) in the DEER experiment that are separated by $r_{12}$ distance in the z-axis.	99
Figure 2.25. Obtaining $F(t)$ from DEER signal.	100
Figure 3.1. The CW EPR spectra of spin labeled WT BtuB, T188C BtuB, the structure of TEMPO-B <sub>12</sub> . Normalized form factors after background subtraction for T188R1 BtuB with TEMPO-B <sub>12</sub> in cells, intact OM and in reconstituted lipid bilayer system.	106
Figure 3.2. Structure of BtuB indicating extracellular side exposed S74, V90, T188, A288, D492 and S533.	107
Figure 3.3. Normalized CW EPR spectra obtained from O/N grown RK5016 cells expressing BtuB without and with L8P after free spin subtraction.	119
Figure 3.4. Normalized CW EPR of A288R1, L8P-A288R1, T188R1, L8P-T188R1, S533R1 BtuB of O/N grown cells.	120
Figure 3.5. Normalized CW EPR without free spin subtraction, obtained from T188R1-BtuB of cells harvested at different inoculation time points.	122
Figure 3.6. Normalized CW EPR spectra of T188C BtuB spin labeled for one and two hours at RT.	124
Figure 3.7. Normalized CW EPR of TEMPO-B <sub>12</sub> binding with T188C BtuB cells.	125
Figure 3.8. Normalized CW EPR spectra of T188R1 BtuB in live <i>E. coli</i> cells resuspended in 100 mM HEPES pH 7.0 and incubated on ice.	127
Figure 3.9. Normalized CW EPR spectra of spin labeled T188R1 BtuB in live cell pellets incubated on ice in the repeat experiment.	128
Figure 3.10. Schematic representation of OMPs clustering in OMP island with	129

heterogeneously distributed LPS within the islands.	
Figure 3.11. Echo signals obtained from the pulsed EPR experiment for ice incubated (20 min and 60 min) cells expressing T188R1 BtuB.	130
Figure 3.12. Normalized CW EPR spectra of T188R1 BtuB in live cells incubated in different wash buffers (MES pH 5.5 and HEPES pH 7.0).	131
Figure 3.13. Normalized CW EPR spectra of L8P-T188R1, A288R1, L8P-A288R1, and D492R1 BtuB after incubation in 100 mM MES buffer, pH 5.5 for an hour on ice.	132
Figure 3.14. DEER data for TEMPO-B <sub>12</sub> bound to T188R1, L8P-T188R1, A288R1, and L8P-A288R1 BtuB in live cells.	133
Figure 3.15 Normalized CW EPR spectra of spin labeled T188C, S74C and V90C BtuB expressing cells collected at OD <sub>600</sub> 0.3 and 0.6.	134
Figure 3.16. Top view of apo BtuB structure showing V90C-T188C and S74-T188 residue sites used for spin labeling in live <i>E. coli</i> cells.	136
Figure 3.17. RK5016 cells expressing V90C-T188C-BtuB spin labeled with range of MTSSL.	137
Figure 3.18. Top view of BtuB surface.	138
Figure 3.19. The normalized CW EPR spectra of WT BtuB treated with 100 mM DTT and spin labeled using excess MTSSL.	139
Figure 3.20. Normalized CW EPR of V90C-T188C BtuB cells treated with 100 mM DTT prior to spin labeling.	140
Figure 3.21. Normalized CW EPR spectra of 10 mM DTT constant and incremental increase and reduction in culture with cells expressing S74C-T188C BtuB.	141
Figure 3.22. Normalized CW EPR spectra of S74C-T188C BtuB of live cells spin labeled on ice.	143
Figure 3.23. Normalized CW EPR spectra of S74C-T188C BtuB of live cells spin labeled at RT.	143
Figure 3.24. Structure of apo BtuB showing residues V90 on core and T188 and G399 on the barrel.	144

Figure 3.25. Normalized CW EPR spectra of V90C-T188C BtuB in live WT, <i>dsbA</i> <sup>-</sup> , <i>dsbB</i> <sup>-</sup> and <i>dsbC</i> <sup>-</sup> cells, spin labeled with MTSSL.	146
Figure 3.26. Normalized CW EPR of V90C-T188C BtuB spin labeling using range of MTSSL per <i>dsbA</i> <sup>-</sup> cell pellet obtained from a 50 mL cell culture at OD <sub>600</sub> around 0.3.	148
Figure 3.27. Normalized CW EPR spectra for spin labeled RK5016 cells and <i>dsbA</i> <sup>-</sup> cells expressing S74C-T188C BtuB using 0.05 mg or 0.1 mg MTSSL and V90C-T188C BtuB using 0.05 or 0.1 mg MTSSL.	149
Figure 3.28. Side view of apo BtuB showing residues S74, V90 and T188 in the core and the barrel.	150
Figure 3.29. Normalized CW EPR spectra of spin labeled WT BtuB (control) and V90C-T188C BtuB expressed in Dsb mutant strains.	151
Figure 3.30. DEER signal of V90R1-T188R1 BtuB.	152
Figure 4.1. Asymmetric bacterial OM consists of an inner leaflet containing phospholipids and outer leaflet containing LPS.	162
Figure 4.2. LPS heterogeneously distributed in OM.	163
Figure 4.3. Echo signals obtained from the pulsed EPR experiment for T188R1 BtuB with spin labeled B <sub>12</sub> analyzed using DEER analysis and DD.	169
Figure 4.4. Echo signals obtained from the pulsed EPR experiment for RK5016 cells expressing T188C BtuB harvested at different growth time points.	170
Figure 4.5. Echo signals obtained from the pulsed EPR experiment for cells expressing T188C BtuB incubated on ice after spin labeling step.	170
Figure 4.6. Echo signals obtained from the pulsed EPR experiment for cells expressing V90C BtuB incubated on ice after spin labeling step.	171
Figure 4.7. Schematic representation of intermolecularly interacting T188R1 spin labeled BtuB in the OM.	172
Figure 4.8. Echo signals obtained from the pulsed EPR experiment for cells expressing T188C BtuB in deuterated main cultures.	173

Figure 5.1. Apo BtuB structure showing top view and side view of apo BtuB.	181
Figure 5.2. The extracted distance distributions from DEER experiments for apo and B <sub>12</sub> added V90R1-T188R1 BtuB in <i>E. coli</i> cells and intact OM.	185
Figure 5.3. Top view of BtuB surface.	186
Figure 5.4. DEER signal of V90R1-T188R1 BtuB in <i>dsbA</i> - cells and incubated on ice and at RT during the washing step.	187
Figure 5.5. Structure of apo BtuB showing residues V90 on core and T188 and G399 on the barrel.	188
Figure 5.6. DEER signals obtained for T188R1-G399R1 BtuB incubated on ice.	189
Figure 5.7. Extracted distance distributions from DEER experiments of V90R1-T188R1 BtuB and T188R1-G399R1 BtuB.	190
Figure 5.8. DEER signal obtained for V90R1-T188R1 BtuB on ice and at RT with varying the substrate concentration (0-100 $\mu$ M B <sub>12</sub> ).	191
Figure 5.9. DEER signals collected from L8P-V90R1-T188R1 BtuB on ice and at RT with and without 100 $\mu$ M B <sub>12</sub> .	193
Figure 5.10. DEER signal of V90R1-T188R1 BtuB expressed in <i>dsbA</i> - cells treated with CCCP (50 $\mu$ M) and KCN (1 mM).	195
Figure 5.11. Top view of apo BtuB showing aspartic acid, glutamic acid and arginine residues in the BtuB core.	196
Figure 5.12. The apo BtuB showing potential ion pair forming residues near V90.	197
Figure 6.1 Structures of BtuB periplasmic view apo and B <sub>12</sub> bound.	204
Figure 6.2. Normalized CW EPR spectra of apo and substrate bound BtuB.	205
Figure 6.3 Conserved ion pairs in BtuB. Structures of BtuB periplasmic view of apo and B <sub>12</sub> bound.	206
Figure 6.4. Normalized EPR spectra of alanine mutations coupled to V10R1 BtuB.	212-213
Figure 6.5. Structure of apo BtuB looking from the periplasmic side.	214
Figure 6.6. Normalized EPR spectra of the alanine mutations coupled to V10R1	215

BtuB.	
Figure 6.7. Spectral subtraction using FreeBee software.	216
Figure 6.8. Structure of BtuB from periplasmic side showing the arginine, lysine, aspartic acid, and glutamic acid residues in the core and the barrel near sites R111 and E419.	220
Figure 6.9. Two mutant cycle analysis of $\Delta^\circ G$ calculated based on the conformational equilibria of the apo and B <sub>12</sub> bound states.	221-222
Figure 6.10. (a) Structure of TEMPO-B <sub>12</sub> and (b) the normalized CW EPR spectra of V10C BtuB bound to TEMPO-B <sub>12</sub> .	224
Figure 6.11. Normalized CW EPR spectra of TEMPO-B <sub>12</sub> binding to the ion pair mutants.	226
Figure A1.1. Structures of colicin E3.	232
Figure A1.2. Structural alignment of apo BtuB and colicin bound BtuB with labeled $\beta$ strands.	233
Figure A1.3. Side view of trimeric OmpF and poly-glycine peptide of colicin T83 fragment is occluding the pore of OmpF.	234
Figure A1.4. Side view of colicin R135 fragment bound to BtuB.	235
Figure A1.5. Structure of colicin R135 bound to BtuB.	242
Figure A1.6 Normalized CW EPR spectra obtained from spin labeled colicin E3R mutants.	243
Figure A1.7. Structural alignment of apo (Ca <sup>2+</sup> bound) BtuB and colicin R135 bound BtuB.	245
Figure A1.8. Echo signal obtained from pulsed EPR experiment for colicin K363R1 incubated with reconstituted T188R1 BtuB.	246
Figure A1.9. Echo signal obtained from pulsed EPR experiment for colicin A370R1 incubated with reconstituted T188R1 BtuB.	246
Figure A1.10. The sequence alignment of colicin E9 and E3 (residue 1-65).	249
Figure A1.11. Schematic representation of colicin E3 binding and translocation events.	250

## List of tables

Table 2.1. Stock and final concentrations of reagents used in PCR mutagenesis	67
Table 2.2. PCR cycle steps and the temperatures and corresponding time for each step used in this thesis work.	68
Table 6.1. The unfolded populations (x) in the apo and substrate bound states for each single and double alanine mutant obtained from the spectral subtraction of the mobile component from the normalized spectra.	217
Table A1.1. The distances between S74 and V90 BtuB in the core domain with colicin residues. The distances are based on the crystal structure 1UJW.	248

### Table of Abbreviations

OM	Outer membrane
Ab	Antibody
Ag	Antigen
ArcAB	Aerobic respiration control system
BAM	$\beta$ -barrel assembly machinery
CM	Cell membrane
CW	Continuous wave
DEER	Double electron-electron resonance
Dsb	Disulfide bond
DegP	Serine endoprotease
DTT	Dithiothreitol
EPR	Electron paramagnetic resonance
ETC	Electron transfer chain
Fnr	Fumarate nitrate regulation protein
Fur	Ferric uptake regulation protein
FkpA	FkpB binding protein A
IM	Inner membrane
IMAC	Immobilized metal affinity chromatography
LPS	Lipopolysaccharides
Lpt	LPS transport machinery
Lol	Lipoprotein localization machinery

MTSSL	S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonylthioate
OG	<i>n</i> -octyl- $\beta$ -D-glucopyranoside
OMPs	Outer membrane proteins
PG	Phosphatidylglycerol
pmf	Proton motive force
POTRA	Polypeptide transport-associated domains
PPI	Protein-protein interactions
PIPE PCR	polymerase incomplete primer extension polymerase chain reaction
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
(p)ppGpp	Guanosine tetraphosphate and guanosine pentaphosphate
ROS	Reactive oxygen species
RNS	Reactive nitrogen species
RT	Room temperature
SAM	Sorting and assembly machinery
SDSL	Site-directed spin labeling
Skp	Seventeen-kilodalton protein
SoxR	Superoxide regulation protein
SurA	Survival protein A
TBDTs	TonB dependent transporters
TOC	Translocon at the outer envelope of chloroplasts
TEMPO-B <sub>12</sub>	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl labeled cyanocobalamin
WT	Wild type

## Acknowledgment

First of all, I am thankful to Professor D. Cafiso, who gave me the opportunity to pursue my graduate studies. Being a great mentor, he allowed me to flourish in science while increasing the expectation bar higher every day. A great example of this is my thesis. When I first joined the Cafiso group as a graduate student, he only asked me to work on one project (Chapter 6) as to obtain the PhD under the circumstances that I joined his team. But, being a great supervisor, he let me work on many projects and to deepen my understanding on membrane proteins, as well as, to become an experienced scientist in the area that I am good in. For all, THANK YOU DAVE.

Thank you to all the current members in Cafiso group, Vanessa Bijak, Jacob Staley, Qian Liang, Dr. Sarah Nyenhuis and David Nyenhuis. I am fortunate, as I joined the Cafiso group at the correct time. To have unlimited support from my dear friends, Sarah, David and Vanessa. Thank you, Vanessa, my yoga buddy, for being there when I needed to calm down from stressful times. Thank you, Sarah and David, for many helpful discussions with EPR and other topics.

I am thankful for my partner in scientific crime, David, who being so supportive throughout the time in our (never ending) joint projects. Easy going with the experiment plans I came up with, which some days were extending long hours into late night, but successful while some turned out to be disappointments or me cramming many experiments into a single day. During all these labor-intensive experiments, he was on board. Also, need to thank Sarah and my husband, Adam Buffington, for supporting our late-night experiments, and patiently staying in the lab for moral support.

I am blessed to have close mentoring by Professor R. Nakamoto, who was and is continuously supervising me and providing helpful discussions and coming to our group

meetings when David and I were presenting our research projects. He never runs out of ideas nor suggestions and is always kind to provide many different bacterial strains needed for my research work.

At the same time, I would like to thank my thesis committee members, Dr. Nakamoto, Dr. Hunt, Dr. Pompano and Dr. Gahlmann for the support and the guidance throughout the thesis defense and Susie Marshall for all the support throughout the years and especially during thesis time.

My time before becoming a grad student in Cafiso lab, was quite a challenge. As now I think it was the best way for me to gain the needed skills, networks and to join them in the most appropriate time line. Thus, there are many wonderful people that I should express my gratitude for. During the time of relocating from Iowa to Virginia, as I was in the middle of UI Chemistry PhD program, UVa chemistry department provided me the opportunity to pursue the graduate studies. I would like to thank Professor Gunnoe and everyone involved in the process to offer me a position in the graduate program within a very short time period. At the same time, I want to thank from the bottom of my heart to my former PhD supervisor, Professor Amnon Kohen, University of Iowa, for being one of the great mentors, as well as, a great supporter who highly respects and values family. His understanding allowed me to smoothly wrap the research work in Kohen lab and to open a new chapter in my scientific life. Furthermore, I am grateful for former Kohen group members, especially Dr. M. Saeed for guiding me during my time in Kohen laboratory.

I am thankful for Dr. J. Patterson for training me and sharpening my molecular biology and biochemistry techniques. Moreover, during the short period of time I was under her mentoring, her strong work ethics, time management and efficient multitasking tactics strongly nourished

me to survive the rough journey I took to reach my scientific accomplishments. Hopefully one day, I can reach her level of efficient working in science. Even though, my journey was bittersome than many who follow the PhD path, I am so fortunate to have follow grad students who were there to support me throughout those times. Many thanks to former and current graduate students in Columbus and Mura groups.

I am so grateful to Professor I. Harrison for guiding me through hard time of my grad life. His understanding, kindness and support along with my keen desire for scientific research allowed me to strongly rise up and to find a new research group. Also, I want to thank, Professor R. Pompano and Professor D. Cafiso for their support and providing new research ideas and kind enough to recruit me into their teams, so I can complete my graduate studies.

Besides that, there is one more person in the department, Dr. Sarah Nyenhuis, who supported me throughout this transition period. So grateful that she is my friend, who assured me that I will survive in an EPR lab. As I still recall, she was sure that I will be collecting data in a month. Surprisingly, it turned out to be so true in Cafiso lab! Your kind words and support helped me to quickly adjust into a completely opposite research field that I have never thought of doing nor been exposed to before.

Besides all these wonderful people, I would like to express my gratitude to those who were and are with me with endless support. I would like to thank from bottom of my heart to my parents for their unconditional love, support, sacrifices and their blessings. All of these strengthened me to never give up the pursuing PhD. Heartful gratitude to my husband, Adam, for being in the grad school roller-coaster for all these years and helping me in every possible way and who was the reason I started this whole adventure in UVa. Thank you. Otherwise, I would have become a radio-oncologist not a structural biologist.



## **Chapter 1. Introduction**

### 1.1 Significance

To survive, bacteria rapidly adopt to continuous external changes in their habitat. One of the best examples is when pathogenic bacterial strains develop drug resistance toward available antibiotics, potentially becoming superbugs. Many of these pathogens are Gram-negative bacteria, which have an outer membrane (OM) in addition to the cytoplasmic membrane. The OM acts as a permeable barrier where OM transport proteins are present and facilitate the import and export of various nutrients, wastes, pathogenic compounds, and antibacterial compounds. One such class of outer membrane proteins (OMPs) is the TonB dependent transporters, TBDTs, which will be the focus of this dissertation. This family of transporters is composed of sophisticated machineries which assemble into gated channels that can regulate specific cargo intake into the host. The role of TBDTs in scarce mineral and vitamin uptake is well recognized. In addition, these TBDTs can also uptake various bacteriocins into the host bacteria.

Structural and functional studies of OMPs, including TBDTs, have utilized many molecular biology, biochemical, and biophysical techniques. A significant amount of this research has focused on understanding these transporters in isolated systems, either in OM preparations or in reconstituted lipid bilayer systems. The structures of some of these OMPs have also been solved at an atomic resolution using crystallography. Even though crystallography provides a platform for understanding possible transport mechanisms, dynamic transitions in the protein are not revealed. The resulting snap-shots from crystallography sometimes result from the use of unnatural reagents during the crystallization process. Consequently, detecting the natural protein

conformations present under physiological conditions becomes an important problem for understanding transport mechanisms.

Imaging techniques, such as fluorescence labeling, have been used to study proteins under physiological conditions, including live bacterial cells. Fluorescent labeling can employ either a fluorophore attached to the OMP of interest or the addition of fluorescently labeled specific antibodies or substrates. However, there are inherent issues with these approaches, such as the size and type of incorporated fluorophores utilized and the requirement to eliminate any naturally interfering fluorophores from the protein of interest. In contrast to these techniques, site-directed mutagenesis coupled to an electron paramagnetic resonance (EPR) based labeling technique is an elegant way to label the OMP of interest. This technique utilizes small nitroxide spin labels linked to the site of interest, which is advantageous instead of linking large fluorophores used in fluorescent techniques.

This dissertation focuses on developing a new physiological approach to utilize live Gram-negative cells (*Escherichia coli*), which express the OMP of interest, coupled to investigations using these EPR techniques. In contrast to previous work on whole cell labeling, which focused on the feasibility of labeling bacterial OMPs, our current work focuses on a straight-forward approach to spin label an OMP of interest in bacteria that are physiologically and microbiologically active for direct analysis by continuous wave and pulsed EPR techniques. There are two main goals for this work. First, to investigate the transport mechanism of TBDTs using a vitamin B<sub>12</sub> transporter, BtuB, as a model system. This transport mechanism is still being debated in the scientific community. By employing physiologically active cells, one can obtain a potentially wide range of unseen intrinsic events associated with the transport system, even evidence of hierarchical OMP assemblies through use of spin labeled BtuB in live bacteria.

Second, to finetune the same approach in order to study opportunistic cargoes, bacteriocins. These antimicrobial compounds are produced by bacteria which target other bacterial strains and can also use many TBDTs to enter cells. Cellular entry mechanisms are still under investigation in the field. In the current research, a spin labeled bacteriocin, colicin E3, was used to screen binding events with its primary receptor, BtuB.

Previous work has demonstrated that both single and double cysteine mutants on BtuB can be labeled in bacterial cells. The optimization approaches described here generate metabolically active cells that are sensitive to periplasmic redox homeostasis and provide evidence for crowding and the existence of recently proposed OMP islands in Gram-negative bacteria. Furthermore, experiments using these cell preparations have led to the discovery of a potentially new conformational state(s) in BtuB that may underlie transport. This can deepen our understanding of intrinsic conformational states of TBDTs, which would otherwise be impossible to detect using isolated systems and inactive or dead cells. As indicated by these findings, our current method enables future structural and functional research work directly in live Gram-negative bacteria expressing the OMP of interest.

## 1.2 Bacterial physiology

Many cellular processes involve various biomolecules (proteins, enzymes, lipids and *etc.*) which function to maintain bacterial life, to grow, and to propagate all within a vast range of habitats. Regulation and maintenance of these key cellular events are synchronized via proper communication between external and internal environments. For example, to enhance the natural survival, specific genes are either activated or deactivated in bacteria.

### 1.2.1 Bacterial cell cycle

During the continuous battle of survival, bacteria undergo rapid modifications in response to varying environmental conditions, such as the availability of nutrients, density of bacterial population and temperature. These external environmental changes are detected by cells, which initiate various signaling pathways to regulate gene expression in order to increase or decrease the levels of specific proteins and/or enzymes. This ultimately promotes a rapid response from bacterial cells to environmental changes. Consequently, influencing the bacterial cell cycle. For example, under nutrient rich media cells grow quickly with a short doubling time, while in nutrient poor media cells grow slowly with a long doubling time.

Slow growing bacteria undergo the traditional cell cycle, which is divided into three phases: B, C and D phase. The period between the division of a new bacterial cell and the initiation of chromosomal DNA synthesis is B phase. The initiation to termination of chromosomal DNA replication is the C phase. Then, the period between the replication termination to cell division is D phase (Figure 1.1). During each cell cycle, a mother cell reaches twice the minimum unit cell size then undergoes one round of DNA replication, followed by chromosomal DNA segregation and cell division. However, in fast growing cells, multiple rounds of chromosomal DNA replication are initiated prior to terminating the first round of replication<sup>1</sup>. These multi-fork replications can have two, four, or eight replication rounds, all within only one round of cell

division per each replication cycle (Figure 1.2). These two main events, the replication and cell division processes, are mediated by various small molecules and proteins, such as: (p)ppGpp, UDP-glucose, enzymes and proteins involved in central carbon metabolism, and fatty acid biosynthesis<sup>2</sup>.

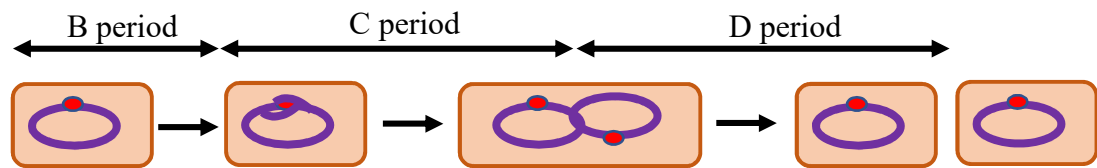


Figure 1.1 Bacterial cell cycle. During each cell cycle, bacterial cells undergo three phases: B, C and D, where the cells grow and propagate by binary division.

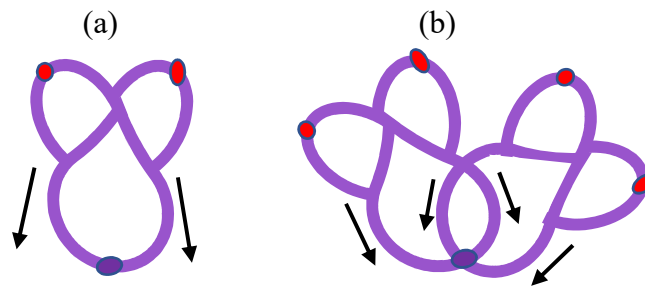


Figure 1.2 Chromosome replication in bacteria. (a) In slow growing bacteria one round of chromosomal replication occurs, while in (b) fast growing cells, multiple rounds of chromosome replications start before terminating the first replication cycle. However, each cell cycle will only have one chromosome replication cycle.

#### (a) Roles of small nucleotides, (p)ppGpp

For each bacterial cell cycle, chromosomal DNA undergoes a single round of replication, which contains: initiation, elongation and termination. The replication is initiated by DnaA binding to a specific site at the origin of replication. The *in vivo* levels of DnaA are regulated by

guanosine tetraphosphate and guanosine pentaphosphate, (p)ppGpp. The levels of (p)ppGpp mediate a significant number of nutrient dependent metabolic pathways, known as a stringent response. Two enzymes, (p)ppGpp synthase (RelA) and (p)ppGpp hydrolase (SpoT), are involved in the regulation of (p)ppGpp. RelA responds to amino acids starvation, while SpoT responds to carbon, nitrogen and fatty acids starvations<sup>3</sup>. In addition to (p)ppGpp mediation of DnaA, there are other mechanisms regulating the levels of DnaA which thereby regulate chromosome replication. These mechanisms include: autoregulation of DnaA level in cells, the SeqA protein which binds to the hemi-methylated chromosome and blocks DnaA binding to the chromosome, and regulation of DnaA activity via DNA initiator-associating protein (DiaA)<sup>1</sup>.

#### (b) Role of UDP-glucose

Soon after the replication process is completed, the two newly synthesized chromosomes are segregated and once the cells have reached twice the minimum unit size, cell division is initiated. During this process, FtsZ (a prototubulin GTPase) assembles into a ring inside the cell membrane to facilitate binding of other proteins involved in cell division. Normally in a cell, FtsZ localizes at the mid-cell region so that cell division will produce two daughter cells with chromosomal DNA. The availability of FtsZ is controlled by the integral membrane protein OpgH, and the interaction between FtsZ and OpgH is enhanced by UDP-glucose. In carbon-rich media, glucose-6-phosphate is produced through the glycolysis pathway and is used as the precursor of UDP-glucose synthesis. Consequently, increasing OpgH bound FtsZ impacts the levels of free FtsZ which ultimately delays cell division and increases the cell's size<sup>1, 3</sup>.

### 1.2.2 Bacterial cell aging and cell death

Binary fission in bacterial cell division allows the progeny to inherit both damaged and undamaged elements<sup>4</sup>. However, similar to all other living organisms, over time bacteria also undergo aging and death. During bacterial growth, cells continuously grow and divide at the expense of nutrients. Nutrient depletion is a key component in bacterial aging and death, causing cells to switch from reproduction to survival. Nutrient depletion, such as amino acid or carbon starvation, activates RelA and SpoT leading to elevated levels of (p)ppGpp in the cell, enhancing the expression of maintenance genes, rather than proliferating<sup>5</sup>. Additionally, bacteria undergo senescence during reproduction.

During division, cells divide from the middle so that each daughter cell inherits one new pole and one old pole. Each round of replication leads to a set of progenitors which inherit older poles. Furthermore, during Gram-negative bacterial growth, outer membrane proteins gradually accumulate at the cell poles. Thus, the cells inheriting continuously the older and older poles have high levels of accumulated proteins in the outer membrane<sup>6</sup>. Newly synthesized peptidoglycan is absent at the poles of Gram-negative bacteria and peptidoglycan at the poles is thought to be less active than that of sidewalls<sup>7</sup>. Thus, over the cycles of replication, the older poles are less adapted than newer poles. Therefore, these cells are slower growing and more prone to death than their counterparts with new poles<sup>6</sup>.

### 1.2.3 Bacterial growth curve

The bacterial growth curve is categorized into four main stages which are typical for batch cultures: a lag phase, a log or exponential growth phase, a stationary phase, and a death phase (Figure 1.3). During the lag phase, cells slowly adapt to the media and growth conditions. While

the cells grow, they do not reproduce or increase in population during the lag phase. Once the cells enter the log phase, they divide and double in population at a constant time interval known as the doubling time. If nutrients are continuously provided, the cells will continue to double without entering the stationary phase. However, if nutrients are not replenished, the available nutrients are depleted while waste and toxins accumulate in the growth culture ultimately leading to cell starvation and slower growth and reproduction. This slowing marks entry into the stationary phase, where the rates of cell growth and death in the culture are equal. Cells enter the death phase over time as the toxins and waste increase and nutrients are depleted, which leads to a higher cell death than growth rate and a decline in the bacterial population. These different growth phases are focused on in Chapter 3 for method optimization process.

Depending on the method of bacterial cultivation, the cells may not reach some of the later phases. In a batch culture or closed system where media is neither added nor removed from the culture, nutrients are depleted and toxins and wastes generated by the cells accumulate. However, this commonly used laboratory cell culture technique has a lower risk for contamination and cells can undergo all the phases of the growth cycle. In contrast, in fed-batch cultures, or semi-closed systems, some nutrients are added in intervals without removing used media. Thus, this method provides nutrients that would inhibit the cells' growth at higher concentrations in small, beneficial amounts throughout the growth. In continuous culture, or an opened system, fresh sterile media is continuously added and used media is removed at the same rate so, the cells are continuously in the log phase, which is also known as steady state growth.

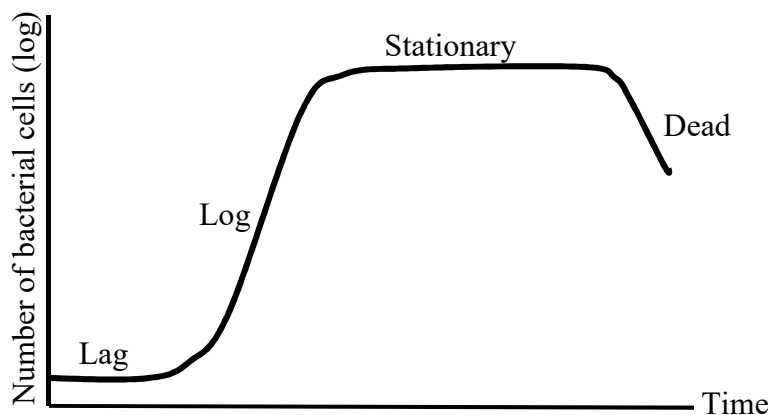


Figure 1.3. Bacterial growth curve. In a typical bacterial growth, during the lag phase cells adapt to new growth conditions and grow slowly. Then cells enter log phase, where cells double at every doubling time. Thus, cells grow exponentially before entering stationary and death phases as the nutrients deplete and accumulate of waste.

#### 1.2.4 Bacterial battle of survival

As the external and internal environmental conditions are continuously fluctuating, bacteria need to continuously and efficiently adopt to every change; thus, allowing them to survive in their habitat. Some changes include: the availability of carbon and other nutrients, such as rare minerals and vitamins, the temperature, the levels of  $O_2$ , and the pH. Bacteria recognize these environmental changes and respond by activating or inactivating various physiological pathways. To respond to these environmental changes at the molecular level, specific sets of genes will be either up-regulated or down-regulated via specific regulators.

##### (a) Global regulators enabling bacterial survival

Global regulators are proteins that individually regulate the expression of multiple operons in response to an environmental change. Some changes include: bacterial metabolic shifts between aerobic and anaerobic environments, catabolite repression, production of reactive oxygen species (ROS), and temperature fluctuations<sup>8-10</sup>.

Under aerobic conditions, oxygen is modified to the highly reactive, and harmful, superoxide anions ( $O_2^-$ ), hydrogen peroxides ( $H_2O_2$ ), and hydroxyl radicals ( $\cdot OH$ ). These toxic ROS can damage many biomolecules, including DNA. Thus, bacterial superoxide dismutase (SOD) reduces superoxide into hydrogen peroxide, which is then converted by catalase into water. SOD expression is regulated by four regulators: the ferric uptake regulation protein (Fur), the fumarate nitrate regulation protein (Fnr), the aerobic respiration control system (ArcAB), and the superoxide regulation protein (SoxR)<sup>9</sup>. Fur also regulates the expression of the siderophore transport system according to cellular  $Fe^{2+}$  concentrations to maintain a constant level of  $Fe^{2+}$ . When cellular levels of iron increase, in order to maintain iron homeostasis, Fur binds with  $Fe^{2+}$  and acts as a repressor on genes involved in both siderophore biosynthesis and iron-siderophore up taking transporters<sup>11, 12</sup>.

Fnr, ArcAB, OxyR and SoxRS are global regulators that are either activated or inactivated in the presence of  $O_2$ . ArcAB is a two-component regulatory system, where ArcB acts as the sensor and ArcA acts as the regulator. ArcAB and Fnr regulate the bacterial transition between aerobic and anaerobic growth conditions<sup>10</sup>. OxyR, a DNA binding transcription factor, is sensitive to the redox conditions in the cells and regulates hydrogen peroxide-inducible genes containing operons<sup>8</sup>. Similar to two-component regulatory systems, SoxR and SoxS in the SoxRS global regulator function as a sensor and the regulator, respectively. During oxidative stress conditions (mainly by superoxide anion,  $O_2^-$ ), SoxR is activated which promotes expression of SoxS, regulating the oxidative stress response genes<sup>13</sup>.

#### (b) Eliminating competition and scavenging scarce nutrients

As nutrient levels deplete over time, bacteria compete for the limited resources. To outcompete other bacteria, bacteria can produce antimicrobial agents (bacteriocins) which target

closely related bacterial species. These lethal compounds eliminate the sensitive strains from the population and leave only killing and resistant strains which reduces population density and thus the competition for the limited nutrient supply <sup>14, 15</sup>.

Aside from the continuous battles for survival, bacteria have also adapted to scavenge rare nutrients, which are essential for life. Iron is highly abundant, however the bioavailability is low<sup>16</sup> so bacteria secrete small organic molecules (siderophores) to the environment to scavenge iron using various resources. These resources can include: iron storing and transporting proteins from higher organisms (humans and plants) and uptake via highly specialized transporters. For example, Gram-negative bacteria express TonB dependent transporters (TBDTs) to uptake various iron and nickel bound siderophores and vitamin B<sub>12</sub><sup>12, 16, 17</sup>.

### 1.3 Gram-negative bacterial membranes

The contents of bacterial cells are encapsulated by a cytoplasmic membrane (CM), which is composed of phospholipids (phosphatidylethanolamine, phosphatidylglycerol, and trace amounts of cardiolipin) as well as membrane proteins (peripheral and integral proteins). Gram-positive bacteria also contain a thick cell wall, which primarily consists of peptidoglycan. However, in Gram-negative bacteria, a second membrane is present, known as an outer membrane (OM). Moreover, the CM and OM are separated within these bacteria by an aqueous space, known as the periplasm (15 -20 nm), and a thin layer of peptidoglycan within the periplasm<sup>18, 19</sup>.

#### 1.3.1 OM lipids

Unlike the cytoplasmic or inner membrane (IM), the OM lipid composition is asymmetric. The inner leaflet of the OM contains phospholipids, while the outer leaflet contains lipopolysaccharides (LPS). In general, LPS contain three components: lipid A, core oligosaccharides, and *O*-antigens (Figure 1.4)<sup>20</sup>. The classical lipid A molecule, present in *E. coli*, contains glucosamine disaccharide which is acylated with an (*R*)-3-hydroxymyristic acid residues at the 2 and 3 positions of both glucosamine units. Each lipid A has six fatty acids, where four fatty acids are acylated directly with the disaccharide and the remaining two molecules are acylated on the fatty acids and linked to a non-reducing glucosamine unit. This creates piggyback structures in an asymmetric manner. In addition to fatty acids, position 1 and 4 of the reducing and the non-reducing glucosamine units are phosphorylated (Figure 1.5)<sup>20</sup>.

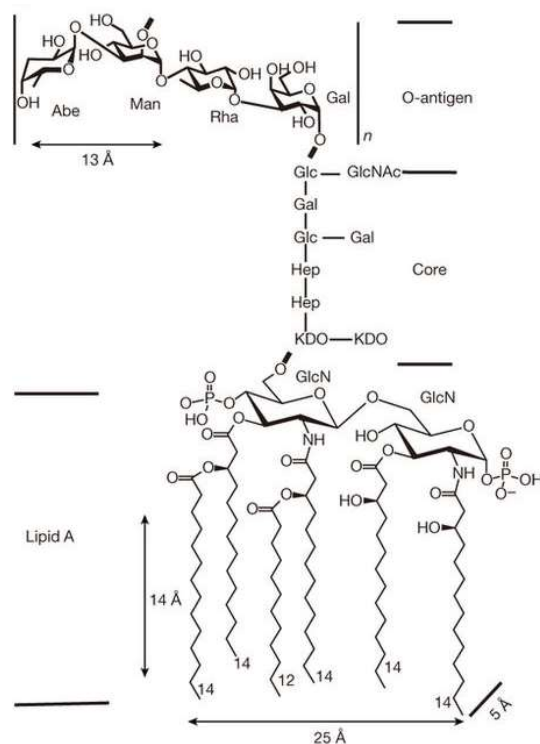


Figure 1.4. The structure of typical LPS molecule consists of lipid A, a core unit, and an *O*-antigen. Figure is from Nikaido. (2003). *Microbiology and Molecular Biology Reviews*. **67**, 593-656.

In contrast to the classical lipid A structure, different lipid A molecular structures can be found within various bacteria. This diversity is due to differences in fatty acids length, number, position of acylation, and end groups of the disaccharide. With respect to the lipid A in *E. coli*, fatty acids of *Neisseria gonorrhoeae* and *Pseudomonas aeruginosa* are symmetrically acylated<sup>21</sup>, <sup>22</sup>. *Porphyromonas gingivalis*, meanwhile, contains lipid A molecules with 2 to 5 acylated fatty acids<sup>23</sup>. *Helicobacter pylori* and *P. aeruginosa* have longer and shorter fatty acid chains when compared to *E. coli*, respectively<sup>22</sup> (Figure 1.5). Furthermore, some bacteria lack phosphorylation at one (*Bacteroides fragilis*) or both (*Rhodococcus vanielli*) ends of the disaccharide of the lipid A or they contain modifications to the phosphate groups<sup>20</sup>.

The second component of LPS, the core oligosaccharide, is connected to position 6 of the non-reducing sugar and can be divided into an inner and outer core. The inner core contains a few 3-deoxy-D-manno-oct-2-ulosonic acid (KDO) and heptose residues, followed by an outer core of hexoses (Figure 1.6). Some inner core residues are phosphorylated and contain phosphoethanolamine or pyrophosphoethanolamine<sup>20</sup>. Five different core oligosaccharides are present in *E. coli* (K-12, R1-R4). Even though *E. coli* B strain core is R1 type, a mutation within the one of the enzymes involved in the biosynthesis pathway truncates the five hexose units to two hexose units in the core oligosaccharide<sup>24</sup>. Finally, the *O*-antigen is a long polysaccharide between 4 to 40 residues which is connected to the outer core hexose<sup>20</sup>.

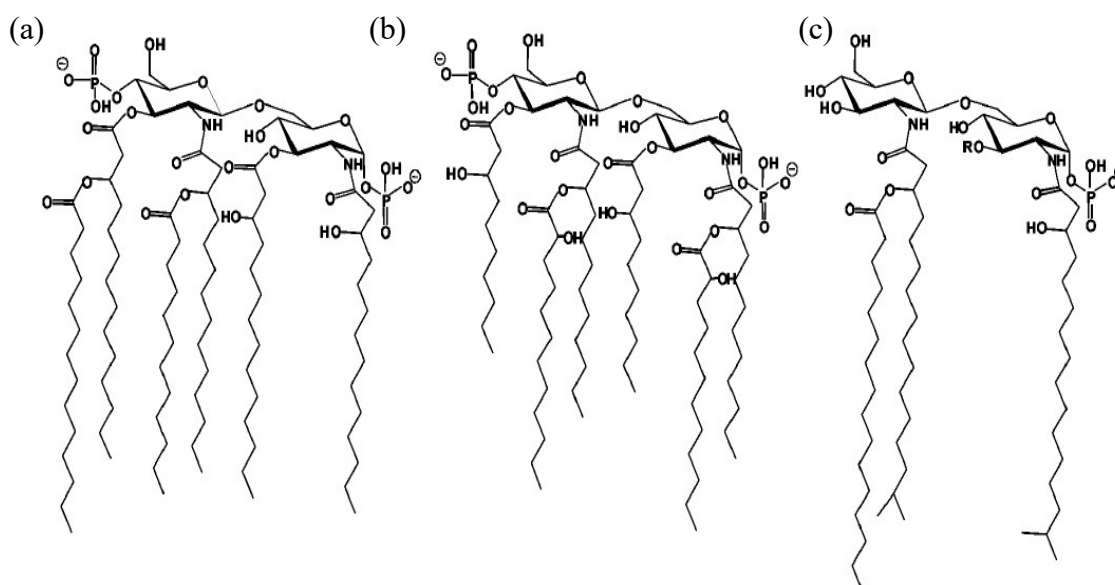


Figure 1.5. Structures of different lipid A molecules. (a) Typical lipid A found in *E. coli*. (b) Lipid A from *P. aeruginosa* and (c) lipid A from *P. gingivalis*. Figures are from Nikaido. (2003). *Microbiology and Molecular Biology Reviews*. **67**, 593-656.

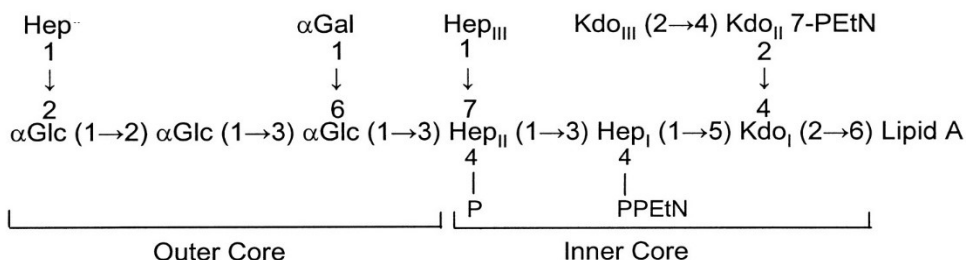


Figure 1.6. Structure of rough oligosaccharide core of *E. coli* K-12 strain which consists of inner and outer core. Figure is from Nikaido. (2003). *Microbiology and Molecular Biology Reviews*. **67**, 593-656.

Many Gram-negative bacteria contain LPS. Most of these bacteria produce the *O*-antigen part of the LPS, however, due to mutations in their genome the most widely used laboratory *E. coli* B strain and K-12 strain derivative bacteria lack *O*-antigen, instead synthesizing rough or R type LPS. Some bacterial groups contain other lipids in addition to or instead of LPS, such as: sphingolipids (*Bacteroides*, *Porphyromonas* and *Sphingomonas*), sulfonolipids (*Cytophaga* and *Flexibacter*), and ornithine lipids (*Paracoccus denitrificans*)<sup>20</sup>.

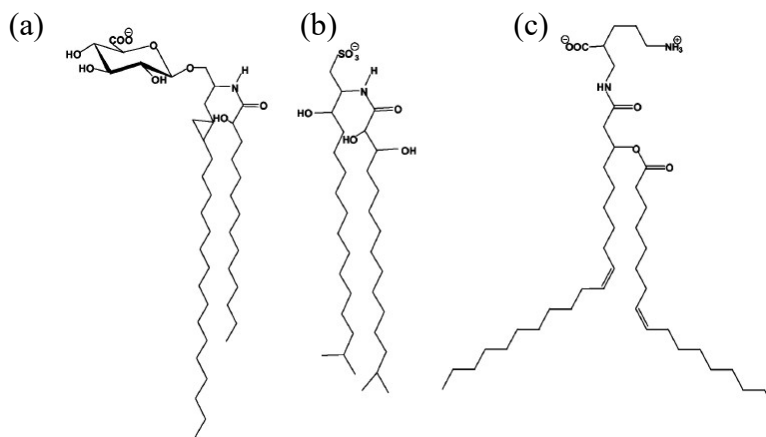


Figure 1.7 The structures of different lipids in the bacterial OM. (a) a sphingolipid from *Sphingomonas* (b) a sulfonolipid from *Flexibacter* and (c) an ornithine lipid present in *P. denitrificans*. Figure is from Nikaido. (2003). *Microbiology and Molecular Biology Reviews*. **67**, 593-656.

LPS in the OM is an effective permeable barrier for both hydrophobic and hydrophilic compounds. Like all other biological membranes, the OM has a hydrophobic core composed of phospholipids and lipid A chains of LPS<sup>25</sup> which creates a barrier to hydrophilic compound transfer. Additionally, the presence of strong lateral interactions between LPS creates a mesh around the cells and provides a sealed barrier for hydrophobic compounds from reaching the lipid core. These strong lateral interactions are mainly facilitated by the interactions between the polyanionic LPS and divalent cations, such as  $Mg^{2+}$  and  $Ca^{2+}$ . In addition to ionic interactions, the presence of many fatty acids (six fatty acids for *E. coli*) per LPS increase the surface area and the presence of H bond donors and acceptors for LPS (4-OH and 2-NH groups within the disaccharide backbone, 2-OH groups within two fatty acids linked to reducing sugar unit, and many other H-bond donors and acceptors in the core oligosaccharide) help to form many H-bonds which create strong lateral interactions between neighboring LPS molecules<sup>20, 25</sup>.

Thus, it appears disadvantageous for bacteria to have an OM which could prevent cellular nutrient uptake and hinder development of drugs targeting pathogenic bacteria. In contrast, these cells also have sophisticated protein machineries, which enable them to uptake/scavenge nutrients, including rare minerals and scarce nutrients, and to remove waste or toxic chemicals from the cell. Additionally, these protein channels can also be targeted to develop highly selective and specific drugs against pathogenic bacteria.

### 1.3.2 OM proteins

The proteins present in the OM are divided into two main categories: (1) integral OM  $\beta$ -barrel proteins, or outer membrane proteins (OMPs) which span the OM and (2) peripheral lipoproteins which are anchored to the inner leaflet. The lipoproteins, such as LolB, BamB-E and

LptE, play vital roles in OM biogenesis. However, many other lipoprotein functions are yet to be discovered. In contrast, OMPs are well studied, but more research is still required in order to reveal information on their structures and functions in cells.

While OMPs are diverse, based on the current structural and functional research they all have a common architecture, where anti-parallel  $\beta$  strands wrap to form a barrel across the OM. Hydrogen bonds form between the first and last  $\beta$  strands to complete the barrel and form a seam. The connecting loops between each  $\beta$  strand face either the periplasmic or the extracellular side and are short turns and long loops, respectively. The structural and functional diversity is due to size and architecture of the domains. The number of  $\beta$ -strands can vary between 8 (OmpX) to 26 (LptD) and sometimes with long loops on both extracellular and periplasmic sides (OmpA). Some OMPs can also assemble into oligomers. For example, OmpF (Figure 1.8a & b) and LamB are trimers and OmpA and OprF can assemble either as monomers (majority conformer) or trimers (minority conformer). In contrast, three TolC (Figure 1.8d & e) subunits link together to form one massive channel spanning across the OM and the periplasm. Some OMPs (BamA and LptD) are also connected to lipoproteins (BamB-E and LptE), while some others can contain core inside the barrel, such as BtuB and FecA (Figure 1.8c)<sup>20, 25, 26</sup>.

At present, these protein transporters are categorized into classes based on their transport properties. Porins are non-specific, diffusion channels that transport nutrients or waste in and out of the cells. The extracellular loops of these channels can undergo rapid mutagenesis, which allow the host to survive during harsh environmental changes, in the presence of antibiotics, and in the presence of bacteriocins. OmpC, OmpF, PhoE, and other homologous porins are known as classical porins<sup>20</sup>. In contrast, there are substrate specific protein channels which are also considered as substrate specific porins by some scientists<sup>27</sup>. LamB (maltose channel) permits

diffusion of maltodextrins, while ScrY (sucrose channel) allows the diffusion of sucrose and some other sugars (glucose, maltose, fructose and lactose)<sup>26</sup>. In contrast, larger nutrients, such as complex carbohydrates, iron chelators, and vitamin B<sub>12</sub>, are transported via gated channels or TonB dependent transporters (TBDTs) which contain core domains within the barrel. During the transport process, the core domain undergoes conformational changes and facilitates the uptake. For example, in *E. coli*, BtuB transports vitamin B<sub>12</sub>, while Cir, FecA, FepA, FhuA, FhuE, and Fiu transport different siderophores bound with iron<sup>17</sup>.

In addition to import machineries, export machineries permit bacteria to transport proteins and polysaccharides out of the cells. Export systems also allow cells to efflux antibiotics and toxic compounds, form pili and flagella, and siderophores for nutrient acquisition. Protein export pathways are divided into Sec or Tat-dependent and Sec or Tat-independent pathways<sup>28</sup>. In Sec-dependent pathways, all exporting proteins cross the IM via Sec translocase and then utilize different channels to cross the OM (type II, type V or autotransporter, chaperone-usheer, two partner secretion pathways). In Tat-dependent pathway, proteins are translocated via TatA, into the periplasm and then excreted by type II secretion system. Sec or Tat-independent pathways include type I and type III secretions pathways. Type IV is more often considered as a Sec or Tat-independent pathway. Within these machineries, export proteins are directly translocated from the cytoplasm to the extracellular side of the cell<sup>29</sup>. One example of this process is, in some bacteria, the type I secretion system that consists of an ABC transporter in the IM, a membrane fusion protein in the periplasm and an OM fusion protein, which forms a channel from the IM and spans the periplasm and the OM in order to secrete hemolysin, colicin V, and metalloprotease<sup>20</sup>.

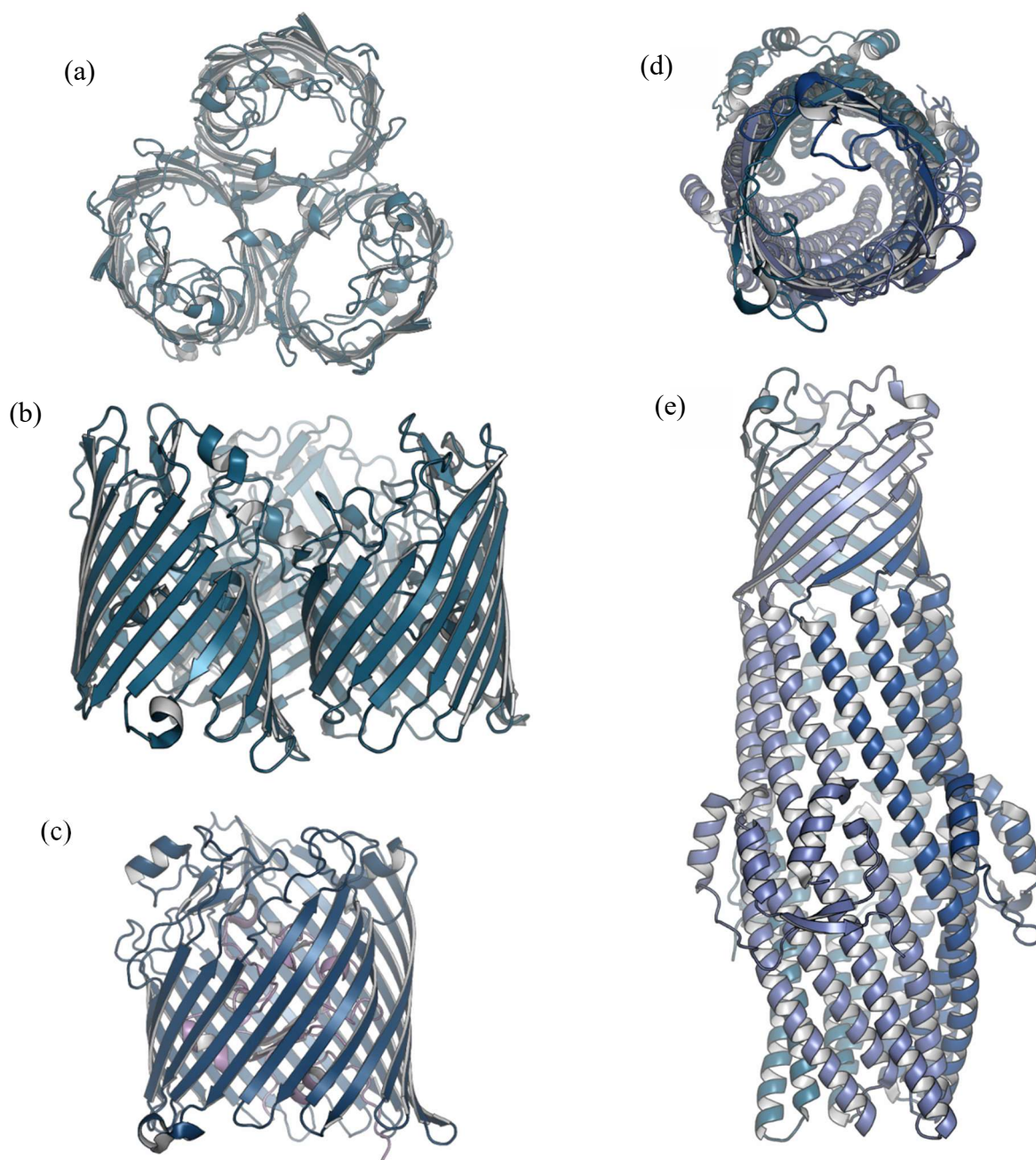


Figure 1.8. Crystal structures of a few OMPs. (a) Top view and (b) side view of **trimeric OmpF** (PDB ID 3K19), (c) side view of FecA (PDB ID 1KMO) in which the **core domain** blocking the pore formed by the **barrel domain**, (d) top view and (e) side view of export channel assembled by three TolC subunits (**A**, **B** and **C**) (PDB ID 1TQQ). Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

## 1.4 OM biogenesis

The vital OM integrity provides the first defense against external environmental changes which control transportation and communication within Gram-negative bacteria's habitat. OM biogenesis is well regulated and is organized in terms of structure and function. Highly specialized pathways are involved in the synthesis, maturation, and translocation of LPS, lipoprotein and OMPs to the OM while ensuring that they are properly inserted into the inner and outer leaflet of the OM. Many protein chaperones, enzymes, and other proteins in the IM, the periplasm, and the OM play vital roles in OM biogenesis, and some of these are discussed below. These processes are focused on in Chapter 3, 4 and 5.

### 1.4.1 LPS synthesis and transport to the OM

During LPS synthesis, lipid A is synthesized in the inner leaflet of the IM, then core sugar units are sequentially added to the lipid A via glycosyl transferase reactions. The lipid A-core unit is then translocated to the periplasmic side of the IM by an ABC transporter, MsbA. Simultaneously, *O*-antigen is synthesized in the inner leaflet of the IM by a range of glycosyltransferases and transported across the IM to the periplasmic side of the IM where the *O*-antigen is then ligated to the core of the LPS. Next, mature LPS are translocated across the periplasm and inserted into the OM via the LPS transport machinery (Lpt)<sup>32, 33</sup>. The Lpt machinery consists of seven proteins (LptA- LptG). During the translocation process, LPS in the inner leaflet of the IM is first extracted by an ABC transporter complex (LptB, LptF and LptG); the complex then ushers LPS across the periplasm via protein filament made by LptC and a few LptAs. Once reaching the OM, LPS cargo is delivered to the LptD-LptE complex, which then inserts the nascent LPS into the outer leaflet of the OM. There are about 200 copies of LptD-E complexes spread throughout the OM of the cell<sup>34</sup>.

LptD-E is an OMP and lipoprotein complex, which inserts LPS into the OM. Based on the crystal structure solved by Qiao and *et al.*, LptD (~550 amino acids) and Lpt E (~150 amino acids) together form a barrel and core architecture<sup>35</sup> similar to TonB dependent transporters (discussed in detail later in this chapter). The N-terminus of LptD forms a  $\beta$ -jellyroll domain (Figure 1.9b), which extends from the OM to the periplasm in order to efficiently deliver LPS from the final LptA on the filament protein. The C-terminus of LptD forms a barrel of 26 anti-parallel  $\beta$  strands, forming a kidney shape pore (Figure 1.9a). To date, it is the largest OM barrel, in which the hydrogen bonds between  $\beta 1$  and  $\beta 26$  strands close the barrel. The barrel is also twisted anticlockwise relative to the periplasmic side<sup>35, 36</sup>.

The tilt angle at N-terminus ( $\beta 1$ ) is  $30^\circ$  and at the C-terminus ( $\beta 26$ ) is  $67^\circ$  which creates an opening between  $\beta 1$  and  $\beta 26$  on the periplasmic side. The extracellular loops (LP) of LptD cover the surface of the pore from the extracellular side, while LP4 and LP8 are inside the barrel and interacting with LptE. LptE forms a roll-like structure, which occludes a significant portion of the pore formed by LptD. The lumen of the barrel is hydrophilic, thus accommodating core oligosaccharide and *O*-antigen, while LptE assists lipid A inside the barrel<sup>36</sup>. Even though the exact OM insertion mechanism is unknown, it is speculated that the lateral opening between  $\beta 1$  and  $\beta 26$  permits LPS insertion into the OM.

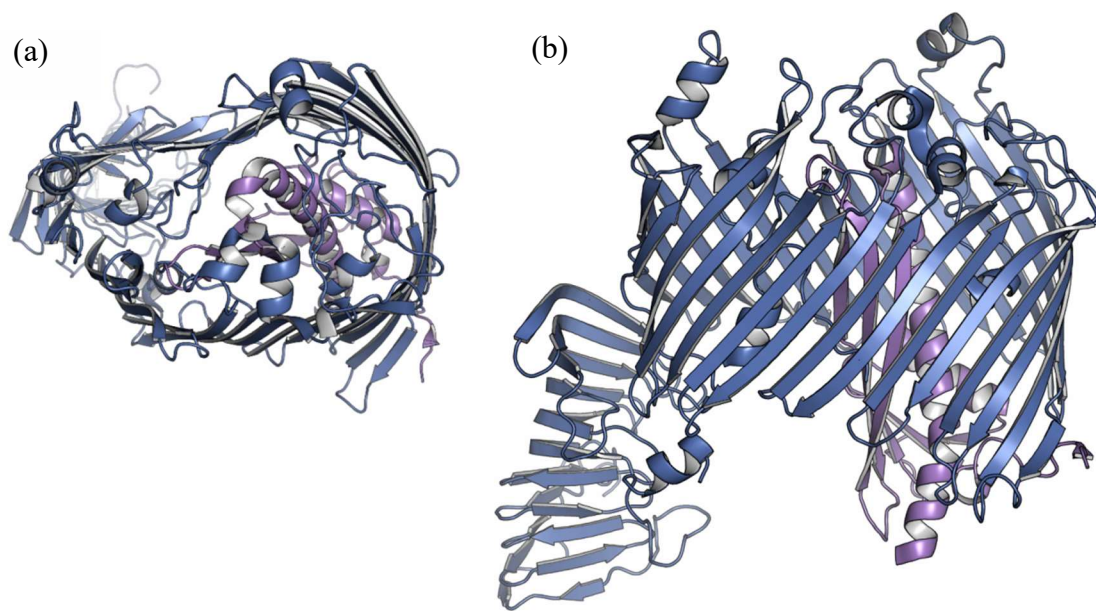


Figure 1.9 (a) Top view and (b) side view of the LptD-E complex. (PDB ID 4Q35). **LptD** formed into a kidney shape barrel, in which, **LptE** forms a core domain. N-terminus of **LptD** is forming a jellyroll-like structure. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

#### 1.4.2 OMP synthesis and transport

There is a vast diversity among OMPs with regard to structure and function. Nevertheless, all of these OMPs follow a similar journey beginning as nascent polypeptide chains synthesized in the cytoplasm which cross the IM, periplasm, and then finally fold into their native conformation and are inserted into the OM. During this journey, hydrophobic regions which form the barrel need to be protected from aggregation in the aqueous cytoplasm and periplasm, while still enabling the transport of hydrophilic stretches of protein (extracellular and periplasmic loops) through the IM or during the OM insertion. Many molecular chaperones facilitate the migration of these nascent unfolded OMPs by binding to these proteins, preventing them from aggregating, then directing them to the folding machineries in the OM<sup>25</sup>.

During the translation, or post-translationally, the nascent polypeptide chain binds to the cytoplasmic chaperone, SecB. SecB directs the bound polypeptide chain to SecYEG, an IM protein, which translocates the polypeptide chain across the IM. The N-terminus of the polypeptide chain contains a signal sequence necessary for IM translocation via SecYEG into the periplasm. Next, periplasmic chaperones bind with the newly translocated unfolded OMPs and transport them to the  $\beta$ -barrel assembly machinery (BAM) complex. Most unfolded OMPs contain a  $\beta$ -signal at the C-terminus in order to be recognized by the BAM complex, which folds the unfolded OMPs into their native conformation and inserts them into the OM (Figure 1.10).

#### 1.4.3 Lipoprotein synthesis and transport

Lipoprotein precursors, prelipoproteins, contain N-terminal signals targeting IM translocation and a lipobox with a conserved cysteine. During the maturation, prelipoprotein is first acylated at the cysteine site producing prolipoprotein which is followed by N-terminal signal sequence cleavage at the modified cysteine. Thus, resulting modified cysteine as the first amino acid of the lipoprotein. In Gram-negative bacteria, lipoproteins are further modified by attaching another fatty acid to the N-terminus of the acylated cysteine<sup>37</sup>.

The matured lipoproteins are localized on the periplasmic side of both the IM and OM. Lipoprotein localization machinery (Lol) assists lipoproteins by translocating OM lipoproteins but retaining the IM lipoproteins. Lol machinery includes LolCDE (an ABC transporter), LolA (periplasmic chaperone), and LolB (OM lipoprotein). LolCDE first extracts lipoproteins from the IM and delivers them to LolA. LolA then delivers the cargo to the OM acceptor protein, LolB, which inserts the lipoprotein into the inner leaflet of the OM (Figure 1.10)<sup>25, 37, 38</sup>.

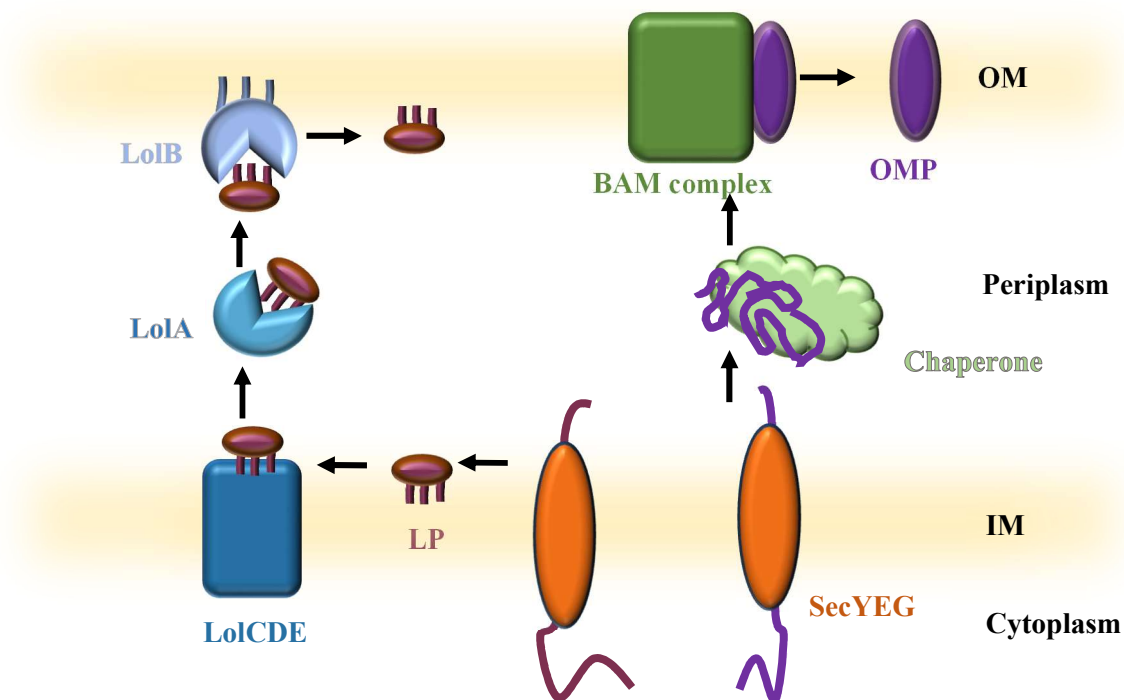


Figure 1.10. Transport of **OMP** and **LP**. Once the polypeptides are exported through the **SecYEG** channel, **nascent OMP precursors** are translocated across the periplasm to OM where they are correctly folded and inserted into OM via the **BAM complex**. The **lipoprotein precursors** are translocated and/or inserted into the proper membrane by **Lol machinery**.

#### 1.4.4 Periplasmic chaperones

Periplasmic chaperones play important roles in OM biogenesis. Some of these chaperones bind with newly transported, unfolded OMPs in order to prevent aggregation and/or to facilitate the transport across the periplasm. LolA is involved in lipoprotein delivery, while other well-known periplasmic chaperones, such as seventeen-kilodalton protein (Skp), Serine endoprotease (DegP), survival protein A (SurA), and FkpB binding protein A (FkpA), are involved in delivery of unfolded OMP to the BAM complex<sup>19, 25</sup>. Each of these chaperones redundantly deliver the cargo to BAM, but they are unique in their structure and function regarding delivery.

Skp and DegP form defined cavities which can cage their substrates. Skp is a homo-trimeric chaperone, which assemble into a jellyfish-like quaternary structure (Figure 1.11a and b). The body of the jellyfish-like structure is formed through closely packed  $\beta$  barrel of the trimer with three long  $\alpha$ -helical tentacles. The body and three tentacles define the internal cavity of Skp which bind with unfolded OMPs and prevent aggregation<sup>39</sup>.

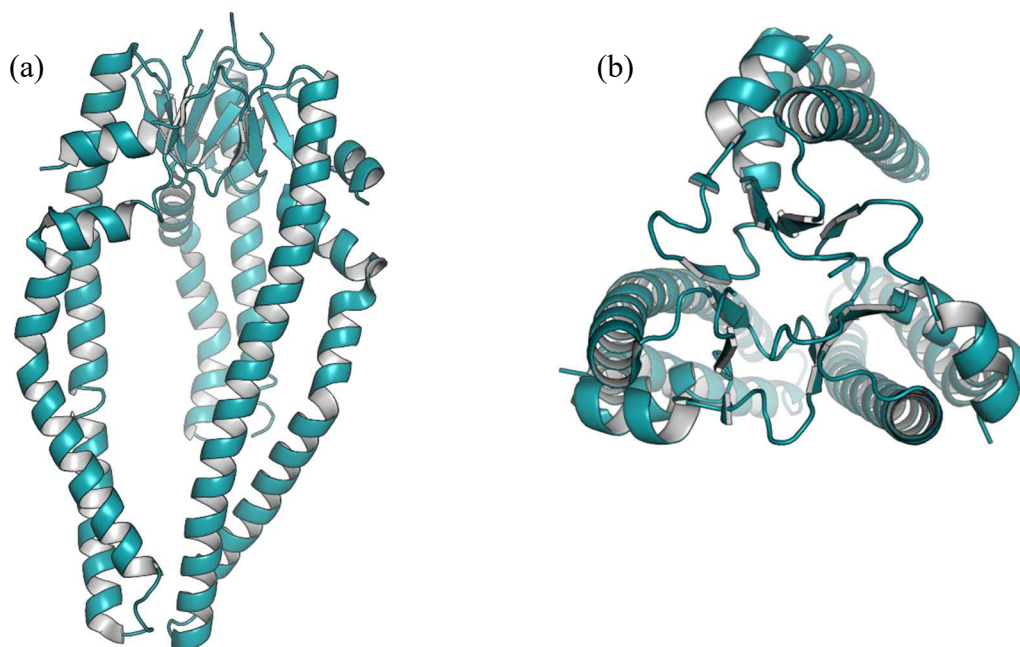


Figure 1.11. Crystal structure of **Skp** (PDB ID 1SG2) (a) side view and (b) top view. Trimeric Skp assemble into jellyfish-like structure. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

In contrast to Skp, DegP is a chaperone with protease activity and a member of the HtrA family. The bacterial members of this family are involved in periplasmic protein quality control. DegP monomers have three domains: protease domain, PDZ1 and PDZ2 domains, which are involved in substrate recognition. In its resting (inactive) state, DegP forms a hexamer with a central cavity. However, in the presence of substrate, the hexamer dissociates into trimers, which

then assemble into 12 and 24-mer cages having an internal protease activity (Figure 1.12). These large cages are the functional protease and refolding (chaperone) forms of DegP<sup>40</sup>.

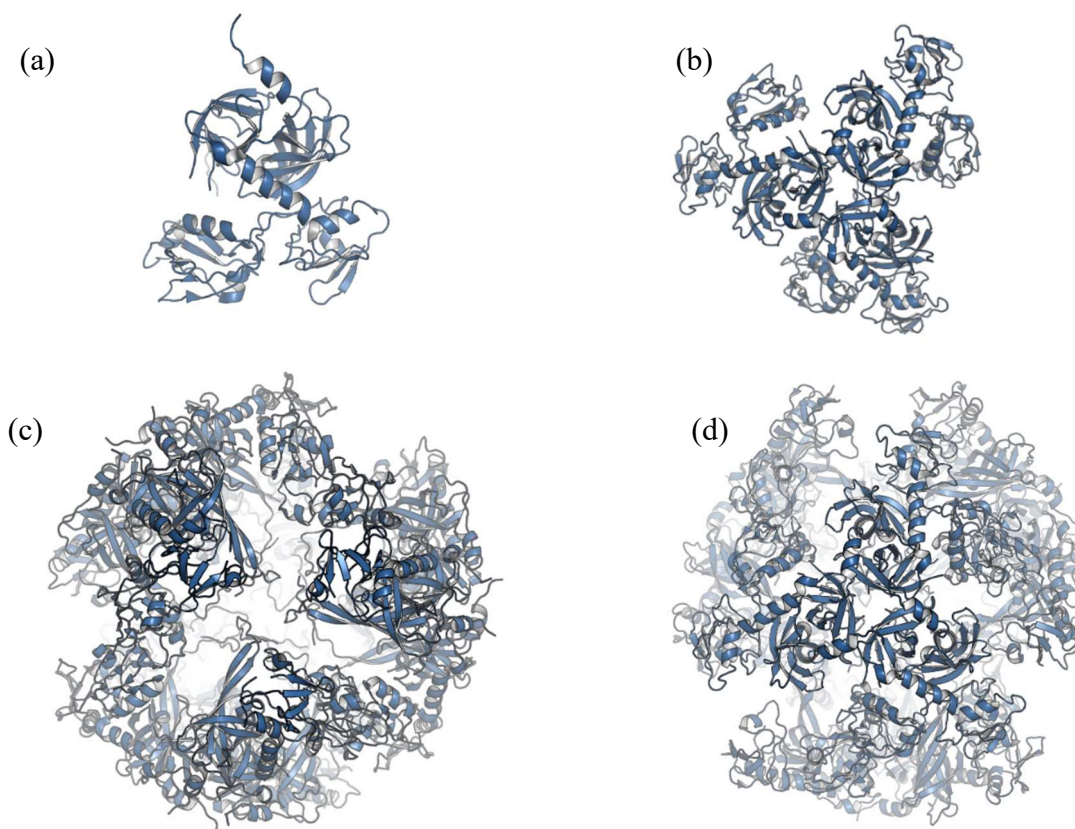


Figure 1.12. Crystal structure of **DegP** (a) monomer (PDB ID 1KY9), (b) trimer and dodecamer (c) top view and (d) side view (PDB ID 2ZLE) with formed cages with internal enzyme activity. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

FkpA is a homo-dimeric chaperone with catalytic activity, being a cis/trans peptidyl-propyl isomerase (PPIase). Each monomer has two domains. The N-terminal domain contains the chaperone activity and forms the dimer with the other monomer (Figure 1.13b), while the C-terminal domain has the PPIase activity<sup>41</sup>. In addition to FkpA, SurA is also a chaperone with PPIase activity. SurA has four domains: N-terminal domain, PPIase domains (P1 and P2) and a

C-terminal domain. Based on the crystal structures, SurA forms a core with N, P1 and C domains, while P2 domain is connected to the core through a linker (Figure 1.13a). SurA preferentially binds to sequences with aromatic (Ar) residues through an Ar-X-Ar motif and is known to undergo conformational changes in order to assist the binding to these aromatic binding regions<sup>42</sup>. To date, SurA is the only chaperone which has been cross-linked with BamA<sup>43</sup>.

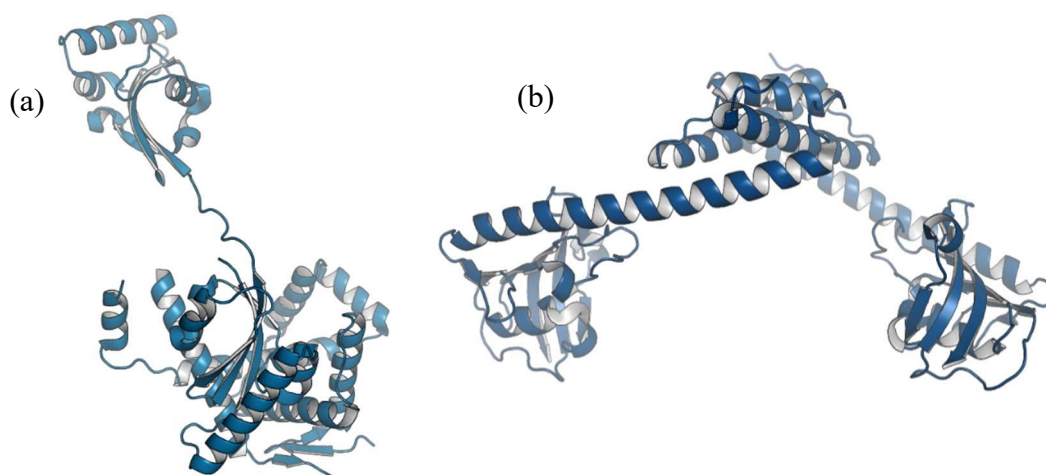


Figure 1.13. Crystal structures of (a) **SurA** (PDB ID 1M5Y) and (b) **FkpA** (PDB ID 1Q6I). Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

LolA is the chaperone which delivers lipoprotein to the OM acceptor, LolB. This monomeric chaperone forms an unclosed  $\beta$  barrel containing 11  $\beta$ -strands and a lid which consists of three  $\alpha$ -helices linked to the concave side of the barrel (Figure 1.14a). LolB has a similar architecture (Figure 1.14b). The cavity formed by LolA is highly hydrophobic and accommodates the acyl chains of lipoproteins. During transport, the lid opens and closes to accommodate and release the cargo<sup>38</sup>.

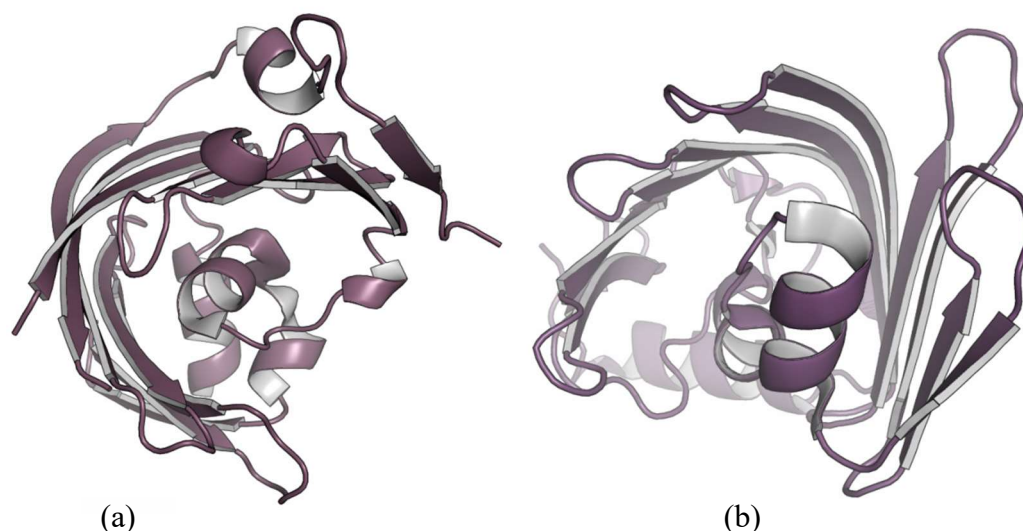


Figure 1.14. The crystal structures of (a) **LolA** (PDB ID 1IWL) and (b) **LolB** (PDB ID 1IWM). Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

#### 1.4.5 $\beta$ -barrel assembly machinery, BAM complex

The BAM complex is required for the OMP's final stage of assembly and insertion of the native conformations into the OM. The BAM complex in Gram-negative bacteria has homologs found in both chloroplasts (translocon at the outer envelope of chloroplasts, TOC) and mitochondria (sorting and assembly machinery, SAM)<sup>44</sup>. This 200 kDa complex consists of five subunits: BamA ( $\beta$  barrel), which is linked to BamB-E (lipoproteins), which assemble into a hat-like structure (Figure 1.15)<sup>45</sup>. BamA is a member of the Omp85 superfamily and is the core of the BAM complex. It contains five N-terminal periplasmic polypeptide transport-associated (POTRA) domains, which provide binding surfaces for the remaining four subunits. Protein-protein interactions (PPI) between either POTRA domains and BamB-E or with potential substrates are thought to be by  $\beta$ -augmentation, where the  $\beta$  strand of one protein are added to  $\beta$

strand of another protein<sup>44</sup>. The  $\beta$  barrel of BamA forms the top of the hat inside the OM, while POTRA domains and BamB-E form the brim at the periplasmic side. In the overall BAM complex, BamB binds to POTRA 2, 3 and 4 domains<sup>46</sup>. BamD binds to both BamC and BamE forming a subcomplex which connects POTRA 1 and 2 with POTRA 5 forming a ring-like structure underneath the  $\beta$  barrel<sup>45</sup>. Additionally, substrate bound SurA interacts with the POTRA 1 domain and delivers cargo to the BAM complex. Even though the BAM complex is composed of these five proteins, only BamA and BamD are vital for bacterial survival<sup>25, 43</sup>.

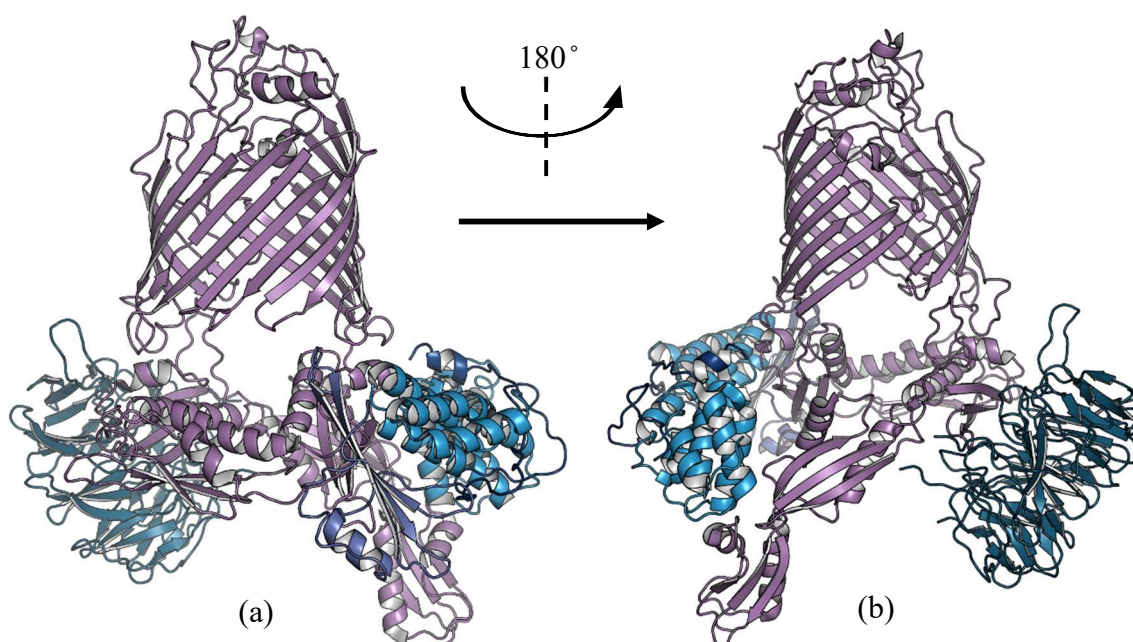


Figure 1.15. The crystal structures of the BAM complex side views (PBD ID 5D0O). (a) The front and (b) rear view of BAM complex, consists of **BamA**, **BamB**, **BamC**, **BamD** and **BamE** subunits. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

BamA is an evolutionary conserved protein. It forms a 16  $\beta$ -stranded barrel, in which the extracellular loops form a dome-like structure on the top (Figure 1.16a) through contact at the 6<sup>th</sup>

extracellular loop with  $\beta$ 14-16 strands. The seam of BamA is formed by poor hydrogen bonding between  $\beta$ 1 and  $\beta$ 16. The barrel's height at  $\beta$ 1 and  $\beta$ 16 is between 9-12 Å, while the opposite side is 20 Å. Thus, it is speculated that lateral opening of BamA can enable insertion of newly folded OMPs into the OM and loop 6 can modulate the lateral gate<sup>18, 25, 43</sup>.

BamB is a  $\beta$ -propeller with eight blades of  $\beta$  sheets and each blade can potentially form  $\beta$ -augmentation (Figure 1.16b). Thus, it has been proposed that these blades are involved in the PPIs between BamA POTRA 2-4, as well as, to increase the substrate binding capacity of BamA<sup>18, 43</sup>. In contrast, BamC (Figure 1.16c) is composed of an unstructured N-terminus followed by folded N- and C-terminal domains that are connected via a long flexible linker. This N-terminus is highly conserved and is required to form the sub-complex with BamD, which formed into a tetratricopeptide repeat (TPR) domain<sup>18</sup>. The TPR domain is composed of five  $\alpha$ -helical pairs, known as TPR motifs (Figure 1.16d), assembling a binding pocket for the C-terminal BAM recognition sequence of the substrates. Intriguingly, this binding pocket is also entirely blocked by the N-terminus of BamC. This has led to speculations that BamC could regulate BamD substrate binding. Besides BamC, BamD also forms a sub-complex with BamE<sup>43</sup>. BamE (Figure 1.16e) has been shown to specifically bind with phosphatidylglycerol (PG) and to enhance OMP insertion into liposomes. Additionally, the lipid binding site also overlaps with the BamD binding site. Therefore, BamE could potentially recruit PG to facilitate the membrane insertion of OMP<sup>44</sup>.

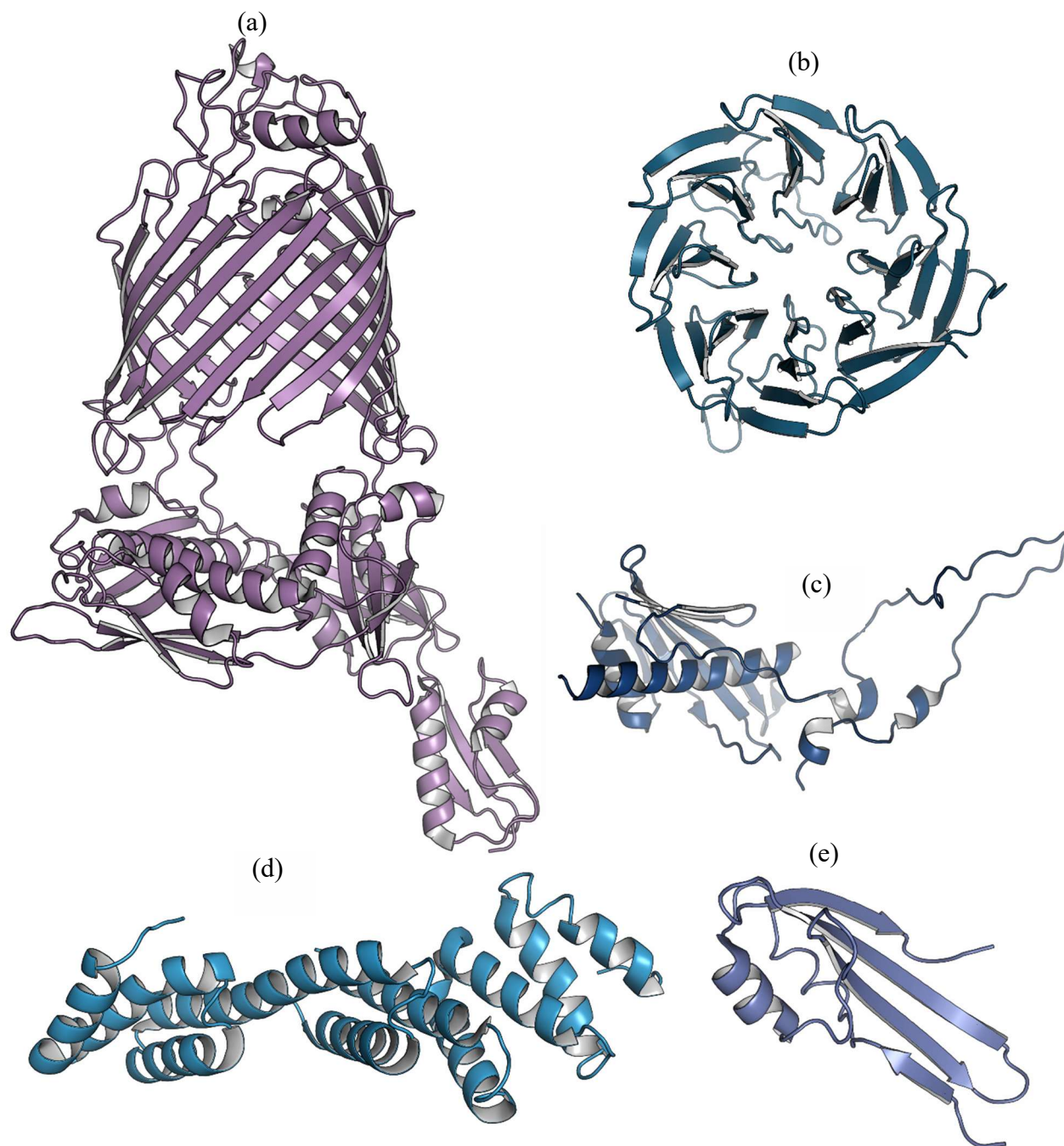


Figure 1.16. The crystal structures of (a) **BamA** (PDB ID 5D0O), (b) **BamB** (PDB ID 3Q54), (c) unstructured N-terminus region followed by the folded N-terminal domain of **BamC** (PDB ID 3TGO), (d) **BamD** (PDB ID 5D0O) and (e) **BamE** (PDB ID 5D0O) from *E. coli*. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

### 1.4.6 OMP folding in the BAM complex

The folding of soluble proteins into their native conformation is driven by burial of hydrophobic residues within the structure, to prevent unfavorable interactions with the aqueous environment. In contrast, in membrane proteins such as OMPs, folding exposes hydrophobic residues to the lipid membrane interior, while hydrophilic residues are facing either the aqueous extracellular or the periplasmic environments. Additionally, some membrane proteins can stabilize hydrophilic residues within their aqueous pore. For OMPs, the BAM complex drives folding to the native structure and OM insertion.

At the BAM complex, the cargo is first delivered to BamD by periplasmic chaperones, such as SurA. The substrate is then bound to POTRA domains and guided into the  $\beta$  barrel of BamA where it is folded into the native conformation. BamD substrate recognition is supposed to be regulated by the N-terminus of BamC. During the folding process, BamB is thought to provide substrate binding surfaces via  $\beta$ -augmentation which increases the POTRA domains' binding capacity. BamE recruits PG and enhances folding and OM insertion. Even though the exact  $\beta$  barrel folding mechanism is unknown, there are a few proposed models to explain the folding.

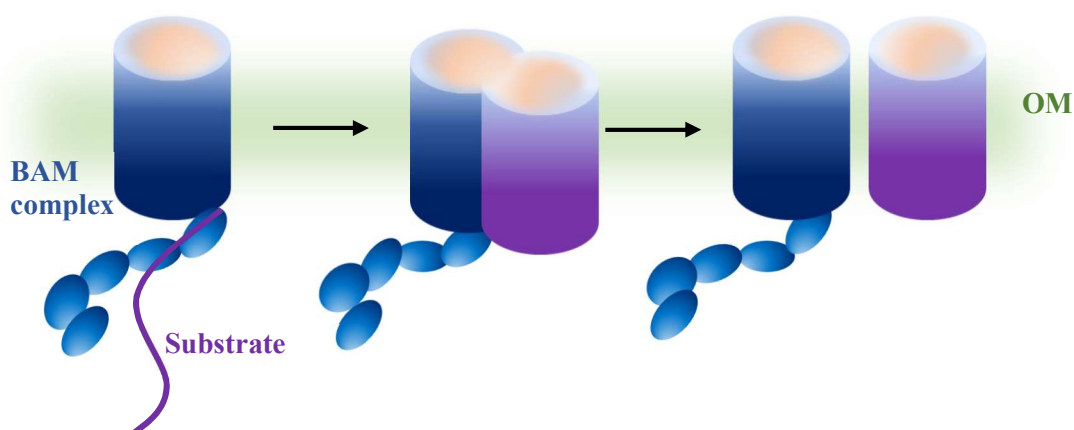


Figure 1.17. Schematic representation of budding model. **BamA** acts as a template and allow the **substrate** to fold. After completely folding, **BamA** releases the **new OMP** into the **OM**.

One such mechanism is the budding model. To date, there are several modified versions of this model. According to the budding model, hydrogen bonds between the 1st and 16th  $\beta$  strands of BamA are broken and act as a folding template, allowing the substrate to form hydrogen bonds between the BamA barrel. This forms a BamA-OMP hybrid. Once the entire peptide chain is folded, it dissociates from the hybrid and both the BamA and the newly folded OMP barrel are sealed and released into the OM (Figure 1.17)<sup>25, 44</sup>. While the exact BAM folding mechanism is yet to be revealed, Höhr and *et al.* have determined the folding mechanism of the BAM homolog, SAM. Based on their putative model, substrate is first passed through the SAM complex to the intermembrane space (IMS) of the mitochondria. The C-terminal signal begins replacing the hydrogen bonds between SAM's  $\beta$  strand 1 and 16, leading to lateral gate opening. Loop 6 of SAM facilitates the insertion of the remaining substrate into the lateral gate via  $\beta$  hairpin-like structures. Once the entire substrate is folded, SAM releases the nascent  $\beta$  barrel into the membrane<sup>47</sup>.

#### 1.4.7 Disulfide bond formation system in bacteria

Cysteines are rarely present in OMPs, however there are exceptions which form native disulfide bonds<sup>48</sup>. For example, LptD of the LPS machinery complex which contains two non-consecutive disulfide bonds, the ferric hydroxamate uptake protein (FhuA) with two consecutive disulfide bonds and the Francisella OMP A (FopA), which contains a single disulfide bond, *etc*<sup>49</sup>. During this post-translational process, oxidoreductase enzymes reduce, oxidize or isomerize the covalent bonds formed between thiol groups of cysteines to protect them from oxidation, and enable proper folding of the precursor protein. However, in bacteria, the oxidoreduction process is compartmentalized into the periplasmic space. Thus, in addition to periplasmic chaperones,

there are specific enzymes involved in disulfide bond formation and isomerization processes. These enzymes are mainly involved in two main pathways: oxidation and isomeration<sup>50, 51</sup>.

(a) Oxidative pathway

DsbA and DsbA homologs, which are located in the periplasm, catalyze disulfide bond formation on precursor proteins. These enzymes are members of the thioredoxin superfamily, containing a thioredoxin (TRX) domain with a CXXC motif (two cysteines are separated by two amino acids) in the active site<sup>48, 50, 52</sup>. DsbA introduces disulfide bonds via disulfide bond exchange and it has a broad substrate specificity<sup>53</sup>. For example, DsbA oxidizes nascent protein precursors of adhesion (subunit of bundle-forming pili, BfpA), mobility (flagellar P-ring motor protein, Flgl), OMPs (FopA), and secretion (enterotoxin, ST<sub>b</sub>)<sup>50</sup>.

Different factors can affect DsbA's oxidative protein folding. When the N-terminus of the unfolded protein chain enters the periplasm via SecYEG, DsbA interact with the first cysteine of the protein and introduces a disulfide bond between the first and the second cysteine entered to the periplasm. Thus, disulfide bonds are often formed in sequential manner. However, the formation of sequential disulfide bonds is impacted by the rate of export. During slow export (co-translational export), DsbA forms more sequential disulfide bonds than rapid export (post-translational export) as the latter is ten times faster than co-translational export (Figure 1.18)<sup>52</sup>.

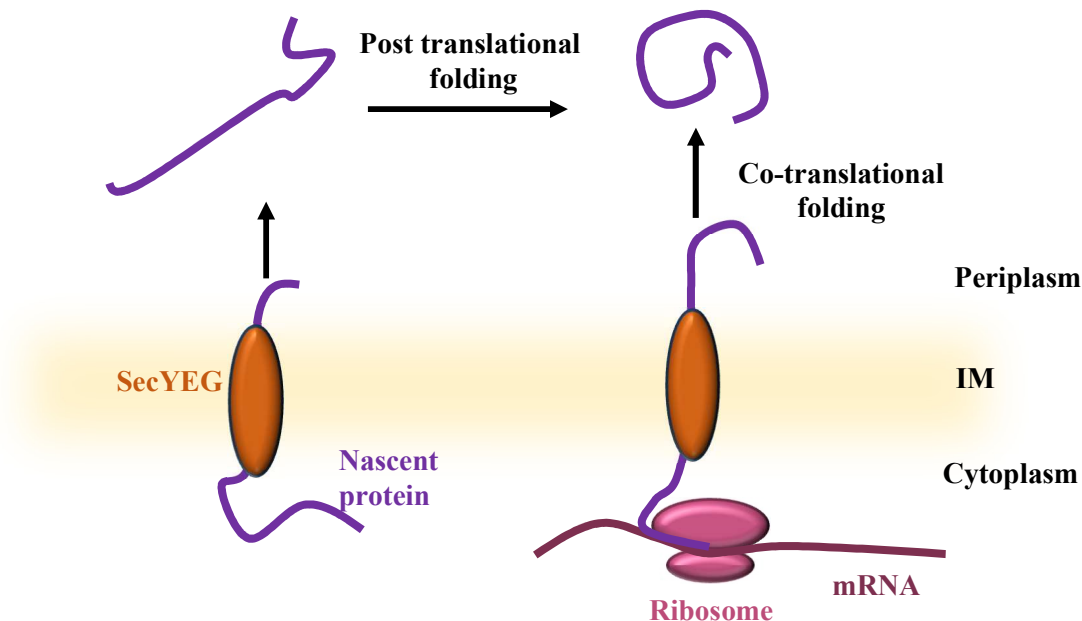


Figure 1.18. Protein export into periplasm as post-translationally and co-translationally and undergo oxidative folding.

During the oxidation process, DsbA is reduced; the IM protein DsbB then re-oxidizes DsbA to its active state. Unlike DsbA, DsbB is substrate specific and does not oxidize disulfide bonds of DsbA substrates. DsbB has four transmembrane helices, a cytoplasmic loop, and two periplasmic loops (P1 and P2). Each periplasmic loop contains a disulfide bond. During DsbA re-oxidation, DsbA passes the electrons to the P2 loop then to P1 loop of DsbB<sup>53, 54</sup>. The electrons are then transferred to either ubiquinone, under aerobic conditions, or to menaquinone (vitamin K2), under anaerobic conditions. From the reduced quinones, electrons are transferred to the electron transfer chain (ETC) (Figure 1.19)<sup>48</sup>.

#### (b) Isomerization pathway

In addition to DsbA being broad substrate specific, mostly it adds disulfide bonds in consecutive manner. Consequently, if a precursor protein requires multiple non-consecutive disulfide bonds, DsbA can form non-native disulfide bonds and lead to protein degradation in the

periplasm<sup>48, 53</sup>. To overcome this issue, the bacterial isomerization pathway corrects these non-native disulfide bonds into native ones. There are two models for how enzymes correct non-native disulfide bonds: the isomerization model and the cycles of oxidation/reduction model (Figure 1.19)<sup>48</sup>. The members of the thioredoxin superfamily, DsbC and its homolog, DsbG, are bacterial isomerases, assembled into V-shaped homodimers containing Trx domains with catalytic sites linked to dimerization domains by a short loop<sup>48, 50</sup>. Isomerization is mainly catalyzed by DsbC, while DsbG is mainly present to protect non-disulfide bond forming cysteines within protein precursors if there are an odd number of cysteines to prevent oxidation of these cysteines which would produce sulfenic acid<sup>55, 56</sup>.

During the isomerization process, DsbC and DsbG are reduced back to their active state by the IM protein DsbD<sup>53</sup>. DsbD has three distinct domains. The transmembrane domain, DsbD $\beta$ , is composed of eight  $\alpha$ -helices which connects to two periplasmic domains at each terminus. The N-terminus folds into the immunoglobulin-like domain, DsbD $\gamma$  and the C-terminus folds into thioredoxin-like domain, DsbD $\alpha$ . During DsbD mediated reduction process, electrons from cytosolic thioredoxins are transferred to DsbD $\beta$ , which are then passed to DsbD $\gamma$ , and finally to Dsb $\alpha$ . Once electrons reach DsbD $\gamma$ , they are passed to either oxidized DsbC or DsbG which reduces the proteins to their active states (Figure 1.19)<sup>54</sup>. However, to date, in Gram-positive bacteria, no DsbC-DsbD system have been identified, leading to the assumption that all multiple native disulfide bonds are formed consecutively<sup>57</sup>.

The conventional Dsb system is a key element in BtuB two cysteine residue labeling process, which is discussed in Chapter 3, as well as in Chapter 5.

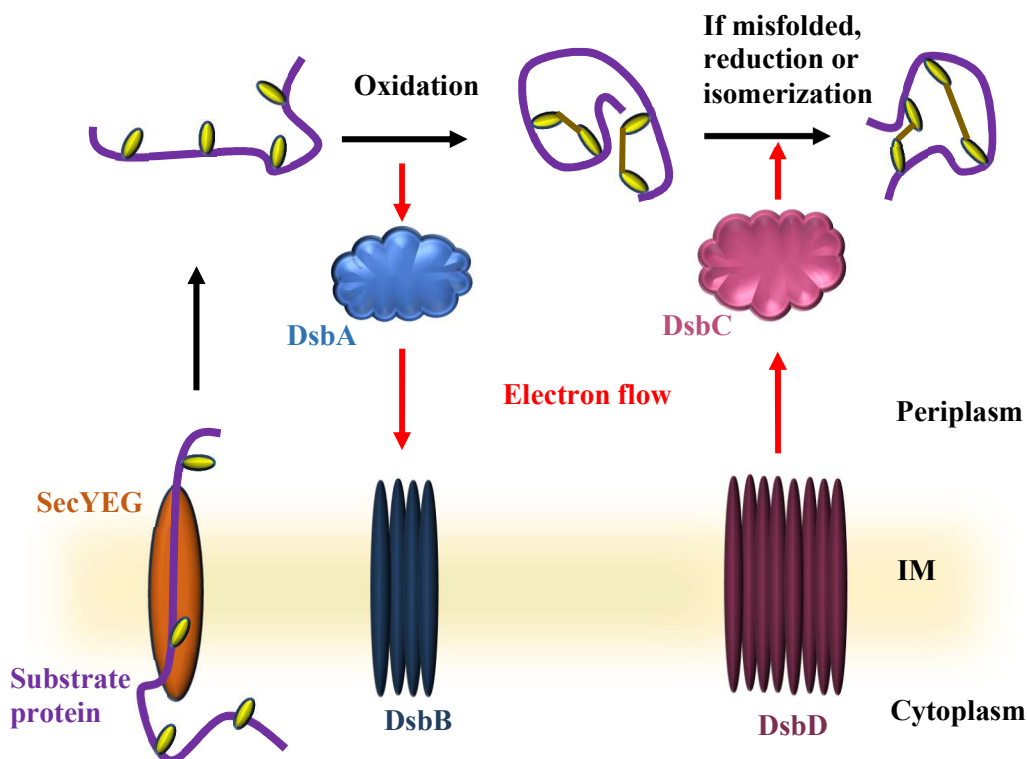


Figure 1.19. The conventional disulfide bond formation pathway in bacteria. The unfolded **substrate protein** enters the periplasm through **SecYEG**. **DsbA** forms **disulfide bonds** in the **substrate protein**. Reduced **DsbA** is re-oxidized by **DsbB** and transfer the electrons to ubiquinone or menaquinone. Mis-oxidized substrates are either reduced or isomerized by **DsbC**. **DsbC** is reduced by **DsbD** which receives electrons from cytoplasmic thioredoxin. The electron flow is indicated by **arrows**.

### (c) Alternative Dsb pathways

Both Gram-positive and Gram-negative bacteria and archaea have alternative oxidative pathways in addition to the conventional DsbA-DsbB oxidative pathway. The presence of conventional and non-conventional oxidative pathways enhances bacterial virulence. These systems facilitate the proper folding of many precursor proteins involved in different virulence pathways and many stages of infection, such as: mobility and adhesion, production and secretion

of toxins, secretion components and bactericidal activity<sup>50, 57</sup>. Thus, the Dsb system is an interesting target system for research in drug development.

To date, several alternative pathways have been identified in different bacterial species. (i) Bacteria containing either DsbA or DsbB, but not both. *Staphylococcus aureus* and *Listeria monocytogenes* encode only DsbA homologs<sup>54</sup>, while *H. pylori* encode only the DsbB homolog<sup>58</sup>. (ii) Some bacteria such as *Streptomyces coelicolor* and *Mycobacterium tuberculosis* encode the DsbA and DsbB analog, vitamin K epoxide reductase and VKOR. DsbA and VKOR are sometimes fused to form a single complex<sup>50, 57</sup>. (iii) Some bacteria do not express either DsbA or DsbB. In *Streptococcus pyogenes* isopeptide bonds have replaced disulfide bonds<sup>50</sup>. (iv) In addition to DsbA-DsbB system, many Gram-negative pathogens encode additional oxidative pathway homologs. For example, pathogenic strains of *E. coli* (F11, UTI89 and 536), *Salmonella*, and *Campylobacter* strains all encode DsbL and DsbI. These are homologous to DsbA and DsbB, respectively. In addition to DsbL-DsbI, *Salmonella enterica* encodes a DsbA-like protein, SugA. In contrast, *N. meningitidis* encodes a single DsbB and three DsbAs. DsbA1 and DsbA2 are lipoproteins anchored to the IM, while DsbA3 is a periplasmic protein<sup>50, 54, 57</sup>. *P. aeruginosa* contains two DsbBs (DsbB1 and Dsb2)<sup>59</sup>. *Legionella pneumophila* encodes two copies of DsbA, DsbB, and DsbD. In *L. pneumophila*, DsbA2 can both reduce and oxidize disulfide bonds. So, in addition to alternative oxidation pathways, DsbA2 in *L. pneumophila* functions similarly to conventional DsbC. Pathogenic bacterial strains lacking DsbC (*Coxiella*, *Rickettsia*) express DsbA with reducing activity<sup>57</sup>. Furthermore, some bacteria have DsbD homologs, CcDA and ScsB in *Rhodobacter capsulatus* and *Salmonella typhimurium*, respectively<sup>60</sup>.

## 1.5 TonB dependent transporters

Compounds that are greater than ~600 Da do not diffuse through OM porins. Thus, active transport is required to uptake such large compounds, as well as, to transport the compounds due to their very low concentrations in the environment. Because there is no proton motive force (pmf) across the OM nor ATP-hydrolyzing proteins, there is no energy source present during such active transport. Nevertheless, active transport through OMPs is facilitated by the pmf in the IM, through the TonB complex (TonB, ExbB and ExbD) and TonB dependent transporters (TBDTs). These TBDTs transport iron, nickel and copper complexes, vitamin B<sub>12</sub>, complex carbohydrates, and opportunistic cargos (bacteriophages and some bacteriocins).

### 1.5.1 Expression and regulation of TBDTs

TBDT expression is regulated either at the transcriptional or the post-transcriptional level. The transcriptional repressor, Fur, regulates iron homeostasis. In the presence of Fe<sup>2+</sup>, Fur binds to DNA and represses the expression of TBDTs, as well as other proteins and enzymes involved in siderophore biogenesis and iron metabolism<sup>11</sup>. This maintains cellular levels of iron. Similar to Fur, Ni<sup>2+</sup> uptake is regulated by NikR, in *H. pylori*. TBDT expression is also regulated by extra cytoplasmic function (ECF)  $\sigma$  factors which are repressed by anti- $\sigma$  factors. The expression of the *fecABCDE* operon is dependent on the ECF  $\sigma$  factor, FecI, enhancing operon transcription. Upon the binding of substrate to FecA, a signal transduction pathway is initiated to activate FecI via FecR. However, FecR is the anti- $\sigma$  factor of FecI and it is also involved in activating FecI. Similarly, the uptake of desferrioxamine and ferrichrome is regulated by FoxR and FiuR in *P. aeruginosa*<sup>12</sup>.

TBDT expression is also regulated by riboswitches and by sRNA. The 5' end of *btuB* mRNA contains a riboswitch that controls expression levels. In the presence of adenosylcobalamin, a product of the vitamin B<sub>12</sub> metabolic pathway, binding to this riboswitch inhibits the expression of *btuB*. The expression levels of *cirA*, *fecA*, and *fepA* are regulated by sRNAs (OmrA and OmrB). OmrA and OmrB both bind to *cirA*, *fecA*, and *fepA* mRNAs, via the RNA chaperone, Hfq, which down regulates expression<sup>12</sup>.

### 1.5.2 Transport of scarce nutrients via TBDTs

Iron is an essential metal for all organisms. Iron acts as a cofactor in numerous enzymes involved in metabolic pathways, such as: DNA replication, transcription, ATP synthesis, photosynthesis, *etc.* However, free iron can also catalyze the production of highly reactive oxygen and nitrogen species (ROS and RNS), which can damage biomolecules in the host cells<sup>17</sup>. Thus, iron is tightly bound to different proteins (lactoferrin, transferrin, heme binding proteins and *etc.*) in cells. In addition to sequestering iron inside proteins, iron also has a very low solubility under aerobic and physiological conditions. Consequently, the bioavailability of iron is very low ( $10^{-18}$ - $10^{-24}$  M)<sup>16</sup>. Thus, bacteria have adapted specific and highly sophisticated methods to scavenge scarce iron and to transport it.

Siderophores are small organic molecules (Figure 1.20) which are secreted by bacteria, fungi, and plants, which are highly specific and high affinity to iron. These iron chelators can range from 150 – 2000 kDa and have higher affinities for Fe<sup>3+</sup> than to Fe<sup>2+</sup>. These chelators high affinity is due to siderophores' metal binding groups, the number of binding sites and their stereochemical arrangements<sup>17</sup>. Many bacteria produce and use several types of siderophores

with different structures<sup>16, 17</sup>. The most common is the hexadentate siderophore which has a higher affinity to siderophores with lower denticity.

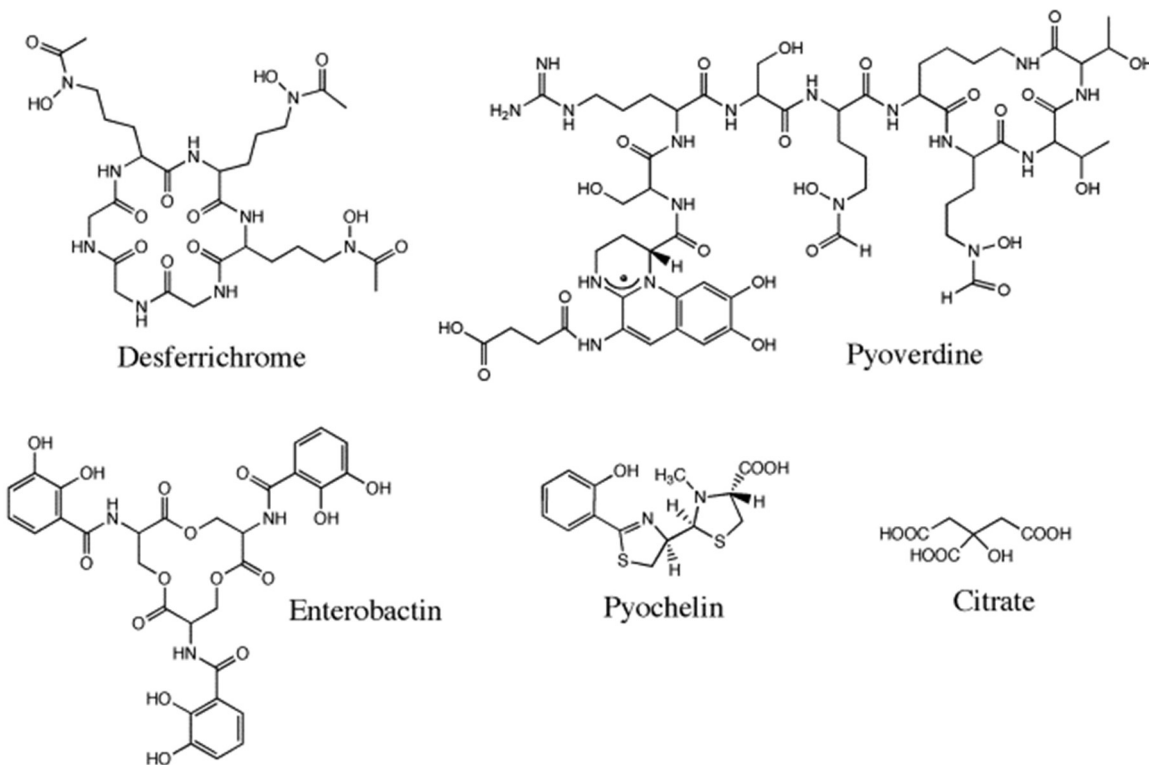


Figure 1.20. Structures of different siderophores synthesized by bacteria<sup>17</sup>. Figure is from Schalk, Mislin, and Brillet. (2012). *Current topics in membranes* **69**, 37-66

In addition to siderophores, pathogenic bacteria express specific TBDTs that can bind with host iron transporting proteins and remove the iron from the host protein and transport it<sup>16</sup>. Furthermore, many bacteria uptake heme as the iron source. Some bacteria can also use hemoproteins; for example: hemoglobin, hemopexin and haptoglobin-hemoglobin.

### 1.5.3 Structure and function of TBDTs

In 1998, the first TBDT (FhuA) crystal structure was solved, since then many other TBDT structures have been solved<sup>17</sup>. Based on the available TBDT crystal structures, TBDTs fold into a C-terminal, 22 standard  $\beta$  barrel, with a pore occluded by an N-terminal core domain. This core regulates the transport across the OM.

#### (a) Crystal structure of FhuA

In *E. coli*, ferrichrome-iron is transported by FhuA. Additionally, FhuA also transports albomycin, bacteriocins (microcin 25 and colicin M) and bacteriophages (T1, T5 and UC-1). According to the structure, 22  $\beta$  stranded barrel domain is formed by amino acids 161-723. The N-terminal core domain is formed by amino acids 1-161 (Figure 1.21a). FhuA has two pockets: a large external pocket and a small periplasmic pocket, which are outlined by the core domain. During the substrate binding, ferrichrome-iron binds to the substrate binding site within the external pocket (Figure 1.21b). Upon substrate binding, FhuA undergoes substrate induced conformational changes. Minor changes are observed within the barrel domain, a couple of periplasmic loops, and in the core. However, a major conformational change is observed with respect to the core domain where the switched helix (residues 24-29) located in the periplasmic pocket unwinds. Both the apo and substrate bound conformations have an unstructured N-terminus from residue 1-19, which also contains the TonB box (residue 7-11)<sup>61</sup>.

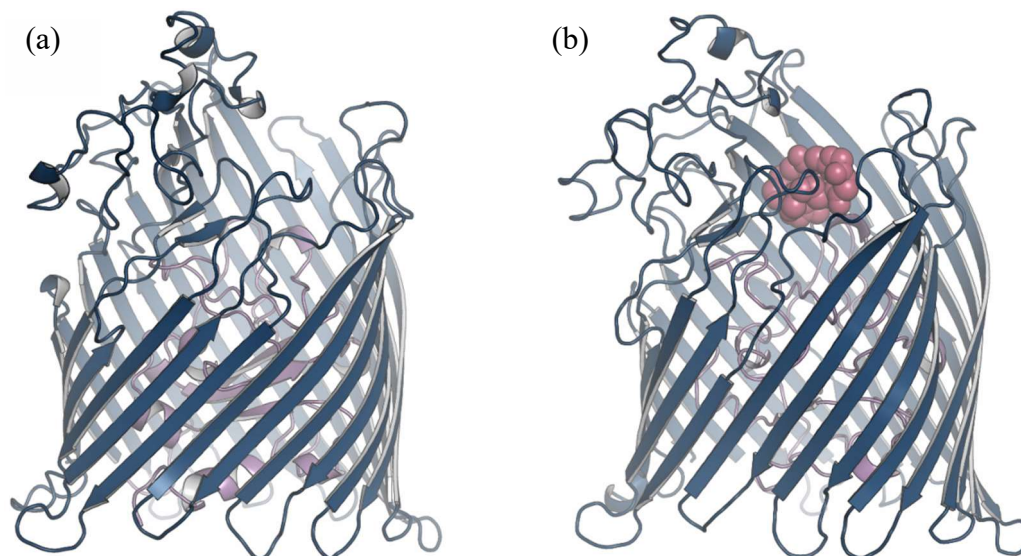


Figure 1.21. Structure of side view of (a) apo FhuA (PDB ID 2FCP) and (b) FhuA bound to **ferrichrome-iron** (PDB ID 1FCP). The **core domain** occludes the **barrel** pore. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

#### (b) Crystal structure of FecA

In 2002, Ferguson and *et al.* solved the crystal structure of FecA, both with and without its substrate, ferric citrate. Based on the structures, it also has the basic architecture with a 22  $\beta$ -stranded barrel domain (amino acids 222 to 741) and a core domain (amino acids 80-221) occluded pore (Figure 1.22a). Additionally, FecA contains a flexible, unstructured, N-terminal domain (amino acids 1-79) located in the periplasm that interact with the FecR receptor (not present in either the apo or substrate bound structures). FecA is about 65 Å in height, but about 30 Å in height of the transporter is extending out to the extracellular side. In contrast to apo structure, FecA undergoes conformational changes in the presence of ferric citrate. The major structural changes are observed in the seventh and eighth extracellular loops which move up and the switch helix unwinds in the ferric citrate bound FecA (Figure 1.22b)<sup>62</sup>.

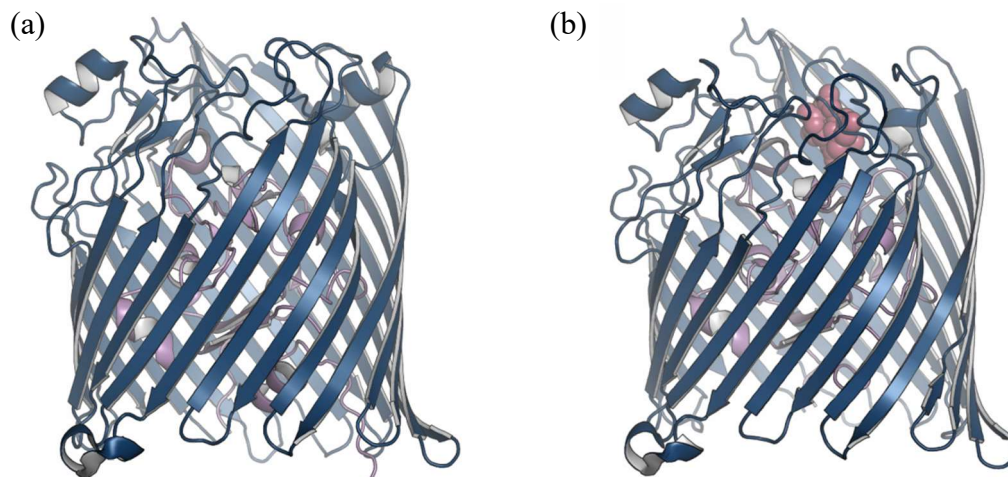


Figure 1.22. Structure of (a) side view of apo FecA (PDB ID 1KMO) and (b) side view of FecA bound to **ferric citrate** (PDB ID 1KMP). The **core domain** occludes the **barrel** pore. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

(c) Crystal structure of BtuB

BtuB has been successfully crystallized as the apo (1NQE), apo with calcium ligand (1NQG), with calcium and CN-Cb (1NQH), with colicin E3 (1UJW) and colicin E2 bound (2YSU) and also BtuB with spin labeled at different residues. Overall, BtuB folds into the characteristic TBDT architecture, with the N-terminal core folding within the 22 standard  $\beta$  barrel. In contrast to FhuA and FecA, BtuB lacks the switch helix, but contains the TonB box (residues 6-12). Based on the apo and calcium bound apo BtuB, calcium partially orders the extracellular loops between strands 3-4, 5-6, and 7-8 which are unstructured in the apo BtuB (Figure 1.23a). In addition to calcium, the substrate (CN-Cb) binding induces conformational changes in BtuB. In the barrel domain, the partially ordered extracellular loops are completely ordered and the loops between the 5-6 and the 7-8 strands directly interact with CN-Cb, while the extracellular loop between 9-10 is disordered. Moreover, there are minor changes in the

periplasmic loops. In the core domain, the TonB box undergoes significant conformational changes shifting position and rotating. However, in this structure, CN-Cb did not induce TonB box extension into the periplasm<sup>63</sup>. Thus, challenged the other experimental work, such as EPR studies done in our lab, where data shows evidence of TonB box unfolding upon substrate binding<sup>64, 65</sup>. However, these differences were due to the use of crystalline reagents (polyethylene glycols, salts and other osmolytes) that shift the TonB box equilibrium towards the folded state<sup>66</sup>. Nevertheless, in 2006 the crystal structure of BtuB bound to calcium, cyanocobalamin (CN-Cb), and TonB (2GSK) was solved which provided long awaited evidence that the TonB box unfolds and extends into the periplasm upon binding with CN-Cb (Figure 1.23b). Furthermore, based on this structure, the TonB box interacts with TonB as well<sup>67</sup>.

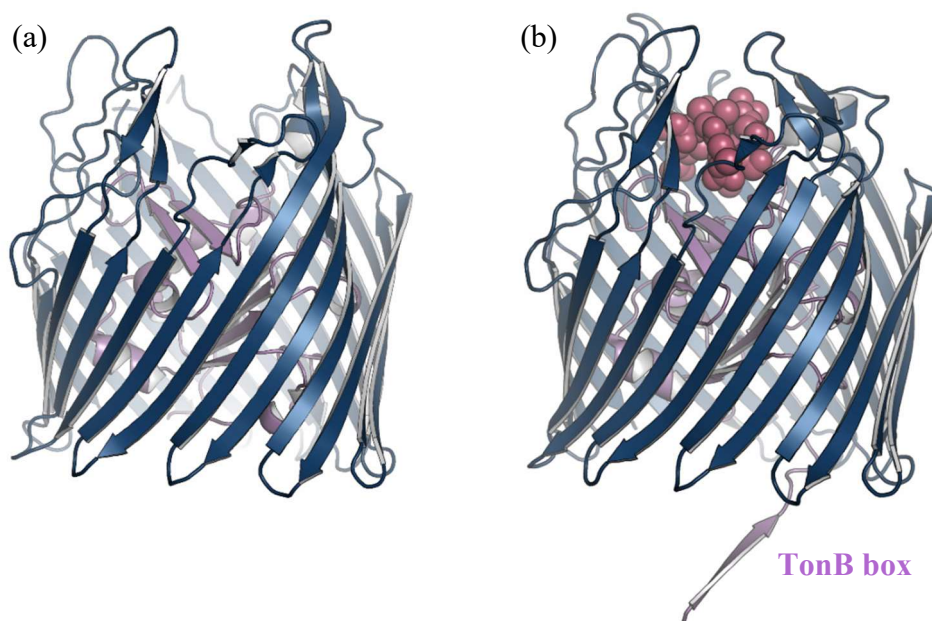


Figure 1.23. Structure of side view of (a) apo BtuB with calcium (PDB ID 1NQG) and (b) BtuB bound to CN-Cb (PDB ID 2GSK). The **core domain** occludes the **barrel** pore. Once the substrate (**CN-Cb**) binds, the **TonB box** unfolds and extends to the periplasm. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

In addition to above, *E. coli* encodes several other TBDTs (CirA, FhuE, FepA and Fiu) which transport siderophore-iron complexes, such as: ferric coprogen, ferric enterobactin, and degraded products of enterobactin<sup>68</sup>. There is a diversity among other bacterial TBDTs. (i) *P. aeruginosa* encodes FpvA and FetA which transport ferric pyoverdine and ferric pyochelin, respectively<sup>17</sup>. These siderophores can also bind with both  $\text{Ni}^{2+}$  and  $\text{Co}^{2+}$  which are then transferred by FpvA and FetA. (ii) Aside from siderophores, pathogenic bacteria, *Neisseria ssp.*, uses TbpA and LbpA which bind with transferrin and lactoferrin, respectively, then transport iron from these proteins<sup>16</sup>. (iii) *Serratia marcescens* and *Haemophilus influenzae* secrete hemophore to bind and uptake heme through HasA and HxuA, respectively. (iv) Hemoglobin in red blood cells is uptaken by TBDTs; HemR from *Yersinia enterocolitica* and HmbR from *N. meningitidis*. (v) Heptoglobin-hemoglobin transported via HpuA from *N. meningitidis*. (vi) Hemopexin is uptaken via HxuD of *H. influenzae* and HemR of *S. marcescens*<sup>16, 69</sup>. (vii) *Shigella dysenteriae* ShuA uptakes heme from methemoglobin<sup>70</sup>. (viii) *H. pylori* encodes FecA3, FrpB2 and FrpB4 to transport nickel into the cell<sup>68, 71</sup>.

#### 1.5.4 Current knowledge in ButB and other TBDTs

In addition to knowledge gained from the crystal structures, other biochemical and biophysical techniques have been employed to investigate TBDTs and their substrate transport mechanism. Since pathogenic bacteria utilize TBDTs to uptake iron from various sources from humans and other hosts, understanding mechanisms of TBDTs, as well as how to target TBDTs to deliver drugs, have drawn special attention from medicine and biotechnology.

Based on the current research progress for TBDTs, experiments have provided evidence for substrate induced conformational changes of TBDTs. These cause interactions with the IM

protein, TonB of the TonB complex (TonB-ExbB-ExbD) and enable transport of the bound substrate into the periplasm. Phenotype assays and uptake assays done using BtuB mutants, support the extracellular loops involvement in substrate binding, as seen in the crystal structure<sup>72</sup>. Based on EPR spectroscopic analysis of BtuB, either in native OM or in reconstituted lipid bilayer systems, has demonstrated that the TonB box of core domain in the periplasmic side unfolds upon binding to the substrate and extend about 20-30 Å into the periplasm<sup>73</sup>. During this substrate induced signal transduction process, the D316-R14 ion pair between the barrel and core acts as a molecular switch facilitating signal propagation and TonB box unfolding<sup>74</sup>. Furthermore, TonB interacts with unfolded TonB box with high affinity and sends a signal to the extracellular surface to partially dissociate the bound substrate eventually. Thus, releasing the substrate to periplasm<sup>75-77</sup>. This is also supported by cross-linking studies and uptake assays done using BtuB mutants<sup>64</sup>. Specific TonB box orientation is also required for proper functional contact with TonB<sup>78</sup>. Recently, intact cells expressing BtuB were used in EPR experiments in order to observe the conformational changes associated with the BtuB extracellular loops and to validate EPR data obtained from OM and lipid reconstituted systems with respect to intact cell systems<sup>79, 80</sup>. Although these measurements validated the results of EPR in native outer membranes and reconstituted systems, the preparations may not have been the most metabolically active and may not be suitable for some future studies. As a result, there was a need to explore other approaches to label intact cell systems (Chapter 3).

## 1.6 Bacteriocins

Bacteriocins are bacterial antibiotics, which target closely related bacterial species. It has been proposed that almost all bacteria produce at least one type of bacteriocin<sup>81</sup>. These proteins and peptides are naturally produced when bacteria undergo starvation and depletion of the available nutrients or due to an over-crowded population. Thus, to compete for survival, both Gram-negative and Gram-positive bacteria produce these toxins, targeting their competitors without lethal impacts on the host cells. Furthermore, these molecules are categorized into different classes: microcins, colicin-like-bacteriocins and tailocins in Gram-negative bacteria, and class I-III in Gram-positive bacteria<sup>81, 82</sup>.

Bacteriocins are natural toxins which target specific bacteria, there are potential applications on food preservation, pharmaceutical drug and healthcare product development, and use as potential anti-cancer agents. During food preservation, bacteriocins can be used as: (i) partially purified, (ii) fermented products, or (iii) protective cultures which produce bacteriocins. For example, *Lactococcus lactis*, *Lactobacillus sakei*, *Leuconostoc carnosum* and *Staphylococcus carnosus* produce nisin, sakacin, leucocin, and carnocin, respectively, that are used in commercially available food products. However, nisin is the only bacteriocin approved by FDA to be used as a food preservative<sup>83</sup>. Additionally, nisin is used as an antimicrobial agent against pathogenic Gram-positive and Gram-negative bacteria. As some pathogenic bacteria are gaining resistance to current available antibiotics, it is important to develop new drugs against these drug-resistance pathogens. One approach is to discover new natural antibiotics and to engineer them with current antibiotics. Consequently, more research on these biomolecules can lead to develop new highly effective antimicrobial agents with high specificity to targeting bacterial species.

### 1.6.1 Bacteriocins from *E. coli*

Compared to other bacteriocins, bacteriocins produced by *E. coli* (colicins) have been widely studied using molecular biology, biochemical and biophysical techniques. These include several solved crystal structures of colicin bound to their receptors and transporters. Based on these studies, colicins are divided into two types: group A and group B. For example, colicin A, E1-E9, K and N are group A colicins, while colicin B, D, M and Ia are group B<sup>14, 15</sup>.

### 1.6.2 Expression of colicins

*E. coli* strains that express colicins are known as colicinogenic strains which contain colicinogenic plasmids (pCol) that have colicin operons. There are two types of pCols, namely: group A type and group B type. Group A pCols are small (6-10 kbp) with about 20 copies per cell, while group B plasmids are large (about 40 kbp) with a single copy per cell. There are between one to three genes in colicin operons. These operons contain a colicin x activity gene (cxa), an immunity gene (cxi), and a lysis gene (cxl), which encode the corresponding colicin, immunity protein which binds with endogenous colicin and protects the host cell and a lysis protein which releases colicin from host cells, respectively<sup>14</sup>. Under normal cellular conditions, colicin operon expression is tightly repressed by an SOS repressor, LexA. However, when SOS responses are triggered they activate RecA which stimulates the autocleavage of LexA and leads to transcription of colicin genes.

### 1.6.3 Structure and functions of group A colicins

All the identified colicins consist of a common architecture: an N-terminus translocation (T) domain, followed by a reception (R) domain, and a C-terminus killing (C) domain<sup>15</sup>. Each of these domains play key roles in invasion and cytotoxicity. During invasion, via R domain, colicins bind with their targeted outer membrane receptors with high affinity. For some colicins,

(Ia, Ib and colicin N) the primary receptors also act as translocators, while for other colicins, another OM receptor mediates the translocation. After binding to the receptor, the T domain transfers the bound colicin from the OM to the periplasm, or through the periplasm. During the translocation process, group A colicins utilize Tol (tolerant) proteins and group B colicins use TonB, ExbB and ExbD. During the invasion step, colicins' C domains kill the invaded cells by one of the two main killing modes: catalytic activity or channel/pore formation, which are present in both group A and group B types. Pore forming colicins bind and insert into the periplasm in order to form channels in the IM and to depolarize the membrane and kill invaded bacteria. Catalytically active colicins can either have DNase, RNase, or peptidoglycan degradation. Currently, colicin M is the only known peptidoglycan precursor degrading colicin that acts in the periplasm, while other enzymatic colicins (nuclease) function in the cytoplasm.

To date, there are few colicin crystal structures available. These include, colicin E3, E9 and N from group A and colicin B, Ia and M from group B. In the colicin E3 structure (PDB ID 1JCH, Figure 1.24a), the first 83 amino acids of the T domain are unstructured and the rest of the domain forms  $\beta$  sheets within two helices. The C-terminal helix of the T domain is 30 Å long and directly connected to the R domain, which assembles into a 100 Å long antiparallel hairpin. The tip of the hairpin contains a BtuB binding site. The R and C domains are connected via a small linker. The C domain consists of six antiparallel  $\beta$  sheets, which have RNase activity. Colicin E3 and E9 have about 90% of identity with respect to T and R domains<sup>84</sup>. However, the colicin E9 is a DNase rather than an RNase. Based on the crystal structure solved by Klein A. and *et.al.*, (PDB ID 5EW5), the colicin E9 N-terminus is unstructured following the T domain, which consists of  $\beta$  strands and loops (Figure 1.24b). The colicin E9R domain forms a 100 Å long coiled coil structure with a hairpin between 376-384 amino acids that bind to BtuB. The C

domain is formed by  $\alpha$  helices,  $3_{10}$  helices and  $\beta$  strands<sup>85</sup>. Relative to these two structures, in the pore-forming colicin N (PDB ID 1A87, Figure 1.24c) structure solved by Vetter I.R. and *et.al.*, (42 kDa, 387 amino acids), the C and R domains have distinct structural features, while the T domain is unstructured. The  $\alpha$  helices in the C domain are formed into a globin-like domain, which includes a 63 Å long helix wrapped by six antiparallel  $\beta$  sheets forming the R domain containing OmpF binding site<sup>86</sup>.

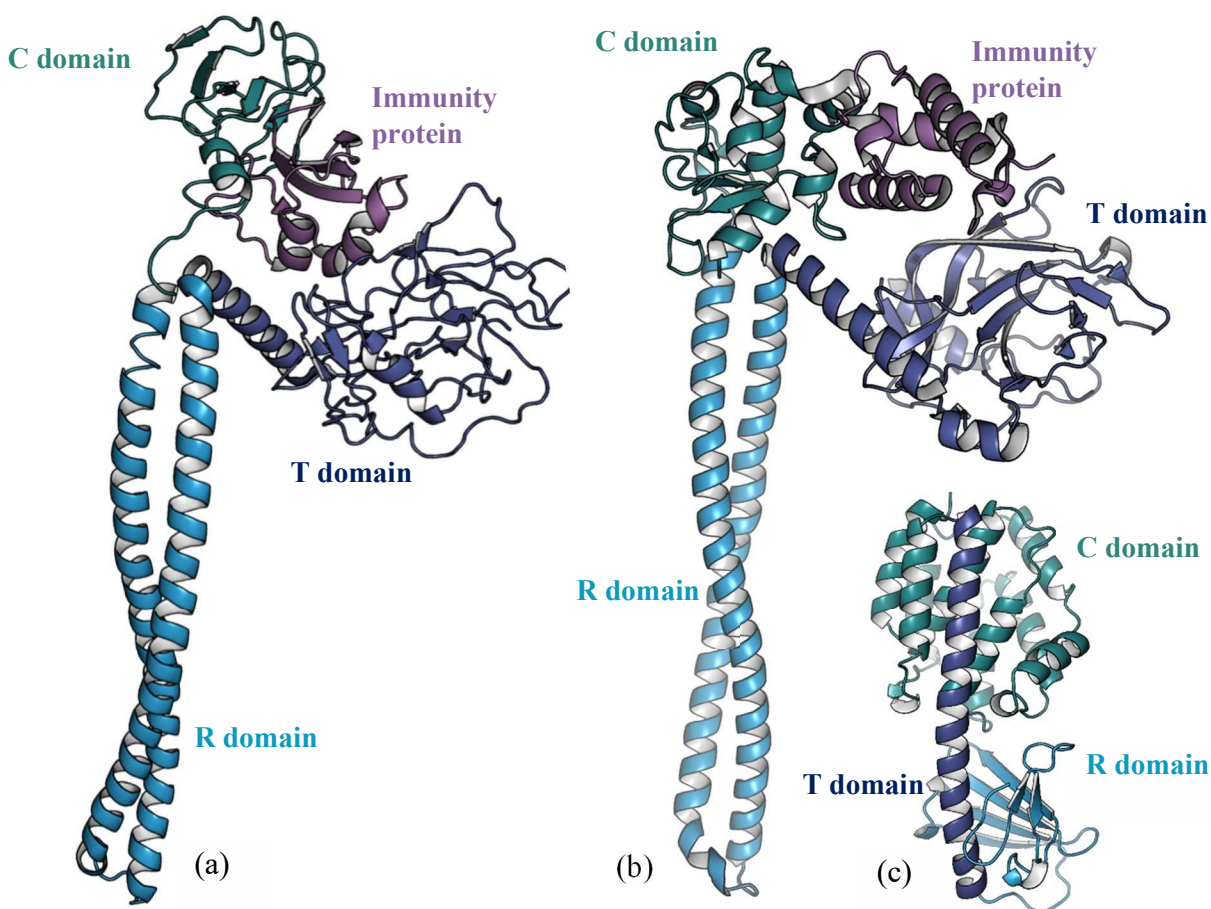


Figure 1.24. Crystal structures of group A colicins. (a) Colicin E3 (PDB ID 1JCH), (b) Colicin E9 (PDB ID 5EW5) and (c) Colicin N (PDB ID 1A87) fold into a **T domain**, **R domain** and **C domain**. Both colicin E3 and E9 are crystallized along with the corresponding **immunity proteins**. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

#### 1.6.4 Translocation mechanism of group A colicins

During bacterial invasion, Group A colicins lacking the TonB box sequence utilize TolA, B, Q and R for translocation. It has been proposed that after binding to their primary OM receptor, the T domain screens for a corresponding translocator in the OM (the fishing pole mechanism)<sup>15</sup>. For example, colicin A, E1-9 use BtuB as their primary binding receptor. Colicin E3R binding with BtuB was studied and is discussed in Appendix A1. The T domains of colicin A, E2-9 are used as fishing hooks to locate OmpF, while colicin E1 screening for export OMP, TolC for import (Figure 1.25)<sup>15</sup>. Searching for a second OMP, rather than BtuB, is also supported by the findings of Cadieux and co-workers, who demonstrated that mutating TonB Box in BtuB impacts the efficiency of B<sub>12</sub> uptake without affecting the sensitivity to colicin E1 and E3<sup>78</sup>. Thus, the interactions between the unfolded TonB box of BtuB and TonB is not required for these colicins.

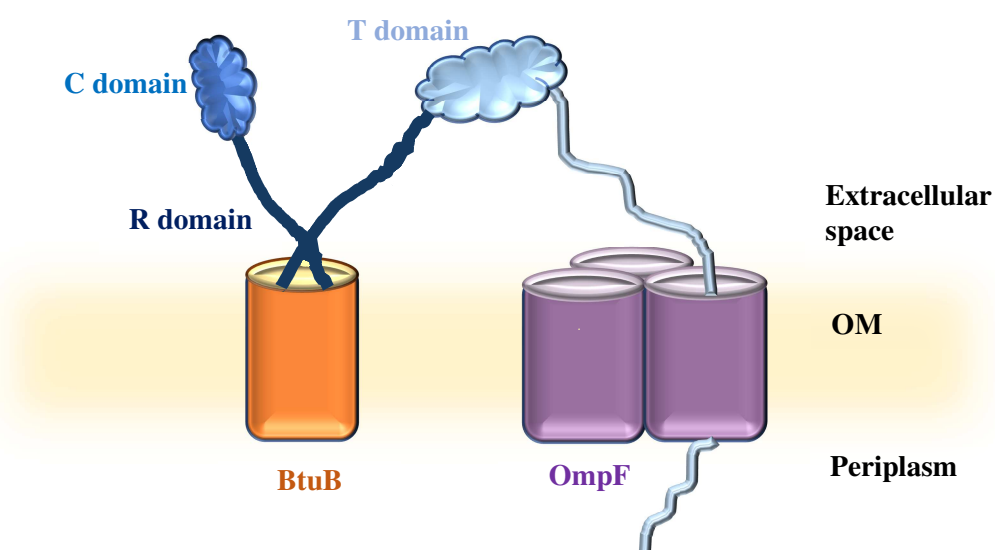


Figure 1.25. Proposed mechanism for translocation of colicin E3 through OM. Once **R domain** binds with **BtuB**, the receptor induces conformational changes lead to dissociation of the **C domain** and **T domain** from the immunity protein. **T domain** of colicin E3 screens for **OmpF** and passes through one **OmpF** unit.

### 1.6.5 Structure and functions of group B colicins

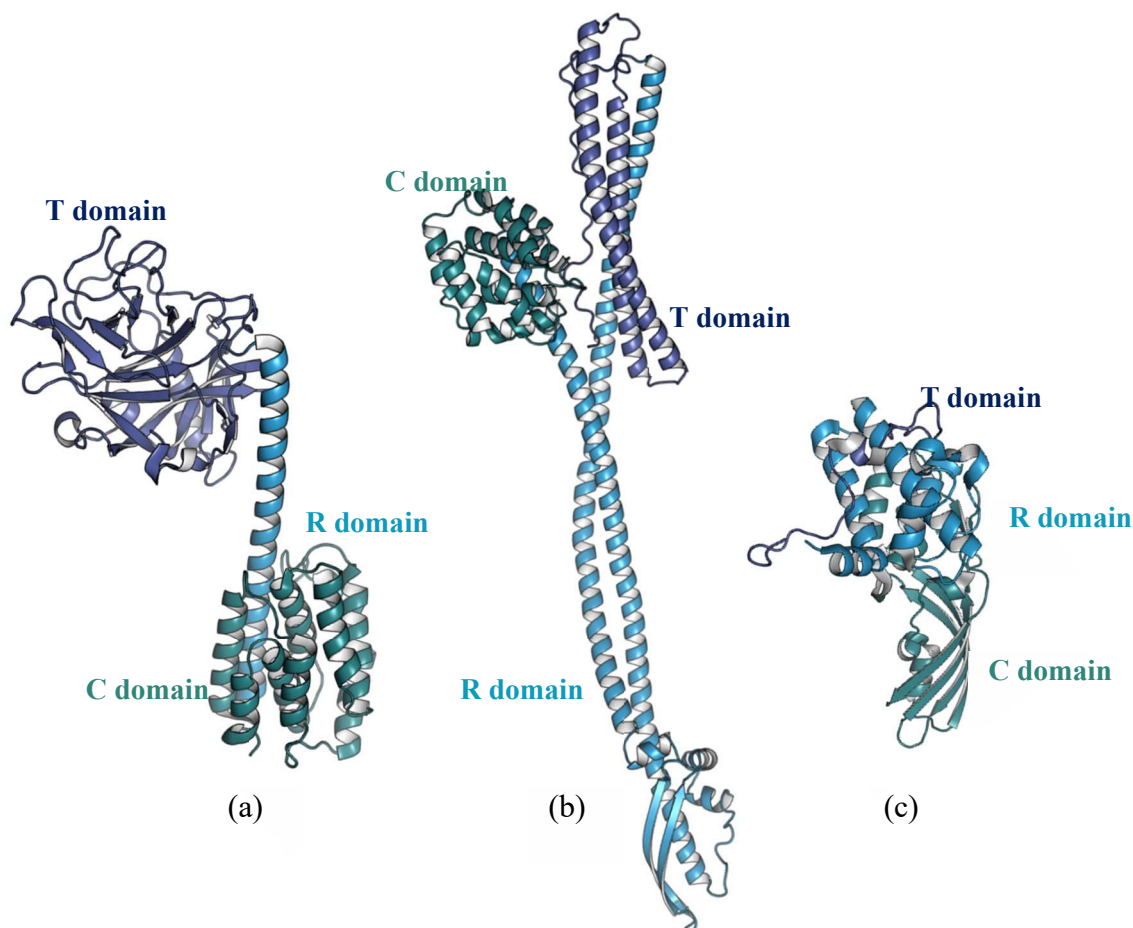


Figure 1.26. Crystal structures of group B colicins. (a) Colicin B (PDB ID 1RH1), (b) Colicin Ia (PDB ID 1CII) and (c) Colicin M (PDB ID 2XMX) fold into a **T domain**, **R domain** and **C domain**. Crystal structures were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>30</sup> and the figures were generated using PyMol Molecular Graphics System version 1.8<sup>31</sup>.

Group B colicins (B, Ia and M) have similar structural features relative to group A colicin structures. Colicin B (Figure 1.26a) is a dumbbell-like structure with two lobes connected via a long  $\alpha$ -helix ( $\sim 74$  Å). The N-terminal lobe assembles into a  $\beta$ -barrel. Unlike the previous colicin examples, the N-terminal lobe contains both the T and R domains, without well separated structural features. Thus, this lobe contains both the TonB box sequence as well as the FepA

binding site. In contrast, the C-terminal lobe is composed of a bundle of  $\alpha$ -helices which assemble the pore-forming C domain<sup>87</sup>. Compared to colicin B, colicin Ia (Figure 1.26b) is a 210 Å long molecule with three well separated and distinctive domains. The T domains' N-terminus contains the TonB box sequence which is followed by long  $\alpha$  helices. The R domain contains a Cir binding site at the hairpin of the long antiparallel  $\alpha$ -helices. The pore-forming C domain is assembled into a bundle of ten  $\alpha$ -helices<sup>88</sup>. In contrast to these two colicins, Colicin M (Figure 1.26c) is a small protein (271 amino acids) which forms into a compact structure containing distinct domains separated by a few amino acid residues. The TonB box containing T domain forms into a loop-like structure, while the R domain is a globule of  $\alpha$ -helices, including a FhuA binding site. Additionally, the C domain forms an opened  $\beta$ -barrel that has phosphatase activity<sup>89</sup>.

#### 1.6.6 Mechanism of group B colicin translocation

Group B colicins, also known as TonB dependent colicins, utilize TonB, ExbB and ExbD proteins for translocation processes. With the exception of the small colicin Js (94 amino acid), all group B colicins have a Ton Box sequence in their N-terminus. Once bound to the TBBDTs, both the receptor's and the bound colicin's TonB boxes interact with two TonBs for translocation<sup>90, 91</sup>. Mutating either of these two TonB boxes impairs translocation without affecting either the receptor binding or the catalytic activity<sup>90, 91</sup>. During this process, colicin Ia uses two copies of CirA for initial binding and for translocation<sup>92</sup>. However, it is still unclear whether or not other TonB dependent colicins also use two copies of receptor proteins for translocation process. For example, colicin M has a compact structure that might not be able to locate a second OMP, so it may employ another translocation mechanism.

## 1.7 References

- [1] Wang, J. D., and Levin, P. A. (2009) Metabolism, cell growth and the bacterial cell cycle, *Nature reviews. Microbiology* 7, 822-827.
- [2] Haeusser, D. P., and Levin, P. A. (2008) The Great Divide: Coordinating cell cycle events during bacterial growth and division, *Current opinion in microbiology* 11, 94-99.
- [3] Vadia, S., and Levin, P. A. (2015) Growth rate and cell size: a re-examination of the growth law, *Current Opinion in Microbiology* 24, 96-103.
- [4] Nyström, T. (2007) A Bacterial Kind of Aging, *PLoS Genetics* 3, e224.
- [5] Nyström, T. (2002) Aging in bacteria, *Current Opinion in Microbiology* 5, 596-601.
- [6] Stewart, E. J., Madden, R., Paul, G., and Taddei, F. (2005) Aging and death in an organism that reproduces by morphologically symmetric division, *PLoS Biol* 3, e45.
- [7] Laloux, G., and Jacobs-Wagner, C. (2014) How do bacteria localize proteins to the cell pole?, *Journal of Cell Science* 127, 11-19.
- [8] Christman, M. F., Storz, G., and Ames, B. N. (1989) OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins, *Proceedings of the National Academy of Sciences of the United States of America* 86, 3484-3488.
- [9] Hassan, H. M., and Sun, H. C. (1992) Regulatory roles of Fnr, Fur, and Arc in expression of manganese-containing superoxide dismutase in *Escherichia coli*, *Proc Natl Acad Sci U S A* 89, 3217-3221.
- [10] Loui, C., Chang, A. C., and Lu, S. (2009) Role of the ArcAB two-component system in the resistance of *Escherichia coli* to reactive oxygen stress, *BMC Microbiology* 9, 183-183.
- [11] Iuchi, S., and Weiner, L. (1996) Cellular and Molecular Physiology of *Escherichia coli* in the Adaptation to Aerobic Environments, *The Journal of Biochemistry* 120, 1055-1063.
- [12] Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010) TonB-dependent transporters: regulation, structure, and function, *Annual review of microbiology* 64, 43-60.
- [13] Palma, M., Zurita, J., Ferreras, J. A., Worgall, S., Larone, D. H., Shi, L., Campagne, F., and Quadri, L. E. N. (2005) *Pseudomonas aeruginosa* SoxR Does Not Conform to the Archetypal Paradigm for SoxR-Dependent Regulation of the Bacterial Oxidative Stress Adaptive Response, *Infection and Immunity* 73, 2958-2966.
- [14] Cascales, E., Buchanan, S. K., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S., and Cavard, D. (2007) Colicin biology, *Microbiology and molecular biology reviews : MMBR* 71, 158-229.
- [15] Jakes, K. S., and Cramer, W. A. (2012) Border crossings: colicins and transporters, *Annual review of genetics* 46, 209-231.
- [16] Schauer, K., Rodionov, D. A., and de Reuse, H. (2008) New substrates for TonB-dependent transport: do we only see the tip of the iceberg?, *Trends in Biochemical Sciences* 33, 330-338.
- [17] Schalk, I. J., Mislin, G. L., and Brillet, K. (2012) Structure, function and binding selectivity and stereoselectivity of siderophore-iron outer membrane transporters, *Current topics in membranes* 69, 37-66.
- [18] Leyton, D. L., Belousoff, M. J., and Lithgow, T. (2015) The beta-Barrel Assembly Machinery Complex, *Methods in molecular biology (Clifton, N.J.)* 1329, 1-16.

- [19] Plummer, A. M., and Fleming, K. G. (2016) From Chaperones to the Membrane with a BAM!, *Trends Biochem Sci* 41, 872-882.
- [20] Nikaido, H. (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited, *Microbiology and Molecular Biology Reviews* 67, 593-656.
- [21] Takayama, K., Qureshi, N., Hyver, K., Honovich, J., Cotter, R. J., Mascagni, P., and Schneider, H. (1986) Characterization of a structural series of lipid A obtained from the lipopolysaccharides of *Neisseria gonorrhoeae*. Combined laser desorption and fast atom bombardment mass spectral analysis of high performance liquid chromatography-purified dimethyl derivatives, *The Journal of biological chemistry* 261, 10624-10631.
- [22] Wilkinson, S. G. (1996) Bacterial lipopolysaccharides--themes and variations, *Progress in lipid research* 35, 283-343.
- [23] Ogawa, T. (1993) Chemical structure of lipid A from *Porphyromonas* (Bacteroides) gingivalis lipopolysaccharide, *FEBS Lett* 332, 197-201.
- [24] Jeong, H., Barbe, V., Lee, C. H., Vallenet, D., Yu, D. S., Choi, S. H., Couloux, A., Lee, S. W., Yoon, S. H., Cattolico, L., Hur, C. G., Park, H. S., Segurens, B., Kim, S. C., Oh, T. K., Lenski, R. E., Studier, F. W., Daegelen, P., and Kim, J. F. (2009) Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3), *Journal of molecular biology* 394, 644-652.
- [25] Konovalova, A., Kahne, D. E., and Silhavy, T. J. (2017) Outer Membrane Biogenesis, *Annu Rev Microbiol* 71, 539-556.
- [26] Koebnik, R., Locher, K. P., and Van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell, *Mol Microbiol* 37, 239-253.
- [27] Ralf, K., P., L. K., and Patrick, V. G. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell, *Molecular Microbiology* 37, 239-253.
- [28] Green, E. R., and Mecsas, J. (2016) Bacterial Secretion Systems – An overview, *Microbiology spectrum* 4, 10.1128/microbiolspec.VMBF-0012-2015.
- [29] Kostakioti, M., Newman, C. L., Thanassi, D. G., and Stathopoulos, C. (2005) Mechanisms of protein export across the bacterial outer membrane, *J Bacteriol* 187, 4306-4314.
- [30] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Research* 28, 235-242.
- [31] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.
- [32] Whitfield, C., and Trent, M. S. (2014) Biosynthesis and export of bacterial lipopolysaccharides, *Annu Rev Biochem* 83, 99-128.
- [33] Maldonado, R. F., Sá-Correia, I., and Valvano, M. A. (2016) Lipopolysaccharide modification in Gram-negative bacteria during chronic infection, *FEMS Microbiology Reviews* 40, 480-493.
- [34] Bishop, R. E. (2014) Structural biology: Lipopolysaccharide rolls out the barrel, *Nature* 511, 37-38.
- [35] Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C., and Huang, Y. (2014) Structural basis for lipopolysaccharide insertion in the bacterial outer membrane, *Nature* 511, 108-111.
- [36] Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., He, C., Zhang, Y., Wang, W., and Dong, C. (2014) Structural basis for outer membrane lipopolysaccharide insertion, *Nature* 511, 52-56.
- [37] Kovacs-Simon, A., Titball, R. W., and Michell, S. L. (2011) Lipoproteins of bacterial pathogens, *Infection and Immunity* 79, 548-561.

- [38] Takeda, K., Miyatake, H., Yokota, N., Matsuyama, S.-i., Tokuda, H., and Miki, K. (2003) Crystal structures of bacterial lipoprotein localization factors, LolA and LolB, *The EMBO Journal* 22, 3199-3209.
- [39] Walton, T. A., and Sousa, M. C. (2004) Crystal Structure of Skp, a Prefoldin-like Chaperone that Protects Soluble and Membrane Proteins from Aggregation, *Molecular Cell* 15, 367-374.
- [40] Ortega, J., Iwanczyk, J., and Jomaa, A. (2009) Escherichia coli DegP: a Structure-Driven Functional Model, *Journal of Bacteriology* 191, 4705-4713.
- [41] Saul, F. A., Arie, J. P., Vulliez-le Normand, B., Kahn, R., Betton, J. M., and Bentley, G. A. (2004) Structural and functional studies of FkpA from Escherichia coli, a cis/trans peptidyl-prolyl isomerase with chaperone activity, *Journal of molecular biology* 335, 595-608.
- [42] Xu, X., Wang, S., Hu, Y. X., and McKay, D. B. (2007) The periplasmic bacterial molecular chaperone SurA adapts its structure to bind peptides in different conformations to assert a sequence preference for aromatic residues, *Journal of molecular biology* 373, 367-381.
- [43] McMorran, L. M., Brockwell, D. J., and Radford, S. E. (2014) Mechanistic studies of the biogenesis and folding of outer membrane proteins in vitro and in vivo: what have we learned to date?, *Archives of biochemistry and biophysics* 564, 265-280.
- [44] Kim, K. H., Aulakh, S., and Paetzel, M. (2012) The bacterial outer membrane beta-barrel assembly machinery, *Protein science : a publication of the Protein Society* 21, 751-768.
- [45] Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., Zhao, Y., and Huang, Y. (2016) Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins, *Nature structural & molecular biology* 23, 192-196.
- [46] Selkrig, J., Leyton, D. L., Webb, C. T., and Lithgow, T. (2014) Assembly of  $\beta$ -barrel proteins into bacterial outer membranes, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1843, 1542-1550.
- [47] Hohr, A. I. C., Lindau, C., Wirth, C., Qiu, J., Stroud, D. A., Kutik, S., Guiard, B., Hunte, C., Becker, T., Pfanner, N., and Wiedemann, N. (2018) Membrane protein insertion through a mitochondrial beta-barrel gate, *Science* 359.
- [48] Berkmen, M. (2012) Production of disulfide-bonded proteins in Escherichia coli, *Protein Expression and Purification* 82, 240-251.
- [49] Ren, G., Champion, M. M., and Huntley, J. F. (2014) Identification of disulfide bond isomerase substrates reveals bacterial virulence factors, *Mol Microbiol* 94, 926-944.
- [50] Heras, B., Shouldice, S. R., Totsika, M., Scanlon, M. J., Schembri, M. A., and Martin, J. L. (2009) DSB proteins and bacterial pathogenicity, *Nature Reviews Microbiology* 7, 215.
- [51] Goemans, C., Denoncin, K., and Collet, J.-F. (2014) Folding mechanisms of periplasmic proteins, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1843, 1517-1528.
- [52] Kadokura, H., and Beckwith, J. (2009) Detecting Folding Intermediates of a Protein as It Passes through the Bacterial Translocation Channel, *Cell* 138, 1164-1173.
- [53] Kadokura, H., and Beckwith, J. (2010) Mechanisms of Oxidative Protein Folding in the Bacterial Cell Envelope, *Antioxidants & Redox Signaling* 13, 1231-1246.
- [54] Inaba, K. (2009) Disulfide Bond Formation System in Escherichia coli, *The Journal of Biochemistry* 146, 591-597.

- [55] Young, Y. J., Jieun, K., Jae, L. S., Sook, K. H., Na, I. H., Hye-Jin, Y., Hoon, K. K., Soon-Jong, K., Woo, H. B., and Won, S. S. (2011) Structural and functional characterization of *Helicobacter pylori* DsbG, *FEBS Letters* 585, 3862-3867.
- [56] Depuydt, M., Leonard, S. E., Vertommen, D., Denoncin, K., Morsomme, P., Wahni, K., Messens, J., Carroll, K. S., and Collet, J. F. (2009) A periplasmic reducing system protects single cysteine residues from oxidation, *Science* 326, 1109-1111.
- [57] Landeta, C., Boyd, D., and Beckwith, J. (2018) Disulfide bond formation in prokaryotes, *Nature Microbiology* 3, 270-280.
- [58] Roszczenko, P., Radomska, K. A., Wywiał, E., Collet, J.-F., and Jagusztyn-Krynicka, E. K. (2012) A Novel Insight into the Oxidoreductase Activity of *Helicobacter pylori* HP0231 Protein, *PLoS ONE* 7, e46563.
- [59] Arts, I. S., Ball, G., Leverrier, P., Garvis, S., Nicolaes, V., Vertommen, D., Ize, B., Tamu Dufe, V., Messens, J., Voulhoux, R., and Collet, J. F. (2013) Dissecting the machinery that introduces disulfide bonds in *Pseudomonas aeruginosa*, *mBio* 4, e00912-00913.
- [60] Cho, S.-H., and Collet, J.-F. (2013) Many Roles of the Bacterial Envelope Reducing Pathways, *Antioxidants & Redox Signaling* 18, 1690-1698.
- [61] Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K., and Welte, W. (1998) Siderophore-Mediated Iron Transport: Crystal Structure of FhuA with Bound Lipopolysaccharide, *Science* 282, 2215.
- [62] Ferguson, A. D., Chakraborty, R., Smith, B. S., Esser, L., van der Helm, D., and Deisenhofer, J. (2002) Structural Basis of Gating by the Outer Membrane Transporter FecA, *Science* 295, 1715.
- [63] Chimento, D. P., Mohanty, A. K., Kadner, R. J., and Wiener, M. C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB, *Nature structural biology* 10, 394-401.
- [64] Cadieux, N., and Kadner, R. J. (1999) Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter, *Proceedings of the National Academy of Sciences of the United States of America* 96, 10673-10678.
- [65] Fanucci, G. E., Cogshall, K. A., Cadieux, N., Kim, M., Kadner, R. J., and Cafiso, D. S. (2003) Substrate-Induced Conformational Changes of the Periplasmic N-Terminus of an Outer-Membrane Transporter by Site-Directed Spin Labeling, *Biochemistry* 42, 1391-1400.
- [66] Flores Jiménez, R. H., Do Cao, M.-A., Kim, M., and Cafiso, D. S. (2010) Osmolytes modulate conformational exchange in solvent-exposed regions of membrane proteins, *Protein science : a publication of the Protein Society* 19, 269-278.
- [67] Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex, *Science* 312, 1396-1399.
- [68] Krewulak, K. D., and Vogel, H. J. (2011) TonB or not TonB: is that the question? This paper is one of a selection of papers published in a Special Issue entitled CSBMCB 53rd Annual Meeting — Membrane Proteins in Health and Disease, and has undergone the Journal's usual peer review process, *Biochemistry and Cell Biology* 89, 87-97.
- [69] Wandersman, C., and Delepelaire, P. (2004) Bacterial iron sources: from siderophores to hemophores, *Annu Rev Microbiol* 58, 611-647.

- [70] Meneghini, L. M., Tripathi, S., Woodworth, M. A., Majumdar, S., Poulos, T. L., and Weiss, G. A. (2017) Dissecting binding of a beta-barrel membrane protein by phage display, *Molecular bioSystems* 13, 1438-1447.
- [71] Haley, K. P., and Gaddy, J. A. (2015) Metalloregulation of *Helicobacter pylori* physiology and pathogenesis, *Frontiers in Microbiology* 6, 911.
- [72] Fuller-Schaefer, C. A., and Kadner, R. J. (2005) Multiple Extracellular Loops Contribute to Substrate Binding and Transport by the *Escherichia coli* Cobalamin Transporter BtuB, *Journal of Bacteriology* 187, 1732-1739.
- [73] Xu, Q., Ellena, J. F., Kim, M., and Cafiso, D. S. (2006) Substrate-Dependent Unfolding of the Energy Coupling Motif of a Membrane Transport Protein Determined by Double Electron-Electron Resonance, *Biochemistry* 45, 10847-10854.
- [74] Lukasik, S. M., David Ho, K. W., and Cafiso, D. S. (2007) Molecular Basis for Substrate-Dependent Transmembrane Signaling in an Outer-Membrane Transporter, *Journal of molecular biology* 370, 807-811.
- [75] Freed, D. M., Lukasik, S. M., Sikora, A., Mokdad, A., and Cafiso, D. S. (2013) Monomeric TonB and the Ton box are required for the formation of a high-affinity transporter-TonB complex, *Biochemistry* 52, 2638-2648.
- [76] Sikora, A., Joseph, B., Matson, M., Staley, J. R., and Cafiso, D. S. (2016) Allosteric Signaling Is Bidirectional in an Outer-Membrane Transport Protein, *Biophysical journal* 111, 1908-1918.
- [77] Cadieux, N., Barekzi, N., and Bradbeer, C. (2007) Observations on the calcium dependence and reversibility of cobalamin transport across the outer membrane of *Escherichia coli*, *The Journal of biological chemistry* 282, 34921-34928.
- [78] Cadieux, N., Bradbeer, C., and Kadner, R. J. (2000) Sequence Changes in the Ton Box Region of BtuB Affect Its Transport Activities and Interaction with TonB Protein, *Journal of Bacteriology* 182, 5954-5961.
- [79] Joseph, B., Sikora, A., Bordignon, E., Jeschke, G., Cafiso, D. S., and Prisner, T. F. (2015) Distance Measurement on an Endogenous Membrane Transporter in *E. coli* Cells and Native Membranes Using EPR Spectroscopy, *Angewandte Chemie (International ed. in English)* 54, 6196-6199.
- [80] Joseph, B., Sikora, A., and Cafiso, D. S. (2016) Ligand Induced Conformational Changes of a Membrane Transporter in *E. coli* Cells Observed with DEER/PELDOR, *Journal of the American Chemical Society* 138, 1844-1847.
- [81] Rea, M. C., Ross, R. P., Cotter, P. D., and Hill, C. (2011) Classification of Bacteriocins from Gram-Positive Bacteria, In *Prokaryotic Antimicrobial Peptides: From Genes to Applications* (Drider, D., and Rebuffat, S., Eds.), pp 29-53, Springer New York, New York, NY.
- [82] Ghequire, M. G. K., and De Mot, R. (2014) Ribosomally encoded antibacterial proteins and peptides from *Pseudomonas*, *FEMS Microbiology Reviews* 38, 523-568.
- [83] Chikindas, M. L., Weeks, R., Drider, D., Chistyakov, V. A., and Dicks, L. M. (2018) Functions and emerging applications of bacteriocins, *Current opinion in biotechnology* 49, 23-28.
- [84] Soelaiman, S., Jakes, K., Wu, N., Li, C., and Shoham, M. (2001) Crystal structure of colicin E3: implications for cell entry and ribosome inactivation, *Mol Cell* 8, 1053-1062.

- [85] Klein, A., Wojdyla, J. A., Joshi, A., Josts, I., McCaughey, L. C., Housden, N. G., Kaminska, R., Byron, O., Walker, D., and Kleanthous, C. (2016) Structural and biophysical analysis of nuclease protein antibiotics, *The Biochemical journal* 473, 2799-2812.
- [86] Vetter, I. R., Parker, M. W., Tucker, A. D., Lakey, J. H., Pattus, F., and Tsernoglou, D. (1998) Crystal structure of a colicin N fragment suggests a model for toxicity, *Structure (London, England : 1993)* 6, 863-874.
- [87] Hilsenbeck, J. L., Park, H., Chen, G., Youn, B., Postle, K., and Kang, C. (2004) Crystal structure of the cytotoxic bacterial protein colicin B at 2.5 Å resolution, *Mol Microbiol* 51, 711-720.
- [88] Wiener, M., Freymann, D., Ghosh, P., and Stroud, R. M. (1997) Crystal structure of colicin Ia, *Nature* 385, 461-464.
- [89] Zeth, K., Romer, C., Patzer, S. I., and Braun, V. (2008) Crystal structure of colicin M, a novel phosphatase specifically imported by *Escherichia coli*, *The Journal of biological chemistry* 283, 25324-25331.
- [90] Braun, V., Patzer, S. I., and Hantke, K. (2002) Ton-dependent colicins and microcins: modular design and evolution, *Biochimie* 84, 365-380.
- [91] Buchanan, S. K., Lukacik, P., Grizot, S., Ghirlando, R., Ali, M. M., Barnard, T. J., Jakes, K. S., Kienker, P. K., and Esser, L. (2007) Structure of colicin I receptor bound to the R-domain of colicin Ia: implications for protein import, *Embo j* 26, 2594-2604.
- [92] Jakes, K. S., and Finkelstein, A. (2010) The colicin Ia receptor, Cir, is also the translocator for colicin Ia, *Mol Microbiol* 75, 567-578.

## Chapter 2. Materials and Main Techniques

### 2.1 Materials

pAG1 plasmid with *btuB* gene was kindly provided by the late Professor R. Kadner, University of Virginia. The pET 17b plasmid containing the wild type (WT) colicin E3R gene (*col E3R*) encoding R343 to E417 with an N-terminus methionine and C-terminal His<sub>6</sub> tag was kindly provided by Professor W. Cramer, University of Purdue, IN.

*E. coli* Dsb knockout strains; the wild type strain RI89 (*araD139 Δ(araABC-leu)7679 galU galK Δ(lac)X74 rpsL thi phoR Δara714 leu<sup>+</sup>*), *dsbA<sup>-</sup>* strain RI90 (RI89 *dsbA::Kan<sup>r</sup>*), *dsbC<sup>-</sup>* strain RI179 is RI89 *ΔdsbC::Cam<sup>r</sup>* and *dsbB<sup>-</sup>* strain RI317 (RI89 *dsbB::Kan<sup>r</sup>*) and RK5016 (*-argH, -btuB, -metE*) were kindly provided by Professor R. Nakamoto, University of Virginia and used for BtuB expression. *E. coli* strains, Top10 and BL21 (DE3) cell lines were purchased from Thermo Fisher Scientific (Waltham, MA) and Agilent technologies (Santa Clara, CA) and were used for plasmid propagation and colicin E3R expression, respectively.

The primers for PIPE mutagenesis were purchased from Integrated DNA Technology (Coralville, IA). *n*-octyl-β-D-glucopyranoside (OG) is from Anatrace (Maumee, OH). 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) is from Avanti Polar Lipids (Alabaster, AL). HiTrap® Q High Performance columns are from GE Healthcare (Chicago, IL).

(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl labeled cyanocobalamin (TEMPO-B<sub>12</sub>) was synthesized and provided by Dr. B. Joseph, University of Frankfurt, Germany. S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonylthioate (MTSSL) and 1-acetyl-2,2,5,5-tetramethyl-Δ<sup>3</sup>-(pyrroline-15N)-3-methyl methanethiosulfonate were purchased from Cayman Chemical (Ann Arbor, MI) and Toronto Research Chemicals (North York, Canada),

respectively. Glass capillaries (0.84 mm O.D. / 0.6 mm I.D.) and quartz capillaries (1.6 O.D.) were ordered from VitroCom (Mountain Lakes, NJ) for EPR analysis.

## 2.2 Molecular biology techniques

Various molecular biology, biochemical and biophysical techniques are used to study the structures and functions of proteins of interest. In order to study these native and mutant proteins at the molecular level, the genes encoding the proteins should be first cloned into appropriate expression vectors and then mutant genes can be generated, via mutagenesis. Hence, the success at the DNA level is important in these studies.

### 2.2.1 Polymerase chain reaction

One of the main molecular biology (MB) techniques is polymerase chain reaction (PCR). PCR involves manipulation and amplification of DNA of interest in a simple and reliable manner<sup>1</sup>. This technique was heavily used during this thesis work as a tool to mutate *btuB* and *colE3R* genes encoding *E. coli* BtuB and colicin E3 receptor domain, respectively. Each PCR cycle consists of 3 main stages. (i) Denaturation, in which the template double stranded (ds) DNA is denatured at around 94-98° C. (ii) This stage is followed by an annealing step where primers bind with the template DNA<sup>2</sup>. (iii) Next, the extension step occurs where the polymerase enzyme (ex. Pfu polymerase) elongates the primers, which are annealed to the template DNA strands<sup>3</sup>. This cycle continues for a designated number of times (cycles) and finally holds the samples at 4° C.

Depending on the PCR application, sometimes the mother plasmid or the template DNA can be removed using enzymatic digestion (DpnI). DpnI is a restriction enzyme that recognizes DNA strands containing methylations and digests at specific methylated sites (Figure 2.1)<sup>4</sup>. Template

DNA is methylated and amplified DNA strands are un-methylated, thus DpnI removes only the methylated template DNA. Next, the PCR amplified samples are analyzed by agarose gel electrophoresis.

During gel electrophoresis, the amplified DNA products migrate toward the anode. The rate of migration depends not only on the size of amplified DNA, but also the percentage of agarose, presence and the concentration of intercalating agents, and DNA conformation (circular, nicked or linear DNA)<sup>1</sup>. This small extra step allows researchers to detect properly amplified PCR products and use only those samples for downstream processes, such as transformation and cloning without wasting available time and resources, such as competent cells.

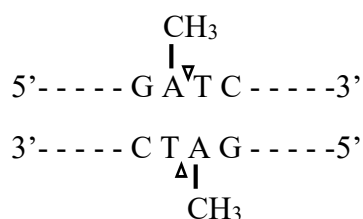


Figure 2.1 DpnI recognition site. The bond cleavage site is represented by  $\Delta$ <sup>4</sup>.

### 2.2.2 Designing primers of PCR

Appropriate primer design is a key element in successful PCR, because primers which anneal only to the specific binding site of interest on the template DNA will generate the desired PCR product. There are many factors to be considered during the primer design process, some criteria to consider are listed below. (i) The base composition, as it is best to have GC content between 40 to 60%. (ii) The length of the primer should be between 18-30 nucleotides. If the primers are too short non-specific amplification can occur, however if the primers are too long it might not be able to screen the template and anneal to the designated DNA sequence within the annealing

time period. (iii) The primer melting temperature should be between 60-64° C, with no significant temperature difference between forward and reverse primer pair. However, there are some DNA regions with very low GC content, these designed primers will have melting temperatures below this range. (iv) The presence of strong DNA secondary structures can inhibit the annealing between the primer and the template DNA. (v) Any self-complimentary sequences in the primer can generate primer dimers, where forward and reverse primers to anneal together instead of annealing to the template DNA<sup>1</sup>.

Primers for V90C

gatggcgtacgcctgaatctggcgggggtgtagtggttctgccgaccttagccagttccct  
D G V R L N L A G V S G S A D L S Q F P

Convert the amino acid codon to the new mutants' codon

gatggcgtacgcctgaatctggcgggggtgtagtggttctgccgaccttagccagttccct  
D G V R L N L A G C S G S A D L S Q F P

Forward primer design

gatggcgtacgcctgaatctggcgggggtgtagtggttctgccgaccttagccagttccct  
D G V R L N L A G C S G S A D L S Q F P

Reverse primer design\*

gatggcgtacgcctgaatctggcgggggtgtagtggttctgccgaccttagccagttccct  
D G V R L N L A G C S G S A D L S Q F P

\*Use IDT OligoAnalyzer tool to obtain reverse primer sequence in 5' to 3' direction

Sequence

5'- GCC TGA ATC TGG CGG GGT GTA GTG GTT CTG CC -3'

Reverse complementary sequence

5'- GGC AGA ACC ACT ACA CCC CGC CAG ATT CAG GC -3'

Figure 2.2 Primer designing for V90C BtuB mutant.

There are different ways of designing primers, the following method is used in this research work to accurately and quickly design the required primers for mutagenesis. Once the mutation site at the protein level is decided, align the amino acid sequence with its DNA sequence and

carefully identify the site to be mutated. For this purpose, Expasy translator website<sup>5</sup> was used to obtain the amino acid translation along with the corresponding nucleotide sequence and each amino acid was exactly aligned with the corresponding three nucleotides in the reading frame (Figure 2.2). All the mutants generated in the present work are point mutations on selected amino acids, either into alanine (A) or cysteine (C). After locating the amino acid(s) to mutate, the corresponding three DNA bases were changed using amino acid codon chart, minimizing changes to the wild type sequence. Next, the coding DNA strand (5' to 3') was used to first design both forward and reverse primers (between 25-35 nucleotides), in such a way so that the primers are either identical or have over hangs. Finally, the coding sequence of reverse primer was converted to its complementary DNA sequence from 5' to 3' direction by using ITD OligoAnalyzer tool<sup>6</sup>. Before ordering the primers, one should check the primer properties using oligonucleotide analyzer (IDT OligoAnalyzer) to determine melting temperature, any potential secondary structures.

### 2.2.3 PCR based cloning and mutagenesis

During cloning, PCR is used to amplify the insert DNA and occasionally to introduce specific restriction digestion sites to the 5' and 3' ends of the insert. Successfully amplified DNA will have over hanging ends with specific end sequences which facilitates proper orientation during ligation within the vector. In contrast to restriction enzyme digested cloning methods, polymerase incomplete primer extension (PIPE) cloning does not use restriction or ligase enzymes. Thus, it is known as enzyme-free cloning. During PIPE cloning, two parallel PIPE reactions, insert PIPE (I-PIPE) and vector PIPE (V-PIPE), are performed to obtain inserts and vectors with over hangs in both 5' and 3' ends to allow insert DNA to anneal with vector DNA.

Finally, the product is directly transformed into competent cells where the ends are repaired and ligated<sup>7, 8</sup>.

Similar to PCR based cloning methods, PCR based mutagenesis can be performed using traditional, primer extension or inversion PCR methods. Successfully mutated DNA are then ligated using DNA ligase to obtain the plasmids containing the desired mutation. However, during the PIPE-mutagenesis, the over hangs of each strand will anneal then once transformed, the ends will be repaired and ligated by *E. coli* enzymes<sup>7, 8</sup>.

pAG1 plasmid containing wild type (WT) *btuB* gene was used to mutate: S74C, V90C, T188C, L8P-T188C, S74C-T188C, V90C-T188C, A288C, L8P-A288C, D492, L8P-D492C, S533C, and L8P-S533C using PIPE PCR mutagenesis and used in Chapter 3. For Chapter 6, pAG1 plasmid containing V10C *btuB* gene was used to mutate ion pair residues into single and double alanine mutants (R36A, R47A, D53A, R69A, R111A, E419A, E465A, D515A, R526A, D548A, R36A-D515A, R47A-D548A, D53A-R526A, R69A-E419A, R69A-E465A, R111A-E419A and R111A-E465A) using PIPE PCR mutagenesis. In appendix A1, pET 17b plasmid containing the wild type (WT) colicin E3R gene (*col E3R*) was used to mutate sites: Q349, S356, K363, A370, T402, A409, and K416 amino acids into cysteines using PIPE PCR mutagenesis. The PCR samples were prepared according to Table 2.1 and the PCR cycle (Table 2.2) used in PIPE PCR mutagenesis with selected annealing temperatures based on the melting temperatures of the primers used (Appendix 4). After incubating the PCR samples with DpnI, samples were analyzed using 1 % agarose gel electrophoresis.

Successfully amplified PCR samples were transformed into Top10 cells according to the manufacturer's instructions (Section 2.3.4),<sup>9</sup> plated on 100 µg/mL ampicillin containing Luria Bertani (LB-Amp) agar plates, and then incubated overnight (O/N) at 37 °C. Single cell colonies

(SCC), originating from transformants, were grown O/N at 37° C in 5 mL of 100 µg/mL ampicillin containing LB cultures then used to purify plasmids according to the manufacturer's instructions (Section 2.2.5)<sup>10</sup>. To identify successfully mutated plasmids, samples were sequenced at Genewiz (South Plainfield, NJ) and the sequenced samples were analyzed according to Section 2.2.8.

Reagent	Stock concentration	Final concentration in individual PCR reaction tube
Template DNA	3 ng/µL	0.06 ng/µL
10 X Pfu Turbo polymerase reaction buffer (100 mM KCl, 100 mM (NH <sub>4</sub> )SO <sub>4</sub> , 200 mM Tris-Cl pH 8.8, 20 mM MgSO <sub>4</sub> , 1 % Triton® X-100 and 1 mg/ml BSA) <sup>11</sup>	10 X	1 X
Forward primer	10 µM	1 µM
Reverse primer	10 µM	1 µM
dNTP mix	10 mM each	200 µM each
Pfu Turbo polymerase	2.5 U/µL	0.05 U/µL
ddH <sub>2</sub> O		To bring final volume to 50 µL

Table 2.1. Stock and final concentrations of reagents used in PCR mutagenesis<sup>11</sup>.

PCR step	Temperature (°C)	Duration
Initial denaturation	95	3:00 min
Subsequent denaturation	95	30 s
Annealing	*	45 s
Extension	68	10:00 min
Number of cycles		**
Final extension	68	5:00 min
Hold	4	Indefinite

Table 2.2. PCR cycle steps and the temperatures and corresponding time for each step used in this thesis work. \* Annealing temperatures were selected based on the melting temperatures of each forward and reverse primer used during mutagenesis (Appendix 4). \*\* The number of PCR cycles were adjusted according to the mutation (either 15 or 20 cycles).

#### 2.2.4 Features in plasmids

In a typical modern-day plasmid vector contains several key elements which facilitate the selection of cells harboring the plasmids, expression of the protein of interest, and aid downstream protein purification (Figure 2.3). (i) Plasmids contain an origin of replication, which promotes plasmid self-replication and propagation in the host cells. (ii) Selective marker genes are present, such as antibiotic resistance genes (ampicillin, kanamycin, and chloramphenicol), which allow researchers to identify the transformed cells from non-transformed cells. (iii) Multiple cloning sites (MCS) encompass multiple restriction sites located in the region after the promoter. During cloning, the vector plasmids are cleaved at MCS using specific restriction enzymes and are used to ligate to the insert DNA of interest. These plasmid vectors include either strong promoters (T7-promoter) or weak promoters, which results in high or low levels of gene expression, respectively. The MCS region also contains sequence tags (His-tag and GFP-

tag) and signal sequences which facilitate downstream processes, such as protein purification or cellular localization of expressed proteins (membrane proteins). (iv) Some plasmids also contain repressor genes which encode repressors (ex. LacI repressor) that inhibit the transcription of the gene of interest. These repressors control the expression of the genes during specific growth phases or times and are regulated by repressor molecules (ex. isopropyl  $\beta$ -D-1-thiogalactopyranoside, IPTG binds to lac repressor protein promoting expression). However, not all the plasmid vectors have this feature and are said to have leaky promoters (ex. pUC8 plasmid)<sup>1, 12</sup>. This leads to gene expression throughout the bacterial growth.

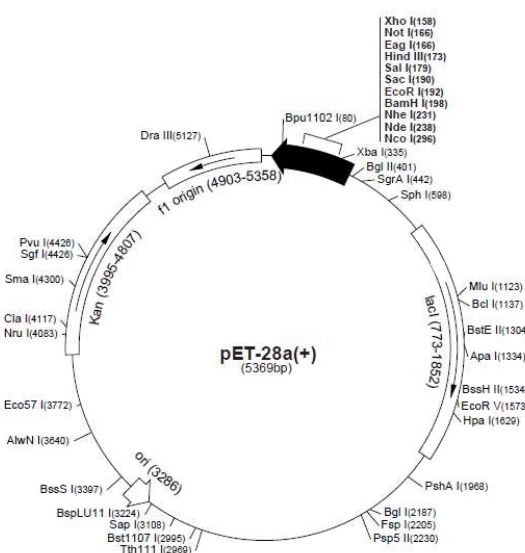


Figure 2.3 Plasmid map of pET28 (a) containing ori, kanamycin resistant gene, *lacI* repressor gene, MCS with T7 promoter and His tags<sup>13</sup>. Figure is from pET-28a-c(+) Vectors, <https://biochem.web.utah.edu/hill/links/pET28.pdf>.

In addition to the above plasmid features, one should also consider the plasmid compatibility and competition for replication processes during co-expression of different proteins using multiple plasmids in the same host. Incompatible plasmids can result in the eventual loss of one

or the other plasmid from the host cells. Thus, plasmids should be selected from different incompatible groups during co-expression<sup>12</sup>.

#### 2.2.5 Plasmid DNA isolation

In the laboratory, contaminant-free isolated plasmids are used as potential templates during PCR mutagenesis, as vectors for cloning, and for sequencing. Plasmid isolation is also used to obtain DNA for downstream processes, including plasmid amplification and subsequent transformation of the sequenced confirmed plasmids into cell lines for expression.

Overall, isolating a pure and high concentrated plasmid sample is critical for many molecular biology and microbiology applications. During this process, cells are grown overnight then used to isolate the plasmid DNA. The volume of overnight (O/N) culture depends on the amount of expected plasmid yield. For example, the typical yield of a mini-prep (1-5 mL) is up to 40 µg, while in a giga-prep (2.5-5 L) yields up to 15 mg<sup>10</sup>. Commercially available kits use solid phase DNA extraction, either via anion exchange resin or a silica membrane.

#### 2.2.6 Detecting DNA quality and quantity

DNA bases absorb UV light with the optimum at 260 nm, thus the concentrations of DNA can be detected using absorbance-based spectroscopy. However, RNA in the samples will also absorb at this wavelength. Here, the ratios between  $A_{260/280}$  can be used to assess the purity of the DNA samples. A 1.8 ratio is representative of pure DNA, while a 2.0 ratio is representative of pure RNA. Another ratio,  $A_{260/230}$ , also provides information about any organic contaminations (eg. Phenol) in the samples. In the presence of organic compounds, this ratio will be lower than 1.8, while in a pure sample it will be between 2.0-2.2<sup>14</sup>.

### 2.2.7 DNA sequencing

DNA sequencing allows the researcher to confirm whether the sample contains the correct version of the DNA fragment of interest. Sequencing can be used to confirm a properly ligated insert in plasmid vectors, to identify successfully introduced mutations into genes, and to identify any unwanted mutations to correct back to wild type. Thus, this check point allows the researcher to confidently identify proper DNA samples to be used for further DNA manipulations, plasmid propagation, and protein expression. At sequencing facilities and companies, a modified version of Sanger sequencing<sup>15</sup> is used. Here, all the four types of dideoxynucleotide triphosphates (ddNTPs) are fluorescently labeled and used in one reaction rather than radio-labelled ddNTPs. During the reaction, ddNTP concentrations are used to terminate the reaction at each base pair, then the products are separated using electrophoresis, and finally the sequence of DNA sample is determined. In our laboratory, plasmids are sequenced at Genewiz (South Plainfield, NJ).

### 2.2.8 Analyzing DNA sequence results

Once the DNA is sequenced, Genewiz provides the DNA sequence file and the trace file, which contains the chromatogram of the sequenced sample. During the analysis process, the nucleotide sequence is first translated to the amino acid sequence using the Expasy translate tool. This tool provides 6 possible translations using the provided DNA sequence as 5' to 3' and 3' to 5' ends, giving the 3 open reading frames for each side (Figure 2.4). Thus, the researcher should select the correct reading frame sequence and directionality based on the protein sequence of interest. These sequences can be aligned using NCBI protein blast for each 5' to 3' reading frame for the forward primed sequences and the 3' to 5' reading frames for the reverse primed

sequences with respect to the wild type sequence (Figure 2.5). The wild type amino acid sequence of BtuB and Colicin E3 is in the Appendix 3. If the sequencing results display poor quality, the trace files (Figure 2.6) should be used to further verify any sequence anomalies to obtain confident sequence results.

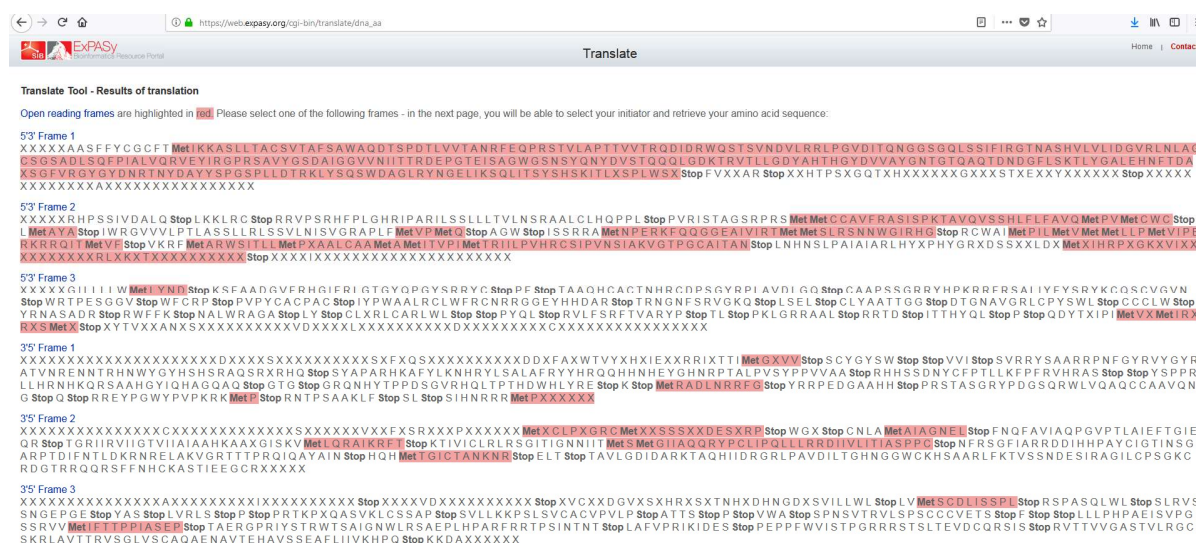


Figure 2.4. Translated DNA sequence results of V90C-BtuB mutant using the Expasy bioinformatics resource portal. For each entered DNA sequence, 3 reading frames for both 5' to 3' and 3' to 5' is displayed and the open reading frames are highlighted in red.

Query 41	SVNDVLRRLPGVDITQNGGSGQLSSIFIRGTNASHVLVLIDGVRLNLAGVSGSADLSQFP 100
Sbjct 41	SVNDVLRRLPGVDITQNGGSGQLSSIFIRGTNASHVLVLIDGVRLNLAGVSGSADLSQFP 100

Figure 2.5. NCBI protein blast is used for sequence alignment between the wild type and the translated plasmid sequence to identify the BtuB mutation. The wild type sequence is the query while the mutation is the subject sequence in the alignment. V90 of wild type is successfully mutated into C90.

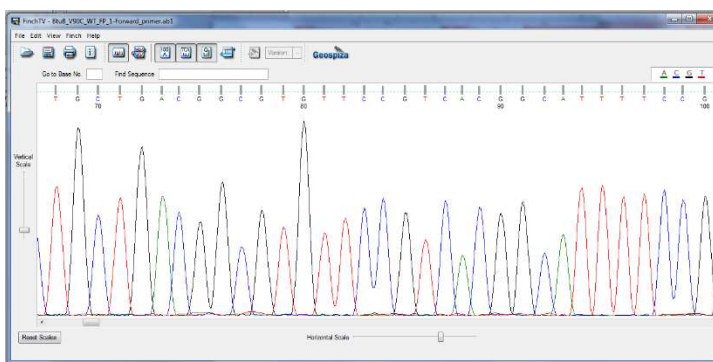


Figure 2.6. A chromatogram of the DNA sequence visualized by Finch TV software. Each ddNTP has a different fluorescence, thus enabling recognition of each base in the sequence by the corresponding peak color.

## 2.3 Microbiological techniques

Once successfully cloned or mutated genes are verified by analyzing DNA sequencing results, the next step is to introduce these expression vectors into suitable protein expression systems. There are different protein expression systems (*E. coli*, yeast, and insect cells), which will be discussed later. The most common expression system is *E. coli*, which has been used to express a diversity of proteins. Many fundamental microbiological techniques are used to introduce these expression vectors and then express the proteins in microbial expression systems.

### 2.3.1 Aseptic techniques and the importance of proper sterile techniques

Microbes are ubiquitous and can easily contaminate the equipment and materials used in different microbiological techniques which are used to grow, select, and screen experiments. Thus, proper aseptic techniques should be utilized to avoid potential contaminations. Some techniques include: disinfecting the working area, minimizing the exposure of liquid media and agar plates to open air, and working near the Bunsen burner (within the sterile field created by

the updraft of the flame)<sup>16</sup>. Autoclaving, filter sterilization, and oven sterilization are all employed in research to sterilize materials and equipment. Flammable or heat sensitive reagents (antibiotics, amino acids, vitamins and drugs) are sterilized using 0.22 µm pore filters, while the autoclave can be used to heat stable reagents and equipment. The autoclave uses a super heating temperature (121° C) and is used under high pressure (15 psi). The duration of autoclaving depends on the autoclaving load and the size, or volume, of reagents. During the oven sterilization process, glassware and metal objects can be sterilized using dry heat (121-170° C) for one to two hour time periods<sup>17</sup>.

### 2.3.2 Competent cells

Competent bacterial cells are capable of directly in-taking foreign DNA. In microbiology, the most commonly used *E. coli* strains are naturally very low competency. Thus, these bacterial strains are first made competent and are then used to introduce foreign DNA, such as plasmids containing the genes of interest, which is known as transformation<sup>12</sup>. There are two types of transformations: heat-shock and electroporation. The competent cells are prepared using two separate methods depending on the desired transformation method<sup>12</sup>.

### 2.3.3 Preparation of competent cells

All the materials should be sterile and aseptic techniques should be used during the preparation of competent cells. In order to gain higher competency, it is best to begin with a single cell colony (SCC) from a freshly streaked plate. Next, pre-cultures are inoculated with the SCC and incubated overnight (O/N). The media culture is then inoculated using the grown O/N

pre-culture and the cells are grown until mid-log phase ( $\sim 0.4$ - $0.6$ ). Cells should be harvested before late-log phase as it is important to use healthy bacterial cells for transformation.

Once the cells are harvested, all the remaining steps are performed on ice. For the electroporation, harvested cells are resuspended a few times in ice-cold water and pelleted to remove any residual components which could interfere with the electroporation. Next, the cells are resuspended in 10% glycerol and stored at  $-80^{\circ}\text{C}$  until the transformation step. In contrast, harvested cells are chemically induced and used for heat-shock transformation. These harvested cells are first incubated in  $\text{MgCl}_2$  and then in  $\text{CaCl}_2$  to increase the membrane permeability. Cells are then stored at  $-80^{\circ}\text{C}$  until the transformation<sup>1, 12</sup>. Once prepared, competent cells are good for 6-12 months, however over time the competency is reduced. All the *E. coli* cell lines used in this research were chemically induced and used for heat-shock transformation.

#### 2.3.4 Transformation

During the transformation process, cells should be handled very gently and carefully to retain the cell's viability. Frozen cells are first thawed on ice, and the plasmid of interest is added then mixed by gently tapping. For heat-shock transformation, after the addition of plasmids, cells are further incubated on ice for 30 min. Next, the cells are incubated at  $42^{\circ}\text{C}$  for 30-45 s then quickly incubated on ice for 2 min (Figure 2.7). This brief temperature elevation (heat-shock) enables the plasmids to penetrate through the cell membrane. During electroporation, cells mixed with the plasmids are exposed to a short, high voltage electric field, which induces transient membrane pores to uptake plasmids into the bacterial cell<sup>9</sup>.

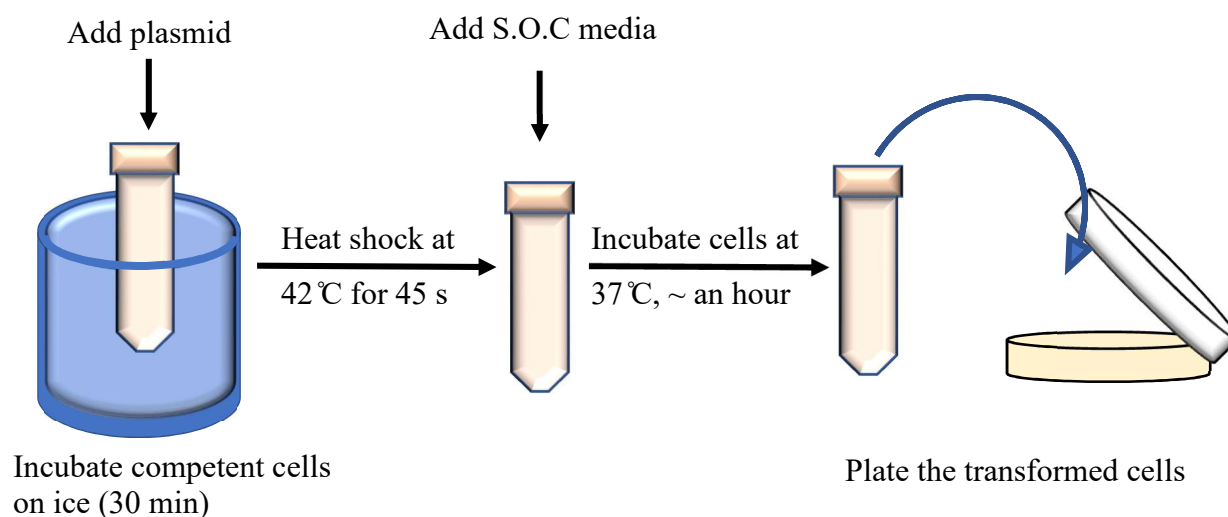


Figure 2.7. Schematic representation of the heat-shock transformation process. On ice, the plasmid is added to the chemically induced competent cells and then incubated on ice for 30 min followed by the heat-shock step. Transformed cells are recovered by incubating in S.O.C. media and finally plated onto an LB agar plate containing selecting compounds which can identify the transformed cells from the non-transformed cells.

After the transformation, cells are grown in antibiotic free growth media to allow cells to recover. During this step, S.O.C media (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, and 20 mM glucose)<sup>1</sup> is added to the cells then incubated for an hour at 37 °C, with shaking at 225 rpm. Finally, recovered cells are plated on agar plates containing either antibiotics or other compounds used to detect the successfully transformed cells (Figure 2.7).

### 2.3.5 Long term preservation of bacterial strains

In the presence of cryoprotectants (ex. Glycerol), cells can be stored long term at  $-80^{\circ}\text{C}$ . During the downstream processes, including protein expression and purification, bacterial cells which contain the gene encoding the protein of interest should be grown, starting from a SCC on an agar plate. Some laboratories start from a plasmid transformation step, while others prepare long term bacterial storage stocks for all the different mutant protein expressing cells. One advantage of these stocks is that one can avoid the use of competent cells and the transformation every single time that expression and purification the protein of interest is required. These bacterial glycerol stocks can be stored for many years and can be used without starting from plasmid level, this process should be done carefully under aseptic conditions to avoid contamination.

Start with inoculating a small culture with a SCC and grow at the optimal temperature ( $37^{\circ}\text{C}$  for *E. coli*) until mid-log phase, where the cells are healthy and active. Transfer small aliquots (250-500  $\mu\text{l}$ ) into sterile screw-top cryovial tubes and mix with sterile 50% glycerol in a 1-to-1 ratio. Store the tubes at  $-80^{\circ}\text{C}$ <sup>8</sup>. When handling  $-80^{\circ}\text{C}$  stored tubes for streaking or for cell cultures, one must avoid frequent temperature fluctuations. Thus, glycerol stocks should be kept on ice, not at room temperature, to retain the viability then quickly returned to the freezer to prevent thawing the frozen sample.

Sequenced confirmed BtuB mutant plasmids were transformed<sup>9</sup> into *E. coli* strain RK5016 or Dsb knockouts, plated on LB-Amp plates, and incubated O/N at  $37^{\circ}\text{C}$ . Resultant SCCs were used to prepare glycerol stocks and stored at  $-80^{\circ}\text{C}$ .

### 2.3.6 Growing bacteria in laboratories

Bacterial cells can be grown in either solid (on agar), semi-solid (reduced agar percentage), or liquid media. Based on the downstream application, the proper nutritional media should be selected to enhance the bacterial cell growth and propagation. In general, cell lines without selection marker genes (ex. antibiotic resistance) are grown in media without those chemicals. However, to specifically enhance the growth of cells with marker genes over other cells, the corresponding chemicals should be added in addition to the other nutrients.

#### 2.3.6.1 Plating method

In order to isolate colonies, mainly a SCC, bacterial cultures need to be grown on solid or semi-solid media which contains the required nutrients and any selective agents required to identify the colonies harboring the plasmids (or genes) of interest. Luria-Bertani (LB) agar (1 % tryptone, 0.5 % yeast extract and 1 % NaCl with 2 % agar) plates are commonly used in both streak and spread plate techniques to obtain a SCC (Figure 2.8). These can be obtained from glycerol stocks and by evenly spreading bacterial culture on the plate, respectively<sup>16</sup>.

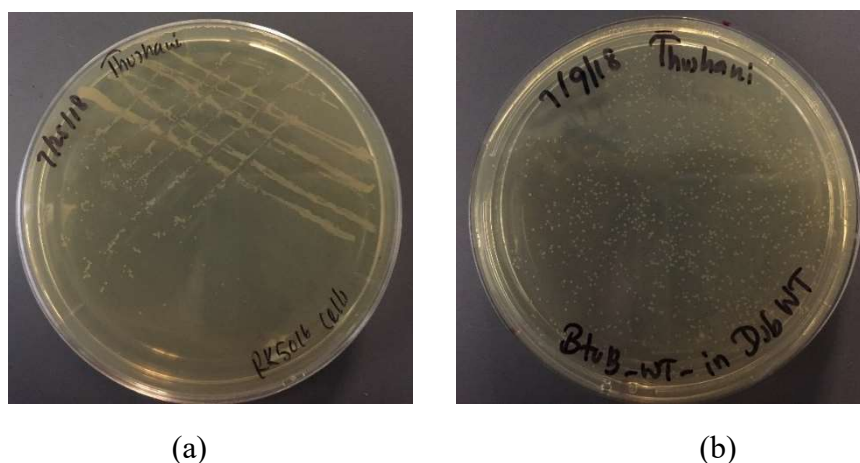


Figure 2.8. Different plating techniques. Bacterial colonies grown on (a) Streaked plate and (b) Spread plate.

### 2.3.6.2 Culturing methods

Cells are grown in liquid cultures in order to express the proteins of interest. There are three main growth culture methods (batch culture, semi fed-batch culture, and few-batch culture) used to grow bacteria of interest for various applications. In the laboratory, cells are mainly grown in batch cultures. Depending on the downstream application, different culture media were used to grow cells. For example, cells were grown in LB media in order to isolate plasmids, to make competent cells and the majority of the time to express the proteins of interest. However, minimal media (MM) is used to provide only the essential nutrients for the growth and is used when the protein of interest must be expressed in the absence of particular nutrients or if the protein is to be labeled with different isotopes. In this research, *E. coli* cells were grown in MM to enhance the expression of BtuB in the absence of B<sub>12</sub>.

During the cell culture, the pre-culture media is first inoculated with a SCC and generally grown O/N at a specific temperature with shaking. This is then used to inoculate the main cultures. These main cultures are also grown at a specific temperature (ex. 37° C for *E. coli*) and are then allowed to express the proteins by addition of inducing agents. Depending on the protein, the incubating temperature is then reduced in order to minimize any protein aggregation within the host cells. However, if the desired plasmid contains a leaky promoter, the protein will be expressed throughout the growth. Finally, the cells are harvested then used for protein purification or other applications. Similar to pre-cultures, mini-cultures can be grown to isolate plasmids for propagation and for sequencing.

## 2.4 Biochemical and biophysical techniques

### 2.4.1 Protein expression and purification

Different protein expression systems can be used depending on the desired application. Such systems should be easy to grow with rapid growth and should be able to produce the protein in large quantities as required. While *E. coli* is the most common expression system, proteins requiring post-translational modifications (PTMs) are expressed in other expression systems, such as yeast and mammalian cell cultures<sup>1</sup>.

Expression of the protein of interest depends on the properties of the vector system used. With respect to *E. coli* expression system proteins, including non-native *E. coli* proteins, are expressed in large scale. It is important to regulate the expression to retain the viability of host cells during the protein expression. Thus, some of these vectors carry genes encoding repressor proteins to block protein expression until induction. Other vector systems can also contain signal sequences which enable direction of the encoded protein into periplasm. For example, proteins containing disulfide bonds need to be translocated to periplasm to promote oxidation, while outer membrane proteins need to be translocated to periplasm in order to deliver to outer membrane. Furthermore, rapid expression of membrane proteins can impact membrane biogenesis. Thus, use of an uninduced protein expression system is recommended<sup>19</sup>.

In contrast to protein studies using intact cell systems, for *in vitro* systems, proteins are purified based on their unique properties including the size, charge, and affinity of the targeted protein. For soluble proteins, proteins can be purified after cell lysis and the removal of cell debris. However, for membrane proteins (IM or OM proteins), the specific membrane should be first isolated and then the proteins within these membranes should be solubilized prior to the purification. Detergents are used in these steps to solubilize membrane proteins into detergent

micelles. There are four main types of detergents, based on structure: ionic, non-ionic, bile acids, and zwitterionic detergents (Figure 2.9). The type of detergent used should be based on the targeted membrane proteins. Some detergents (sodium dodecyl sulfate, SDS) will denature the target protein, while others are mild and comparatively non-denaturing (OG)<sup>20, 21</sup>. Once the membrane proteins are solubilized, they can be purified using chromatography techniques similar to soluble protein purification.

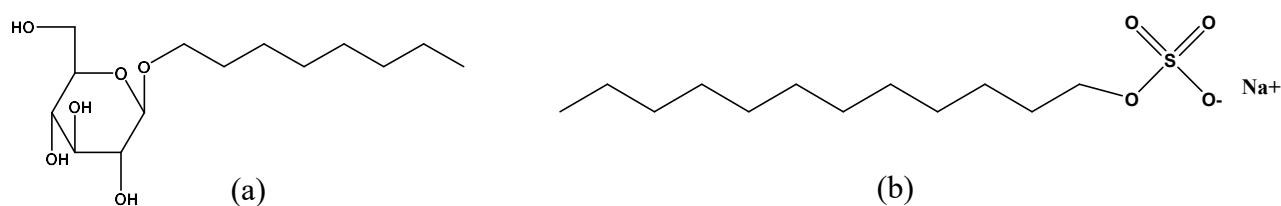


Figure 2.9. Structures of (a) n-octyl-β-D-glucopyranoside and (b) sodium dodecyl sulfate (SDS).

There are different chromatographic methods which are used to purify the proteins of interest from host proteins. (i) In size-exclusion chromatography, proteins are separated by size<sup>22</sup>. (ii) In chromatofocusing, proteins are separated based on their isoelectric points (pI). During the separation, proteins migrate through a generated pH gradient which affects the rate of migration according to the pI values of sample proteins<sup>23</sup>. (iii) Affinity chromatography is used to purify proteins (Figure 2.10) via reversible binding of targeted proteins to specific immobilized ligands<sup>24</sup> (Figure 2.11). This method was used to purify colicin E3R (Appendix A1).

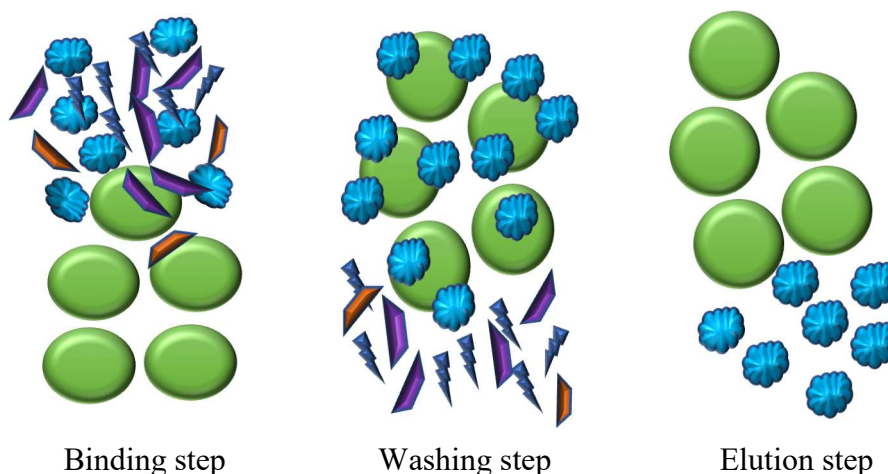


Figure 2.10. Schematic representation of affinity chromatography purification. **Targeted proteins** bind with the **affinity resin**. During the washing, unbound proteins are eluted. Next the **bound proteins** are eluted during the elution step.

At the molecular level, vectors with different affinity tags can be used to aid the downstream purification. Generally, these proteins are fused with an affinity tag that can be cleaved after protein purification step<sup>1, 24</sup>. (a) Glutathione-S transferase (GST) tag fused proteins are purified by using glutathione resin. The immobilized proteins are eluted using free glutathione containing buffers. (b) Maltose binding protein (MBP) fused proteins are purified using amylose resins. (c) His<sub>6</sub> tagged proteins are purified using either Ni<sup>2+</sup>, Cu<sup>2+</sup>, or Co<sup>2+</sup> immobilized resins. The bound proteins are then eluted using a competing chelating agent (ex. Imidazole). In addition to His<sub>6</sub> tagged proteins, native metal binding proteins can also occasionally bind to the resin. After the affinity purification, if needed, the tag can be cleaved and untagged proteins can be separated from the tag. In intein-mediated purification with an affinity chitin binding tag (IMPACT), the fused proteins are immobilized on chitin resin and then eluted by cleaving the tag using either thiol containing buffer or by lowering pH from 8.5 to 6. This method avoids the additional tag removal and protein purification steps<sup>25</sup>.

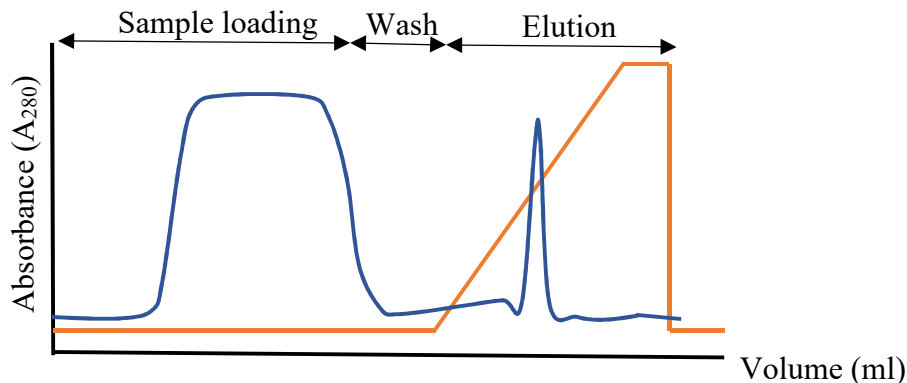


Figure 2.11. **Elution profile** of an IMAC protein purification using a gradient elution. When loading the sample, unbound proteins come out along with the flow through. This is followed by the washing step to remove any trace unbound and or non-specifically bound proteins. Next, **elution buffer percentage** is gradually increased to 100%. During this time bound proteins are eluted from the resin.

(iv) In ion exchange chromatography, the proteins are separated according to the surface charge at a given pH. If the pH of the buffer is higher than the pI of the protein, it will have net negative charge. However, in lower pH, it will have net positive charge (Figure 2.12). Thus, proteins will bind to positively charged resin (anion exchanger) at higher pHs than the pI and to negatively charged resin (cation exchanger) at lower pHs than the pI<sup>26</sup>.

During ion exchange chromatography, depending on the stability of the target protein and based on the remaining sample proteins, one can either use anion or cation exchange resin to purify the target protein. For example, quaternary ammonium (Q) and diethylaminoethyl (DEAE) are anion exchange resins, while sulfopropyl (S) and carboxymethyl (CM) are cation exchange resins. Here, Q and S columns are strong ion exchangers with wide pH ranges, while DEAE and CM are weak ion exchanges with smaller pH range<sup>26</sup>. After purifying the target protein, the chemicals used during the purification are removed using either desalting or buffer exchange so that these ions do not impact downstream processes. In addition to these steps,

purified membrane proteins in detergents need to be reconstituted into lipid systems as they are more appropriate for membrane protein. During BtuB purification (Chapter 6) Q column was used to purify solubilized BtuB in OG micelles (Section 2.4.1) and reconstituted into lipid bilayers.

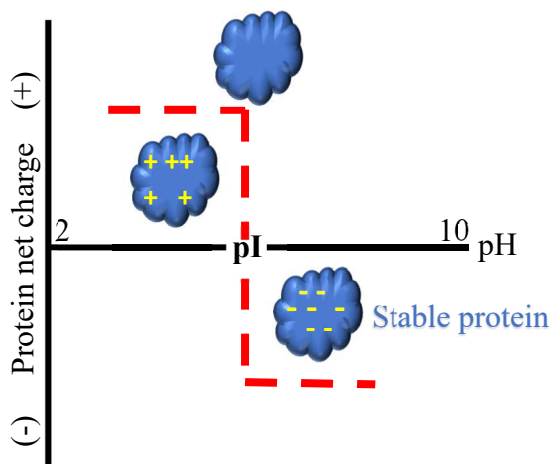


Figure 2.12. The net charge of **protein** changes depending on the pH of the surroundings. Beyond the protein **stability range** the **proteins** will denature. When the pH is equal to the pI, the net charge of **protein** is zero and above and below the pI value the **proteins** will have a net charge, either positive or negative.

#### 2.4.2 Reconstitution of membrane proteins into lipid systems

After the target membrane proteins are purified within detergent micelles, they are incorporated into different phospholipid (Figure 2.13) systems, such as: Langmuir monolayers, supported lipid bilayers, liposomes, and nanodiscs (Figure 2.14)<sup>27</sup>. Many membrane protein studies have been performed in liposomes. There are two main methods of reconstituting the membrane proteins into liposomes, either via direct incorporation of proteins into liposomes or via detergent-mediated reconstitution (detergent solubilization and dilution). In the detergent-mediated approach, the proteins in detergent are first solubilized in lipids in order to obtain an isotropic solution of detergent-protein-lipid and detergent-lipids (Figure 2.15). Next, there are two main approaches to reconstitute membrane proteins into the liposomes (proteoliposomes).

Either, by diluting this mixture to bring the detergent concentration below the detergent's critical micelle concentration (cmc) which destabilizes the detergent micelles transfers the membrane proteins into liposomes<sup>21</sup> or by the removal of detergents from the reconstitution mixture, gradually incorporating the protein into the liposomes (proteoliposomes).

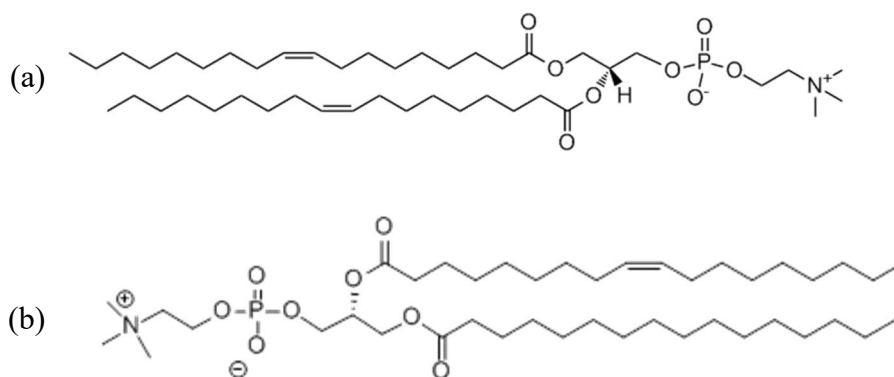


Figure 2.13. Structures of (a) 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC)<sup>28</sup> and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC)<sup>29</sup>.

There are different ways to remove detergent from the reconstitution mixture, including dialysis, adsorption to hydrophobic resins, or by chromatography techniques (size exclusion, ion exchange and affinity)<sup>20, 21</sup>. Detergent removal via dialysis is more suitable for detergents with high cmc values (such as OG), because below the cmc, micelles are dissociated into monomers and removed from the reconstitution mixture<sup>21, 30</sup>. In order to completely remove detergents, dialysis is performed using a detergent free buffer over a several days period with a few buffer exchanges (a considerable time period is needed to reach the equilibrium during the dialysis). For detergents with a low cmc, hydrophobic resins (bio-beads) are used to effectively remove the detergents which can interact with the hydrophobic surface of the resin. This detergent removal method is fast. Following the detergent removal, the added resin can be removed by centrifugation or filtration<sup>21, 30</sup>.

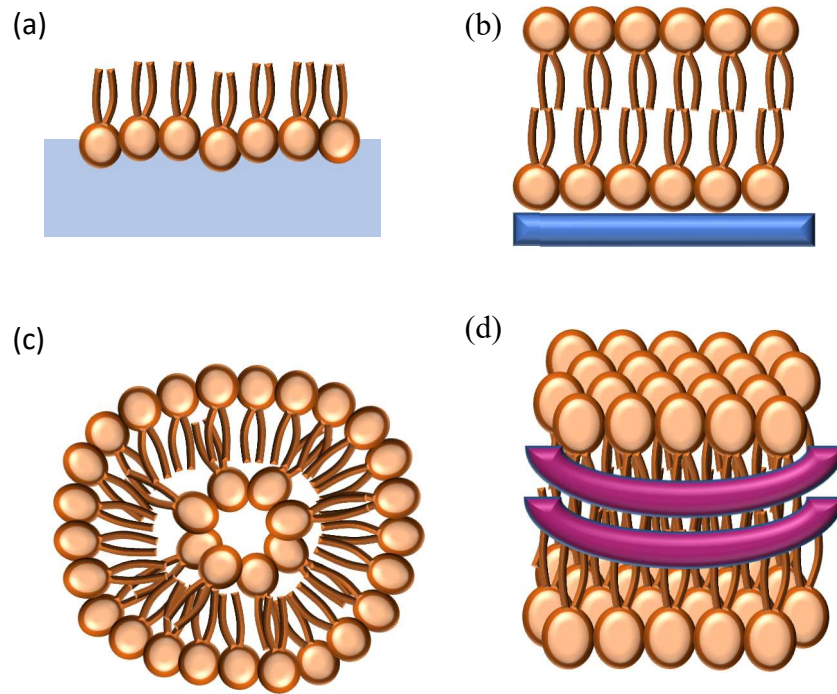


Figure 2.14 Different lipid systems. (a) Lipid monolayer, (b) supported lipid bilayer system, (c) liposomes and (d) nanodiscs.

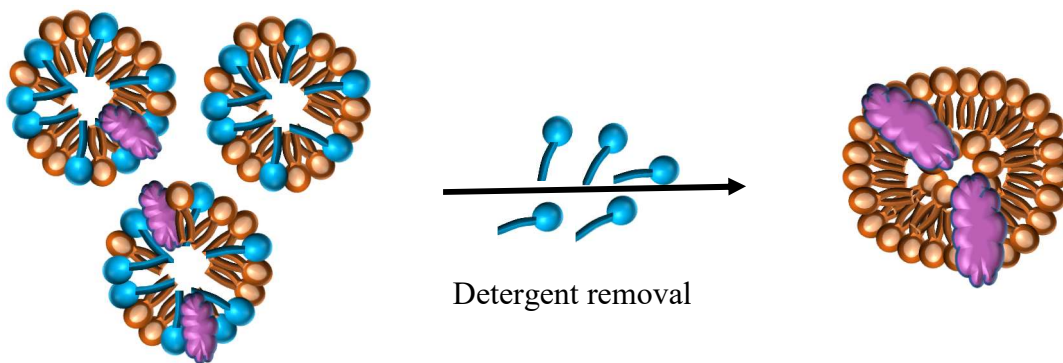


Figure 2.15. Detergent mediated reconstitution of membrane proteins. Purified **membrane proteins** in micelles are solubilized with **lipids** in the excess of **detergents**. The **detergents** from the resultant micelles (**protein-lipid-detergent** and **lipid-detergent**) are then removed by a suitable detergent removal method to facilitate the reconstitution of **membrane proteins** into the **liposomes**.

## 2.5 EPR spectroscopy as a tool to study protein systems

Protein structure and function can be studied at atomic or molecular level using different biophysical techniques, such as: NMR, x-ray crystallography, circular dichroism, and isothermal titration calorimetry. However, with respect to membrane proteins studies, many of these biophysical techniques are limited due to the size limitations and stability-required native-like lipid systems<sup>31</sup>. Recent developments of EPR technique have expanded the applications of biomolecules studied, including: DNA, RNA, and proteins (both soluble, as well as, membrane proteins)<sup>32</sup>. Since the technique does not depend on the tumbling of the labeled protein, the protein system is not size limited (membrane proteins and protein complexes)<sup>33, 34</sup>. In addition, the technique detects the unpaired electrons spin systems (paramagnetic), which are rare in biological systems. Thus, protein systems can be studied via artificially introduced paramagnetic centers without background interference.

Over the years different methods have been developed to introduce spin labels onto specific sites of the protein. (i) The site-directed spin labeling (SDSL) method can mutate a selected site on the protein into a cysteine residue at the DNA level. This is followed by expression of the mutated gene in a selected expression system. After isolating the cysteine mutated protein, this site is modified using a specific EPR spin label via cysteine chemistry to produce a disulfide bond between the spin label and the protein<sup>35</sup>. Thus, native free cysteines should be mutated either to alanine or serine to avoid the labeling of those residues. (ii) Spin labeled, synthesized peptide fragments can be incorporated into protein fragments via intein-mediated protein ligation (IPL) method producing semi-synthesized proteins with spin labels<sup>36</sup>. (iii) Spin labeled amino acids can also be combined using nonsense repressor technology, where tRNAs attached to spin labeled amino acids are used to incorporate these labels at specific sites during the translation

process<sup>37</sup>. (iv) Similar to attaching fluorescent probes onto biomolecules, there are also spin labels that can be attached via click chemistry<sup>38</sup>.

EPR spectroscopy can be used with these labeled proteins to obtain information, such as: secondary structures, structural changes associated with protein dynamics, protein-protein interactions, distances between two spin labeled sites and protein backbone dynamics. Nitroxides with different functional groups (Figure 2.16) are the most commonly used paramagnetic labels ( $R_2NO$ ) in EPR. Compared to other reactive spin labels, the methanethiosulfonate spin label (MTSSL) is the most widely used in SDSL to spin label selected sites on proteins<sup>32, 35</sup>, because it is small, very selective, quickly reacting and structurally similar to a tryptophan side chain (Figure 2.17)<sup>34</sup>. After spin labeling, excess free label is removed and spin labeled protein samples are analyzed using EPR spectroscopy.

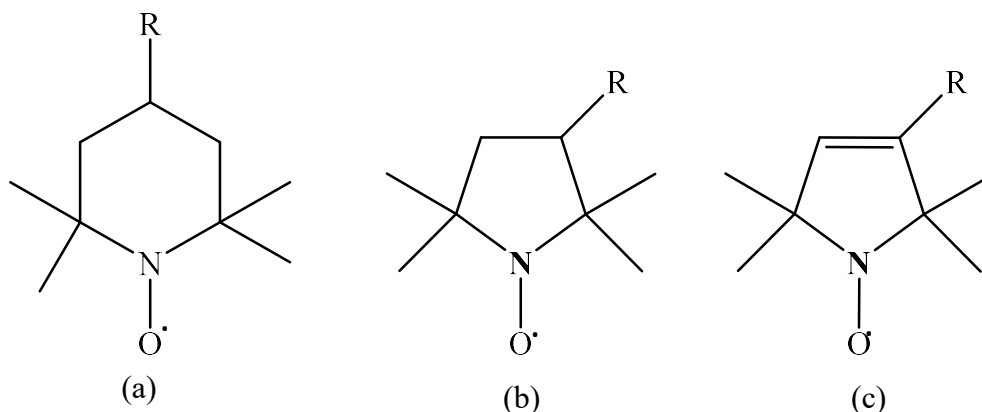


Figure 2.16. Structures of some nitroxide spin labels. (a) 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO), (b) 2,2,5,5-tetramethylpyrrolidine-1-oxyl (PROXL) and (c) 2,2,5,5-tetramethylpyrroline-1-oxyl (dehydro-PROXYL)<sup>39</sup>.

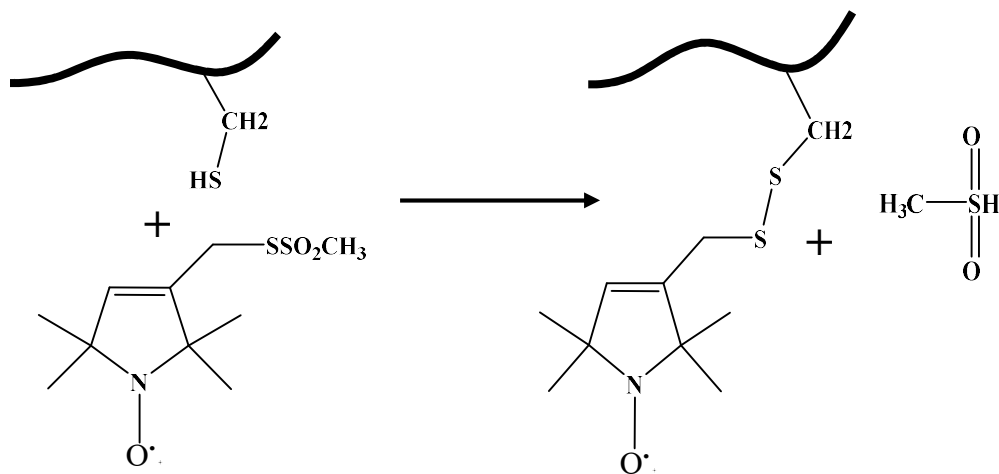


Figure 2.17. Spin labeling at the sulfhydryl group of a cysteine residue via disulfide bond formation with MTSSL generates an R1 side chain.

### 2.5.1 EPR theory

Similar to other spectroscopic techniques, EPR detects the energy absorbance between different energy states. The electron spin has an angular momentum,  $S$  that gives rise to a magnetic moment ( $\mu$ ), which is given by:

$$\mu = h\gamma_e S = -g\beta_e S \quad (2.1)$$

where  $\gamma_e$  is the electron's gyromagnetic ratio,  $\beta_e$  is the Bohr magneton ( $\beta_e = \frac{e\hbar}{2m_e}$ ,  $e$  and  $m_e$  are the charge and mass of the electron, respectively) and  $g$  is the  $g$ -factor. The  $g$  value of a free electron is 2.0023193043617. When the electron spin is exposed to an external magnetic field ( $B_0$ ), the spin energy state is split into two states because the magnetic quantum number ( $m_s$ ) is  $\pm 1/2$ . The energy difference between these two states depends on the applied magnetic field.

$$E = \pm g \beta_e B_0 \quad (2.2)$$

The electromagnetic radiation at the energy difference ( $\Delta E$ ) will induce transitions between electron spin states where,

$$\Delta E = g\beta_e B_0 = h\nu \quad (2.3)$$

$h$  is the Planck constant and  $\nu$  is the applied frequency. The frequency at which absorption takes place is known as resonance or Larmor frequency. In EPR, the magnetic field is varied while the frequency of the electromagnetic radiation (microwave energy) is held constant, and absorption occurs when the resonance condition (Eq. 2.4) is reached. Furthermore, EPR detects the net magnetic moment per unit volume, which is known as macroscopic magnetization ( $M$ ), in the sample. At equilibrium, without a microwave field, spin populations are distributed between the two energy states according to Boltzmann distribution, where  $n_\alpha$  is the population in  $m_s = +1/2$  and  $n_\beta$  is the population in  $m_s = -1/2$ .

$$\frac{n_\alpha}{n_\beta} = e^{\left(-\frac{\Delta E}{kT}\right)} \quad (2.4)$$

$k$  is the Boltzmann constant and  $T$  is the temperature. The resulting excess polarization,  $P$ , is given as,

$$P = \frac{n_\alpha - n_\beta}{n_\alpha + n_\beta} \quad (2.5)$$

At a steady magnetic field in  $z$  axis, the equilibrium magnetization is,

$$M_0 = \frac{1}{V} \sum_i \mu_i = \frac{1}{2} N \hbar \gamma_e P e_z \quad (2.6)$$

Where  $N$  is the total number of spins ( $n_\alpha + n_\beta$ ). In the external field, a torque is generated by the motion of the magnetic moment.

$$\hbar \frac{dS}{dt} = \mu B \quad (2.7)$$

For the spins in the sample,

$$\hbar \frac{dM}{dt} = M \gamma_e B \quad (2.8)$$

The displacement of the magnetization vector ( $M$ ) from the  $z$ -axis leads to motional precession of the magnetization vector rotating at a precession frequency ( $\Omega_s$ ) equal to the Larmor frequency or resonance frequency. When the sample is off-resonance, the offset of Larmor frequency and the applied microwave frequency and resulting magnetization vector rotates in  $xy$ -plane instead of  $z$ -axis<sup>39</sup> (Figure 2.18). Over time, the magnetization vector returns to thermal equilibrium through longitudinal relaxation ( $T_1$ ) and by the loss of coherence in the transverse plane by transverse relaxation ( $T_2$ ).

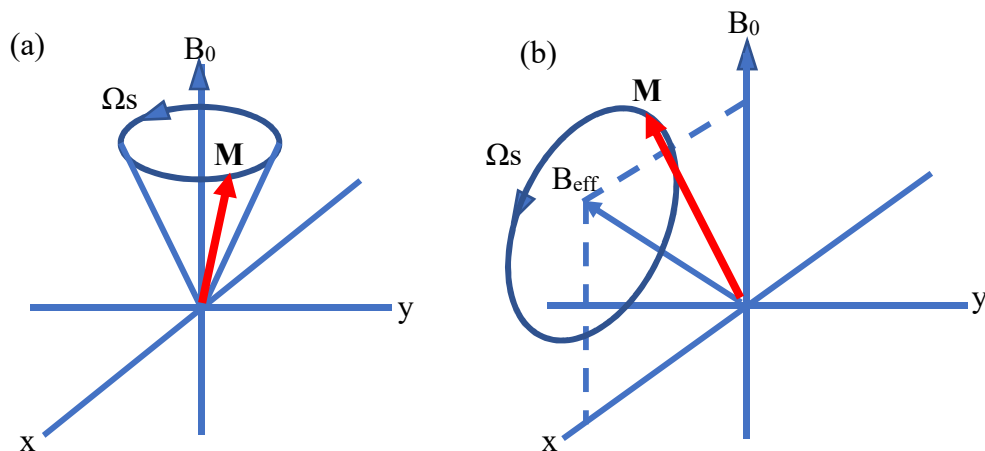


Figure 2.18. The rotating frames of on-resonance and off-resonance magnetization vectors.

In an external magnetic field, electron spins are affected by different interactions, such as: the electron Zeeman interaction and the hyperfine coupling. (i) The electron Zeeman interaction results from the interaction between the electron spin and the external magnetic field, while (ii) the hyperfine coupling results from the interaction of the electron with nearby nuclear spins. The electron and nuclear spins interact via dipolar-dipolar interactions and Fermi contacts. In

addition to these main interactions, electron spins may weakly interact with other nearby electron spins. Similar to nuclear and electron dipolar-dipolar interactions, two electron spins can also weakly interact via dipolar-dipolar interactions. In addition, Heisenberg exchange coupling may occur when two unpaired electrons are exchanged between their overlapping spin orbitals as two compounds undergo a collision.

The unpaired electron spin of a nitroxide in dilute solution is influenced by both the electron Zeeman interaction and the hyperfine coupling. Thus, the spin Hamiltonian can be described as,

$$H_{N0} = H_{EN} + H_{HF} \quad (2.9)$$

$$H_{N0} = \beta_e B_0^T g S + S^T A I \quad (2.10)$$

where T is the dipolar coupling tensor, S is the spin operator, A is the hyperfine coupling tensor and I is the spin operator of the coupled nucleus. Hence, the energy of a fast-rotating isotropic spin is given by,

$$E_{N0} = g_{iso} \beta_e B_0 m_s + a_{iso} m_s m_I \quad (2.11)$$

In the nitroxide system, the electron spin quantum number,  $m_s = \pm \frac{1}{2}$ , and  $^{14}\text{N}$  nucleus spin quantum number,  $m_I = 1$ . Based on the selection rules of  $\Delta m_s = 1$  and  $\Delta m_I = 0$ , three allowed transitions are observed for unpaired electron in nitroxide system (Figure 2.19).

$$\Delta E = h\nu = g_{iso} \beta_e B_0 + a_{iso} m_I \quad (2.12)$$

### 2.5.2 Nitroxide spin label detection using CW EPR

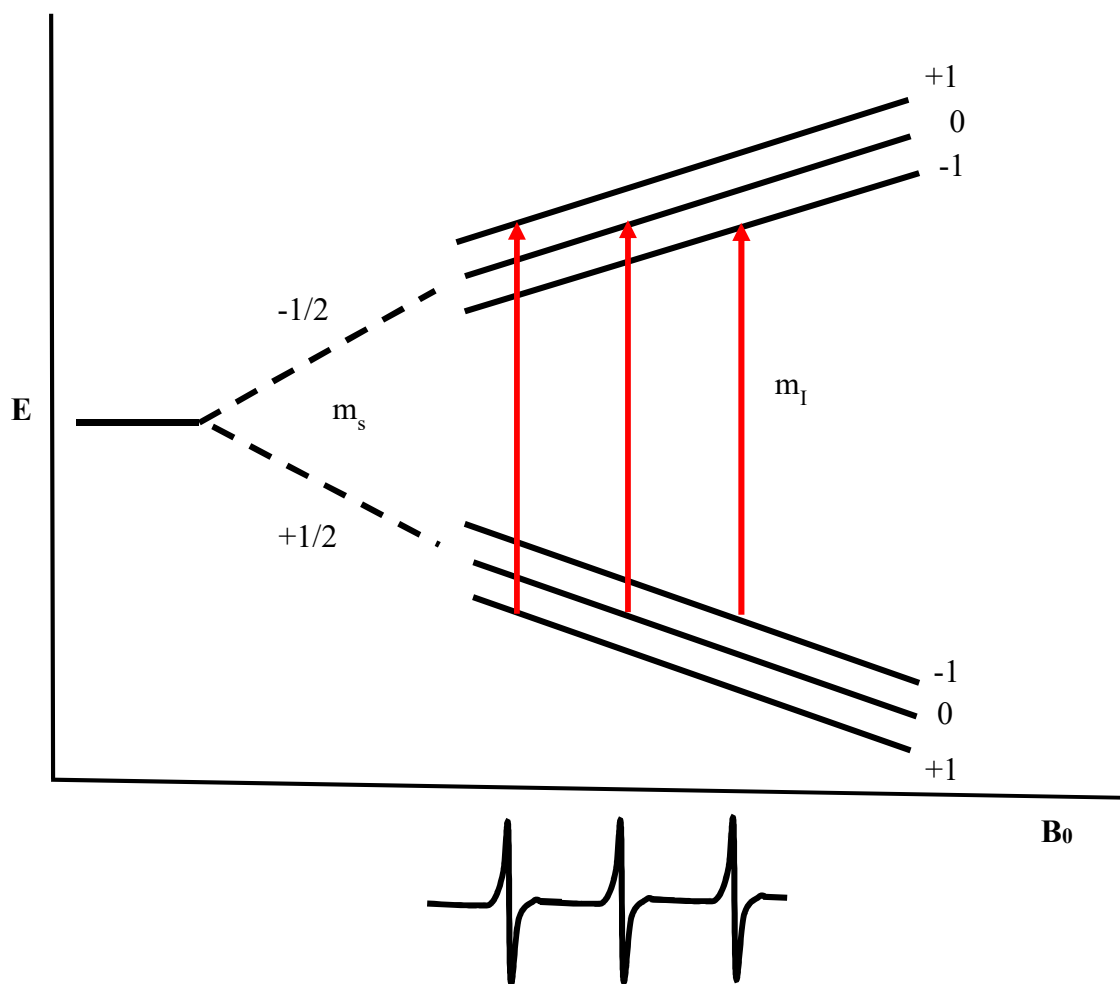


Figure 2.19. The energy diagram for spin system with  $m_s = \pm \frac{1}{2}$  and  $m_I = 1$ , where the energy states are split by the electron Zeeman interactions and the hyperfine couplings. The three allowed transitions are observed as the first derivative in CW EPR signal and each allowed transition is between  $m_I = +1, 0$  and  $-1$  with equal energy difference observed at low, central and high magnetic fields, respectively.

In nitroxides, the  $2p_z$  orbital is along the  $z$ -axis, while the N-O bond is along the  $x$ -axis and  $y$ -axis is perpendicular to  $xz$ -plane<sup>39</sup> (Figure 2.20). The continuous wave (CW) EPR spectra of nitroxide spin collected along each axis ( $x$ ,  $y$ , and  $z$ ) consists of three equal lines with the central

line at the principle g tensor ( $g$ ) element and the spacing between each line corresponds to the hyperfine tensor ( $A$ ). The unpaired electron is mainly localized in the  $2p_z$  orbital ( $z$ -axis), therefore the maximum hyperfine coupling is observed in spectra recorded along the  $z$ -axis<sup>32, 40</sup>.

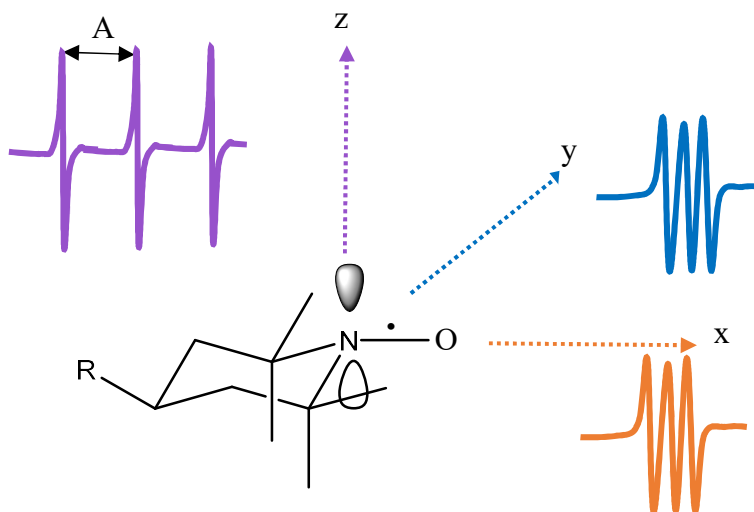


Figure 2.20. Coordinates of nitroxide spin and the CW EPR spectra recorded along  $x$ -,  $y$ - and  $z$ -axis. The central line of each spectra is corresponding to principle  $g$ -value and the distance between two lines are corresponding to  $A$ .

Both nitroxide dynamics and the local environment affect the motion of the spin label. With respect to the dynamics, the EPR lineshapes and the magnetic interactions are influenced by rotational diffusion within the sample. For a fast rotating spin, hyperfine couplings are averaged to an isotropic value,  $a_{iso}$ , giving rise to three equal EPR lines, whereas if rotation is prohibited, it results in an anisotropic powder spectrum. In between these two extremes, spectra can be impacted by the time period of rotational diffusion. Depending upon the rotational correlation time ( $\tau_c$ ), spins can be either in the fast motion regime, the intermediate, the slow motion regime, or at the rigid limit (Figure 2.21). Starting from isotropic to anisotropic EPR spectra, the lineshape (height and widths of lines) change due to the increase of anisotropic contribution.

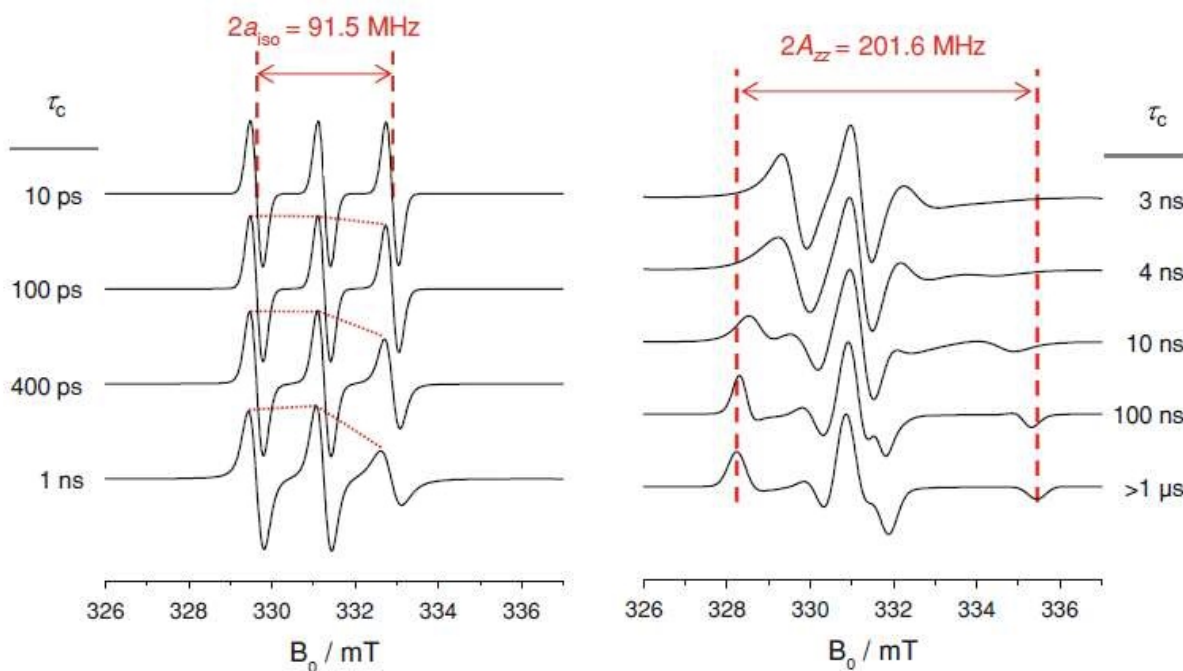


Figure 2.21. EPR spectra at obtained at fast (10-400 ps), intermediate (1-100 ns) and rigid-limit ( $> 1 \mu s$ ). Figure is from Junk (2012) *Assessing the Functional Structure of Molecular Transporters by EPR Spectroscopy*, Springer-Verlag Berlin Heidelberg, New York.

Furthermore, the local environment can also affect the spin label's motion. The energy difference between the SOMO and LUMO are altered according to the polarity of the solvent. The polarity ultimately impacts the EPR lineshape by changing both  $g$  value and the hyperfine coupling constants. Overall, this permits extraction of information related to a spin attached protein system, with respect to secondary structural elements, conformational changes, and protein-protein interactions (Figure 2.22)<sup>41-43</sup>.

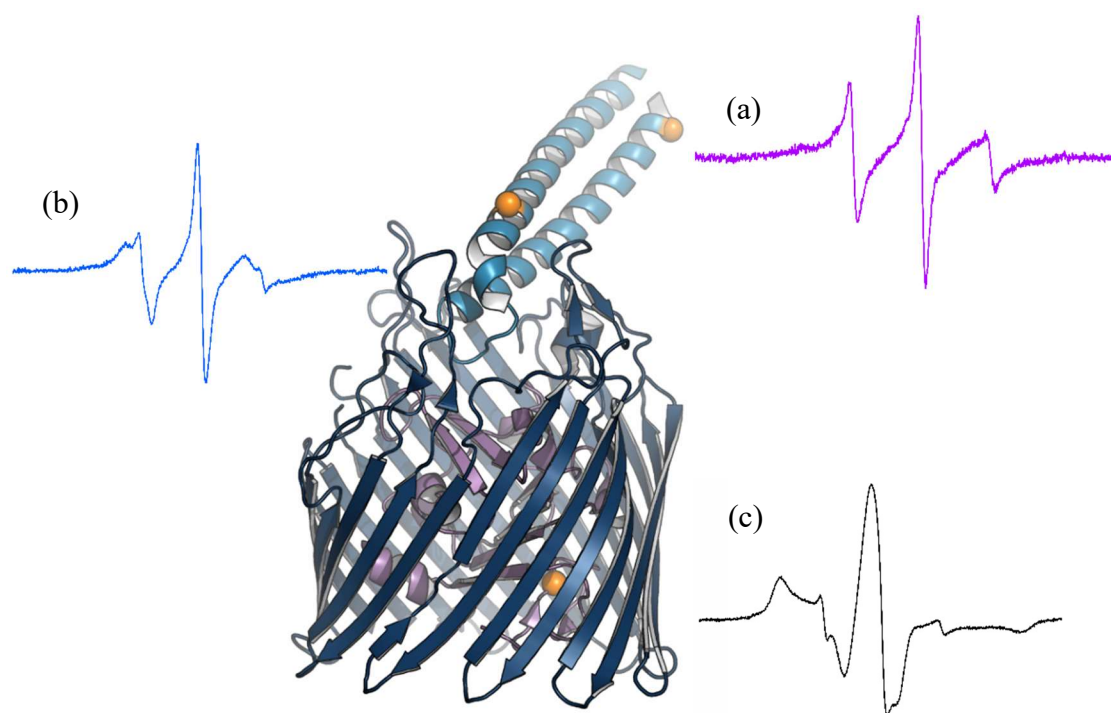


Figure 2.22. EPR lineshape analysis. CW EPR spectra obtained for **colicin receptor domain** bound to **BtuB**. (a) **K416R1** mobility ( $\alpha$ -helix) is uninhibited upon binding with BtuB and it contains a large nitroxide spin population in fast motion. (b) BtuB loops interact with **A370R1** ( $\alpha$ -helix) and restrict the spin mobility. Thus, the fast-moving component of **A370R1** is reduced. (c) **V10R1** of BtuB is folded inside **BtuB barrel** and it is with tertiary contact. Thus, reduce the rotational diffusion and the resultant EPR spectrum is broad and low amplitude.

### 2.5.3 Double electron-electron resonance (DEER)

Double electron-electron resonance (or DEER) is a technique to measure distances between pairs of electron spins, and it is a technique based upon an electron spin-echo (ESE). With strong excitation pluses, there is a certain time period in which the signals cannot be recorded (dead time) due to ring-down in the resonator<sup>39</sup>. In order to overcome this issue, a pulse sequence can be used to form an ESE to regain the initial magnetization<sup>44</sup>. However, due to spin-spin

relaxation, the size of the echo decreases as the time-period over which echo is generated is increased. The echo amplitude decreases exponentially as the echo time is increase where the characteristic decay time is referred to as the phase memory time,  $T_m$ <sup>44</sup>.  $T_m$  values that are not sufficiently long will not allow the ESE to be recorded with long echo times. Unwanted echoes generated by these pulses can be removed by phase cycling, as it permits selection of a single coherent transfer pathway of the echo of interest and maximizes the amplitude, while removing others. Furthermore, this pulse EPR technique has been developed to produce distance distributions between two coupled electron spins (dipole-dipole interactions) by eliminating other interactions.

#### 2.5.4 Four pulse DEER

This four pulse EPR technique is commonly used to obtain distance distributions in both soluble<sup>45</sup>, as well as, membrane proteins<sup>46</sup>. By coupling this technique with the double SDSL (DSDSL) method, two selected residues can be spin labeled to extract information regarding the distance distributions and the relative interacting populations between these pairs. Thus, enabling to gain more information regarding conformational changes. However, when selecting the residues to spin label for these studies, the sites should be carefully selected to have accessible rotamers at which the label can be attached. Because the rotamers of the residues are affected by nearby residues, as well as other neighboring molecules in the sample (OM or intact cells), some sites cannot be labeled properly<sup>33</sup>.

Four pulse DEER allows one to avoid dead-time issues<sup>47</sup> by coupling an independent second pulse (pump pulse) with the initial pulse (observer pulse) sequence<sup>44</sup> (Figure 2.23). By applying the  $\pi/2$  pulse, the magnetization vector in z-axis is shifted to xy-plane where different spins relax

at different rates. Thus, the  $\pi$  pulse is applied in order to invert the spins and to enforce the magnetization vector in xy-plane which gives rise to first (Hahn) echo. As the spins relax, after a certain time, a second  $\pi$  pulse is applied to re-enforce the echo in xy-plane. A period of time following the first echo, a pump pulse ( $\pi$ ) is applied to re-enforce another set of spins independently from the spins re-enforced by observer pulse. Thus, allowing two re-enforced spins to interact via dipole coupling.

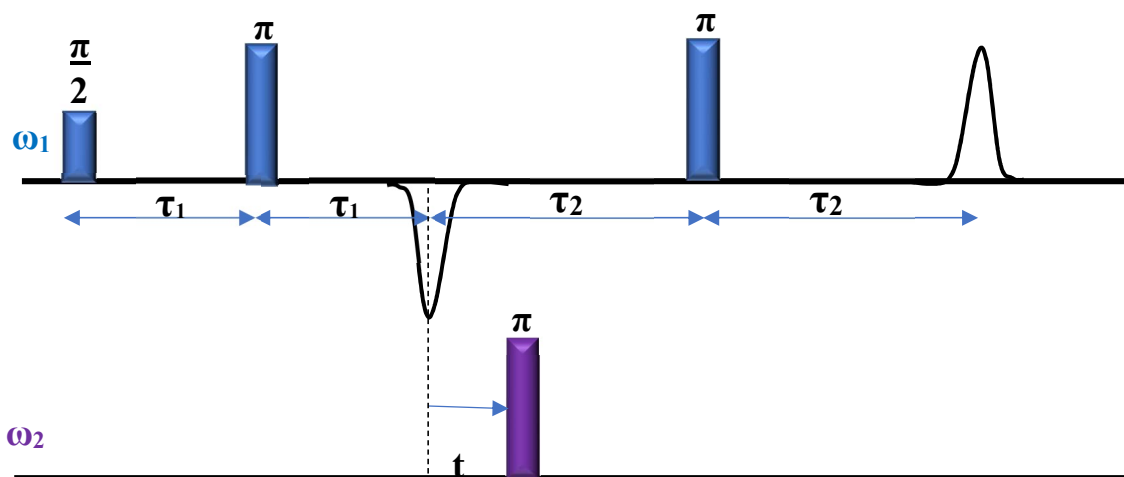


Figure 2.23. Four-pulse sequence of the DEER experiment. Primary echo sequence with observer frequency ( $\omega_1$ ) is refocused by applying pump frequency ( $\omega_2$ ) at time  $t$  after the undetected first echo. Times  $\tau_1$  and  $\tau_2$  are kept constant, while  $t$  can be varied.

Since the pump pulse is not closer to the detecting pulse, it provides zero dead-time<sup>44</sup>. Currently, DEER can be used to measure distances between 2-6 nm and for more favorable situations, it can measure up to 8 nm<sup>31, 32, 48</sup>. These distances are based on the dipole-dipole coupling between unpaired electron spins (S1 and S2) on the labeled protein or the molecule (Figure 2.24).

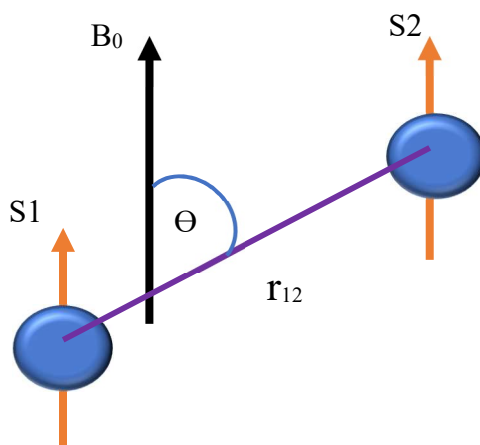


Figure 2.24. Dipole-dipole coupling between two electron spins (S1 and S2) in the DEER experiment that are separated by  $r_{12}$  distance in the z-axis.

When two spins (S1 and S2) are separated by distance  $r_{12}$  and the dipolar coupling constant (a) is given as,

$$a(r, \theta) = \omega_d(1 - 3 \cos^2 \theta) \quad (2.13)$$

where the dipolar frequency ( $\omega_d$ )

$$\omega_d \equiv 2\pi\nu_d = \frac{\gamma_e^2 \hbar}{r^3} \quad (2.14)$$

In order to obtain distance data, strong interactions are first removed to detect the weak electron dipole-dipole coupling. One way of achieving this is by refocusing the S1 spin interactions by echo experiment. Next, the coupling between S1 and S2 is re-introducing by applying an inversion pulse on S2 spin<sup>31</sup>. It is also important to separate the background from the signal. The collected DEER signal is combination of spin pair interactions (intramolecular) and interactions between other spin labeled molecules in the sample (intermolecular). Thus, the echo

signal,  $V(t)$ , is composed of a spin pair interaction contribution,  $F(t)$ , and a background contribution,  $B(t)$  where  $B(t)$  needs to be subtracted to obtain  $F(t)$  (Figure 2.25a). After fitting the background,  $B(t)$ ,  $F(t)$  can be obtained by dividing  $V(t)$  by  $B(t)$  (Figure 2.25b) and  $F(t)$  is converted into distance distribution  $P(r)$ <sup>31</sup>.

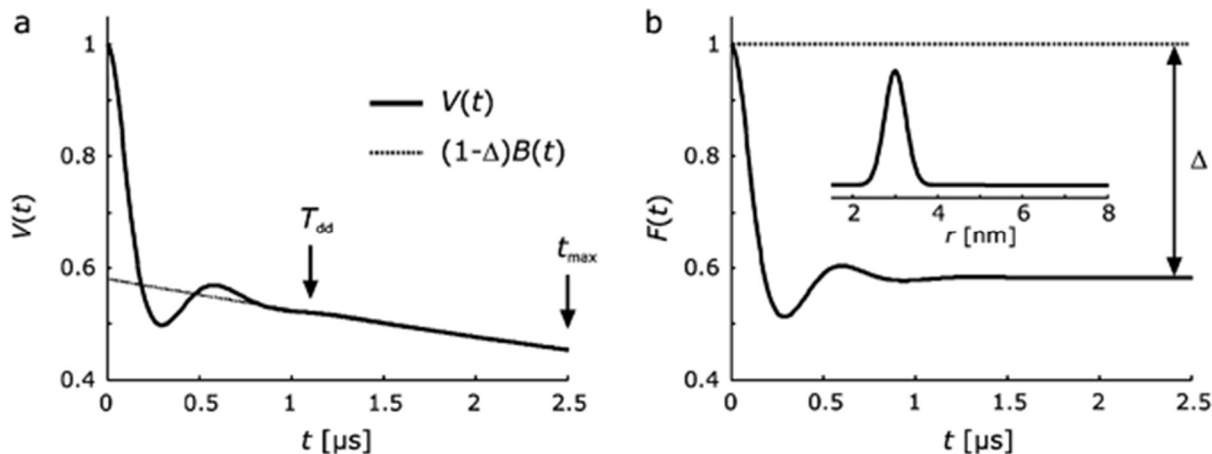


Figure 2.25. Obtaining  $F(t)$  from DEER signal. (a) The echo signal is observed until  $t_{max}$ .  $F(t)$  has a damped oscillation shape, which completely decays at  $T_{dd}$ . The rest of the signal is  $B(t)$ . Thus,  $B(t)$  is fitted in  $V(t)$ . (b)  $F(t)$  is obtained after removing  $B(t)$  from  $V(t)$  and re-normalized to time zero and  $F(t)$  is used to obtain the corresponding Gaussian distribution<sup>31</sup>. Figure is from Jeschke, and Polyhach. (2007), *Physical chemistry chemical physics : PCCP* **9**, 1895-1910.

## 2.6 References

- [1] Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning A Laboratory Manual*, Third ed., Cold Spring Harbor Laboratory Press, New York.
- [2] Rychlik, W., Spencer, W. J., and Rhoads, R. E. (1990) Optimization of the annealing temperature for DNA amplification in vitro, *Nucleic Acids Research* 18, 6409-6412.
- [3] Lundberg, K. S., Shoemaker, D. D., Adams, M. W., Short, J. M., Sorge, J. A., and Mathur, E. J. (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*, *Gene* 108, 1-6.
- [4] Mierzejewska, K., Siwek, W., Czapinska, H., Kaus-Drobek, M., Radlinska, M., Skowronek, K., Bujnicki, J. M., Dadlez, M., and Bochtler, M. (2014) Structural basis of the methylation specificity of R.DpnI, *Nucleic Acids Research* 42, 8745-8754.
- [5] Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Res.* 31, 3784-3788.
- [6] Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almabrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntaggart, N. O., Sailor, C. A., Dawson, R. B., and Peek, A. S. (2008) IDT SciTools: a suite for analysis and design of nucleic acid oligomers, *Nucleic Acids Research* 36, W163-W169.
- [7] Klock, H. E., Koesema, E. J., Knuth, M. W., and Lesley, S. A. (2008) Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts, *Proteins* 71, 982-994.
- [8] Klock, H. E., and Lesley, S. A. (2009) The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis, *Methods in molecular biology (Clifton, N.J.)* 498, 91-103.
- [9] One Shot® TOP10 Competent Cells [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/oneshottop10\\_man.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/oneshottop10_man.pdf), Accessed on 02 August 2018.
- [10] DNA purification and analysis Maximize sample yield, purity, and integrity, <https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/DNA-Reference-Guide-08-2017-Final.pdf>, Accessed on 10 July 2018.
- [11] Certificate of Analysis: Pfu DNA Polymerase:, <https://www.promega.com/-/media/files/resources/protocols/product-information-sheets/g/pfu-dna-polymerase-protocol.pdf>, Accessed on 20 July 2018.
- [12] Lodish H, B. A., Zipursky SL, et al. (2000) *Molecular Cell Biology*, 4th ed., W. H. Freeman, New York.
- [13] pET-28a-c(+) Vectors, <https://biochem.web.utah.edu/hill/links/pET28.pdf> Accessed on 07 July 2018.
- [14] Gallagher, S. R. (2001) Quantification of DNA and RNA with absorption and fluorescence spectroscopy, *Current protocols in cell biology Appendix 3*, Appendix 3D.
- [15] Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors, *Proceedings of the National Academy of Sciences* 74, 5463-5467.
- [16] Sanders, E. R. (2012) Aseptic Laboratory Techniques: Plating Methods, *Journal of Visualized Experiments : JoVE*, 3064.
- [17] Bykowski, T., and Stevenson, B. (2008) Aseptic Technique, *Current Protocols in Microbiology* 11, A.4D.1-A.4D.11.

- [18] Creating Bacterial Glycerol Stocks for Long-term Storage of Plasmids, <https://www.addgene.org/protocols/create-glycerol-stock/>, Accessed on 15 July 2018.
- [19] Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels, *Journal of molecular biology* 260, 289-298.
- [20] Rigaud, J.-L., Pitard, B., and Levy, D. (1995) Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins, *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1231, 223-246.
- [21] Seddon, A. M., Curnow, P., and Booth, P. J. (2004) Membrane proteins, lipids and detergents: not just a soap opera, *Biochimica et biophysica acta* 1666, 105-117.
- [22] Sepsey, A., Bacskey, I., and Felinger, A. (2014) Molecular theory of size exclusion chromatography for wide pore size distributions, *Journal of chromatography. A* 1331, 52-60.
- [23] Guo, H., Li, X., and Frey, D. D. (2014) Development of chromatofocusing techniques employing mixed-mode column packings for protein separations, *Journal of chromatography. A* 1323, 57-65.
- [24] Hage, D. S., and Matsuda, R. (2015) Affinity chromatography: a historical perspective, *Methods in molecular biology (Clifton, N.J.)* 1286, 1-19.
- [25] Mitchell, S. F., and Lorsch, J. R. (2015) Protein Affinity Purification using Intein/Chitin Binding Protein Tags, *Methods in enzymology* 559, 111-125.
- [26] Duong-Ly, K. C., and Gabelli, S. B. (2014) Chapter Eight - Using Ion Exchange Chromatography to Purify a Recombinantly Expressed Protein, In *Methods in enzymology* (Lorsch, J., Ed.), pp 95-103, Academic Press.
- [27] Shen, H.-H., Lithgow, T., and Martin, L. L. (2013) Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities, *International Journal of Molecular Sciences* 14, 1589-1607.
- [28] 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine, <https://www.anatrace.com/Products/Lipids/LIPIDS/D518>, Accessed on 10 July 2018.
- [29] 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine, <https://www.anatrace.com/Products/Lipids/LIPIDS/P516>, Accessed on 10 July 2018.
- [30] Schubert, R. (2003) Liposome Preparation by Detergent Removal, In *Methods in enzymology*, pp 46-70, Academic Press.
- [31] Jeschke, G., and Polyhach, Y. (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance, *Physical chemistry chemical physics : PCCP* 9, 1895-1910.
- [32] Jeschke, G. (2012) DEER distance measurements on proteins, *Annual review of physical chemistry* 63, 419-446.
- [33] Fanucci, G. E., and Cafiso, D. S. (2006) Recent advances and applications of site-directed spin labeling, *Curr Opin Struct Biol* 16, 644-653.
- [34] Klare, J. P., and Steinhoff, H. J. (2009) Spin labeling EPR, *Photosynthesis research* 102, 377-390.
- [35] Berliner, L. J., Grunwald, J., Hankovszky, H. O., and Hideg, K. (1982) A novel reversible thiol-specific spin label: papain active site labeling and inhibition, *Analytical biochemistry* 119, 450-455.

- [36] Becker, C. F., Lausecker, K., Balog, M., Kalai, T., Hideg, K., Steinhoff, H. J., and Engelhard, M. (2005) Incorporation of spin-labelled amino acids into proteins, *Magnetic resonance in chemistry : MRC 43 Spec no.*, S34-39.
- [37] Cornish, V. W., Benson, D. R., Altenbach, C. A., Hideg, K., Hubbell, W. L., and Schultz, P. G. (1994) Site-specific incorporation of biophysical probes into proteins, *Proc Natl Acad Sci U S A* 91, 2910-2914.
- [38] Kolb, H. C., Finn, M. G., and Sharpless, K. B. (2001) Click Chemistry: Diverse Chemical Function from a Few Good Reactions, *Angewandte Chemie (International ed. in English)* 40, 2004-2021.
- [39] M.J.N.Junk (2012) *Assessing the Functional Structure of Molecular Transporters by EPR Spectroscopy*, Springer-Verlag Berlin Heidelberg, New York.
- [40] Hubbell, W. L., and Altenbach, C. (1994) Investigation of structure and dynamics in membrane proteins using site-directed spin labeling, *Current Opinion in Structural Biology* 4, 566-573.
- [41] Fanucci, G. E., Cogshall, K. A., Cadieux, N., Kim, M., Kadner, R. J., and Cafiso, D. S. (2003) Substrate-Induced Conformational Changes of the Periplasmic N-Terminus of an Outer-Membrane Transporter by Site-Directed Spin Labeling, *Biochemistry* 42, 1391-1400.
- [42] Columbus, L., Kalai, T., Jeko, J., Hideg, K., and Hubbell, W. L. (2001) Molecular motion of spin labeled side chains in alpha-helices: analysis by variation of side chain structure, *Biochemistry* 40, 3828-3846.
- [43] Crane, J. M., Mao, C., Lilly, A. A., Smith, V. F., Suo, Y., Hubbell, W. L., and Randall, L. L. (2005) Mapping of the docking of SecA onto the chaperone SecB by site-directed spin labeling: insight into the mechanism of ligand transfer during protein export, *Journal of molecular biology* 353, 295-307.
- [44] Borbat, P. P., and Freed, J. H. (2013) Pulse Dipolar Electron Spin Resonance: Distance Measurements, In *Structural Information from Spin-Labels and Intrinsic Paramagnetic Centres in the Biosciences* (Timmel, C. R., and Harmer, J. R., Eds.), pp 1-82, Springer Berlin Heidelberg, Berlin, Heidelberg.
- [45] Borbat, P. P., McHaourab, H. S., and Freed, J. H. (2002) Protein Structure Determination Using Long-Distance Constraints from Double-Quantum Coherence ESR: Study of T4 Lysozyme, *Journal of the American Chemical Society* 124, 5304-5314.
- [46] Jeschke, G., Wegener, C., Nietschke, M., Jung, H., and Steinhoff, H.-J. (2004) Interresidual Distance Determination by Four-Pulse Double Electron-Electron Resonance in an Integral Membrane Protein: The Na(+)/Proline Transporter PutP of Escherichia coli, *Biophysical journal* 86, 2551-2557.
- [47] Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W. (2000) Dead-time free measurement of dipole-dipole interactions between electron spins, *Journal of magnetic resonance (San Diego, Calif. : 1997)* 142, 331-340.
- [48] Schiemann, O., and Prisner, T. F. (2007) Long-range distance determinations in biomacromolecules by EPR spectroscopy, *Quarterly reviews of biophysics* 40, 1-53.

## **Chapter 3: Direct spin labeling approach for EPR analysis using live bacterial cells to explore outer membrane transporters**

### **3.1 Abstract**

In the outer membrane (OM), the active transport of TonB dependent transporters (TBDTs) is facilitated by coupling the proton motive force (pmf) in the inner membrane (IM) through the TonB complex. However, our current knowledge regarding these transporters is mainly based on the studies done in isolated systems. These isolated systems do not truly mimic the physiological OM nor contain the vital protein players which are required to maintain properly functioning TBDTs. Thus, we aimed to develop an approach in which actively metabolizing bacterial cells expressing TBDT of interest can be directly spin labeled and studied using electron paramagnetic resonance (EPR) spectroscopy.

Here, *Escherichia coli* (*E. coli*) cells expressing BtuB was used as our model system to optimize this approach. In our final protocol, we successfully spin labeled single or double cysteine residues spaced in the extracellular side of BtuB. These single and double mutants are expressed in RK5016 and in Dsb knockout *E. coli* strains respectively. In our approach, we achieved labeling by harvesting the cells in the early-log phase in order to maximize cell viability, and we used carefully chosen cell lines and adjusted the labeling protocol so that the reductive potential in the periplasm was not dramatically altered. The approach allows us to examine the structural state of BtuB under conditions where transport takes place and allows us to explore the role of the lipopolysaccharides in altering OMP structure and the organization of OMPs into domains or islands that are believed to be a central feature of OM structure.

### 3.2 Introduction

The outer membrane (OM) in Gram-negative bacteria, acts as a selective barrier in which outer membrane proteins (OMPs) enable transport across the OM. As described in Chapter 1.5, the active transport of TonB dependent transporters (TBDTs) require the inner membrane (IM) proton motive force (pmf) and the IM TonB complex<sup>1, 2</sup>. During the last few decades, a vast amount of research has been performed in order to understand some of TBDT family members (Chapter 1.5.3 and Chapter 1.5.4). These studies have provided insight into the TBDT structure and function, but they have usually been performed in purified protein in detergent or reconstituted liposomes, and in a few cases in isolated OM preparations. The available structures display TBDTs in low energy states and do not reveal transient or excited conformations of the protein<sup>3, 4</sup>. In contrast to crystallography, site directed spin labeling (SDSL) coupled to electron paramagnetic resonance (EPR) is a powerful tool for studying intrinsic states of TBDTs (Chapter 2.5).

Another limitation using isolated systems (intact OM or lipid vesicles) is the lack of key proteins and other biomolecules that are involved in the TBDT transport and regulation available in live bacterial cells. As discussed in Chapter 1.3 and Chapter 1.4, bacterial membranes and OM biogenesis are both extremely complex. Each biomolecule plays a vital role in maintaining cell viability, and insuring that TBDTs are in functional states. Determining the transport mechanism of TBDTs is important, as this knowledge can be utilize in many fields including the development of new pharmaceuticals<sup>2</sup>. So, one should attempt to design methods in order to monitor intrinsic events associated with the transport process in live, healthy, and active cells, minimizing alterations to the protein and to the cells.

A few years ago, our laboratory working in collaboration Dr. Joseph in the Prisner group at the University of Frankfurt, Germany, developed a new strategy using *Escherichia coli* (*E. coli*) cells expressing BtuB coupled to SDSL-EPR<sup>5, 6</sup>. These initial studies demonstrated that it was possible to directly spin label BtuB and generate good-quality EPR data from cells. The spin labeled side chain at position 188 (T188R1) is located on the second extracellular loop in BtuB (Figure 3.2) in cells was analyzed using continuous wave (CW) EPR in its apo and B<sub>12</sub> bound states in the presence of Ca<sup>2+</sup> (Figure 3.1a). Next, both CW and pulsed EPR were used with T188R1 to explore binding of the spin labeled substrate, (2,2,6,6-Tetramethylpiperidin-1-yl)oxyl labeled cyanocobalamin, TEMPO-B<sub>12</sub> (Figure 3.b, c, and d)<sup>5</sup>.

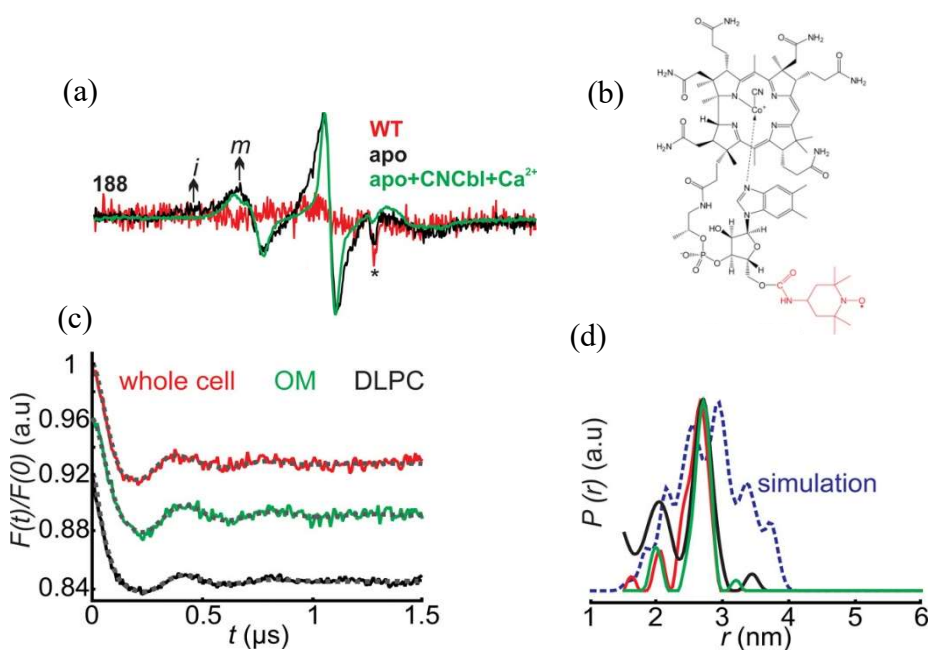


Figure 3.1 (a) CW EPR spectra of spin labeled WT BtuB, T188C BtuB apo and B<sub>12</sub> bound in the presence of Ca<sup>2+</sup>. (b) The structure of TEMPO-B<sub>12</sub>. (c) Normalized form factors after background subtraction for T188R1 BtuB with TEMPO-B<sub>12</sub> in cells, intact OM and in reconstituted lipid bilayer system (DLPC) and (d) the corresponding extracted distance distributions. The normalized distance distributions are within the MMM rotamer distribution predictions<sup>5</sup>. Figures are from Joseph *et al.* (2015) *Angewandte Chemie (International ed. in English)*. **54**, 6196-6199.

Starting with this platform of cell preparation, our original goal was to investigate protein dynamics associated with substrate binding in the binding pocket of BtuB. More specifically, to determine whether the B<sub>12</sub> binding site might be altered in the transport defective mutant, L8P BtuB, relative to wild type BtuB. Experiments were designed to probe different distances between spin labeled sites (T188, A288, D492, and S533) on the extracellular side of BtuB in cells (Figure 3.2) relative to bound TEMPO-B<sub>12</sub> using pulsed EPR. In order to successfully obtain DEER results, one has to overcome two main hurdles: (i) Spin labeling the cysteine residue on the extracellular site efficiently with minimum or no background labeling and (ii) obtaining good modulation depth in the DEER experiment, which result from dipole-dipole interactions between significant populations of the protein's and the bound substrate's spins.

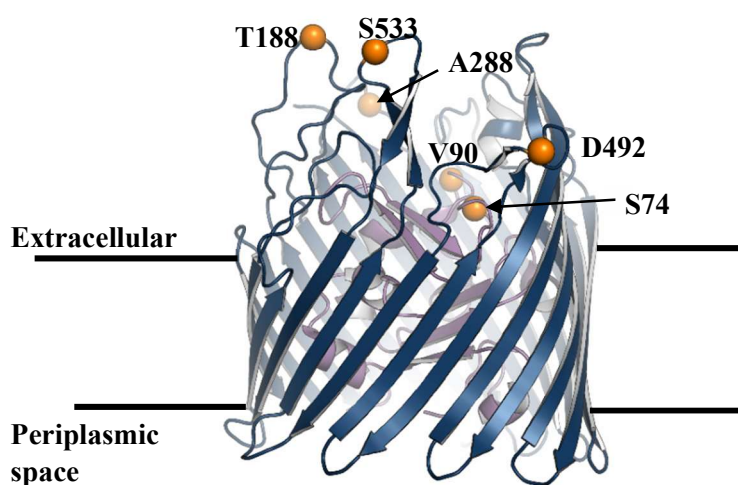


Figure 3.2. Structure of BtuB (1NQG) indicating extracellular side exposed S74, V90, T188, A288, D492 and S533.

However, our initial attempts repeating the initial protocol were unsuccessful. We concluded that it was necessary to develop an altered protocol that would generate a preparation where the metabolic state of the cells was optimized. Thus, the current work was heavily focused on method optimization in order to overcome the two hurdles mentioned above for single residue labeling and next for labeling pairs of cysteine residues labeling on the extracellular surface of BtuB.

### 3.3 Methods

#### 3.3.1 Growing RK5016 cells expressing BtuB mutants

Minimal media (MM) pre-cultures (100 mM phosphate buffer, 8 mM  $(\text{NH}_4)_2\text{SO}_4$ , and 2 mM sodium citrate) supplemented with 200  $\mu\text{g}/\text{mL}$  ampicillin, 0.2 % w/v glucose, 150  $\mu\text{M}$  thiamine, 3 mM  $\text{MgSO}_4$ , 300  $\mu\text{M}$   $\text{CaCl}_2$ , 0.01 % w/v methionine, and 0.01% w/v arginine were directly inoculated using small aliquots (5-10  $\mu\text{L}$ ) of the corresponding mutant glycerol stocks. These were then used to inoculate MM main cultures.

#### 3.3.2 Spin labeling of O/N grown cells expressing BtuB

##### 3.3.2(a) Screening for spin labeling time period

RK5016 cells expressing D492C and L8P-D492C BtuB were grown O/N at 34° C in nutrient supplemented 1 L minimal media cultures. The cells were harvested by centrifugation at 6000 rpm for 10 min at 4° C. The cell pellets were resuspended in 10 mL of 50 mM HEPES pH 8.0 and 150 mM NaCl (high salt buffer). A freshly prepared MTSSL (36 mM) was added to each cell resuspension and incubated at 34° C for 20, 30, and 40 min in the dark with gentle mixing. At the end of the incubation, cells were pelleted (4500 rpm for 10 min at 4° C) and the surfaces

of cell pellets were washed three times with high salt buffer. Next, the cells were loaded into a glass capillary and analyzed by CW EPR in order to determine spin labeling efficiency.

The above experiment was repeated with RK5016 cells expressing A288C, L8P-A288C, T188C, L8P-T188C, S533C, L8P-S533C BtuB. RK5016 cells either with or without WT BtuB expression were used as control samples. Based on the spin labeling efficiencies determined from previous experiments, L8P-T188C and L8P-S533C cell resuspensions were incubated for 40 min and the controls were incubated for 20 min with the MTSSL spin label.

### 3.3.2(b) Impact of spin labeling at high and low salt buffers

O/N grown RK5016 cells expressing T188C BtuB were harvested (4x 50 mL samples) and resuspended in 10 mL of 50 mM HEPES pH 8.0 with (high salt buffer) and without (low salt buffer) 150 mM NaCl. Resuspended cells were spin labeled with 36 mM MTSSL for an hour at room temperature (RT) in the dark with gentle mixing. At the end of the incubation, cells were collected and surface washed, to remove excess unbound MTSSL then analyzed by CW EPR.

Low salt spin labeling was repeated with O/N grown RK5016 cells expressing L8P-T188C, A288C, L8P-A288C, and S533C BtuB. The cells from 50 mL culture samples were harvested by centrifugation at 4500 rpm for 10 min at 4° C. The cells were spin labeled, spin labeling efficiency was analyzed after removing the free spin label.

### 3.3.2(c) TEMPO- B<sub>12</sub> binding to BtuB single mutants

Another set of T188C and S533C BtuB expressed cells were incubated in low salt buffer without MTSSL for an hour. Harvested cells were used to study the binding of spin labeled vitamin B<sub>12</sub> (TEMPO-B<sub>12</sub>) with the final concentrations ranging from 0 – 200 µM. Spin labeled cells of T188C and S533C BtuB (T188R1 and S533R1) were then mixed with 30 µM TEMPO-B<sub>12</sub> and used to prepare DEER samples.

### 3.3.3 Screening the spin labeling efficiency at different growth time points

Main minimal media culture (1 L) was inoculated with the O/N grown pre-culture of RK5016 cells expressing T188C BtuB and a 50 mL of aliquot was taken every 90 min. Cells were harvested at 4500 rpm for 10 min at 4° C. Cell pellets were then resuspended in low salt buffer and freshly prepared MTSSL (1 mg per sample) was added and spin labeled for an hour at RT. At the end, cells were harvested (centrifugation at 4500 rpm for 10 min at 4° C), cell pellets were washed, and then analyzed by CW EPR. DEER samples were then prepared for pulsed EPR analysis.

### 3.3.4 Spin labeling BtuB mutants using cells grown to mid-log phase of bacterial growth curve

#### 3.3.4(a) Spin labeling cells harvested at mid-log phase

O/N grown pre-cultures of T188C BtuB expressing RK5016 cells were used to inoculate minimal media main culture (250 mL) and incubated at 34° C until about 0.6 OD<sub>600</sub>. Then the cells (50 mL of cell culture samples) were harvested and resuspended in 10 mL of 100 mM HEPES pH 8.0 (resuspension buffer). Freshly prepared MTSSL (1 mg per sample) was added to a set of samples and others were incubated without MTSSL for 1 and 2 hrs. At the end of the incubation, cells were harvested, cell pellets were washed, and then analyzed by CW EPR. Cell pellets were then mixed with TEMPO-B<sub>12</sub> (5 – 100 µM) and used to prepare samples for the DEER experiment.

#### 3.3.4(b) Minimizing the reduction of spin labeled BtuB

T188C expressing RK5016 cells were grown until 0.6 OD<sub>600</sub> and cells from 100 mL culture samples were harvested and resuspended into 10 mL of resuspension buffer, pH 7.0.

These were spin labeled using freshly prepared MTSSL stock (1 mg per sample) and incubated for an hour. After the incubation, cells were collected and resuspended either into 100 mM HEPES pH 7.0 or in 100 mM MES pH 5.5 and then incubated on ice for an hour. Cells were harvested and analyzed for spin labeling efficiency, then samples for DEER experiment were prepared.

The above protocol was repeated with L8P-T188C, A288C, L8P-A288C, D492C, S533C, and L8P-S533C BtuB mutant expressing cells. Aliquots of spin labeled cell pellets were first analyzed by CW EPR and then incubated with 32.5  $\mu$ M TEMPO-B<sub>12</sub> on ice for 20 min then samples for the DEER experiment were prepared.

### 3.3.4(c) Impact on early log phase vs mid log phase grown cells on spin labeling

RK5016 cells expressing T188C and WT BtuB were grown until 0.3 or 0.6 OD<sub>600</sub>. The cells were harvested from 100 mL culture samples and spin labeled (1 mg of MTSSL) as above where 100 mM MES pH 5.5 was used during the ice incubation step. Processed samples were then analyzed for labeling efficiency.

### 3.3.5 Spin labeling the cysteine mutants in BtuB hatch domain

Cells expressing S74C and V90C were grown until OD<sub>600</sub> 0.3 and 0.6. Cells were harvested from 100 mL samples and resuspended into the resuspension buffer. 3 mg of MTSSL per sample was added and spin labeled at RT for an hour in the dark. Cells were then harvested and resuspended into wash buffer for an hour on ice. Finally, cells were harvested and analyzed by CW EPR. The experiment was next repeated with two washing steps with 30 min incubation periods on ice instead of a single wash step. Additionally, MES buffer prepared in D<sub>2</sub>O was used during the final washing step instead of double deionized H<sub>2</sub>O (ddH<sub>2</sub>O).

### 3.3.6 Spin labeling cysteine double mutants between BtuB barrel and the hatch using live *E. coli*

#### 3.3.6(a) Optimization of MTSSL spin label

RK5016 cells expressing V90C-T188C BtuB were grown until OD<sub>600</sub> 0.3 and cells were harvested and resuspended into resuspension buffer. A series of MTSSL (0.1, 0.2, 2 and 3 mg) per cells harvested from 100 mL cell cultures were added, spin labeled, cells were processed as above and analyzed the labeling efficiency.

#### 3.3.6(b) DTT treatment of BtuB in *E. coli* prior to spin labeling step

V90C-T188C BtuB expressing cells grown until OD<sub>600</sub> 0.3 and 0.6. 100 mL culture samples were then harvested. Prior to the spin labeling step, a set of cells were incubated with 100 mM DTT at 37° C for 10 min in order to reduce any potential disulfide bonds between the BtuB double mutants. At the end of the incubation, the cells were washed three times with resuspension buffer in order to remove leftover DTT from the cell resuspension. A control set were incubated without DTT. Next, 3 and 5 mg of MTSSL were added to each sample collected at OD<sub>600</sub> 0.3 and 0.6, respectively. Control, WT BtuB expressing cells grown until OD<sub>600</sub> 0.3 from 100 mL cell cultures were treated with and without 100 mM DTT and spin labeled using 3 and 5 mg MTSSL. At the end of the spin labeling, cells were processed and analyzed.

#### 3.3.6(c) Cells grown in DTT supplemented media

S74C-T188C BtuB expressing cells were inoculated in 250 mL of minimal media with overnight cultures and grown for about 2 hours until the starter culture was adopted into the new growth media (OD<sub>600</sub> was around 0.03). Then the media was supplemented with 2.5 mM DTT and OD<sub>600</sub> was monitored for cell doubling. Once the cultures reached OD<sub>600</sub> 0.06, more DTT was added to bring the final concentration to 5 mM. After next doubling around 0.12, the DTT concentration was increased up to 10 mM. Once the OD<sub>600</sub> reached 0.3, cells from 100 mL

culture samples were harvested and washed with resuspension buffer three times to remove any residual DTT within the cells. Then 3 mg of MTSSL were used for spin labeling. Cells grown without DTT were used as controls. After the labeling step, cells were washed and analyzed by CW EPR.

Another set of cell cultures were supplemented with a constant concentration of DTT throughout the growth rather than the gradual increase of the DTT. About 2 hours after the cells were inoculated in the main cultures, media was supplemented with 0, 10, 20, and 30 mM DTT final concentrations and then grown until OD<sub>600</sub> 0.3. Only the 0 and 10 mM DTT supplemented cultures grew to OD<sub>600</sub> 0.3, while the 20 mM and 30 mM DTT supplemented cultures did not grow. Thus, the cells from 100 mL cultures with 0 and 10 mM DTT containing cultures were harvested, washed to remove DTT, and used to spin label with 3 mg MTSSL per sample. After the removing free MTSSL, samples were analyzed using CW EPR.

### 3.3.7 Screening for spin labeling time and the temperature for BtuB double mutants

RK5016 cells expressing S74C-T188C BtuB were grown until OD<sub>600</sub> 0.3. Cells were then harvested and resuspended into 1 mL of resuspension buffer rather than 10 mL in order to increase the spin label accessibility using lower amounts of MTSSL. A 0.5 or 1 mg of MTSSL per cells harvested from 100 mL culture was added and then incubated for 20, 40, or 60 min at RT and in ice with gentle mixing. At the end of the incubation, cells were harvested and the unbound spin label was removed by washing then analyzed by CW EPR.

### 3.3.8 Spin labeling the TBDT double mutants expressed in Dsb knockout system

#### 3.3.8(a) Long term preservation of Dsb knockout strains expressing BtuB mutants

First, glycerol stocks of each *E. coli* Dsb knockout strain was prepared and stored in  $-80^{\circ}\text{C}$  for long term storage. SCC from a freshly streaked plate for each bacterial strain was used to prepare competent cells<sup>7, 8</sup> which were then stored at  $-80^{\circ}\text{C}$ .

Sequence confirmed T188C-S74C and T188C-V90C BtuB plasmids with ampicillin (Amp) resistance genes were transformed into all four Dsb strains and plated on LB-Amp plates to identify the transformants. The glycerol stocks of Dsb WT, *dsbA*<sup>-</sup>, *dsbB*<sup>-</sup> and *dsbC*<sup>-</sup> expressing T188C-S74C and T188C-V90C were prepared and stored at  $-80^{\circ}\text{C}$  for long term storage.

#### 3.3.8(b) Spin labeling of BtuB double mutants

*dsbB*<sup>-</sup> and *dsbC*<sup>-</sup> cells expressing T188C-V90C were grown until OD<sub>600</sub> reached 0.3 and 0.6. Then the cells (100 mL of culture) were harvested and resuspended into 10 mL of resuspension buffer. The cells were pelleted again to remove any residual growth media and then resuspended in 1 mL of resuspension buffer. A freshly prepared stock of MTSSL (1 mg) was then added and incubated at RT for an hour. The cells were pelleted and resuspended in washing buffer (1 mL), incubated on ice for 30 min, and then centrifuged. The cells were resuspended again in washing buffer prepared in D<sub>2</sub>O followed by incubating on ice for 30 min. Harvested cells were analyzed using CW EPR. The experiment was repeated with Dsb WT and *dsbA*<sup>-</sup> cells expressing the same double mutant, but the cells were only grown until OD<sub>600</sub> 0.3. Finally, the DEER samples were prepared using the samples with sufficient spin labeling.

### 3.3.8(c) Optimization of spin labeling in Dsb knockout system

After harvesting *dsbA*<sup>-</sup> cells expressing T188C-V90C (50 mL cell culture x 4) to remove residual cell culture media, the cell pellets were resuspended into either 1 mL or 0.5 mL of resuspension buffer depending on the volume of the spin label being added to the cells. 1 mg/100 uL fresh MTSSL stock was prepared and 0.5, 0.25, 0.1, or 0.05 mg of MTSSL was added to each resuspended cell culture. After labeling samples were processed as above. At the end, harvested cells were analyzed using CW EPR and DEER samples were prepared using the samples with sufficient spin labeling.

### 3.3.8(d) Spin labeling of T188C-S74C and T188C-V90C BtuB

RK5016 cells and *dsbA*<sup>-</sup> cells expressing BtuB double mutants were grown until OD<sub>600</sub> 0.3 and spin labeled, as above, using 0.1 mg or 0.05 mg MTSSL per 50 mL cells culture at OD<sub>600</sub> 0.3. At the end of the incubation, cell pellets were washed twice with wash buffer prepared in H<sub>2</sub>O and then D<sub>2</sub>O. Cells were pelleted and apo and B<sub>12</sub> added cells were analyzed using CW EPR. Properly spin labeled samples were used to prepare DEER samples.

### 3.3.8(e) Spin labeling of WT and T188C-V90C BtuB expressed in Dsb knockout strains

*dsbA*<sup>-</sup>, *dsbB*<sup>-</sup>, *dsbC*<sup>-</sup> and Dsb WT cells expressing WT BtuB and V90C-T188C BtuB were grown until OD<sub>600</sub> 0.3 and used to spin labeled as above and analyzed using CW EPR. Efficiently spin labeled samples were used to prepare DEER tubes.

### 3.3.9 EPR analysis of spin labeled BtuB

#### 3.3.9(a) CW-EPR analysis

During the experiments, samples were loaded into glass capillary tubes and the CW-EPR spectra were collected using an X-band operating Bruker EMX spectrometer with an ER 4123D dielectric resonator, a 100 G sweep width, 1 G modulation, and 2 mW of incident microwave power. The collected CW-EPR spectra (one scan spectra) were normalized, phase corrected, and plotted using in-house programs (Dave plot CW and CWPhase by David Nyenhuis).

#### 3.3.9(b) DEER experiments

After cells were processed, cell pellets (15 or 16  $\mu\text{L}$ ) were mixed with 4  $\mu\text{L}$  of deuterated glycerol (20 % v/v). A 1 or 2  $\mu\text{L}$  of TEMPO-B<sub>12</sub> (final concentration  $\sim$  30 or 60  $\mu\text{M}$ ) was added to produce substrate bound samples. Mixed samples were loaded into quartz capillary tubes and flash frozen either using liquid nitrogen or a dry ice isopropanol bath. For DEER samples prepared post 3.3.5(b), after TEMPO-B<sub>12</sub> was added, samples were incubated on ice for about 10 min then loaded into the tubes and flash frozen. Samples were either run or stored at  $-80^\circ\text{C}$  until the DEER experiment. DEER data was collected using a dead time free four pulsed DEER sequence on a Bruker ELEXSYS E580 at Q-band with a dielectric resonator, EN5107D2 dielectric resonator and a 10 W Bruker AmpQ. During the data collection 32 ns and 36 ns lengths were used for the observe and pump  $\pi$  pulses, given with a 75 MHz frequency offset between the observe and pump pulse.

DEER data collected for single residue labeled BtuB were analyzed using DD version 7b<sup>9</sup>. After removing the 2+1 pulse artifact by overlapped excitation bands of pump and observer pluses, the data were fitted to a Gaussian model. Model based fitting was used to fit the background and Gaussian distributions to initial time data. The Akaike and Bayesian information

criterion values (AIC and BIC) were used to evaluate the quality of fits. The data for BtuB labeled at two sites were analyzed using LongDistances version 771 (Christian Altenbach, UCLA) and truncated to 2  $\mu$ s to minimize the contribution of OMP clustering effect on the background. After background subtraction, the distance distributions were generated using Gaussian model-based co-fitting scheme similar to DD. The chi-squared values were used to assess the quality of fits. The analyzed data were plotted using DavePlot (by David Nyenhuis) through a Plotly and Dash frameworks developed by the Plotly corporation.

### 3.3.10 Generated BtuB structures

BtuB crystal structure (1NQG)<sup>3</sup> was obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>10</sup> and figures were generated using PyMol Molecular Graphics System version 1.8<sup>11</sup>

### 3.4 Results and discussion

At the beginning of this project, based on the published literature<sup>5, 6</sup>, the first attempts were made to spin labeled BtuB cysteine mutants in *E. coli* cells. The levels of BtuB in the cells are regulated by different factors<sup>1</sup>. In the absence of B<sub>12</sub>, the expression levels of chromosomal BtuB is low (200-300 per cell<sup>12</sup>). Thus, to overexpress BtuB, minimal media were used to grow cells for all of the experiments. Additionally, the *E. coli* RK5016 cell line was used, as this strain is a *btuB* knockout which allows high-level expression of only the transformed *btuB* gene<sup>13</sup>. Thus, during BtuB expression in RK5016 cells, plasmid *btuB* gene expression levels will not be compromised by any chromosomal *btuB* expression.

However, we could not reproduce the results and obtain a good DEER signal according to the protocol. Briefly, in this protocol, O/N grown cells from 1 L culture were pelleted and resuspended into 50 mL of 50 mM Tris, pH 7.5, with 60 mM NaCl, and 0.5 % glucose. Next cells were spin labeled with 20  $\mu$ M MTSSL, for an hour at RT followed by pelleting, washing the cell pellets' surface (a couple of times), and resuspending into 2.5 mL of buffer and analyzing using both CW-EPR and DEER (Figure 3.1). Thus, we optimized the protocol to improve the spin labeling efficiencies to obtain sufficient interested spin pair interactions during the DEER experiment and to obtain a good DEER signal<sup>5, 6</sup>.

#### 3.4.1(a) Optimization of O/N grown cells' spin labeling

The protocol was first optimized using O/N grown cells expressing different BtuB barrel mutants either with or without the L8P mutation. During our initial optimizations, the spin labeling time and the buffer compositions were screened to identify the optimum conditions for these O/N grown cells. Based on the CW-EPR spectra obtained, the 20 min and 40 min

incubation periods had better spin labeling efficiencies for D492C and L8P-D492C BtuB, respectively. Thus, L8P-cysteine mutants were incubated for 40 min, while the control samples with and without WT BtuB and other mutants were incubated for 20 min (Figure 3.3).

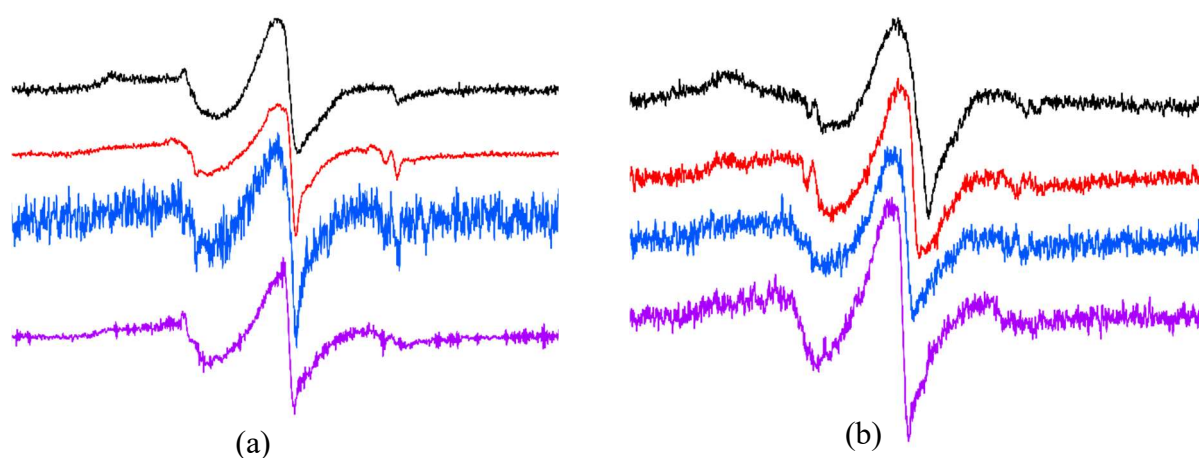


Figure 3.3 Normalized CW-EPR spectra obtained from O/N grown RK5016 cells expressing BtuB (a) without and (b) with L8P after free spin subtraction. The cells were spin labeled in high salt buffer and analyzed for spin labeling efficiency of (a) **D492R1**, **T188R1**, **A288R1**, and **S533R1** BtuB and (b) **L8P-D492R1**, **L8P-T188R1**, **L8P-A288R1**, and **L8P-S533R1** BtuB.

In addition to selecting the incubation time period, buffer composition was screened with NaCl (high salt buffer, 150 mM NaCl in 50 mM HEPES pH 8.0) and without NaCl (low salt buffer, 50 mM HEPES pH 8.0). Based on the CW EPR spectra (Figure 3.3 and Figure 3.4) equal spin labeling efficiency was detected under both buffer conditions. Thus, future experiments were performed in low salt buffer. At these sites, the spin labels have limited mobility (Figure 3.3 and Figure 3.4).

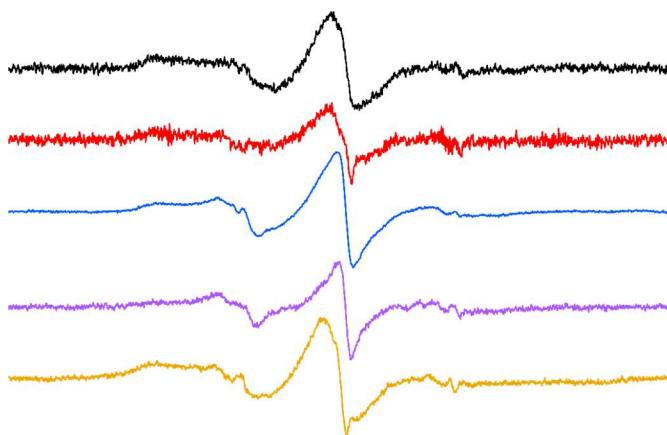


Figure 3.4. Normalized CW EPR of **A288R1**, **L8P-A288R1**, **T188R1**, **L8P-T188R1**, **S533R1** BtuB of O/N grown cells spin labeled in low salt buffer.

During these optimizations, we continuously used the cells expressing T188C, A288C, D492C, and S533C either with or without the L8P mutation. Even though the consistent labeling efficiency was obtained for cells in low salt buffer for both transport defective and the control BtuB mutants, the pulsed EPR data obtained for each sample (T188R1, L8P-T188R1, A288R1, L8P-A288R1) in the presence of TEMPO-B<sub>12</sub> did not display significant modulation depths. Thus, we decided to methodically develop a better protocol for spin labeling the cells. Since T188C BtuB has been studied previously in the O/N grown intact cells, as well as in the isolated systems (also during our so far optimizations T188C BtuB labeling had better spin labeling), we selected T188C BtuB to develop our method for intact cell spin labeling. Then, the successful, finalized protocol would be used for the other BtuB mutant labeling in the future.

### 3.4.1(b) Spin labeling at different growth points

As described in Chapter 1.2.3, bacterial cells can be grown either in batch, fed-batch, or continuous cultures. The most common laboratory method is using a batch cultures where the

cell cultures will go through all four growth phases: lag, log, stationary, and death phase. The time periods that bacteria spend on each growth phase depends on the growth medium, cell strains, and even on the expressed protein. For RK5016 cells expressing BtuB in minimal media, typically, cells enter stationary phase in about 8 hrs and after O/N cultures, the cells are in the late stationary or death phase.

As indicated above, our attempts to follow the published protocol did not yield satisfactory results and we focused harvesting cells from the appropriate growth phase in order to optimize the physiology and labeling of the cells (Chapter 1.2.3). Thus, we first screened the spin labeling efficiencies at different growth time points with cells expressing T188C BtuB. During the growth assay, cell culture aliquots from the same main culture (Section 3.3.3) were collected starting from the inoculation. However, at very early time points did not produce enough cell pellets for spin labeling. Thus, in this experiment, samples collected from the late log phase to the death phase were used for spin labeling. Because the procedure was cumbersome, instead of collecting samples at OD<sub>600</sub> points, samples were collected at every 90 min. Based on the CW-EPR spectra, each time point contains a mobile and an immobile population in each field (Figure 3.5).

As of the current protocol, we were washing the final spin labeled cell pellets' surface prior to CW analysis. Thus, residual free MTSSL gave rise to a small population of a narrow, high amplitude mobile component, while the population of the spin is bound and has slow or restricted dynamics results in the broad, low amplitude immobile component in each spectrum. Additionally, the lineshape of immobile component in the low field was gradually changing potentially due to changes in the spin dynamics and the population over the time. Based on the DEER experiments, for each time point in the growth assay, the modulation depths were insignificant (< 4%) and the resultant distance distribution probabilities did not contain the short

distance in the range between 2 to 3 nm which was expected for the T188R1 BtuB site with TEMPO-B<sub>12</sub>.

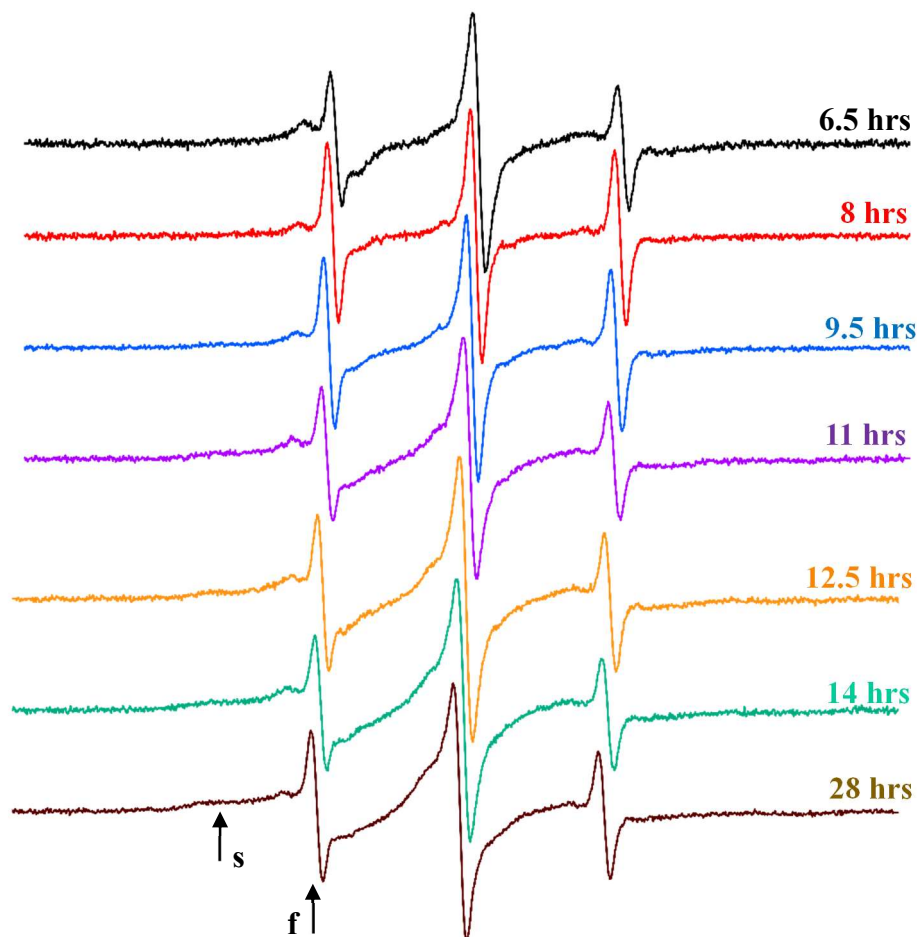


Figure 3.5 Normalized CW-EPR without free spin subtraction, obtained from T188R1-BtuB of cells harvested at different inoculation time points (6.5, 8, 9.5, 11, 12.5, 14, and 28 hrs from the inoculation) in the late log, stationary, and dead phase of the growth curve. The slow (s) and fast (f) populations in the low-field lines indicated by arrows.

Even though we were able to spin label the T188C BtuB site at each growth time point with similar labeling efficiencies, the cells that were used during this experiment were likely not healthy. When cells begin entering the stationary phase, the cells undergo various physiological and morphological changes. Some of these changes include: (i) modifications in the cell

envelope which increases the cross-linking between the OM and peptidoglycan layer (in the periplasm), increase in the peptidoglycan layer and reduction of the number of OMPs, (ii) reduction of cell division and size dwarfing due to self-digestion and degradation, (iii) reduction of growth rates, metabolism of carbohydrates, and synthesis of proteins and phospholipids, (iv) synthesis of secondary metabolites, bacteriocins, and toxins, and (v) repression of the aerobic metabolism, in order to protect the cells from reactive oxygen species (ROS) and to ration cellular reserves<sup>14</sup>. During the stationary phase, the cells are dying at an equal rate as the cells are doubling. Over time cells enter the death phase, where toxins and waste accumulate and nutrients are depleted which leads to an increased rate of death. Furthermore, dead cells and the accumulation of other biomolecules such as nucleic acids, polysaccharides and proteins lead to the formation of biofilms. Consequently, the viable cells in both the stationary and death phases are under stress.

In the previous work<sup>5, 6</sup> we suspect that some population of the cells are dead or dying with the remaining cells under stress and leading to the formation of biofilms. Since there may be some cell lysis, one might expect a significant level of non-specific labeling, as was seen previously for measurements on intact OM samples<sup>15</sup>. Pulse EPR measurements on intact OM preparations provide a viable method to explore protein structure in a more native environment, but under conditions where the IM machinery is absent. Since our objective was to characterize transport and OMPs under conditions where cells were as metabolically active as possible, we focused our protocol on cells harvested at earlier growth phases.

### 3.4.1(c) Mid-log phase cell labeling

As explained above, cells in both stationary and death phases are not ideal for our method development. Hence, we aimed to use cells in the exponential (log) phase. The cells were collected when the  $OD_{600}$  reached 0.6. Typically,  $OD_{600}$  at 0.6 is the mid-log phase and the cells are healthy and actively undergoing the bacterial cell cycle (explained in Chapter 1.2.1). Thus, this is an ideal growth phase to be targeting for spin labeling metabolically active bacterial cells.

Harvested cells at  $OD_{600}$  0.6 were spin labeled with MTSSL for one or two hours in low salt buffer at RT in order to optimize the labeling time period for the current protocol, so that a sufficient time period is given for spin labeling the T188C-BtuB site in live cells. Based on the normalized CW EPR spectra (Figure 3.6), spin labelling either for an hour or two has very similar efficiencies, with some broad component in the 2 hr spin labeled spectrum. Thus, for the future method optimizations, an hour incubation was selected to spin label BtuB cysteine mutants.

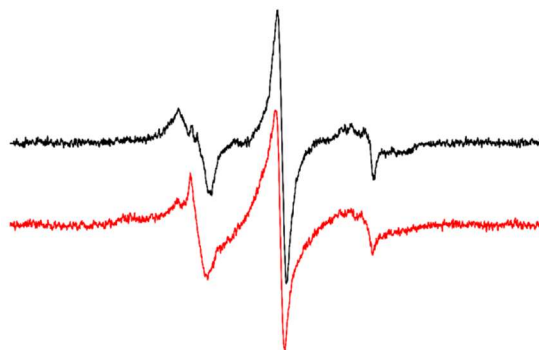


Figure 3.6. Normalized CW EPR spectra of T188C BtuB spin labeled for **one** and **two** hours at RT.

Simultaneously, unlabeled cells were used to assess a convenient range of TEMPO-B<sub>12</sub> addition to spin labeled cells for DEER, optimizing the dipole-dipole coupling between the T188R1 and the bound TEMPO-B<sub>12</sub> in order to obtain modulation depths greater than 6 % in the

echo signal. The observed CW EPR spectrum of TEMPO-B<sub>12</sub> is first narrow high amplitude due to freely diffusing substrate then upon binding with BtuB, the spin dynamics are restricted producing broad low amplitude spectra due to immobile substrate (anisotropic). When TEMPO-B<sub>12</sub> is incubated with the cells expressing BtuB, first, the substrate will bind with BtuB with high affinity and then it will be transported in to the cells. Upon substrate transport, the spin label will then be reduced in the periplasm<sup>6</sup>.

Different TEMPO-B<sub>12</sub> concentrations (5-100  $\mu$ M) were incubated with unlabeled cells in order to determine the suitable substrate concentration to be used for the pulsed EPR experiments. When the concentrations were 30  $\mu$ M or below, TEMPO-B<sub>12</sub> was completely bound and or transported into the cells. Once the TEMPO-B<sub>12</sub> is transported into the cells and all the BtuB are saturated with TEMPO-B<sub>12</sub>, the excess unbound substrate is left freely diffusing. Thus, the EPR spectra for 50  $\mu$ M and higher TEMPO-B<sub>12</sub> concentrations were narrow and high amplitude with overshadowed populations of immobile TEMPO-B<sub>12</sub> (Figure 3.7).

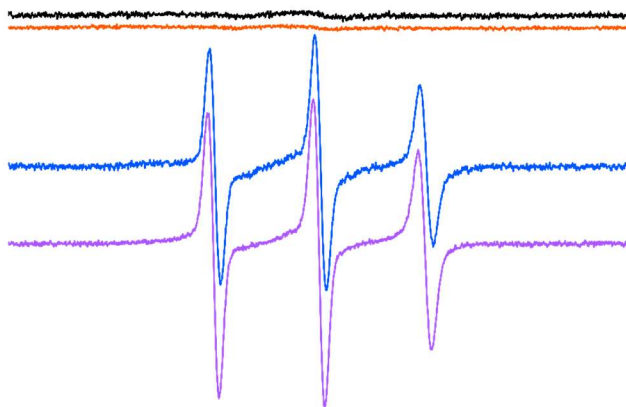


Figure 3.7. Normalized CW EPR of TEMPO-B<sub>12</sub> binding with T188C BtuB cells. The cells were incubated with **20  $\mu$ M**, **30  $\mu$ M**, **50  $\mu$ M**, **100  $\mu$ M**, TEMPO-B<sub>12</sub> concentrations and analyzed by CW-EPR.

Based on the above results, DEER samples were prepared using ~30 and 50  $\mu\text{M}$  final TEMPO-B<sub>12</sub> concentrations as it was critical to maintain TEMPO-B<sub>12</sub> saturated BtuB, but not in far excess concentrations of unbound substrate, because we are interested in detecting the interactions between the spin labeled BtuB and the BtuB bound spin labeled substrate. When there are excessive concentrations of unbound spin label substrate the interactions between these molecules will dominate in the background of the pulsed EPR echo signal and it will be hard to subtract the background without affecting the actual echo signal corresponding to the labeled BtuB and bound substrate.

#### 3.4.1(d) Minimizing the spin reduction during sample processing

Even though we were able to optimize the spin labeling of the cells and determined an appropriate TEMPO-B<sub>12</sub> concentration for the DEER experiment, low modulation depths indicated that there was not enough dipolar coupling between the T188R1 BtuB and the TEMPO-B<sub>12</sub>. This could potentially be due to spin reduction of either the T188R1 or the TEMPO-B<sub>12</sub> during the sample processing and/or due to transport of the TEMPO-B<sub>12</sub> into the cells. During the thus far optimized protocol (section 3.3.4), after the spin labeled cells were harvested, the surfaces of the cell pellets were washed and then used for CW and pulsed EPR sample preparation. Furthermore, these cell pellets were kept at RT until flash freezing.

Even though the cells were collected at the mid-log phase, when the cells are healthy and metabolically active, however some of the cells are aged due to continuous inheritance of old poles and are more susceptible to death than their counterparts with new poles<sup>16, 17</sup> (Chapter 1.2.2). Thus, some bacterial cells continuously die during the spin labeling step and spin labeled cell pellet processing at RT. The longer the cell pellets are left at RT, more cell lysis will occur. As the cells begin lysing, high levels of ascorbate are released into the extracellular space, which

reduces the T188R1 and TEMPO-B<sub>12</sub> spins<sup>18, 19</sup>. Subsequently, current sample processing caused significant reduction of both the spin labeled populations of BtuB and the bound substrate, thus the DEER samples did not have enough dipolar-dipolar coupling to produce a significant modulation depth.

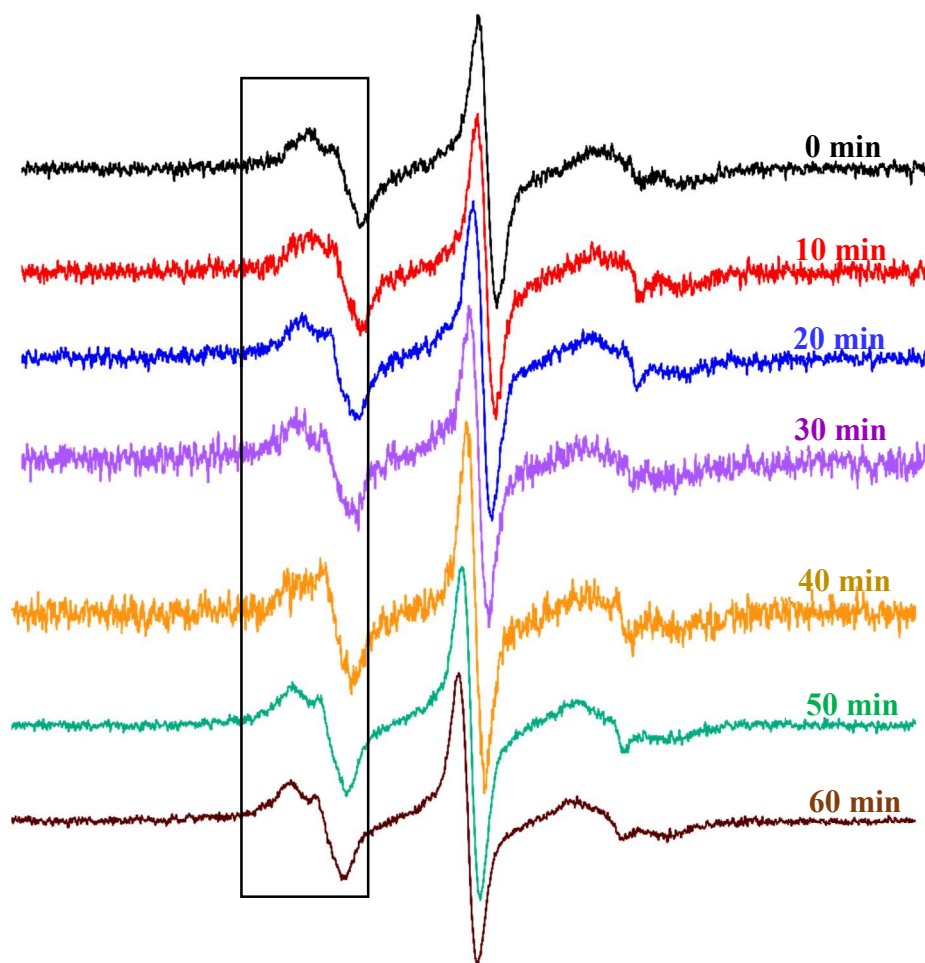


Figure 3.8. Normalized CW EPR spectra of T188R1 BtuB in live *E. coli* cells resuspended in 100 mM HEPES pH 7.0 and incubated on ice for 0, 10, 20, 30, 40, 50, and 60 min time points during the wash step.

We screened conditions in order to maintain cell viability throughout the entire process, as well as, to eliminate any residual ascorbate, or any other reducing agents released, to minimize

the spin reduction. Thus, after spin labeling, the harvested cell pellets were resuspended in 100 mM HEPES pH 7.0 to remove residual reducing agents, as well as, to optimize the removal of any leftover unreacted MTSSL label. When the cells (either in pellets or in resuspension) are incubated on ice, cell viability is maintained more than that of RT, because on ice the metabolic rate is slowed. Therefore, the resuspended cells were incubated on ice over a range of time points (0-60 min) in order to identify the sufficient ice incubation time period(s). At the end of the on-ice incubation, the pelleted cells were analyzed with CW EPR. Based on the EPR spectra (Figure 3.8), each sample was similarly spin labeled and also similar to the previously reported T188R1 BtuB site with  $\text{Ca}^{2+}$  and the substrate bound. Furthermore, each spectrum (Figure 3.8) is narrow due to T188R1 spin mobility on the BtuB loop two. The mobile component is indicated by the box in Figure 3.8 which consists of two different spin dynamic populations. Moreover, these lineshapes fluctuated over the time course for ice incubation. Thus, in the live cells on ice, BtuB loop two is actively moving with two different spin mobilities without reducing the T188R1.

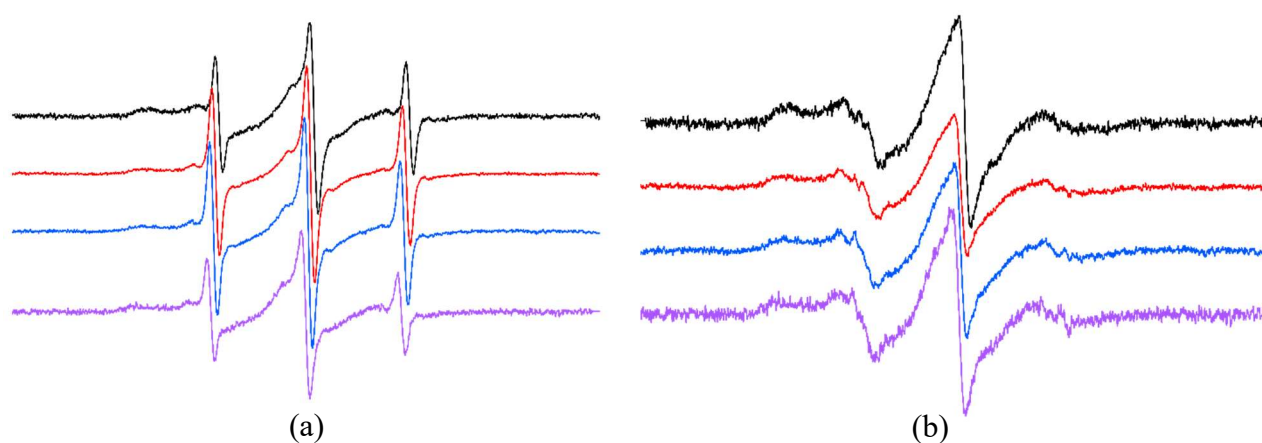


Figure 3.9. Normalized CW EPR spectra of spin labeled T188R1 BtuB in live cell pellets incubated on ice (a) before and (b) after free-spin subtraction in the repeat experiment using FreeBee software by Christian Altenbach. Resuspended spin labeled cells were incubated on ice for 0, 20, 30, and 40 min followed by analysis by CW EPR.

Surprisingly, when this experiment was repeated, the recorded EPR spectra (Figure 3.9) were similar to the previously reported apo BtuB (Figure 3.1a)<sup>5</sup>. Two different EPR lineshapes obtained in each experiment (resembling reported BtuB apo with and without  $\text{Ca}^{2+}$ ) which could be due to different bacterial physiological processes. In live cells, many physiological processes can influence the spin labeling process, as well as, the spin dynamics on the labeled residue site. For example, the OM composition and the BtuB's spatial organization (Figure 3.10) within the OM (discussed in detail in Chapter 4).

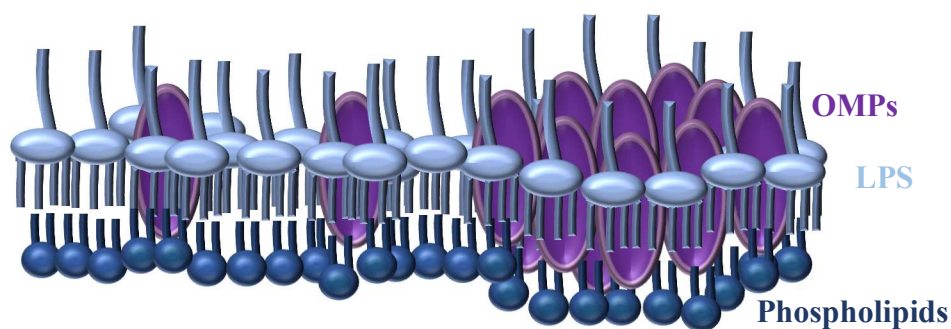


Figure 3.10. Schematic representation of **OMPs** clustering in OMP island with heterogeneously distributed **LPS** (in the outer leaf-let) within the islands.

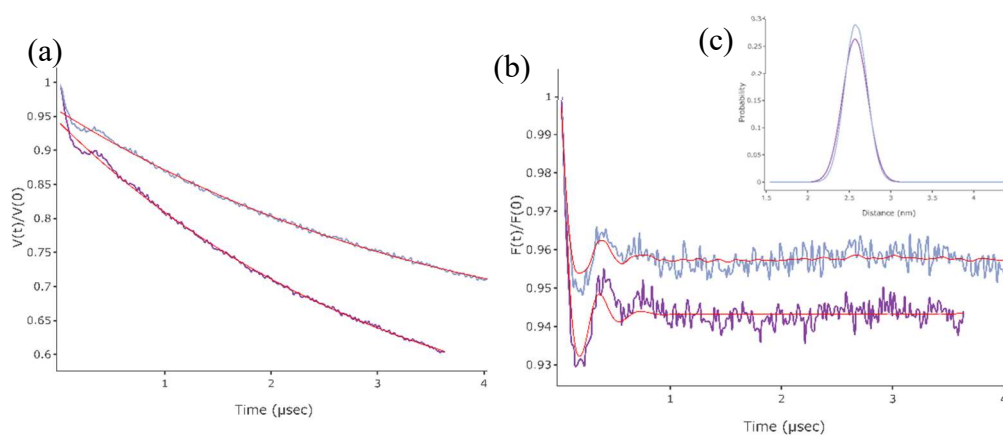


Figure 3.11. Echo signals obtained from the pulsed EPR experiment for ice incubated (20 min and 60 min) cells expressing T188R1 BtuB. (a) The echo signals collected from V90R1-T188R1 BtuB DEER experiment, (b) the form factor after subtracting the background and (c) the resultant normalized distance distribution.

The samples which had EPR spectra for T188R1 BtuB similar to ones with  $\text{Ca}^{2+}$  were used in the pulsed EPR experiment. Based on the echo signals for DEER experiments, the 20 min and 60 min ice incubated cell samples had ~6 and 8 % modulation depths, respectively (Figure 3.11). These were the first higher modulation depths ( $\leq 6$  % modulation depth) observed. In both samples, the short-range distance distribution (between 2-3 nm) corresponded to the T188R1 and the TEMPO- $\text{B}_{12}$  within the BtuB binding pocket. Therefore, for the future live cells labeling, spin labeled cells were incubated on ice for an hour.

$\text{B}_{12}$  transport by cells expressing native BtuB in the late log or early stationary phase of the growth curve was studied a few decades ago. Based on those uptake assays,  $\text{B}_{12}$  uptake is biphasic with a fast initial phase followed by a slow secondary phase<sup>20</sup>. Furthermore, both the initial and the secondary uptake rates of radiolabeled- $\text{B}_{12}$  were significantly reduced at lower temperatures. The maximum uptake rates at  $15^\circ \text{C}$  were reduced by about 25 and 75%,

respectively. At a pH below the optimum pH (around 6-6.6) a significant initial uptake rate drop was also observed (about 50% reduction at pH 5.5) and the uptake rate was diminished at pH 5.0 and below<sup>20</sup>. However, the fast-initial phase could be the binding of B<sub>12</sub> instead of transport.

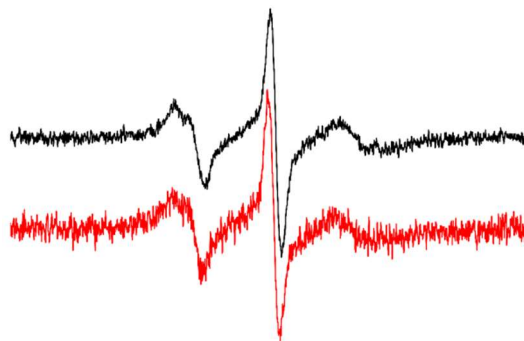


Figure 3.12. Normalized CW EPR spectra of T188R1 BtuB in live cells incubated in different wash buffers (**MES pH 5.5** and **HEPES pH 7.0**) for an hour on ice prior to collecting cells for CW-EPR analysis.

Consequently, sample processing was again optimized so the spin labeled cells were resuspended in lower pH (5.5) instead of pH 7.0 during the wash step. At the lower pH, the pmf in the IM is altered. Reduction of pH leads to activate acid responses (ARs) to maintain the cytoplasmic pH homeostasis. Thus, *E. coli* (neutrophilic bacteria) efflux protons, activate strategies to consume protons and decrease the expression of ATPase synthase up taking protons<sup>21-23</sup>. This can affect the TonB complex in the IM and ultimately inhibit or reduce the uptake of bound TEMPO-B<sub>12</sub>. Thus enabeling a significant population of bound TEMPO-B<sub>12</sub> in the BtuB binding pocket without uptaking. Based on the normalized CW EPR spectra, at both pH 5.5 and 7.0, the efficiency was the same while the EPR spectra were similar to the Ca<sup>2+</sup> bound apo BtuB EPR lineshape (Figure 3.12). Furthermore, in DEER signals, the modulation

depth for the pH 5.5 sample was around 8 % (Figure 3.14). Thus, we used this new, optimized protocol to spin label the other residues (A288C and D492C).

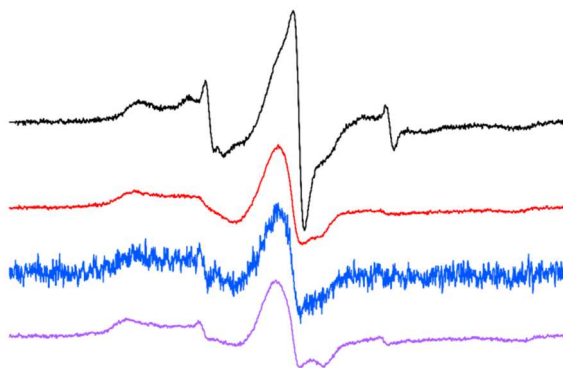


Figure 3.13. Normalized CW EPR spectra of **L8P-T188R1**, **A288R1**, **L8P-A288R1**, and **D492R1** BtuB after incubation in 100 mM MES buffer, pH 5.5 for an hour on ice.

Based on the normalized CW EPR spectra (Figure 3.13), the L8P-T188R1 BtuB EPR spectrum is a composite of mobile and immobile components, while the other three spin label mutants have broad and low amplitude lineshapes, due to the restricted spin mobility at each residue site (A288 and D492). Even with our finalized single cysteine spin labeling protocol, we were unable to obtain good a DEER signal for the D492R1 site on BtuB. The DEER data collected for the T188R1 and the A288R1 sites with and without L8P, the modulation depths were about 5 % for both L8P-T188R1 and A288R1, while 3 % for L8P-A288R1 (Figure 3.14). These comparative low modulations depths within the current optimized protocol are potentially due to inefficiencies of the BtuB residue labeling perhaps due to steric hindrance at the labeled site. Thus, it is important to select residue sites with accessible rotamers. For example, T188R1 has a rotamer population pointing upward while rotamers for others are pointing away from the barrel to the sides. These rotamer positions are likely, hard to spin labeled in the cells where the OMPs and LPS can be densely packed within islands.

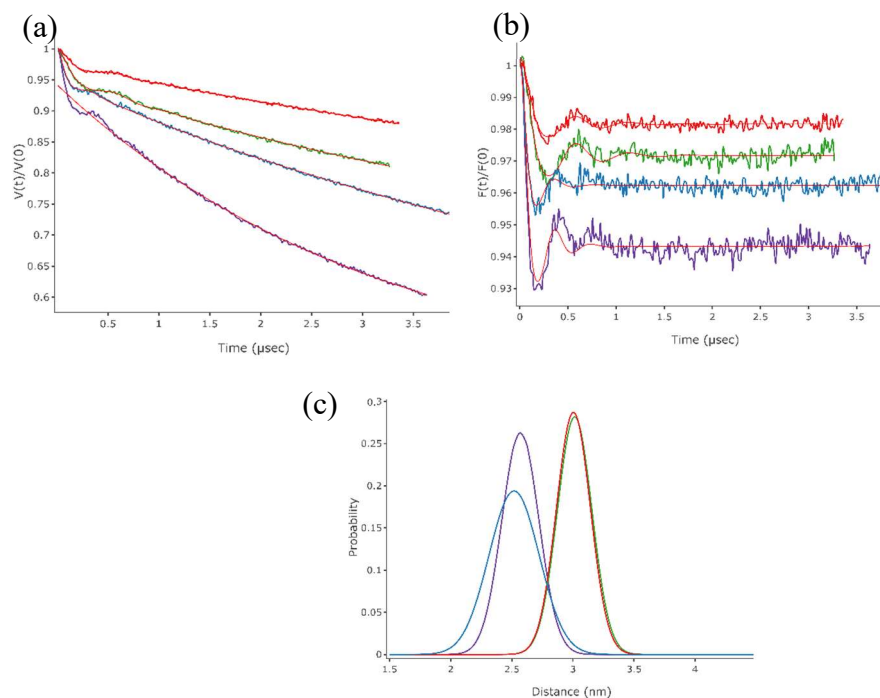


Figure 3.14. DEER data for TEMPO-B<sub>12</sub> bound to **T188R1**, **L8P-T188R1**, **A288R1**, and **L8P-A288R1** BtuB in live cells. (a) The echo signals collected from each DEER experiment, (b) the form factors after subtracting the background and (c) the resultant normalized probability distance distribution.

Furthermore, modulation depths for both the L8P-T188R1 and the L8P-A288R1 are lower than the T188R1 and the A288R1. This could be potentially the effects of expressing transport defective mutants in RK5016 cells. However, the resultant normalized short distance distributions, at  $\sim 2.5$  Å and  $\sim 3.0$  Å, were identical for both with and without L8P mutants with T188R1 and A288R1 BtuB, respectively. Thus, there is no significant difference in the position of the substrate within the binding pocket in the transport defective mutant and the wild type BtuB.

Since our protocol was optimized for the cells in the mid-exponential phase, we also studied the possibility of efficient spin labeling at the early-log phase using the T188C BtuB site. As

explained above (also in Chapter 1.2.2), even in the mid-log phase there is a percentage of cells which are older than the other cells due to continuous accumulation of old poles. Thus, we wanted to harvest cells that are a few generations younger than what we collect at the mid-log phase, but within the log phase. Once cells are inoculated in the main culture, they first enter the lag phase and then the early-log phase ( $OD_{600} \sim 0.3$ ). Therefore, we modified our method so that after the initial pre-culture inoculation, cells will undergo at least three cell cycles to reach early-log phase, then they will be harvested and spin labeled. According to EPR spectra, at both log phases, similar efficiencies were observed for T188C BtuB (Figure 3.15.a). Thus, early log phase cells were used in the remaining experiments.

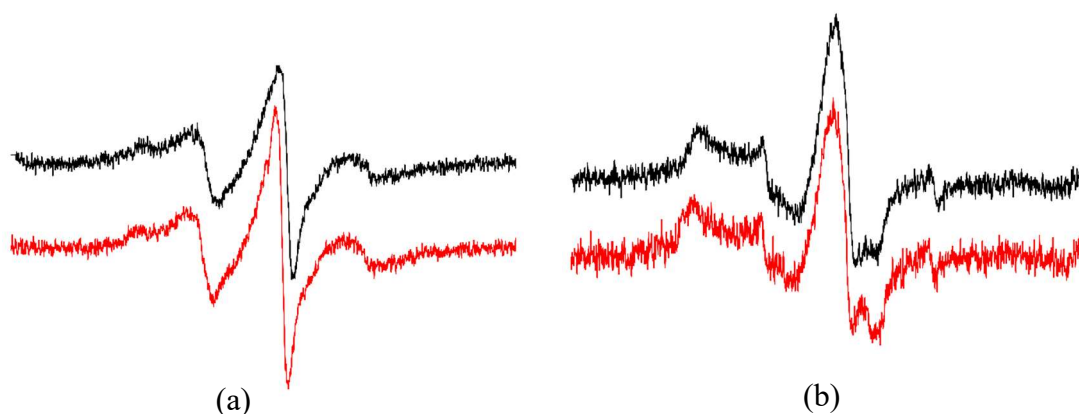


Figure 3.15 Normalized CW EPR spectra of (a) T188C BtuB cells collected at OD **0.3** and **0.6** (b) **S74C** and **V90C** BtuB expressing cells collected at OD 0.3 and spin labeled using the finalized protocol for the single residue labeling in live cells.

In addition to spin labeling we also focused on spin labeling BtuB core residues which are facing the extracellular side. This way we can either use our modified method as a common method to spin label pre-selected barrel and core residues within live cells. First, we attempted to spin label S74C and V90C within the core (Figure 3.15.b). Cells at early log phase ( $OD_{600}$  around 0.3) for V90C and S74C BtuB were spin labeled and analyzed. The corresponding

spectra were a combination of both mobile and immobile components at each residue site with restricted mobile population, probably due to the limited dynamics at the core of BtuB.

Aside from the optimization of the single residue spin labeling of the extracellular loops and the core of BtuB within live cells and investigating substrate binding within the binding pocket of transport defective BtuB, we also focused on a potential hierarchical organization of OMPs into domains using single spin labeled BtuB (which is the focus in Chapter 4). In the finalized version of the single residue labeling in live healthy cells, the main cultures were inoculated with pre-cultures with appropriate initial cell densities to allow the inoculated cells to undergo a minimum of three cell doublings to reach an OD<sub>600</sub> of 0.3. Harvested cells were spin labeled with MTSSL ( $\leq 1$  mg per cells harvested from 100 mL culture aliquot) in 100 mM HEPES pH 7.0 at RT for an hour with gentle mixing, followed by resuspending the pelleted spin labeled cells in 100 mM MES pH 5.5 and incubating on ice for an hour. Finally, cells were pelleted and analyzed by CW EPR and DEER samples were prepared.

As we were able to spin label a couple of core residues, we then moved to the second phase of the spin labeling live cells project in order to improve our finalized method for spin labeling two residues between the BtuB core and the barrel (T188C-S74C and T188C-V90C BtuB). Thus, the successful outcome will be used in future studies for the investigation of potential intermediate states of BtuB associated with the substrate transport process.

### 3.4.2. Method optimization of cysteine double mutants spin labeling

Even though, the protocol was optimized for spin labeling single cysteine BtuB mutants S74C, V90 and T188C (Figure 3.16) expressed on RK5016 cells, it needed to be further optimized for double cysteine BtuB mutants (S74C-T188C and V90C-T188C) as the level of expression of BtuB double mutants in RK5016 cells, as well as, spin labeling efficiencies are potentially differed than that of the single BtuB mutants.

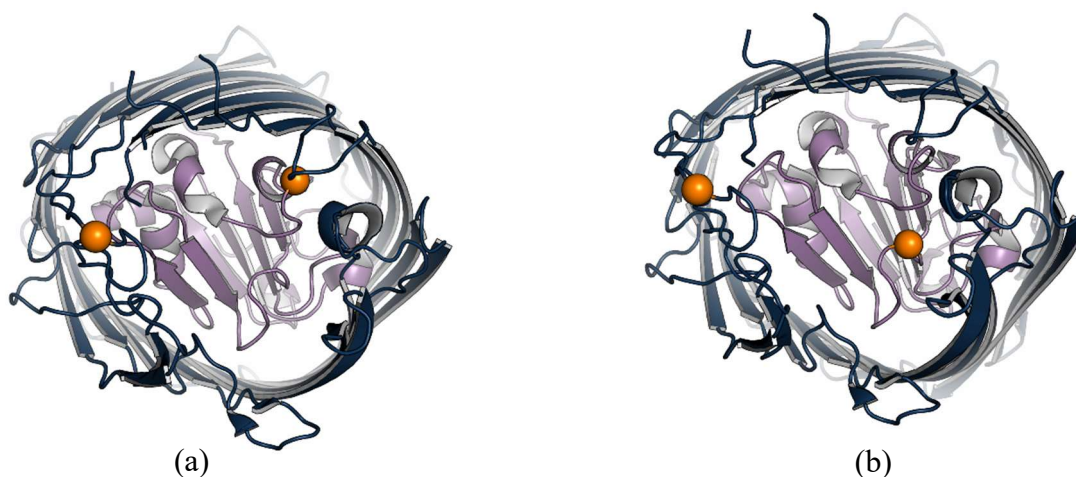


Figure 3.16. Top view of apo BtuB (1NQG) structure showing (a) V90C-T188C and (b) S74C-T188 residue sites used for spin labeling in live *E. coli* cells.

#### 3.4.2 (a) Optimization of spin label, MTSSL

During MTSSL labeling step, disulfide bonds are formed between the cysteine residues and MTSSL and also slowly between the spin labels (intramolecular). Thus, to identify the appropriate amount(s) of spin label required for significantly labeling the cysteine double mutants, cells were incubated with series of MTSSL (0.1-3 mg) concentrations. Based on the EPR spectra (Figure 3.17), at the lower levels of MTSSL label (0.2 and 0.1 mg of label for cells harvested from 100 mL cell culture) spin labeling efficiency was low and the noise was high. However, there was a very small fraction that was spin labeled. In contrast, with higher levels of

MTSSL (2 and 3 mg per cells harvested from 100 mL cell culture), the extra label did not improve the labeling efficiency. The spectra were narrow high amplitude due to small fraction of unreacted free spin leftover even after the washing step with some background labeling (Figure 3.17). However, the individual cysteine mutants (V90C and T188C BtuB) were efficiently spin labeled with 1 mg of MTSSL.

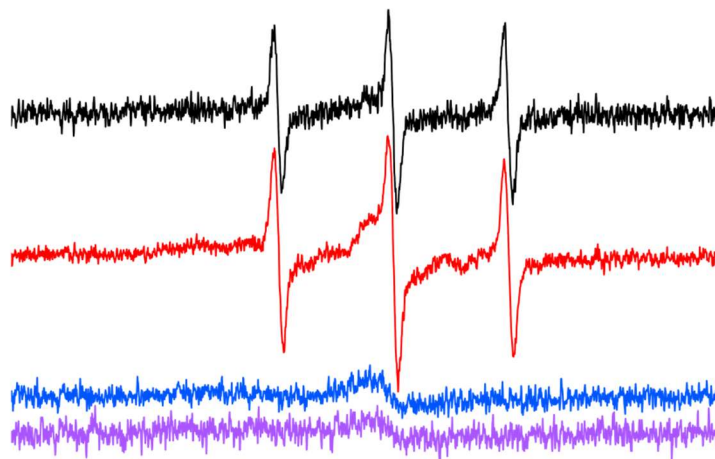


Figure 3.17. Normalized CW EPR spectra of RK5016 cells expressing V90C-T188C-BtuB spin labeled with **3 mg**, **2 mg**, **0.2 mg** and **0.1 mg** of MTSSL per cells harvested from 100 mL of culture at OD600 0.3.

Thus, the next logical question to address was whether the introduced cysteine mutants within the barrel at extracellular loop 2 and the core domains might undergo intramolecular disulfide bond formation. Analyzing the crystal structure 1NQG, in which the BtuB is bound with calcium ions, loop 2 is ordered, it is structurally highly unlikely that disulfide bonds form between both the core domains mutants (S74C and V90C) and the loop 2 mutant (T188C). The two distances between T188 to S74 or V90 are around 3 nm, and based on the BtuB overall surface, loop 2 which has T188, cannot move into proximity to form an intramolecular disulfide bonds with the core (Figure 3.18).

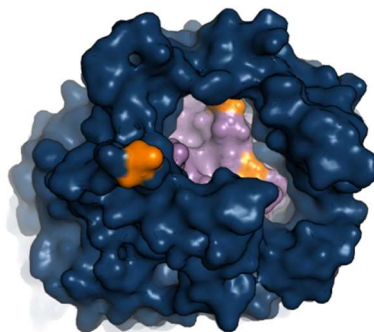


Figure 3.18. Top view of BtuB (1NQG) surface. The **T188C** accessibility to the BtuB **core** is limited by the presence of other **extracellular loops**' dynamics within the **barrel**.

But there is the potential that during the folding of BtuB mutants, at the BAM complex, the core domain that is folded in the periplasm could form cross-linking with the extracellular loops' cysteine residue (Chapter 1.4.6). Once the  $\beta$  strands 3 and 4 of BtuB formed the hydrogen bonds, while loop 2 is exposed to the extracellular side, the loop 2 and core dynamics could enable the formation of a core-loop intramolecular disulfide bond. Thus, we first attempted to reduce the disulfide bonds prior to the spin labeling step.

### 3.4.2(b) intramolecular disulfide bond reduction via DTT treatment of cells

We first attempted to reduce the disulfide bonds which have already formed between the barrel and the core using DTT. High level of DTT was used to efficiently reduce the disulfide bonds in a short period of time, as longer incubations would kill the cells. However, based on the CW-EPR spectra for the control samples (WT BtuB), when the cells were incubated with excess amounts of MTSSL (3 and 5 mg), a small fraction of non-specific labeling occurs. The corresponding EPR spectra are a combination of mobile and immobile components of nitroxide spin dynamics, thus, most likely trace amounts of free spin are also left during the washings which produce the mobile component and the immobile component is due to the small

population of background labeling. However, when the DTT treated control samples were labeled with MTSSL, increased populations of immobile and mobile components appeared (Figure 3.19). Thus, DTT was either likely adsorbed to the OM or some traces were left within the cell pellets and being labeled by the MTSSL, consequently, giving higher background labeling. Another possibility is during the DTT treatment some of the cells were lysed and the cellular debris is also being labeled.

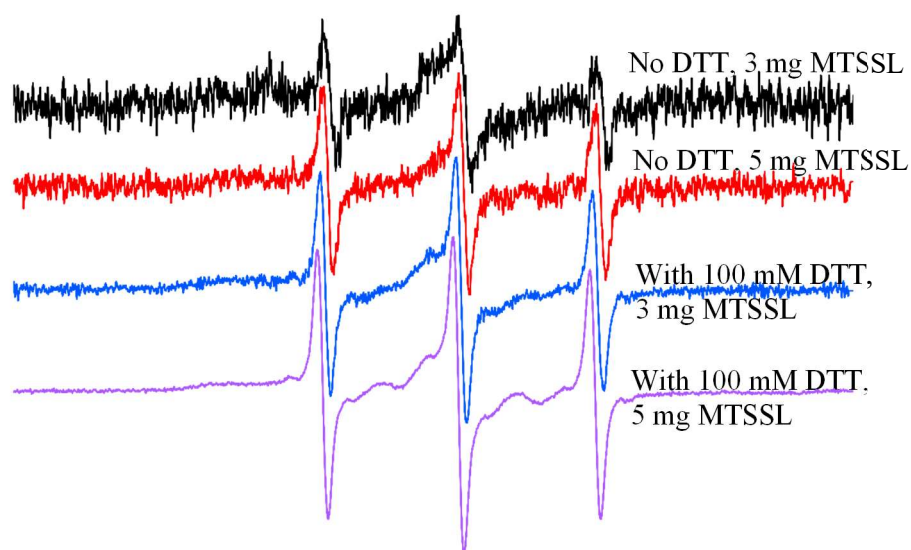


Figure 3.19 The normalized CW EPR spectra of WT BtuB treated with 100 mM DTT and spin labeled using excess MTSSL (3 and 5 mg). No DTT cells were incubated with **3 mg** and **5 mg** MTSSL and cells treated with DTT were spin labeled with **3 mg** and **5 mg** of MTSSL.

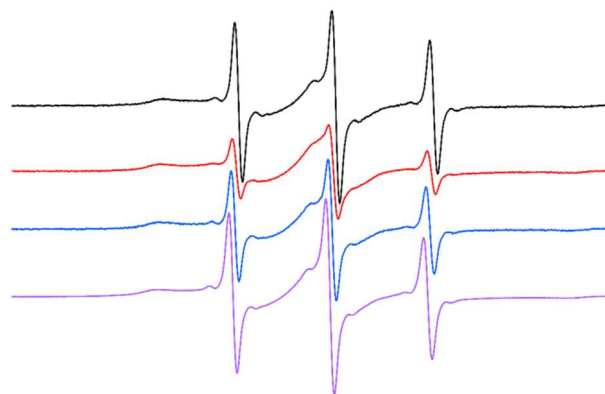


Figure 3.20. Normalized CW EPR of V90C-T188C BtuB cells treated with 100 mM DTT prior to spin labeling. Cells at **OD<sub>600</sub> 0.3 no DTT**, **OD<sub>600</sub> 0.3 with DTT**, cells at **OD<sub>600</sub> 0.6 no DTT**, and **OD<sub>600</sub> 0.6 with DTT**.

With respect to the WT BtuB control samples, V90C-T188C BtuB had higher levels of spin labeling within the EPR spectra with a large immobile population and a mobile component (Figure 3.20). However, the DEER samples prepared from these cells did not produce a significant DEER signal. Thus, the attempts of reducing the potential disulfide bonds on BtuB in the OM were unsuccessful. This could be due to the fact that DTT cannot access the bond to reduce it after it is formed. Hence, the original CW EPR analyzed samples with spin labeling are likely caused by leftover, OM bound DTT, and higher levels of off-target labeling in the absence of the BtuB cysteines for labeling.

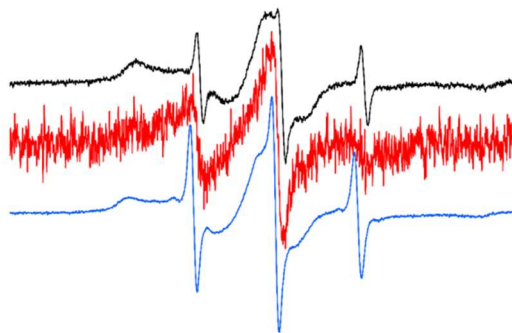


Figure 3.21. Normalized CW EPR spectra of 10 mM DTT constant and incremental increase and reduction in culture (**control- no DTT**, DTT supplemented as **fed-batch addition**, and **batch culture**) with cells expressing S74C-T188C BtuB.

Parallel set of cell cultures were grown in fed-batch or batch cultures with DTT to minimize disulfide bond formation during the folding at the BAM complex or after folding BtuB. The cells incubated with 20 and 30 mM DTT in the batch cultures did not grow due to the high levels of DTT, above the toxic levels for *E. coli*<sup>24</sup>. Based on the CW EPR spectra (Figure 3.21), in the no DTT treated sample, when the cells were incubated with excess amounts of MTSSL, off-target labeling occurs which results in the high amplitude broad spectrum with relatively small population of mobile component due to lack of targeted labeling sites on BtuB. In contrast, when the cells were grown in fed-batch cultures, the resultant spectrum is low in amplitude and noisy due to the low levels of potential cysteine labeling sites in BtuB. When compared to fed-batch cultures grown in the presence of incremental increase of DTT, up to 10 mM, the samples prepared from batch cultures with 10 mM DTT, had high amplitude spectra with both mobile and immobile components. Again, no significant echo signal appeared in the DEER experiment. This high level of spin labeling for cells grown in the batch cultures is then most likely due to DTT

adsorbed to OM throughout the cell growth and also due to enhanced off target labeling which was not seen in the cells grown in fed-batch culture when the DTT was increased along with increasing cell density. However, we had no success with either method of reducing BtuB core-loop disulfide bonds.

Hence, we revised the spin labeling procedure and screened a range of labeling conditions using 0.5 and 1 mg MTSSL instead of the amounts used in the aforementioned experiments. Our main goal was to minimize the off-target labeling by reducing the spin labeling time and lowering the incubation temperature at high pH, thereby, enhancing the specific labeling of S74C-T188C BtuB. Based on the CW-EPR spectra all the analyzed samples contain a combination of high amplitude narrow and broad low amplitude components due to the presence of both fast moving and slow-moving spin label in the samples. However, the population of mobile component reduces as the spin labeling time increases (Figure 3.22a and Figure 3.23a). Thus, over the time MTSSL reacts with more cysteine sites in the samples. During the washing step, unreacted MTSSL is significantly removed from the processed samples. Thus, the mobile population was more reduced in the spectra than the corresponding pre-washed samples' spectra (Figure 3.22 and Figure 3.23) with the exception of 40 min incubated samples in both RT and on ice spin labeling. This could be due to sample handling error during the washing step. Based on the post-washed samples, at both RT and on ice, cells were efficiently spin labeled. A significant population of immobile component, similar to the previous DTT reduction experiments' samples CW EPR spectra also appeared, due to non-specific spin labeling.

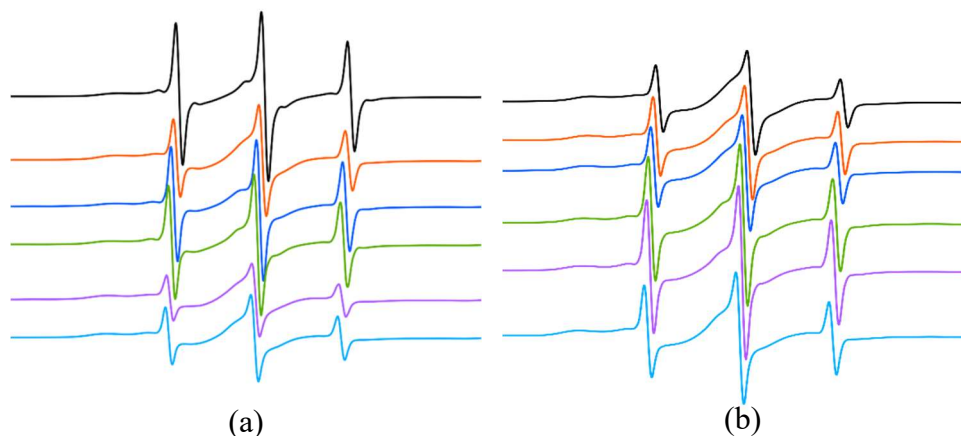


Figure 3.22. Normalized CW EPR spectra of S74C-T188C BtuB of live cells spin labeled on ice (a) before and (b) after incubation in wash buffer on ice for an hour. Cells were incubated with 1 mg of MTSSL for 20, 40, and 60 min and with 0.5 mg MTSSL for 20, 40, and 60 min.

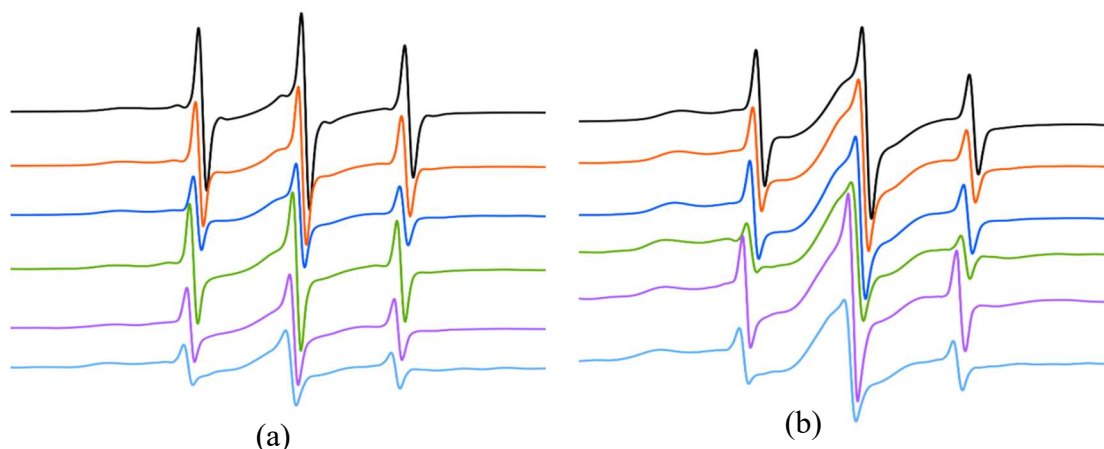


Figure 3.23. Normalized CW EPR spectra of S74C-T188C BtuB of live cells spin labeled at RT (a) before and (b) after incubation in wash buffer on ice for an hour. Cells were incubated with 1 mg of MTSSL for 20, 40, and 60 min and with 0.5 mg MTSSL for 20, 40, and 60 min.

Thus, overall, we were still unable to spin label double cysteine BtuB mutants between the core and the extracellular loops. In contrast, two cysteine mutants on the extracellular loop sites (T188C of loop two and G399C of loop seven) BtuB were spin labeled and analyzed using the

original protocol and distance distributions were obtained by pulsed EPR<sup>6</sup>. Therefore, it is highly likely that one can easily spin label two extracellular sites which are on the loops, probably within a certain loop dynamic range which would prevent the formation of disulfide bonds. Analyzing the  $\text{Ca}^{2+}$  bound apo BtuB crystal structure (1NQG) in which the loop two is ordered due to  $\text{Ca}^{2+}$ , both V90 and G399 residues are relatively in close proximity (Figure 3.24), thus it was first puzzling why we cannot spin label two sites between the loops and hatch domains.

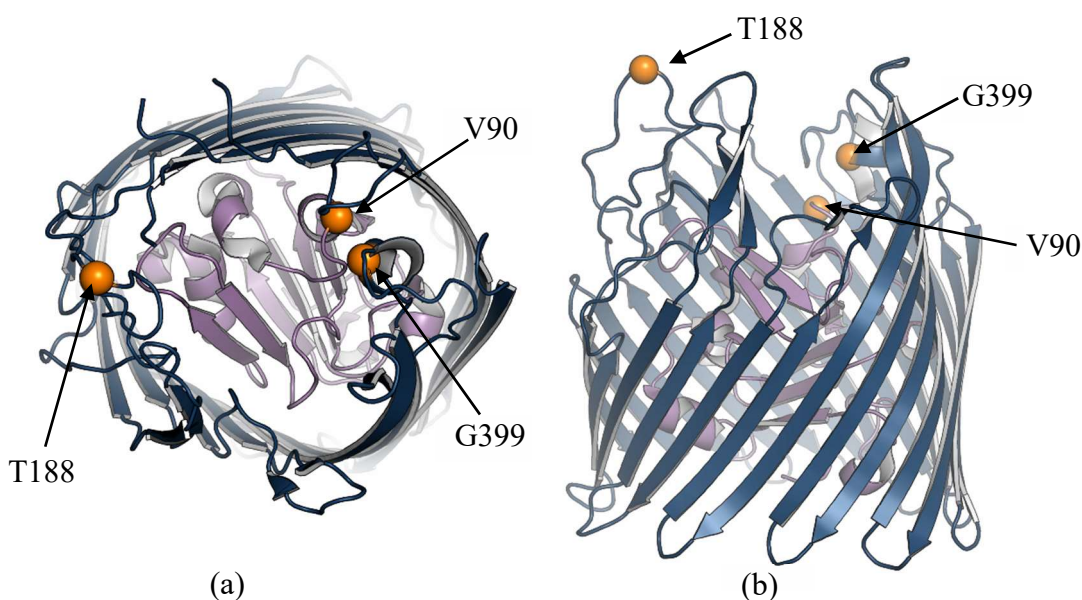


Figure 3.24. Structure of apo BtuB (1NQG) (a) top view and (b) side view showing residues **V90** on **core** and **T188** and **G399** on the **barrel**.

Next, we considered potential impacts of the key players both in the periplasm and the OM, from the entrance of the nascent OMP precursors, through the SecYEG channel to the nascent folded OMP inserted by BAM complex<sup>25, 26</sup>. In order to successfully spin label between the BtuB core and loop residues (also potentially for other TBDTs) the disulfide bond formation during any part of the journey of OMP precursors should be then considered. The disulfide bonds within

OMP can potentially form during BAM folding, but if this is the case, it will be difficult to overcome using a BAM knockout strain, as this would impact the entire OMP biogenesis and bacterial physiology.

Instead of the BAM complex, the bacterial disulfide bond system (Dsb) in the periplasm (described in the Chapter 1.4.7), is the key element which regulates disulfide bond formation in these proteins. Since Dsb system contains four main proteins (DsbA, DsbB, DsbC and DsbD) that are involved in oxidative and isomerization pathways, our main interest was to identify a potential Dsb mutant (knockout) strain to use for future studies. Different Dsb mutant *E. coli* strains are constructed and have been studied mainly by Professor J. Beckwith, Harvard Medical School, MA<sup>27, 28</sup>. Based on the studies, Dsb knockouts are viable and have different genotypes. For example, strain RI89 is wild type, strain RI90 is *dsbA*<sup>-</sup>, strain RI179 is *dsbC*<sup>-</sup>, and strain RI317 is *dsbB*<sup>-</sup>.

DsbA and DsbB are involved in the oxidative pathway while DsbC is involved in isomerization pathway (Chapter 1.4.7)<sup>29</sup>. Thus, we utilized Dsb mutant strains, *dsbA*<sup>-</sup> and *dsbB*<sup>-</sup> to prevent the formation of intramolecular disulfide bonds between S74C-T188C BtuB or in V90C-T188C BtuB. In the current BtuB constructs only contain two cysteines, so there is no mismatched multiple disulfide bond formation. So, *dsbC*<sup>-</sup> and Dsb WT (here after referred as WT) strains were employed as controls in our studies. However, *dsbC*<sup>-</sup> will be highly useful for studying OMPs with more than two cysteine residues.

### 3.4.2(c) Spin labeling BtuB using Dsb knockouts

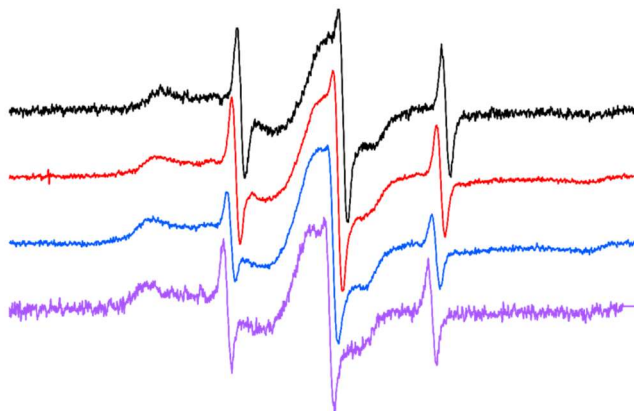


Figure 3.25. Normalized CW EPR spectra of V90C-T188C BtuB in live **WT**, *dsbA*<sup>-</sup>, *dsbB*<sup>-</sup> and *dsbC*<sup>-</sup> cells, spin labeled with MTSSL.

An examination of the CW EPR spectra collected for V90C-T188C BtuB expressed in each cell line (Figure 3.25) indicated that we were still labeling non-specific targets. Since, the disulfide bond formation is impacted either at the oxidation or the isomerization pathways, the knockout cells endogenously create the potential to spin label all native cysteines present in other proteins (mainly in the OM and periplasm) in addition to the BtuB. Thus, potentially leading to background labeling in each knockout cell line. But the WT also had a similar CW EPR spectrum. Therefore, the current labeling conditions were indeed still producing background labeling, potentially in addition to labeling the cysteine double mutants of BtuB.

### 3.4.2(d) Optimization of spin labeling BtuB in Dsb knockouts

Our main goal was to use live healthy bacterial cells; therefore, we had to focus on a significant fraction of bacterial physiology in order to understand how to develop an effective method to spin label healthy cells, minimizing physiological alterations and obtain the required data from EPR. Overdosing the cells with more and more spin labels in order to label the target

sites on BtuB, also negatively impacted the bacterial reduction potentials, probably in both the periplasm and the cytoplasm, which could provoke the bacterial SOS systems controlled by global regulators (Chapter 1.2.4).

During each experiment, the required amounts of spin label depend on the expression levels of the BtuB cysteine mutants. But, it is more convenient to use a set amount of label during each cell culture experiments, rather than trying to quantify the level of BtuB in each sample prior to the spin labeling step. More specifically, we need to have sufficient quantities to only spin label OMPs, without impacting the periplasm and its reduction potential. Additionally, to enable the specific labeling of over-expressed BtuB cysteine mutants and to minimize off target OMP labeling from the Dsb knockout systems. In Dsb knockout strains, the level of over-expressed BtuB (in minimal media) is the result of both chromosomal *btuB* (native WT BtuB) and plasmid *btuB* (non-native BtuB) either encoding WT or a mutant BtuB. In *E. coli*, there are only about 200-300 copies of native BtuB per cell<sup>12</sup>, while the non-native will be expressed at about 1000-fold higher levels<sup>30</sup>. But, the expression levels of the cysteine double mutant within BtuB in the Dsb knockouts could also be different than the non-native WT BtuB or the single cysteine mutant BtuB expression levels.

Interestingly, we were able to successfully use 1 mg of MTSSL during the single cysteine BtuB mutants labeling without issues with off target labeling, while the same amount of MTSSL resulted in off target labeling for the cysteine double mutant labeling. Thus, *dsbA*<sup>-</sup> cells grown until OD<sub>600</sub> 0.3, were spin labeled with a range of MTSSL (0.05, 0.1, 0.2 and 0.5 mg) and analyzed by CW-EPR (Figure 3.26). These label quantities were used in terms of MTSSL per cells pelleted from 50 mL of culture media at OD<sub>600</sub> around 0.3.

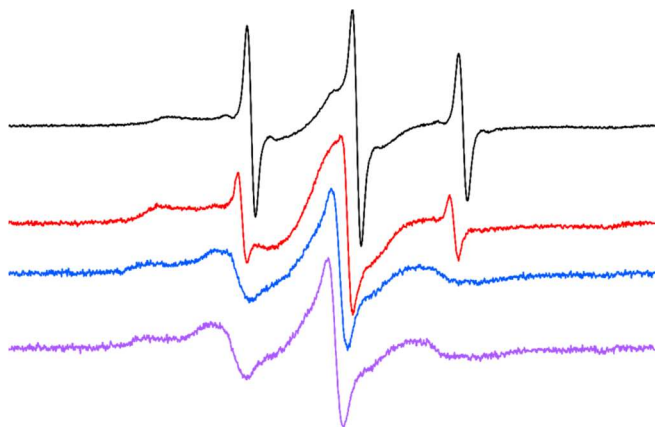


Figure 3.26. Normalized CW EPR of V90C-T188C BtuB spin labeling using **0.5**, **0.25**, **0.1** and **0.05** mg of MTSSL per *dsbA*<sup>-</sup> cell pellet obtained from a 50 mL cell culture at OD<sub>600</sub> around 0.3.

According to the CW EPR spectra, samples were spin labeled, even at the small quantities of MTSSL. At the higher levels of label (0.5 and 0.25 mg), with respect to the above-mentioned cell density, leading to off-target labeling (Figure 3.26) and the EPR spectra were similar to the previous spectra obtained with 1 mg or more spin label (Figure 3.25). In contrast to the higher quantities, when the label added was low (0.1 and 0.05 mg), the CW EPR spectra of V90R1-T188R1 BtuB obtained were a combination of T188R1 and V90R1 BtuB CW EPR lineshapes, more toward the lineshape of the T188R1 BtuB site. Thus, we were finally able to specifically spin label BtuB cysteine double mutants in the core and the extracellular loop!

As explained previously, in this labeling protocol, we cannot simultaneously quantify the levels of BtuB in order to add the optimal amount of spin label. Hence for this assay, it is more convenient to use less spin label, but enough to obtain sufficient labeling, so that we can use the samples for DEER experiments, but still avoid excess label, causing off target labeling. Therefore, we selected 0.05 and 0.1 mg of MTSSL per cells harvested from a 50 mL of cell

culture at OD<sub>600</sub> around 0.3. The quantities were adjusted to maintain equivalent amounts of MTSSL for given cell density.

The spin labeling of the RK5016 cells expressing either S74C-T188C or V90C-T188C BtuB using 0.05 and 0.1 mg of MTSSL was inefficient (Figure 3.27). The corresponding EPR spectra was noisy and low amplitude due to slow moving nitroxide spin labels probably bound to trace amounts of expressed BtuB with free cysteines or due to background labeling. With respect to labeling of RK5016 cells, spin labeling of S74C-T188C or V90C-T188C BtuB expressed in *dsbA*<sup>-</sup> cells was significantly higher for both the 0.05 mg and 0.1 mg label additions (Figure 3.27). Based on the EPR spectra corresponding to the cells labeled with 0.1 mg MTSSL, this quantity was selected for future experiments. Proportional amounts were used with higher volumes of *dsbA*<sup>-</sup> cell pellets.

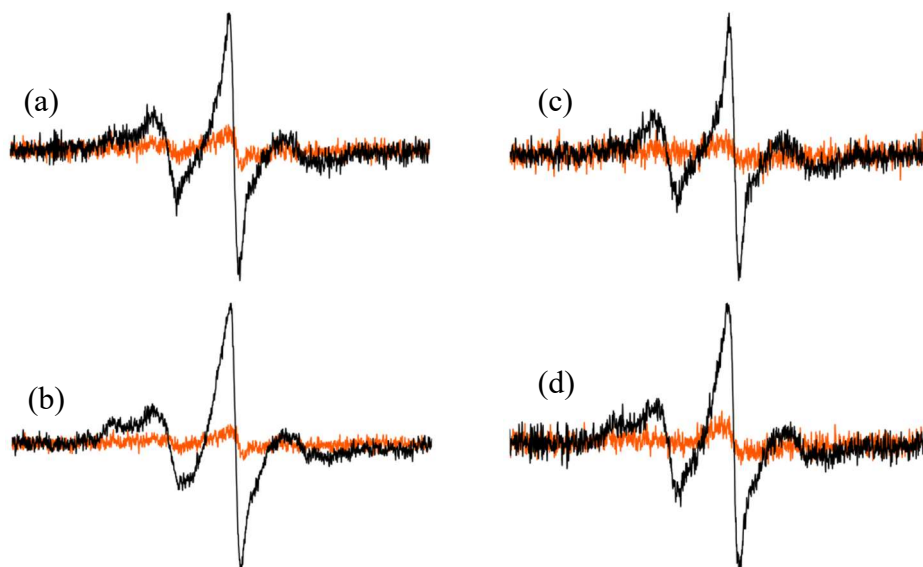


Figure 3.27. Normalized CW EPR spectra of S74R1-T188R1 BtuB spin labeled with (a) 0.05 mg and (b) 0.1 mg MTSSL and V90R1-T188R1 BtuB spin labeled with (c) 0.05 mg and (d) 0.1 mg MTSSL that were expressed in *dsbA*<sup>-</sup> and **RK5016** cells.

The spin labeling of the S74C-T188C BtuB site, was less efficient than the V90C-T188C BtuB site. This could potentially be because the S74C site is less accessible to label relative to the V90C site (Figure 3.28). Thus, for future experiments we need to screen every mutant option for the dynamics of each residue on BtuB, as they are all impacted differently in live cells; thus, these residues could be less accessible for reasons that cannot be concluded from the crystal structures and rotamer libraries.

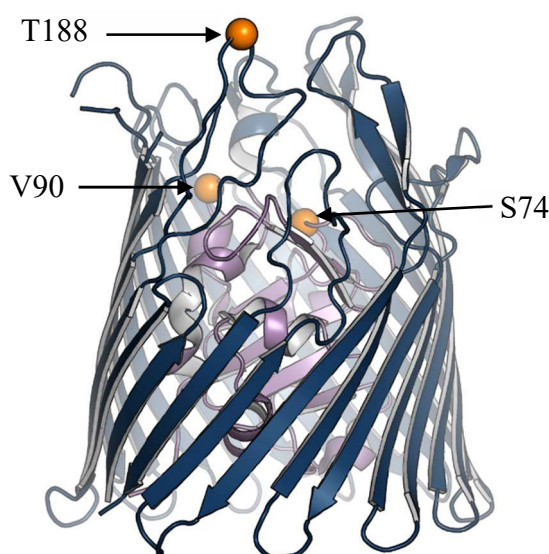


Figure 3.28. Side view of apo BtuB (1NQG) showing residues **S74**, **V90** and **T188** in the **core** and the **barrel**. **S74** residue is comparatively buried in the structure relative to the **V90** residue.

Spin labeling of Dsb mutant strain cells expressing V90C-T188C BtuB and WT BtuB (control) were also studied using the finalized method. According to CW EPR spectra, when WT BtuB is expressed in each cell line, there is a very small amount of background labeling even with 0.1 mg of MTSSL (Figure 3.29 a). The off-target labeling is comparatively more in *dsbA*<sup>-</sup> and *dsbB*<sup>-</sup> as lack of either DsbA or DsbB in these cells prevents the disulfide bond formation in other protein substrates in the periplasm. Thus, leading these unreacted cysteine residues of these

proteins to be spin labeled as well. Both *dsbC*<sup>-</sup> and Dsb WT expressing either WT or V90C-T188C BtuB had almost equal off-target labeling. Both these cell lines have active DsbA and DsbB in the periplasm and in the IM, respectively. Thus, V90C-T188C BtuB (two cysteine residues) can still form disulfide bonds. Therefore, these two cell lines have small scale off-target labeling. In contrast, both *dsbA*<sup>-</sup> and *dsbB*<sup>-</sup> expressing V90C-T188C BtuB have efficiently spin labeling (Figure 3.29) as lack of either DsbA or DsbB in the cells prevent to disulfide bond formation between V90C-T188C BtuB. Thus, the free cysteines in BtuB in these two cell lines can be easily spin labeled without disturbing the periplasmic reductive potential.

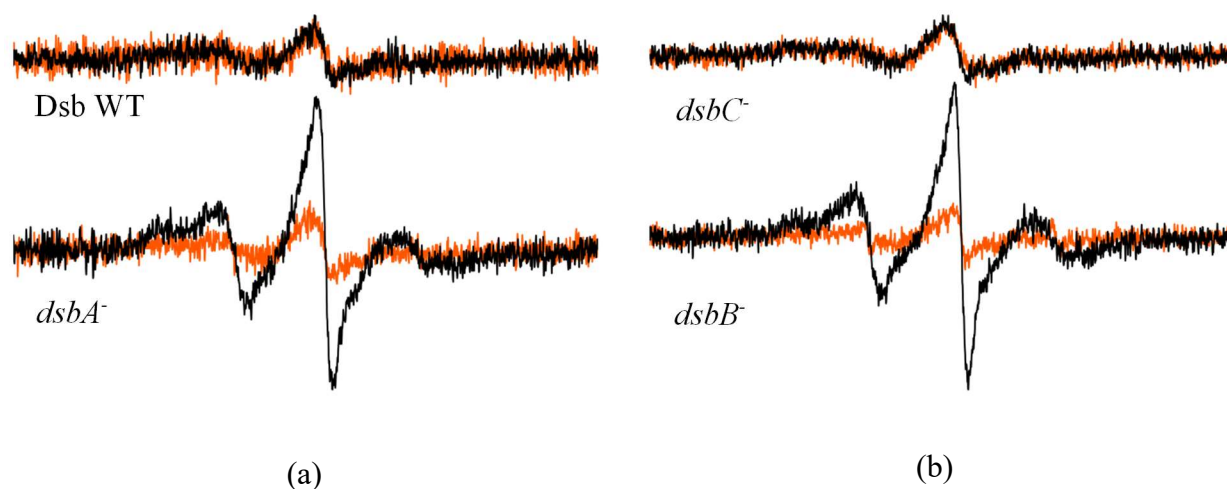
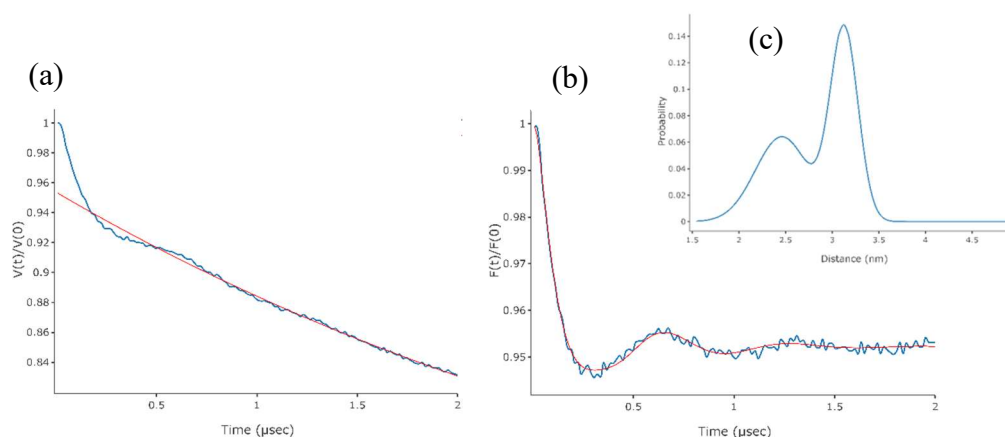


Figure 3.29. Normalized CW EPR spectra of spin labeled WT BtuB (control) and V90C-T188C BtuB using 0.1 mg MTSSL. (a) Spin labeling of BtuB in Dsb WT and in *dsbA*<sup>-</sup> (**V90R1-T188R1** and **WT**). (b) Spin labeling of BtuB in *dsbB*<sup>-</sup> and *dsbC*<sup>-</sup> (**V90R1-T188R1** and **WT**).

V90R1-T188R1 spin labeled in *dsbA*<sup>-</sup> cells was then used in the DEER experiments, for the first time we were able to get a DEER signal with a good modulation depth (Figure 3.30). Based on the resultant normalized probability distance distribution, there are two main distance populations around  $2.4 \pm 0.4$  nm and  $3.1 \pm 0.4$  nm while the crystal structure distance only appears at 3.0 nm. These distance distributions were reproducible in repeated experiments for

independently grown live cell samples. Therefore, we started to explore a potential new conformational state, resulting in a shorter distance between the loop two of BtuB barrel and the core.

At the beginning of the discussion for two cysteine residues spin labeling between the BtuB barrel and the core, it was impossible for the BtuB loop two to reach closer to the core due to other loops of BtuB occluded the core (Figure 3.18). However, this new, reproducible short distance distribution (Figure 3.30c), could potentially show an intrinsic state present in the live cells that brings this loop closer to the core (discussed in Chapter 5, exploring intrinsic state(s) project).



3.30. DEER signal of **V90R1-T188R1** BtuB. (a) The echo signals collected from V90R1-T188R1 BtuB DEER experiment, (b) the form factor after subtracting the background and (c) the resultant normalized distance distribution.

### 3.5 Applications of finalized protocols

Our current method is focused on spin labeling single and double cysteine residues exposed to the extracellular space in BtuB, a member of TBDTs. Finalized versions of the optimized protocol were successfully incorporated into different applications, which will be presented in the following chapters. Chapter 4 is focused on spatial organization of OMPs in OM using single cysteine residue labeling. Chapter 5 is focused on exploring new conformational states associated with the BtuB transporter mechanism. In contrast, Chapter 6 and Appendix 1 are focused on employing healthy active cells, processed in a similar fashion to be used to explore binding of spin labeled B<sub>12</sub> with unlabeled BtuB in order to understand the roles of conserved ion pairs in BtuB and to study the conformational changes associated with colicin E3 receptor domain binding, respectively.

### 3.6 Factors to consider for future spin labeling of OMPs in live cells

In addition to above applications, we aim to optimize the protocol for the OMPs which are cloned into plasmids that need to be induced for expression. As explained in Chapter 2, the *btuB* gene is expressed by a leaky promoter which does not need an inducer. But many OMPs encoding genes need inducers, such as IPTG or arabinose, in order to activate expression (FhuA and FecA). However, strong promoter expression, sometimes leads to OMP aggregation. There are special bacterial cell lines designed to overcome this issue. However, it will be a challenge to have these special cell lines with Dsb knockouts required to spin label cysteine double mutants. One current solution, which we will test in the future, is to grow the transformed cells, with the desired OMP encoding genes without adding the inducing agent. Even though expression is controlled, there will be a basal level of expression throughout the bacterial growth which could potentially be enough to spin label the cell and study using EPR.

Besides that, some OMPs are expressed under specific conditions because their gene expressions are regulated by cellular demand. For example, the expression of TBDTs is firmly controlled at various steps of gene expression (Chapter 1.5.1). Similarly, OmpC and OmpF expression depends on the osmolarity of the cell culture<sup>31</sup>. Therefore, the cells should be grown in main cultures which promote expression of the desired OMP and the current protocol steps (for example, pH of the buffers and salt concentrations) will need to be adjusted in order to accommodate expression of the OMP of interest.

In addition to studying *E. coli* BtuB, if other members of the TBDTs expressed in *E. coli* needed to be spin labeled in live cells, the current protocol can most likely be used with the above-mentioned modifications. However, the main difficulty will arise with respect to analyzing the spin labeled TBDTs in the substrate bound form using pulsed EPR. *E. coli* expresses TBDTs up-taking siderophore-iron complexes such as ferric coprogen and ferric enterobactin<sup>32</sup>. The iron ( $\text{Fe}^{3+}$ ) present in the complexes will reduce the phase-memory time because it enhances the relaxation rates<sup>33</sup>, thus it will be more difficult to obtain the distances between residues which are closer to the  $\text{Fe}^{3+}$  when compared to sites which are more distant from  $\text{Fe}^{3+}$ . However, some TBDTs can bind with the iron free siderophore or  $\text{Fe}^{2+}$  which can be used in such cases. Also, one can try to grow the cells in  $\text{D}_2\text{O}$  in order to increase the poor phase-memory time. Besides that, other TBDTs which bind heme and other iron bound proteins (lactoferrin or transferrin), will be able to study with a similar method to the BtuB without affecting the phase-memory time in pulsed EPR.

Even though the protocol is focused on *E. coli* cells, if the current protocol is to be used for other Gram-negative bacteria, there are some factors which need to be considered. Because, similar to studying OMPs in live *E. coli*, sometimes it is desirable to study OMPs in their native

host cells instead of in common laboratory *E. coli* strains, as their OM biogenesis is potentially different than *E. coli*. Some of these include LPS structure in the outer leaflet, the Dsb system present, required growth conditions for selected bacterial species and conditions which facilitate the expression of the desired OMP without altering the other physiological pathways.

There is a diversity of LPS in the outer leaflet<sup>34</sup> of the OM in Gram-negative bacteria (Chapter 1.3.1) because they adapt to their habits and to changing temperature. Even though the laboratory used *E. coli* strains have rough LPS, many Gram-negative bacteria, including pathogenic bacteria, also contain *O*-antigen in the LPS. When using bacterial strains with LPS with *O*-antigen, one must screen for many spin labeling sites that are exposed on the extracellular side, as the long LPS chains could interfere with labeling at the selected site, based on the lengths of *O*-antigens and the structures of lipid A present. Thus, allowing identification of the best residue sites on living cells to be labeled, which could be quite different than selecting sites to label in isolated systems.

As discussed in Chapter 1.4.7, bacteria have alternative Dsb pathways and different DsbA and DsbB homologs (for example DsbL and DsbI in *Salmonella*, and *Campylobacter* strains). These sometimes appear with multiple copies present per cell<sup>35</sup>. Thus, during direct studies of such bacterial species, one should focus on which Dsb homolog use the OMP of interest as a substrate, and then if available, try to use the corresponding Dsb mutant strains.

The expression of native BtuB levels is low (~200-300 when over-expressed). Thus, even in the Dsb knockout strains, the low levels of native BtuB expression did not affect the expression of the cysteine mutants nor did it cause dilution of the mutant BtuB in the OM. But there are some OMPs which are expressed in higher levels. For example, OmpF and OmpA express at about  $10^5$  copies per cell<sup>36</sup>. If the cysteine mutants of these OMPs are expressed, along with the

native OMP in the cells, the levels of mutant OMPs will be diluted by the levels of native OMPs. Depending on the investigation (ex. exploring the conformational states), spin dilution can reduce the background contribution in the DEER experiment. But for exploring spatial organizations, the spin dilute can be an issue.

Moreover, the OMPs assembled into multimers, such as OmpF (trimer), LamB (trimer) and TolC assembled in three subunits, will require careful planning for the in cell labeling. With respect to these OMPs, single cysteine residue labeling in one subunit will results in three spin labels within the assembled OMP. Thus, in DEER experiment spin interactions will provide a good signal. However, if the expression levels are high (especially for OmpF, as it has higher expression levels in the cells) it can potentially affect the DEER experiment results during the background subtraction. This could potentially overcome with use of cell lines expressing the native OMP which will dilute the mutant OMP levels in OM.

### 3.7 References

- [1] Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010) TonB-dependent transporters: regulation, structure, and function, *Annual review of microbiology* 64, 43-60.
- [2] Schalk, I. J., Mislin, G. L., and Brillet, K. (2012) Structure, function and binding selectivity and stereoselectivity of siderophore-iron outer membrane transporters, *Current topics in membranes* 69, 37-66.
- [3] Chimento, D. P., Mohanty, A. K., Kadner, R. J., and Wiener, M. C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB, *Nature structural biology* 10, 394-401.
- [4] Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex, *Science* 312, 1396-1399.
- [5] Joseph, B., Sikora, A., Bordignon, E., Jeschke, G., Cafiso, D. S., and Prisner, T. F. (2015) Distance Measurement on an Endogenous Membrane Transporter in E. coli Cells and Native Membranes Using EPR Spectroscopy, *Angewandte Chemie (International ed. in English)* 54, 6196-6199.
- [6] Joseph, B., Sikora, A., and Cafiso, D. S. (2016) Ligand Induced Conformational Changes of a Membrane Transporter in E. coli Cells Observed with DEER/PELDOR, *Journal of the American Chemical Society* 138, 1844-1847.
- [7] Sambrook, J., and Russell, D. W. (2001) Molecular Cloning A Laboratory Manual, Third ed., Cold Spring Harbor Laboratory Press, New York.
- [8] Lodish H, B. A., Zipursky SL, et al. (2000) Molecular Cell Biology, 4th ed., W. H. Freeman, New York.
- [9] Stein, R. A., Beth, A. H., and Hustedt, E. J. (2015) A Straightforward Approach to the Analysis of Double Electron-Electron Resonance Data, *Methods in enzymology* 563, 531-567.
- [10] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Research* 28, 235-242.
- [11] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.
- [12] Di Masi, D. R., White, J. C., Schnaitman, C. A., and Bradbeer, C. (1973) Transport of Vitamin B(12) in Escherichia coli: Common Receptor Sites for Vitamin B(12) and the E Colicins on the Outer Membrane of the Cell Envelope, *Journal of Bacteriology* 115, 506-513.
- [13] Cadieux, N., and Kadner, R. J. (1999) Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter, *Proceedings of the National Academy of Sciences of the United States of America* 96, 10673-10678.
- [14] Navarro Llorens, J. M., Tormo, A., and Martinez-Garcia, E. (2010) Stationary phase in gram-negative bacteria, *FEMS Microbiol Rev* 34, 476-495.
- [15] Hubbell, W. L., Cafiso, D. S., and Altenbach, C. (2000) Identifying conformational changes with site-directed spin labeling, *Nature structural biology* 7, 735-739.
- [16] Stewart, E. J., Madden, R., Paul, G., and Taddei, F. (2005) Aging and death in an organism that reproduces by morphologically symmetric division, *PLoS Biol* 3, e45.

- [17] Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J., and Taddei, F. (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation, *Proceedings of the National Academy of Sciences of the United States of America* 105, 3076-3081.
- [18] Keana, J. F., Pou, S., and Rosen, G. M. (1987) Nitroxides as potential contrast enhancing agents for MRI application: influence of structure on the rate of reduction by rat hepatocytes, whole liver homogenate, subcellular fractions, and ascorbate, *Magnetic resonance in medicine* 5, 525-536.
- [19] Karthikeyan, G., Bonucci, A., Casano, G., Gerbaud, G., Abel, S., Thomé, V., Kodjabachian, L., Magalon, A., Guigliarelli, B., Belle, V., Ouari, O., and Mileo, E. (2018) A Bioresistant Nitroxide Spin Label for In-Cell EPR Spectroscopy: In Vitro and In Oocytes Protein Structural Dynamics Studies, *Angewandte Chemie* 130, 1380-1384.
- [20] Di Girolamo, P. M., and Bradbeer, C. (1971) Transport of vitamin B 12 in Escherichia coli, *J Bacteriol* 106, 745-750.
- [21] Seputiene, V., Daugelavicius, A., Suziedelis, K., and Suziedeliene, E. (2006) Acid response of exponentially growing Escherichia coli K-12, *Microbiological research* 161, 65-74.
- [22] Bearson, B. L., Lee, I. S., and Casey, T. A. (2009) Escherichia coli O157 : H7 glutamate- and arginine-dependent acid-resistance systems protect against oxidative stress during extreme acid challenge, *Microbiology (Reading, England)* 155, 805-812.
- [23] Krulwich, T. A., Sachs, G., and Padan, E. (2011) Molecular aspects of bacterial pH sensing and homeostasis, *Nature reviews. Microbiology* 9, 330-343.
- [24] Kirakosyan, G., Bagramyan, K., and Trchounian, A. (2004) Redox sensing by Escherichia coli: effects of dithiothreitol, a redox reagent reducing disulphides, on bacterial growth, *Biochemical and biophysical research communications* 325, 803-806.
- [25] Kim, K. H., Aulakh, S., and Paetzel, M. (2012) The bacterial outer membrane beta-barrel assembly machinery, *Protein science : a publication of the Protein Society* 21, 751-768.
- [26] Plummer, A. M., and Fleming, K. G. (2016) From Chaperones to the Membrane with a BAM!, *Trends Biochem Sci* 41, 872-882.
- [27] Rietsch, A., Belin, D., Martin, N., and Beckwith, J. (1996) An in vivo pathway for disulfide bond isomerization in Escherichia coli, *Proc Natl Acad Sci U S A* 93, 13048-13053.
- [28] Rietsch, A., Bessette, P., Georgiou, G., and Beckwith, J. (1997) Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin, *J Bacteriol* 179, 6602-6608.
- [29] Kadokura, H., and Beckwith, J. (2010) Mechanisms of Oxidative Protein Folding in the Bacterial Cell Envelope, *Antioxidants & Redox Signaling* 13, 1231-1246.
- [30] Fuller-Schaefer, C. A., and Kadner, R. J. (2005) Multiple Extracellular Loops Contribute to Substrate Binding and Transport by the *Escherichia coli* Cobalamin Transporter BtuB, *Journal of Bacteriology* 187, 1732-1739.
- [31] Yoshida, T., Qin, L., Egger, L. A., and Inouye, M. (2006) Transcription Regulation of ompF and ompC by a Single Transcription Factor, OmpR, *Journal of Biological Chemistry* 281, 17114-17123.
- [32] Krewulak, K. D., and Vogel, H. J. (2011) TonB or not TonB: is that the question? This paper is one of a selection of papers published in a Special Issue entitled CSBMCB 53rd Annual Meeting — Membrane Proteins in Health and Disease, and has undergone the Journal's usual peer review process, *Biochemistry and Cell Biology* 89, 87-97.

- [33] Lyubenova, S., Siddiqui, M. K., Vries, M. J., Ludwig, B., and Prisner, T. F. (2007) Protein-protein interactions studied by EPR relaxation measurements: cytochrome c and cytochrome c oxidase, *The journal of physical chemistry. B* 111, 3839-3846.
- [34] Nikaido, H. (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited, *Microbiology and Molecular Biology Reviews* 67, 593-656.
- [35] Landeta, C., Boyd, D., and Beckwith, J. (2018) Disulfide bond formation in prokaryotes, *Nature Microbiology* 3, 270-280.
- [36] Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M., Garrod, T. J., Sansom, M. S. P., Piehler, J., Baumann, C. G., and Kleanthous, C. (2015) Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria, *Nature* 523, 333.

## **Chapter 4. Exploring live Gram-negative bacterial outer membrane proteins' hierarchical organization into domains**

### **4.1 Abstract**

In Gram-negative bacteria, outer membrane proteins (OMPs) can organize into clusters or islands in the outer membrane. During the bacterial cell cycle, the mid-cell region is the hotspot for new OMP islands while the poles are crowded with inherited OMPs that are gradually moved to poles as a result of cell division. The hierarchical clustering of OMPs has been investigated mainly by single cell fluorescence microscopy and MD simulations. However, these studies are not directly based on OMP labeling.

In contrast, OMP organization can be explored using direct spin labeled OMPs and pulsed electron paramagnetic resonance (EPR). Based on the pulsed EPR results from spin labeled BtuB in live *Escherichia coli* cells indicates that there is a long-distance component resulting from intermolecular interactions between nearby BtuBs. The distance distributions are broad, potentially due to additionally interacting BtuB molecules, which may or may not be arranged linearly. Currently, we cannot exactly define the underlying physiological processes and contacts associated with these intermolecular interactions observed in live cells. However, we speculate that this could be a composite of intermolecular interactions of BtuBs at the poles (the final destination of the inherited OMPs') and the mid-cells (where the new OMPs are actively inserting).

## 4.2 Introduction

Gram-negative bacteria contain an outer membrane (OM) which is asymmetric with respect to both lipid composition and membrane protein distribution (Chapter 1.3.1 and Chapter 1.3.2)<sup>1</sup>. As OM is directly in contact with external environment, OM biogenesis is critical for Gram-negative bacterial survival. At the OM's outer leaflet, the lipopolysaccharides (LPS) create a hydrophilic mass around the cell which repels any hydrophobic compounds<sup>2</sup>. The outer membrane proteins (OMPs) enable the import and export of nutrients and wastes, pathogenic and virulent proteins across the OM. Many enzymes, chaperones, and other proteins within the IM, periplasm, and the OM play key roles in OM biogenesis (Chapter 1.4).

During LPS biogenesis, matured LPS are translocated by LPS transport (Lpt) machinery and inserted into the OM (Chapter 1.4.1)<sup>3</sup>. As discussed in Chapter 1.3.1, there is a diversity among Gram-negative bacterial LPS as an adaptation to enhance survival in their habitats<sup>1</sup>. The nascent OMP precursors are delivered to the  $\beta$ -barrel assembly machinery (BAM) by periplasmic chaperones (Chapter 1.4.2, and Chapter 1.4.4) where the nascent OMP precursors are folded and inserted into the OM (Chapter 1.4.6, and Figure 4.1)<sup>2, 4</sup>. The Lol machinery is involved in the lipoprotein biogenesis process (Chapter 1.4.3)<sup>5</sup>.

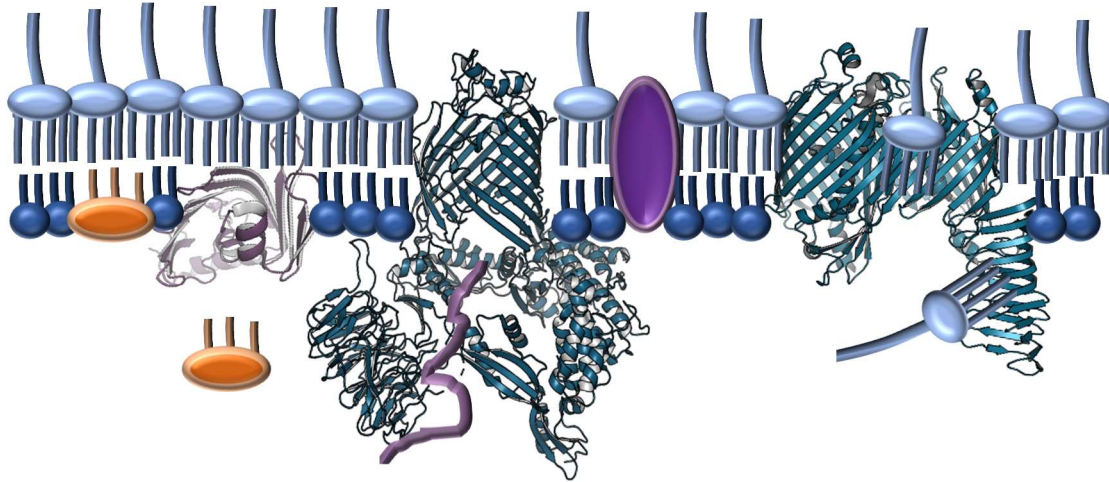


Figure 4.1. Asymmetric bacterial OM consists of an inner leaflet containing **phospholipids** and outer leaflet containing **LPS**. Periplasmic chaperones deliver **lipoproteins (LPs)** to LolB (PDB ID 1IWM) of the lipoprotein localization machinery, which inserts the **LPs** into the inner leaflet. The **OMPs** span across the OM. Periplasmic chaperones deliver **OMP precursors** to the **BAM** complex (PDB ID 5D0O), where the **precursors** are folded and inserted into the OM. **LptD-LptE** (PDB ID 4Q35) inserts **LPS** into the outer leaflet of the OM.

In recent years, new evidence has suggested OMPs can form higher levels of organization in the OM. These proteins are assembled into OM domains, or more specifically, into OMP islands within the sea of LPS. The work done by Rassam *et. al.* using single molecule fluorescence microscopy and molecular dynamics (MD), indicates that the OMPs are assembled into  $\sim 0.5 \mu\text{m}$  diameter islands. These restrict the mobility of the resident OMPs, which could range in number from 100 to 1000 OMPs<sup>6</sup>. Furthermore, this limited mobility is not due to (i) covalently or non-covalently formed connections with the peptidoglycan layer by lipoprotein, Lpp or OMPs, OmpA or Pal. (ii) It is also not due to LPS in the OM nor due to the high levels of OmpA or OmpF ( $\sim 10^5$  copies per cells). (iii) It is also not caused by the proton motive force (pmf) across

the IM nor due to the IM proteins (IMP) TonB or TolA, which are involved in OMP transport processes<sup>6</sup>. It is then likely due to the protein-protein interactions between OMP islanders.

During the bacterial cell cycle, a new cell envelope is assembled within the mid-cell region. Thus, new OMP islands are assembled in the mid-cell region while old OMP islands (inherited from the mother cell to the daughter cells) are localized at the poles<sup>6</sup>. Thus, over multiple cell cycles, there is continuous accumulation of OMP islands at the poles, which leads to bacterial aging and death (Chapter 1.2.2)<sup>7</sup>. Within the OMP island, one or multiple BAM complexes insert the nascent OMP islanders. As new OMP islands are assembled in the mid-cell, the initial BamA of the BAM complex (an OMP) is believed to be inserted into the OM with the aid of the lipoproteins in the BAM complex. After the initial assembly of the BAM complex, the insertion of nascent OMP precursors occurs in the mid-cell. As LptD is also a BAM precursor, within the islands the LPS insertion machinery (Lpt) also assembles. Thus, in an OMP island, both OMPs and LPS form aggregates and protein-protein interactions are prominent with neighboring OMPs. Consequently, LPS are then heterogeneously distributed within the islands, with some OMPs not being completely surrounded by LPS<sup>8</sup>.

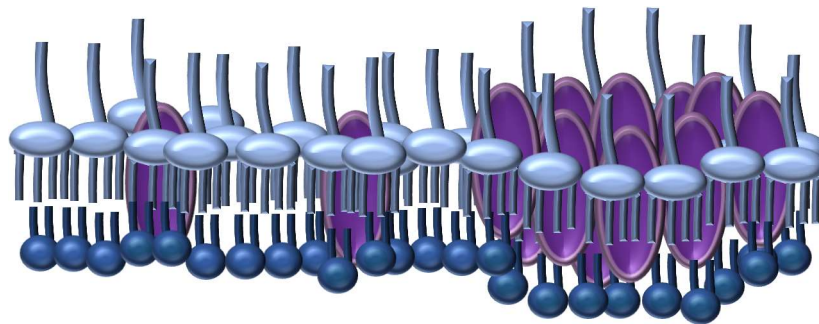


Figure 4.2 LPS heterogeneously distributed in the OM, where some OMPs are clustered together in some regions while other OMPs are surrounded by LPS in the outer leaf-let.

Furthermore, spatial assemblies have demonstrated influence on the lateral mobility of inner membrane proteins (IMPs) during the bacteriocin translocation process. Even though the OMP islanders have restricted mobilities, IMPs laterally diffuse in the IM. However, during the colicin E9 invasion, BtuB is used as the primary receptor which then screens OmpF (Chapter 1.6.4). Once bound to OmpF, the translocator domain passes through OmpF and interacts with TolB in the periplasm which recruits TolA (IMP) in the Tol-Pal system. This restricts the mobility of the IMP and consequently localizes Tol-Pal IMPs to align with colicin E9 bound OMPs within the islands<sup>9</sup>.

Current knowledge on OMP islands is mainly based on single cell fluorescence microscopy and MD stimulations. During fluorescent labeling, intact *Escherichia coli* (*E. coli*) cells were stained with fluorescently labeled antibodies which detect BamA and BamC, while fluorescently labeled colicin E9 and colicin Ia were added to detect their primary OMP receptors (BtuB and Cir, respectively)<sup>6, 8, 9</sup>. Thus, these results are not directly based on OMP labeling which could impact the support for spatial organization of OMPs into clusters, or OMP islands.

With this regard, site directed spin labeling (SDSL) coupled to pulse EPR is a novel way to explore the higher organization of OMPs within the OM. Currently, we have optimized an approach for spin labeling single cysteine residues either at a pre-selected extracellular loop or in the core region exposed to the extracellular side of BtuB using live, healthy *E. coli* cells. Thus, spin labeled BtuB organization in the OM can be explored via pulsed EPR to detect spin coupling between neighboring BtuBs within certain vicinities in the live cells.

### 4.3 Materials and Methods

#### 4.3.2 (a) Mutagenesis and long-term preservation of V90C and T188C BtuB expressing cells

V90C and T188C *btuB* gene containing pAG1 plasmids, constructed for the method optimization project (Chapter 3), were transformed into RK5016 cells and glycerol stocks were prepared for long term storage. During the cell culture, the pre-cultures were directly inoculated with small aliquots (5-10  $\mu$ L) of the corresponding mutant's glycerol stocks.

#### 4.3.3 Spin labeling of live *E. coli* cells

Main cell cultures (100 mM phosphate buffer, 8 mM  $(\text{NH}_4)_2\text{SO}_4$ , and 2 mM sodium citrate) was supplemented with 200  $\mu$ g/mL ampicillin, 0.2 % w/v glucose, 150  $\mu$ M thiamine, 3 mM  $\text{MgSO}_4$ , 300  $\mu$ M  $\text{CaCl}_2$ , 0.01 % w/v methionine, and 0.01% w/v arginine then directly inoculated by pre-cultures. As an extension to our single residue spin labeling method optimization process (Chapter 3), cells expressing T188C BtuB were grown under numerous conditions then used for MTSSL labeling.

Briefly, (a) a set of T188C BtuB expression RK5016 cells were grown for different time periods (6.5 – 28 hrs), where the cells are in late log, stationary, or in the death phase (Chapter 3.3.3), (b) cells expressing V90C and T188C were grown until mid-log and early log phase (Chapter 3.3.4 and 3.3.5). All cells were spin labeled and processed under various conditions: (i) with and without ice incubation after the spin labeling step, (ii) an ice incubation step at pH 7.0 and 5.5, (iii) two ice incubation steps instead of one and (iv) a final incubation in MES buffer prepared in  $\text{D}_2\text{O}$ . At the end, spin labeling efficiencies were analyzed by CW EPR and used to prepare DEER samples.

#### 4.3.4 Spin labeling cells grown in deuterated culture media

RK5016 cells expressing T188C were grown in deuterated pre-cultures O/N. Pre-cultures were used to inoculate the deuterated main cultures, using 0.2 % w/v deuterated glycerol for a carbon source instead of glucose. The cells were grown until OD<sub>600</sub> of 0.3, cells from 100 mL cultures were harvested and used for spin labeling using 1 mg MTSSL. After spin labeling for an hour at RT, pelleted cells were suspended into 100 mM MES, pH 5.5 prepared in D<sub>2</sub>O (MES-D<sub>2</sub>O) and incubated on ice for 30 min. The washing step was repeated one more time and finally, cells were harvested and analyzed by CW-EPR. Then samples were prepared for pulsed EPR analysis.

RK5016 cells expressing T188C were grown in minimal media pre-pre-cultures O/N, then used to inoculate deuterated pre-cultures, and incubated for 8 hrs. Pre-cultures were then used to inoculate deuterated main cultures, using 0.2 % w/v deuterated glycerol for the carbon source instead of glucose. The cells were grown until OD<sub>600</sub> ~0.5, spin labeled, and processed as above.

#### 4.3.5 Spin dilution of T188C BtuB and WT BtuB spin labeling

Cells expressing T188C and WT BtuB were grown until OD<sub>600</sub> ~0.6 then spin labeled. Cells were harvested from 100 mL of main culture, which expressed T188C BtuB, then spin labeled using 1:1 diluted MTSSL spin with 1-acetyl-2,2,5,5-tetramethyl- $\Delta^3$ -(pyrroline-15N)-3-methyl methanethiosulfonate, while a WT BtuB (control) was spin labeled with 1 mg of MTSSL. At the end of the incubation, cells were processed as above and analyzed using CW and pulsed EPR.

#### 4.3.6 CW and pulse EPR data collection of spin labeled BtuB in *E. coli*

Processed cells were directly loaded into glass capillaries and CW EPR measurements were collected using a Bruker EMX spectrometer at X-band with an ER 4123D dielectric resonator

using a 100 G sweep width, 1 G modulation, and 2 mW of incident microwave power at RT. All the collected spectra were normalized, phase corrected, and plotted using in-house build programs (DavePlot CW and CWPhase by David Nyenhuis).

Next, DEER samples were prepared with spin labeled cells (16  $\mu$ L) mixed with 4  $\mu$ L (20%) deuterated glycerol and loaded into quartz capillaries. Loaded samples were frozen using either liquid nitrogen or a dry-ice isopropanol bath. DEER data was collected using a dead-time free, four pulse DEER sequence on a Bruker ELEXSYS E580 at Q band (EN5107D2 dielectric resonator and 10 W Bruker AmpQ) with 32 ns and 36 ns lengths for the observe and pump  $\pi$  pulses, with a 75 MHz frequency offset between observe and pump pulse.

#### 4.3.7 Pulse EPR data analysis

DEER data collected for single residue labeled BtuB were first analyzed using the DEER analysis 2015 software<sup>10</sup>. After removing the 2+1 pulse artifact, the background was subtracted and distance distributions were generated from the dipolar factor using Tikhonov regularization. DEER data were then analyzed using DD version 6C<sup>11</sup>. Once the 2+1 pulse artifact was removed, data were fitted into a Gaussian model. Model based fitting was then used to co-fit the background and the Gaussian distributions. Fits were evaluated using Akaike and Bayesian information criterion values. The analyzed data were plotted using DavePlot (by David Nyenhuis) through Plotly and Dash frameworks developed by the Plotly Corporation.

#### 4.3.8 Generated pymol images

Crystal structures of BtuB, LolB, BAM complex and LptD-E (1NQG, 1IWM, 5D0O, and 4Q35)<sup>12-14</sup> were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>15</sup> and the images were created using PyMol Molecular Graphics System version 1.8<sup>16</sup>.

#### 4.4 Results and Discussion

The echo signal,  $V(t)$ , collected during the DEER experiments contains both the form factor,  $F(t)$ , that results from either intra- or intermolecular spin interactions of interest, and the background,  $B(t)$ , which is due to all other spin interactions during the DEER experiment (Chapter 2.5.4)<sup>17</sup>. Therefore, once the echo signal is collected, the background signal needs to be properly subtracted in order to obtain a proper  $F(t)$ . There are a several software packages to analyze DEER data, such as: DEER Analysis by Jeschke, DD by Hustedt, and LongDistances (LD) by Altenback<sup>11</sup>. These software programs use different methods to obtain distance distributions resulting from the intra- or intermolecular spin interactions of interest. Model-based or model-free fitting methods can be used to extract distance distributions using either Gaussian distributions or Tikhonov regularization, respectively. Typically, we use Tikhonov regularization with DEER analysis and Gaussians with DD. LD performs well with both Gaussian and Tikhonov regularization<sup>11</sup>.

During our method development for live cell labeling project, the collected DEER data resulted in a long-distance component in addition to the short distance component, which corresponded to the distance between the spin labeled BtuB (T188R1) and spin labeled B<sub>12</sub> (Figure 4.3). This long-distance signal appeared regardless of the dimensionality or form of the intramolecular background form  $B(t)$ , and was not an artifact of improper background correction. In most cases, a model-based Gaussian fitting using DD produced slightly better results than the model free fit using DEER Analysis. Therefore, DEER data was analyzed by DD.

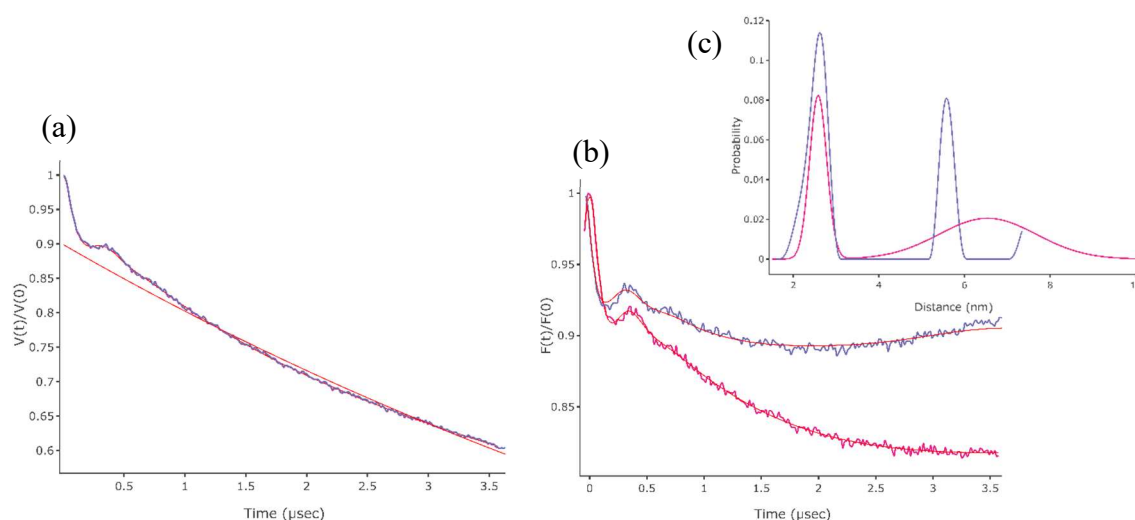


Figure 4.3 Echo signals obtained from the pulsed EPR experiment for T188R1 BtuB incubated with spin labeled B<sub>12</sub> analyzed using **DEER analysis** and **DD**. (a) The echo signal was collected from the T188R1 BtuB DEER experiment, (b) the form factor is after subtracting the background using **DEER analysis** and **DD** and (c) the resultant normalized distance distribution is extracted from each analysis method.

This long-distance component was detected under various conditions, such as: cells harvested from different growth curve phases (early-log, mid-log, late-log, and stationary phase). Spin labeled (T188R1 BtuB) cells were processed on ice to reduce spin reduction and to remove residual unreacted MTSSL at pH 7.0 or 5.5, with one or two washing steps. These were spin labeled either on the second extracellular loop, T188R1 or the core, V90R1 of BtuB (Figure 4.4, Figure 4.5, and Figure 4.6). In these T188R1 BtuB samples, the long-distance was around  $65 \pm 15$  Å. This distance was not present in the WT BtuB control sample, because no DEER signal was obtained during this experiment. Thus, the long-distance component is not a result of off-target labeling of other OMPs on the cells, but is a result of the spin labeled BtuB interacting with neighboring BtuBs in the OM of live bacterial cells.

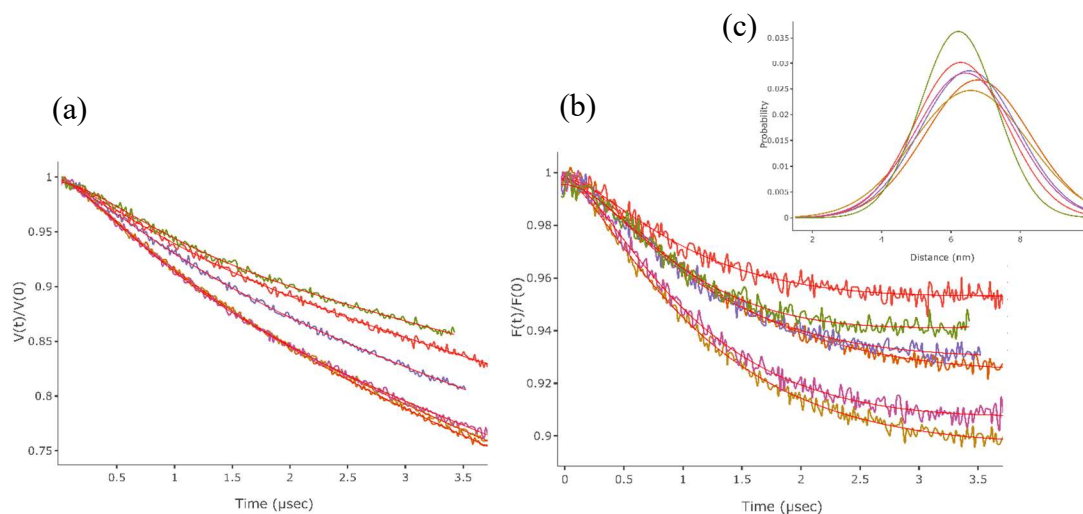


Figure 4.4. Echo signals obtained from the pulsed EPR experiment for RK5016 cells expressing T188C BtuB. (a) The echo signals were collected for cells collected at different growth time points (6.5 hrs, 8 hrs, 9.5 hrs, 11 hrs, 12.5 hrs, and 14 hrs) representing late-log, stationary, and late-stationary/death phases of the growth curve, (b) the form factor is after subtracting the background using DD, and (c) the resultant normalized distance distribution.

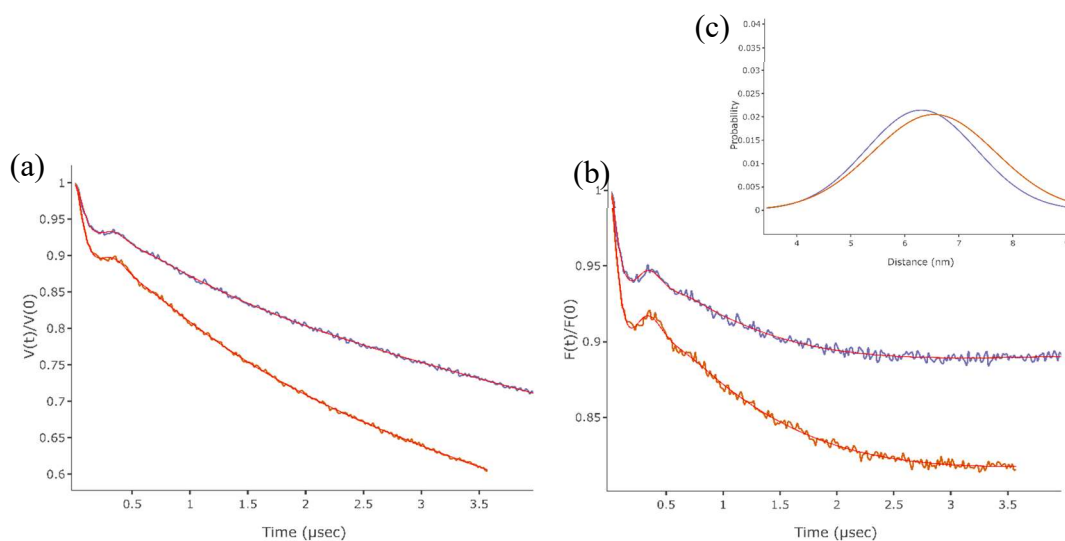


Figure 4.5. Echo signals obtained from the pulsed EPR experiment for cells expressing T188C BtuB incubated on ice after the spin labeling step. (a) The echo signals are collected for cells collected at mid-log phase that were incubated on ice for 20 and 60 min after the

spin labeling step, (b) the form factor is after subtracting the background using DD, and (c) the resultant normalized distance distribution.

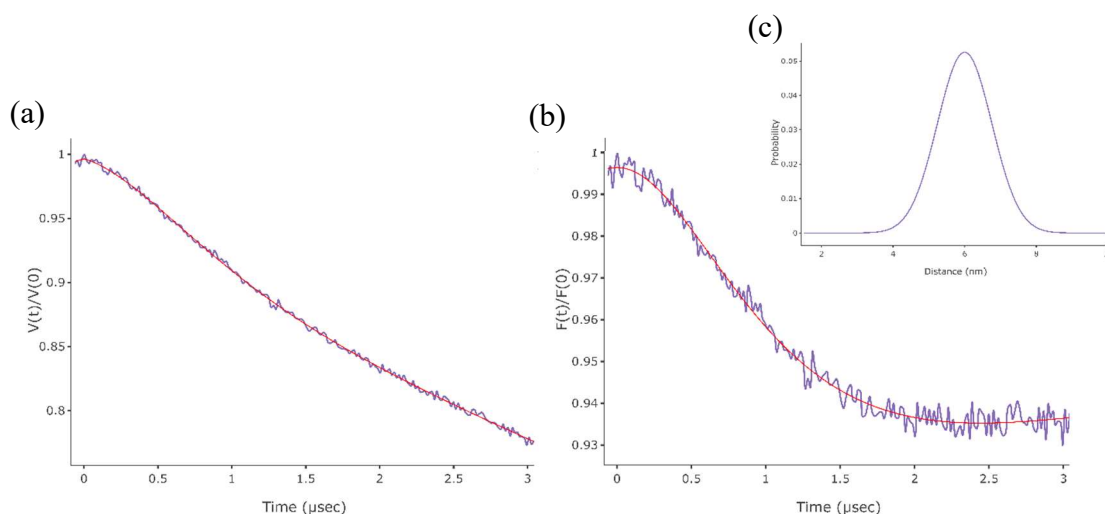


Figure 4.6. Echo signals obtained from the pulsed EPR experiment for cells expressing V90C BtuB incubated on ice after the spin labeling step. (a) The echo signals are collected for cells collected at early-log phase that were incubated on ice for 60 min after the spin labeling step, (b) the form factor is after subtracting the background using DD, and (c) the resultant normalized distance distribution.

The observed broad long-distance component results from intermolecular interactions between the T188R1 nitroxide on loop 2 of nearby BtuB within the OM. Since we were able to obtain DEER signals resulting from intermolecular BtuB interactions, the BtuB are not randomly distributed within the OM. Similar broad, long-distance distributions around 65 Å, ranging from around 4 nm to 9 nm were observed in many samples that were prepared under various spin labeling and processing conditions. The fact that BtuBs is not randomly distributed in the OM is consistent with the proposal that BtuB is present in clusters. Whether there are specific BtuB-interactions mediating these clusters or non-specific interactions due to crowding is not presently clear. Based on the BtuB crystal structure, the dimensions of the elliptical cross-section in the

lumen is around  $42 \text{ \AA} \times 37 \text{ \AA}$ <sup>18</sup>, while T188 of the BtuB on loop two is positioned near the luminal edge. This resultant distribution ranges from 2-8 nm for two BtuBs to interact. Thus, the distance is mainly around  $65 \text{ \AA}$  and may be due to BtuBs interacting with each other in a linear fashion along the major axis as has been proposed. However, the long-distance is broad. Thus, there are populations of intermolecular interactions that are potentially due to deviations of the linearly interacting BtuBs as the levels of BtuB increase within the interacting OMPs population (Figure 4.7).

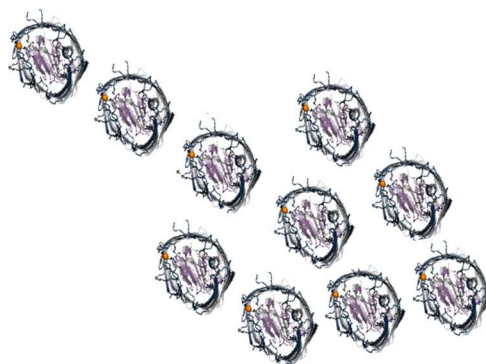


Figure 4.7 Schematic representation of intermolecularly interacting T188R1 spin labeled BtuB in the OM.

In addition to detecting the intermolecular interactions of BtuB with respect to loop two, these interactions were also detected using V90R1 on the BtuB core (Figure 4.4 d). Based on the DD analysis, the main distance distribution is around  $60 \text{ \AA}$ . Thus, the obtained broad range associated with the main distance is not just due to dynamics within loop two, but it is most likely due to deviations in the linear arrangement of the interacting BtuBs on the OM and any peripheral interactions.

To increase the phase memory time ( $T_m$ ) during the DEER experiments, samples can be deuterated to reduce spin relaxation times which improves the sensitivity and extends the

reliability of the obtained distance range. Therefore, cells expressing the protein of interest can be grown in deuterated media and the purified deuterated protein of interest can be used in DEER experiments<sup>19</sup>. Similarly, RK5016 cells expressing T188C BtuB were grown in minimal media prepared with D<sub>2</sub>O and deuterated glycerol as a carbon source instead of glucose. In contrast to purification of deuterated proteins for pulsed EPR, we wanted to directly use the cells grown in deuterated media as the OMP domain organization is a phenomenon detected only in live cells. During the cell culturing, cell growth is significantly slower when the cells use deuterated media within their metabolic pathways and incorporate the deuterated sources into their newly synthesized biomolecules. Thus, the cell cultures had long lag phases as the initial cell population was adjusting to deuterated minimal medium. Each round of cell doubling gradually converted the cells' OM (both LPS and OMPs) into a deuterated OM and surrounded them in a D<sub>2</sub>O environment.

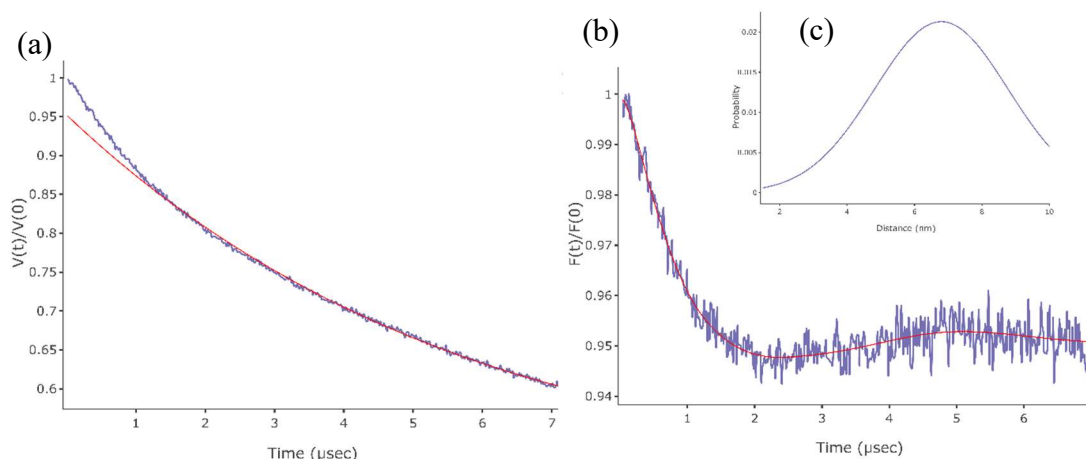


Figure 4.8. Echo signals obtained from the pulsed EPR experiment for cells expressing T188C BtuB in deuterated main cultures (second experiment). (a) The echo signals collected were for cells collected at mid-log phase that were spin labeled and incubated on ice for 60 min after the spin labeling step, (b) the form factor after subtracting the background using DD and (c) the resultant normalized distance distribution.

The samples prepared by the cells grown in deuterated media were to ultimately increase the phase memory time in the DEER experiment to improve the data quality and to obtain a reliable long-distance component. We were able to increase the phase memory from around 3-4  $\mu\text{s}$  to 6  $\mu\text{s}$  (in the first experiment) and 7  $\mu\text{s}$  (in the second experiment) (Figure 4.8). Distance distributions extracted from these longer  $T_m$  are more reliable and the resultant distance distribution was broader. The broader distribution could potentially be due to combination of both longer  $T_m$  (7  $\mu\text{s}$ ) and cells being adapted to deuterated growth media as the OM organization is modified to accommodate biomolecules with deuterium instead of hydrogen (including lipids and proteins). Therefore, the intermolecular interactions and the clustering of nearby BtuBs on the OM could be potentially altered in deuterated cultures.

Since the DEER data was significantly improved for the second experiment over the first one, the corresponding sample had a higher population of interacting spin labeled BtuB in the OM, potentially because the cells were grown longer than the previous one. During the cell culture, cells in the second experiment were grown starting with pre-pre-cultures instead of pre-cultures, thus, the cells used in the second experiment were better adapted to deuterated medium than the first experiment and had undergone additional cell cycles resulting in more OMPs in their OM than the others (mostly at the poles).

At the moment, we cannot exactly explain the underlying OM physiological process(es) driving the wide distance distributions seen for interacting BtuBs in the OM of live cells; nonetheless, we can speculate based upon the results of these DEER experiments. With the birth of new bacterial daughter cells, new OM materials are continuously synthesized into the mid-cell region of the cells, while inherited biomolecules migrate towards the bacterial poles<sup>20, 21</sup>. In these mid-cell regions, BAM complexes are actively inserting<sup>22</sup> new OMPs into the OM. In a typical

bacterium, BtuB expression is low ( $\sim 200$  to  $300$  copies per cells<sup>23</sup>), while general porins expression is high ( $\sim 10^5$  copies of OmpF and OmpA per cell<sup>6</sup>). However, when over-expressed using the pAG1 plasmid, the levels of BtuB expression are enhanced by about 1000-fold<sup>24</sup>. Therefore, there are higher numbers of newly inserted BtuB in OM, resulting crowding on the OM and potentially causing regulation of overall OMP expression by the BAM complex and/or the LPS insertion into the OM by the LptD-E complex in order to maintain OM biogenesis.

LPS is heterogeneously distributed within the OM, with patches where OMPs are without LPS<sup>8</sup>. Thus, the presence of LPS could potentially result in longer distances which are deviating from the main long-distance. Moreover, the broad distance distribution could be a combined result of the BtuB intermolecular interactions occurring at the poles where the inherited OMPs are resting in the crowd, which could result in differently oriented BtuBs, which interact with nearby BtuBs and the actively inserting BtuBs by BAM in the mid-cell.

#### 4.5 Future directions

During cell cycle, BAM inserts new OMPs into the mid-cell regions which produce new OMP islands while old islands are inherited by previous generations which end up at the poles. This leads to bacterial aging and death<sup>20</sup>, even at the log phase, which is the active and healthy phase within the bacterial growth curve (Chapter 3.4). Therefore, the distances extracted from the DEER experiments can be due to crowded OMPs at the poles or they can be due to both new OMP islands and inherited old OMP islands (clusters) in the OM. So, it will be interesting to identify and, if possible to quantify, the contribution of inherited OMPs from previous generations to newborn daughter cells, which are gradually and continuously aggregating at the poles, then by the OMPs assembled into OMP islands from the BAM complexes at the mid-regions. Our sample preparation could be modified to separately spin label old and new BtuB in

the domains either at the poles or at the mid-cell, thus, allowing us to explore potential clustering at each region of cells.

Additionally, spin dilution experiments either by co-expressing both WT and T188C or V90C BtuB in the cells or using non-paramagnetic spin label along with MTSSL to label T188C or V90C BtuB could be used to further determine potential BtuB clustering. In the first method, this could allow us to arbitrarily insert both WT and T188C BtuB into OMP domains. However, it is difficult to assume that each corresponding *btuB* gene containing the plasmid will be in equal ratios within the cells. In contrast, spin dilution can be also achieved using the appropriate ratios of spin label MTSSL and non-paramagnetic form of MTSSL during the labeling step to effectively dilute the population of spin labeled BtuB. All with the goal of determining the possibility of overcoming the long-distance component and to support that the long distance is a result of clustered OMPs by protein-protein interaction in the OM. Furthermore, it would be interesting to investigate T188C or V90C BtuB with and without the barrel residues that are involved in protein-protein interactions, thus, to determine whether in cells these residues are involved in BtuB clustering.

#### 4.6 References

- [1] Nikaido, H. (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited, *Microbiology and Molecular Biology Reviews* 67, 593-656.
- [2] Konovalova, A., Kahne, D. E., and Silhavy, T. J. (2017) Outer Membrane Biogenesis, *Annu Rev Microbiol* 71, 539-556.
- [3] Whitfield, C., and Trent, M. S. (2014) Biosynthesis and export of bacterial lipopolysaccharides, *Annu Rev Biochem* 83, 99-128.
- [4] Plummer, A. M., and Fleming, K. G. (2016) From Chaperones to the Membrane with a BAM!, *Trends Biochem Sci* 41, 872-882.
- [5] Kovacs-Simon, A., Titball, R. W., and Michell, S. L. (2011) Lipoproteins of bacterial pathogens, *Infection and Immunity* 79, 548-561.
- [6] Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M., Garrod, T. J., Sansom, M. S. P., Piehler, J., Baumann, C. G., and Kleanthous, C. (2015) Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria, *Nature* 523, 333.
- [7] Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J., and Taddei, F. (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation, *Proceedings of the National Academy of Sciences of the United States of America* 105, 3076-3081.
- [8] Chavent, M., Duncan, A. L., Rassam, P., Birkholz, O., Hélie, J., Reddy, T., Beliaev, D., Hambly, B., Piehler, J., Kleanthous, C., and Sansom, M. S. P. (2018) How nanoscale protein interactions determine the mesoscale dynamic organisation of bacterial outer membrane proteins, *Nature Communications* 9, 2846.
- [9] Rassam, P., Long, K. R., Kaminska, R., Williams, D. J., Papadakos, G., Baumann, C. G., and Kleanthous, C. (2018) Intermembrane crosstalk drives inner-membrane protein organization in *Escherichia coli*, *Nature Communications* 9, 1082.
- [10] Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data, *Applied Magnetic Resonance* 30, 473-498.
- [11] Stein, R. A., Beth, A. H., and Hustedt, E. J. (2015) A Straightforward Approach to the Analysis of Double Electron-Electron Resonance Data, *Methods in enzymology* 563, 531-567.
- [12] Takeda, K., Miyatake, H., Yokota, N., Matsuyama, S.-i., Tokuda, H., and Miki, K. (2003) Crystal structures of bacterial lipoprotein localization factors, LolA and LolB, *The EMBO Journal* 22, 3199-3209.
- [13] Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y., Wang, W., and Dong, C. (2016) Structural basis of outer membrane protein insertion by the BAM complex, *Nature* 531, 64.
- [14] Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C., and Huang, Y. (2014) Structural basis for lipopolysaccharide insertion in the bacterial outer membrane, *Nature* 511, 108-111.
- [15] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Research* 28, 235-242.
- [16] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.

- [17] Jeschke, G., and Polyhach, Y. (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance, *Physical chemistry chemical physics : PCCP* 9, 1895-1910.
- [18] Cherezov, V., Yamashita, E., Liu, W., Zhalnina, M., Cramer, W. A., and Caffrey, M. (2006) In Meso Structure of the Cobalamin Transporter, BtuB, at 1.95 Å Resolution, *Journal of molecular biology* 364, 716-734.
- [19] Ward, R., Bowman, A., Sozudogru, E., El-Mkami, H., Owen-Hughes, T., and Norman, D. G. (2010) EPR Distance Measurements in Deuterated Proteins, *Journal of magnetic resonance (San Diego, Calif. : 1997)* 207, 164-167.
- [20] Stewart, E. J., Madden, R., Paul, G., and Taddei, F. (2005) Aging and death in an organism that reproduces by morphologically symmetric division, *PLoS Biol* 3, e45.
- [21] Laloux, G., and Jacobs-Wagner, C. (2014) How do bacteria localize proteins to the cell pole?, *Journal of Cell Science* 127, 11-19.
- [22] Kleanthous, C., Rassam, P., and Baumann, C. G. (2015) Protein–protein interactions and the spatiotemporal dynamics of bacterial outer membrane proteins, *Current Opinion in Structural Biology* 35, 109-115.
- [23] Di Masi, D. R., White, J. C., Schnaitman, C. A., and Bradbeer, C. (1973) Transport of Vitamin B(12) in Escherichia coli: Common Receptor Sites for Vitamin B(12) and the E Colicins on the Outer Membrane of the Cell Envelope, *Journal of Bacteriology* 115, 506-513.
- [24] Fuller-Schaefer, C. A., and Kadner, R. J. (2005) Multiple Extracellular Loops Contribute to Substrate Binding and Transport by the *Escherichia coli* Cobalamin Transporter BtuB, *Journal of Bacteriology* 187, 1732-1739.

## **Chapter 5. Exploring novel conformational state(s) of spin labeled BtuB in *Escherichia coli* cells.**

### 5.1 Abstract

The members of TonB dependent transporter (TBDT) family such as BtuB, undergo substrate-induced conformational changes where a transmembrane signal propagates through salt bridges between the barrel and the core of the protein, resulting in the unfolding of the TonB box. The unfolded TonB box extends into the periplasm and interacts with the TonB complex which facilitates the active transport. Here, we have used our direct spin labeling strategy to explore previously unseen conformational states of BtuB through the labeling of exposed extracellular sites and pulsed electron paramagnetic resonance (EPR) spectroscopy.

In this work, we detect a short distance component in the distance distribution obtained from V90R1-T188R1 that has not previously been identified in BtuB. The population of this short distance component increases in the presence of substrate, while a longer-distance population (corresponding to the protein crystal structure) decreases. This population is not altered in a transport-defective mutant (L8P) and it is not altered by agents that uncouple electron transport from ATP synthesis. Further experiments will be required to determine if this state is or is not energized by TonB. Moreover, future experiments that incorporating ion pair mutants to trap BtuB in different conformational states involved in the signal transduction and/or substrate transport process, could deepen our understanding of the BtuB transport mechanism.

## 5.2 Introduction

Outer membrane proteins (OMPs) undergo conformational changes during the substrate transport across the outer membrane (OM)<sup>1-3</sup>. One class of these OMPs is the TonB dependent transporters, TBDTs (Figure 5.1)<sup>4</sup> which actively transport specific substrates (Chapter 1.5)<sup>5</sup>. During the transport process, TBDTs undergo substrate induce conformational changes<sup>6, 7</sup> such as BtuB TonB box unfolding into the periplasm and interacting with TonB<sup>8, 9</sup>.

As discussed in Chapter 3, most of the structural data on TBDT has been obtained on isolated systems where the inner membrane machinery to drive transport is not present<sup>4, 8-10</sup>. In contrast, direct studies of TBDT members in live cells can provide a better understanding of the structure and function related to the TBDT substrate transport mechanism. Thus, we use our new direct spin labeling methodology (Chapter 3) to gain structural information under conditions where BtuB undergoes transport in *Escherichia coli* (*E. coli*).

During our final attempts to spin label two cysteine residues of BtuB (Chapter 3, V90C-T188C BtuB) we were able to obtain a significant DEER signal resulting from the spin interactions between sites V90R1 and T188R1 on BtuB. Based on the extracted distance distributions, the two main distance populations are around 24.4 Å (short distance) and 31.2 Å (long distance) where the short distance had not previously been identified in BtuB in the isolated systems. This appears to represent a novel conformational state for BtuB where the C-terminal side of the core is brought into close proximity to loop two.

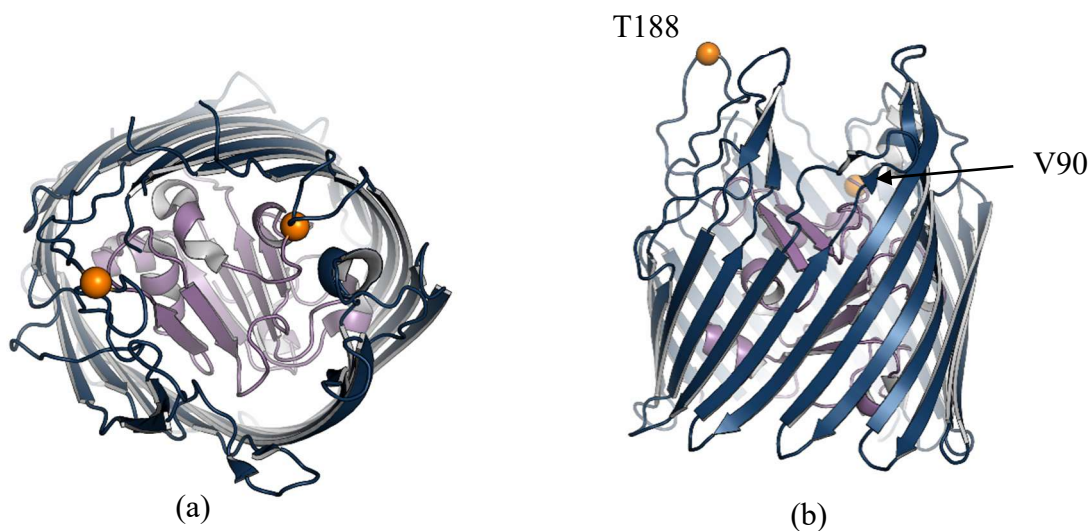


Figure 5.1. Apo BtuB (1NQG) structure showing (a) top view and (b) side view of apo BtuB. The residues **V90** and **T188** in the **core** and the **barrel** were used for spin labeling in live *E. coli* cells.

### 5.3 Methods

#### 5.3.1 Mutagenesis and long-term preservation of V90C-T188C BtuB expressing cells

V90C-T188C *btuB* gene containing pAG1 plasmids, constructed for direct spin labeling on live cells method project (Chapter 3), were transformed into RK5016 cells and Dsb mutant strains and glycerol stocks were prepared for long term storage. During the cell culture, the pre-cultures were directly inoculated with small aliquots (5-10  $\mu$ L) of corresponding mutant glycerol stocks.

#### 5.3.2 Exploring conformational state in V90R1-T188R1 BtuB

Dsb knockout strain, *dsbA*<sup>-</sup> expressing V90C-T188C BtuB were grown in main cultures (250 mL) until OD<sub>600</sub> of 0.3. The cells from 100 mL culture aliquot were harvested by centrifugation at 4000 rpm for 10 min at 4° C. Residual culture media was removed by re-suspending the resultant cell pellets into 10 mL of resuspension buffer, RB (100 mM HEPES, pH7.0) then the

cells were pelleted again. Next the cell pellets were resuspended into 1 mL of RB and 0.2 mg of MTSSL from a freshly prepared 1 mg/100  $\mu$ L stock was added (final concentration of 0.1 mg/cell pellet of 50 mL main culture at OD<sub>600</sub> 0.3) and incubated at room temperature for an hour with gentle mixing in the dark.

At the end of the incubation, spin labeled cells were harvested at 4500 rpm for 5 min at 4° C and each cell pellet was resuspended into (1 mL x 2) washing buffer, WB (100 mM MES, pH 5.5). One set of resuspended cells from each strain was incubated on ice for 30 min, followed by re-pelleting the cells, and re-suspending in 1 mL of WB prepared in D<sub>2</sub>O. These were then incubated on ice for another 30 min. The other set was incubated at room temperature (RT) for a 15 min incubation instead of on ice. Finally, the cells were harvested and analyzed using CW-EPR. Based on the spin labeling efficiency, DEER samples were prepared for both the apo and B<sub>12</sub> (100  $\mu$ M final concentration) added.

### 5.3.3 Impact on new conformational state in BtuB with changing B<sub>12</sub> concentration

*dsbA*<sup>-</sup> cells expressing V90R1-T188R1 BtuB were grown in main cultures (2 x 250 mL) until OD<sub>600</sub> ~ 0.3 and cells were pelleted in 200 mL from each culture (50 mL x 4 x 2), spin labeled (as above) with freshly prepared MTSSL (0.1 mg/cell pellet of 50 mL main cell culture at OD<sub>600</sub> 0.3), and processed. At the end the labeling efficiency was analyzed by CW-EPR. Next, the cell pellet aliquots of both RT and on ice incubated samples were mixed with B<sub>12</sub> (final concentrations of 0, 1, 5, 20, 30, 60, and 100  $\mu$ M) and then with deuterated glycerol (20% v/v). These samples were quickly loaded into DEER tubes and flash frozen by liquid N<sub>2</sub>. Samples were stored in the -80° C freezer in dry ice until running the DEER experiments.

### 5.3.4 Investigating the presence of new conformational states in T188R1-G399R1, L8P-V90C-T188C BtuB mutants expressed in *E. coli* and V90R1-T188R1 in intact OM

*dsbA*<sup>-</sup> cells expressing either T188C-G399C or L8P-V90C-T188C BtuB were grown (as above) and cells were collected (50 mL x 2) and spin labeled with freshly prepared MTSSL (0.1 mg /cell pellet from a 50 mL main culture at OD<sub>600</sub> 0.3). After spin labeling, cells were processed, analyzed by CW-EPR, and DEER samples for apo and 100  $\mu$ M B<sub>12</sub> were prepared. Samples were stored in dry ice in the -80°C until running pulsed EPR.

Intact OM from cells expressing V90C-T188C BtuB was purified as described in Chapter 6 and spin labeled. Samples (apo and 300  $\mu$ M B<sub>12</sub>) were analyzed using CW EPR and pulsed EPR.

### 5.3.5 Impact on new conformational state in V90R1-T188R1 BtuB by CCCP and KCN

*dsbA*<sup>-</sup> cells expressing V90C-T188C BtuB were spin labeled (as above) using cells harvested from a 400 mL main culture at OD<sub>600</sub> ~ 0.3. After spin labeling, pellets were resuspended in 4 mL of WB and aliquoted into four tubes. Carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) and potassium cyanide (KCN) (50  $\mu$ M and 1 mM final concentrations, respectively) were added to two tubes each, then one set of tubes (CCCP and KCN added) was incubated at RT and the other set on ice for 10 min. Cells were harvested and resuspended in WB with D<sub>2</sub>O and incubated again for 30 min. Finally, samples were used to prepare DEER samples.

### 5.3.6 EPR analysis of spin labeled BtuB

#### 5.3.6(a) CW-EPR analysis

During the experiments, samples were loaded into glass capillary tubes and the CW-EPR spectra were collected using an X-band operating Bruker EMX spectrometer with an ER 4123D dielectric resonator, a 100 G sweep width, 1 G modulation, and 2 mW of incident microwave

power. The collected CW-EPR spectra were normalized, phase corrected, and plotted using in-house programs (DavePlot CW and CWPhase by David Nyenhuis).

### 5.3.6(b) DEER experiments

After cells were processed, cell pellets (15 or 16  $\mu\text{L}$ ) were mixed with 4  $\mu\text{L}$  of deuterated glycerol (20 % v/v final concentration). 1 or 2  $\mu\text{L}$  of  $\text{B}_{12}$  was added to produce substrate bound samples. Mixed samples were loaded into quartz capillary tubes and quickly flash frozen using liquid nitrogen. Samples were either run or stored at  $-80^\circ\text{C}$  in dry ice until running the DEER experiment. DEER data was collected using a dead-time free four-pulsed DEER sequence on a Bruker ELEXSYS E580 at Q-band, with an EN5107D2 dielectric resonator, and a 10 W Bruker AmpQ. During the data collection 32 ns and 36 ns pulse lengths were used for the observe and pump  $\pi$  pulses, given with a 75 MHz frequency offset between the observe and pump pulse.

DEER data were analyzed using LongDistances version 771 (Christian Altenbach, UCLA)<sup>11</sup>. Traces were truncated to 2  $\mu\text{s}$  to remove the contribution of OMP islands in the background signal<sup>12</sup>. After background subtraction, the distance distributions were generated using a Gaussian model-based co-fitting. The chi-square values were used to evaluate the quality of fittings. The data were plotted using DavePlot (by David Nyenhuis) through a Plotly and Dash frameworks developed by the Plotly corporation. The predicted DEER distributions for T188R1-G399R1 BtuB apo (PDB ID 1NQG) and  $\text{B}_{12}$  (PDB ID 1NQH) bound were generated using *in silico* labeling and simulations using MMM 2017.2 with the standard library for R1 label<sup>13</sup>.

### 5.3.7 Generated BtuB structural images

BtuB crystal structure (1NQG)<sup>4</sup> was obtained from the Protein Data Bank (PDB), [www.rcsb.org](http://www.rcsb.org)<sup>14</sup>. All the BtuB structural images were generated by using the PyMol Molecular Graphics System version 1.8<sup>15</sup>.

## 5.4. Results and Discussion

After optimizing the spin labeling for two cysteine residues within BtuB expressed in *dsbA*<sup>-</sup> cells, we were able to obtain echo signals corresponding to the V90R1 and the T188R1 spin couplings from the DEER experiment. In the extracted distance distribution, two main populations were present, one with a shorter distance of  $2.4 \pm 0.4$  nm and second longer distance at  $3.1 \pm 0.4$  nm. These two distance distributions represent two different conformational states. The longer distance lines up best with that expected based on the crystal structure (1NQG). The shorter distance represents a state where T188R1 of loop two and V90R1 in the core domain have moved closer. This movement could either be due to movement of loop 2 or movement of the core. As the dramatic structural change with substrate is either not seen or is largely depressed in isolated OM preparations where the substrate is known to stay bound to BtuB, indicating that changes in label rotameric states are not the source of the distance changes seen upon substrate binding (Figure 5.2).

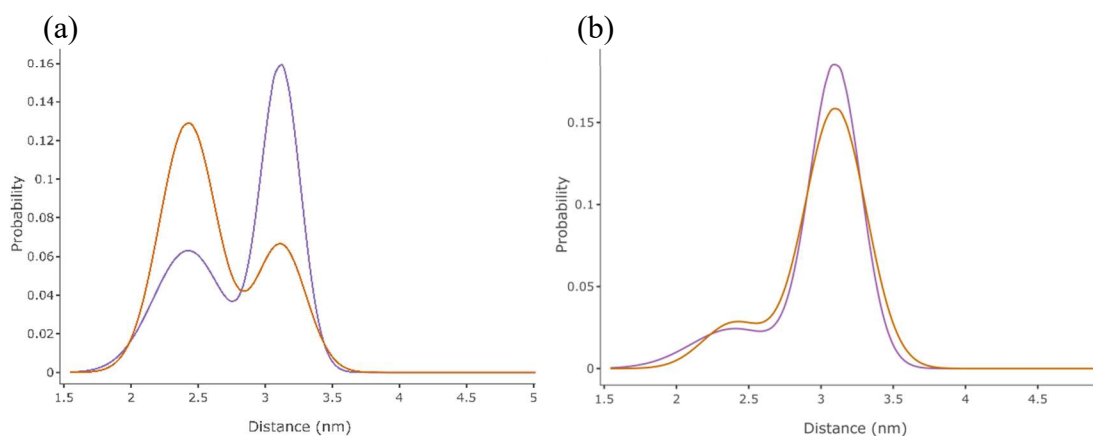


Figure 5.2. The extracted distance distributions from DEER experiments for **apo** and **B<sub>12</sub>** added V90R1-T188R1 BtuB in (a) *E. coli* cells and (b) intact OM.

According to the crystal structure, the other extracellular loops on BtuB should prevent loop two from moving inward and closer to the core (Figure 5.3). Furthermore, as discussed below distance measurements from site 399R1 to 188R1 do not show evidence for two populations or changes in the position of 188R1 with substrate addition.

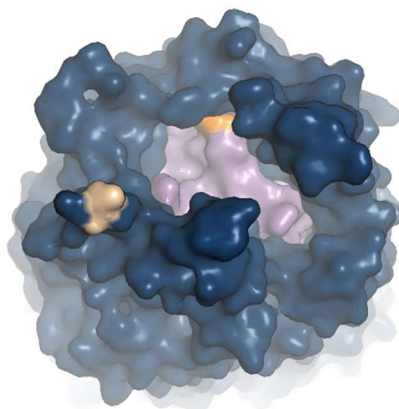


Figure 5.3. Top view of BtuB surface. The T188C accessibility to V90C in BtuB core is limited by the presence of other extracellular loops' dynamics within the barrel.

The two states seen in Figure 5.2(a) are detected with spin labeled samples incubated on ice, as well as at RT during the washing step (Figure 5.4), and the percent population of short distance was similar under each condition. Interestingly, when the substrate was added in excess (100  $\mu$ M B<sub>12</sub>), the short distance population increased and the long-distance population was significantly decreased in the RT samples relative to those on ice. Thus, in the presence of B<sub>12</sub>, the population of this new BtuB conformational state increases. This could potentially be a key state associated with the substrate transport mechanism.

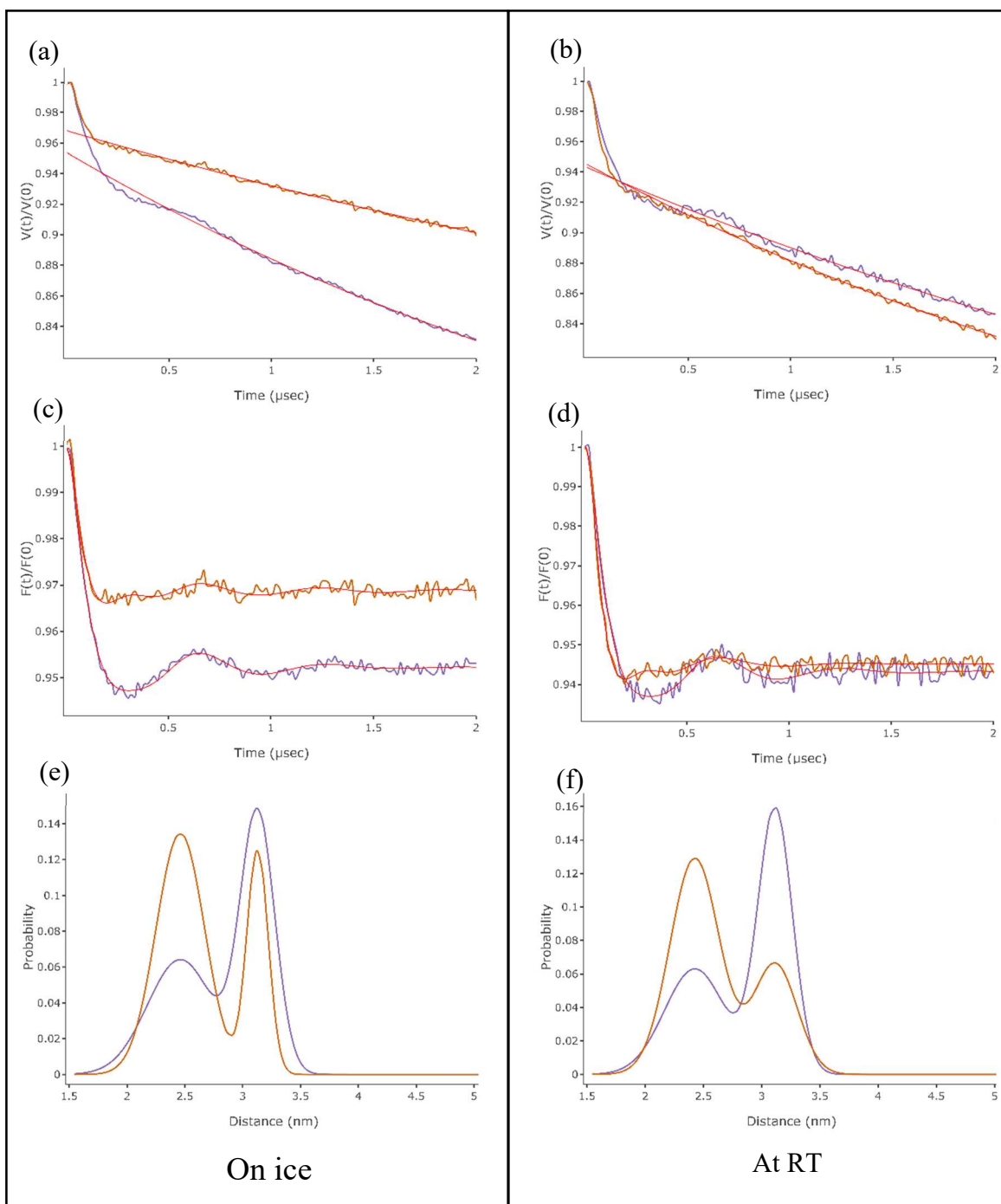


Figure 5.4. DEER signal of V90R1-T188R1 BtuB spin labeled in *dsbA*<sup>-</sup> cells and incubated on ice and at RT during the washing step. (a) and (b) are the echo signals collected, (c) and (d) are the background subtracted form factors and (e) and (f) are the extracted distance distributions for **apo** and **B<sub>12</sub>** added V90R1-T188R1 BtuB.

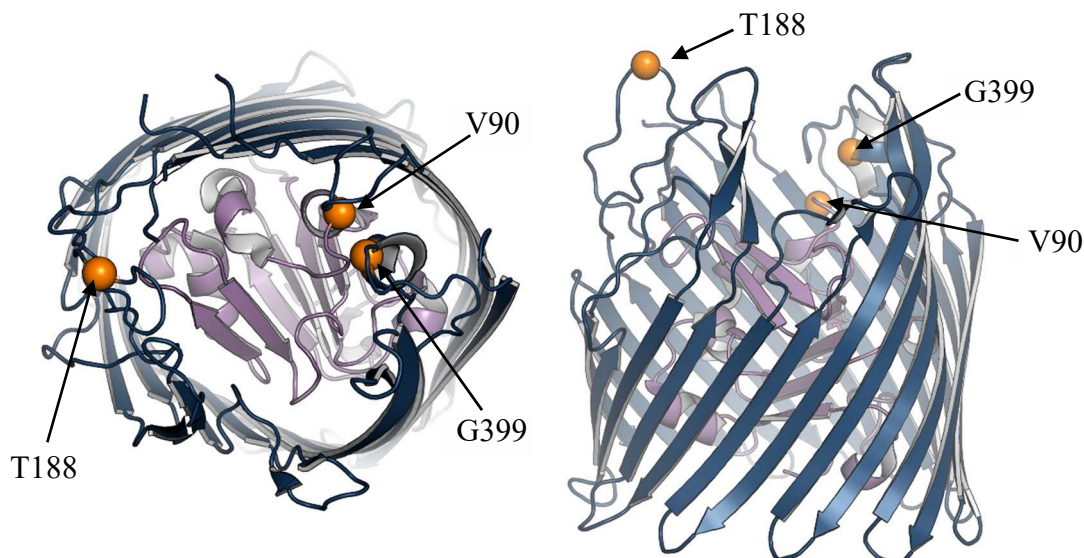


Figure 5.5. Structure of apo BtuB (1NQG) (a) top view and (b) side view showing residues **V90** on **core** and **T188** and **G399** on the **barrel**.

In order to determine whether T188R1 on loop two of BtuB in live *E. coli* cells moves toward the core or whether the C-terminal side of the core moves to the exterior surface, we used another cysteine double mutant, T188C in the loop two and G399C in the loop seven. G399 in the loop seven was selected because in the crystal structure (1NQG) it is in the vicinity of V90 in the core domain. Thus, if the resultant short distance distribution is because loop two T188R1 moves inward, then it should also be detected using the T188R1-G399R1 BtuB pair (Figure 5.5).

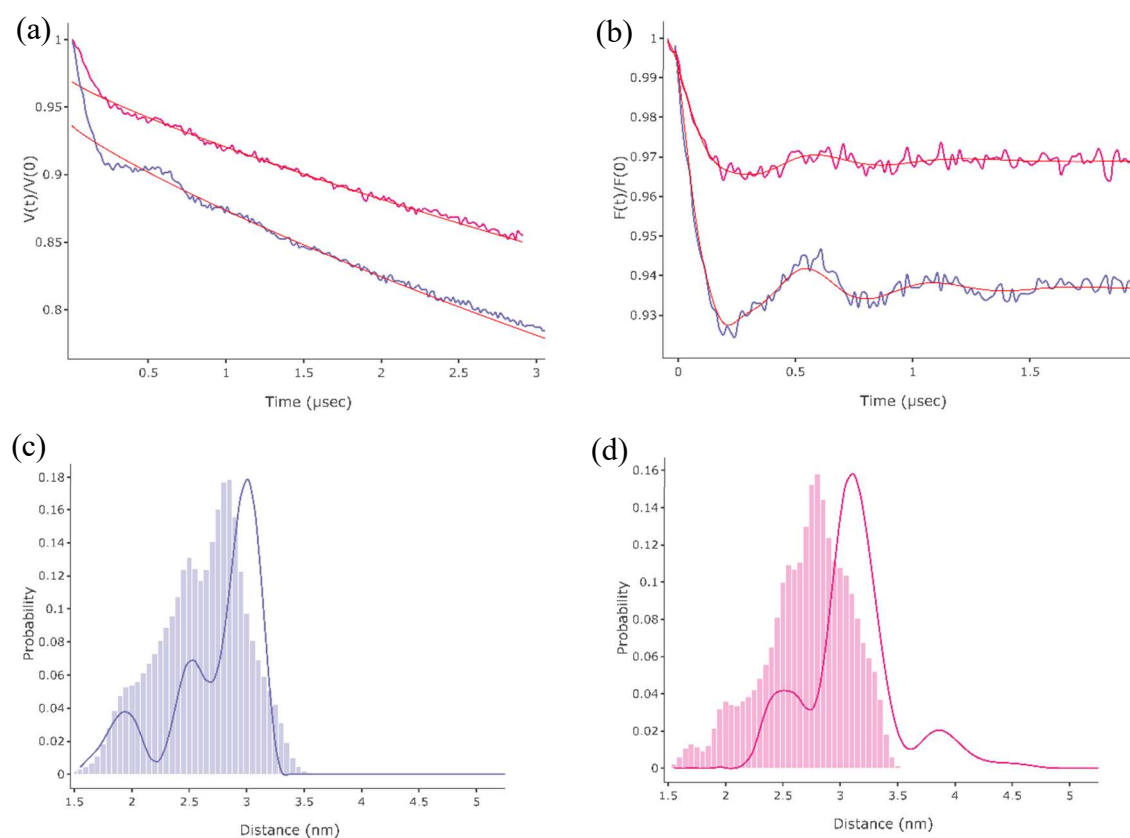


Figure 5.6. DEER signals obtained for T188R1-G399R1 BtuB incubated on ice. (a) The echo signals collected, (b) the background subtracted form factors. (c) and (d) are the extracted distance distributions for **apo** and **B<sub>12</sub>** added V90R1-T188R1 BtuB and the corresponding histograms of distance distributions using MMM simulation for BtuB apo (1NQG) and B<sub>12</sub> bound (1NQH) crystal structures, respectively.

Based on the DEER data along with the MMM rotamer predictions, the extracted distance distributions are within the predicted range for the apo T188R1-G399R1 BtuB (Figure 5.6c). However, the DEER data shows a larger distance distribution population near 30 Å than the MMM prediction. With respect to B<sub>12</sub> bound sample's MMM predicted distance distribution, the DEER data show a larger population around 3.1 nm to 3.5 nm range (Figure 5.6 d). Therefore,

the short distance BtuB conformational component seen in the V90R1-T188R1 BtuB is not due to loop two of BtuB barrel moving inward in live *E. coli* cells.

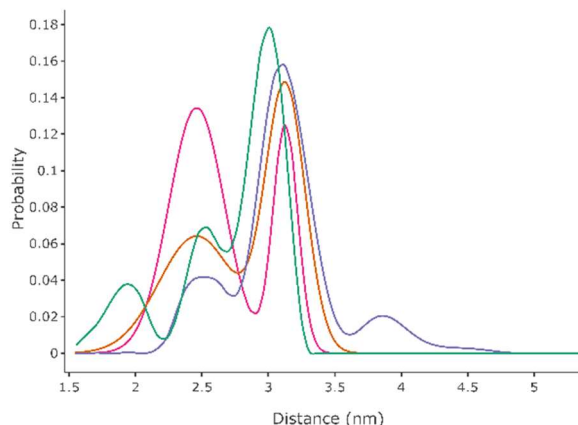


Figure 5.7. Extracted distance distributions from DEER experiments of V90R1-T188R1 BtuB (**apo** and **B<sub>12</sub> bound**) and T188R1-G399R1 BtuB (**apo** and **B<sub>12</sub> bound**).

Taken together, from the extracted distance distributions for V90R1-T188R1 and T188R1-G399R1 BtuB mutants bound to B<sub>12</sub>, the short distance population increases in V90R1-T188R1 BtuB compared to that around 3.1 nm, while in T188R1-G399R1 BtuB, the distance distribution around 3 nm in apo only shifts to ~ 3.1 nm (Figure 5.7). Thus, we speculate that there is a population of BtuB in cells, where a portion of the C-terminus of core is shifted up towards the extracellular space. However, at this time it is difficult to predict whether loops move outward (most likely loop seven) allowing V90R1 of the core to move upward.

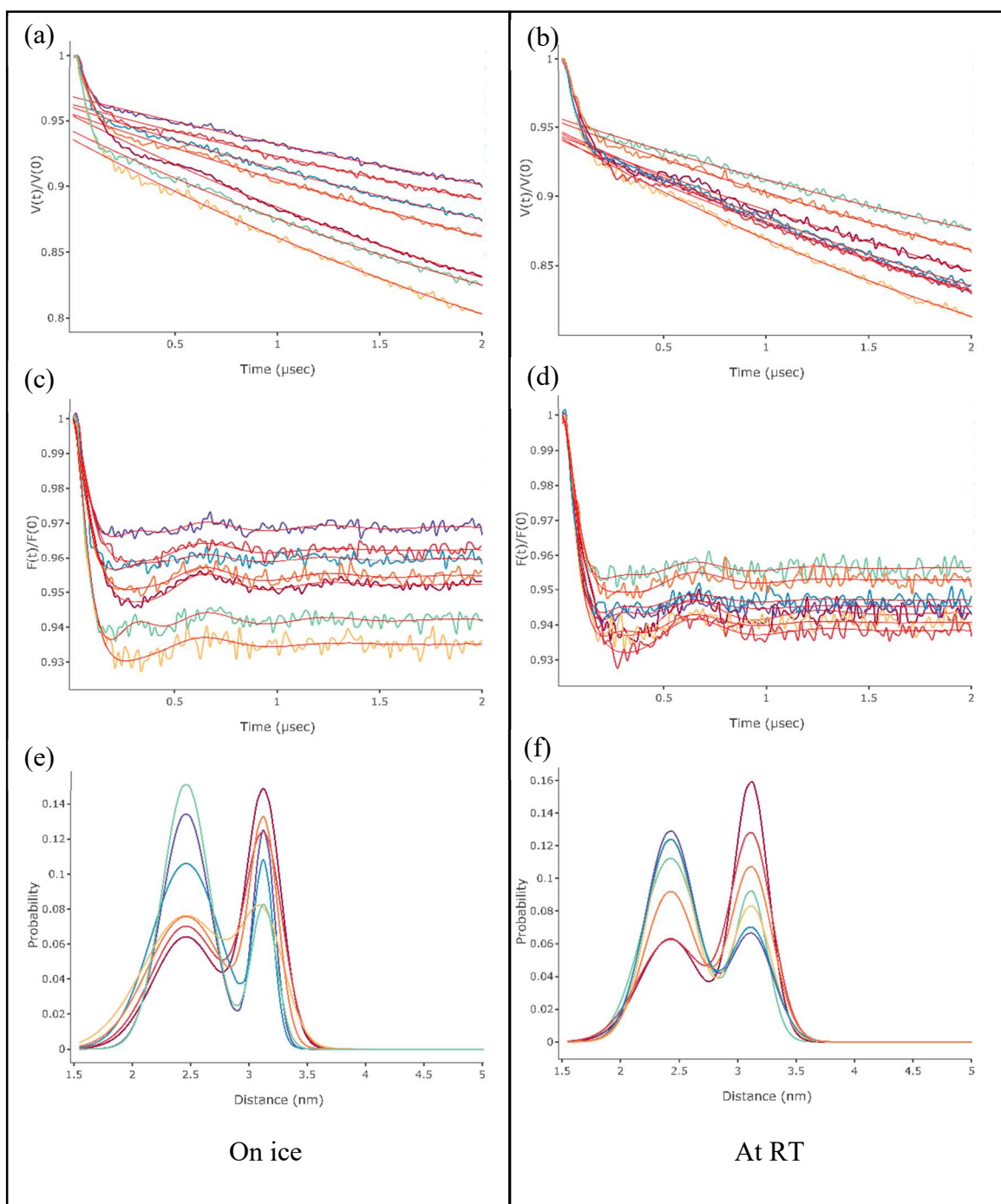


Figure 5.8. DEER signal obtained for V90R1-T188R1 BtuB on ice and at RT with varying the substrate concentration (0, 5, 10, 20, 30, 60 and 100  $\mu\text{M}$   $\text{B}_{12}$ ). (a) and (b) are the echo signals collected, (c) and (d) are the background subtracted form factors and (e) and (f) are the extracted distance distributions for 0, 5, 10, 20, 30, 60 and 100  $\mu\text{M}$   $\text{B}_{12}$  added V90R1-T188R1 BtuB.

The population of short distance representing V90R1-T188R1 conformational state increases in the presence of B<sub>12</sub>. As the B<sub>12</sub> concentrations were increased with V90R1-T188R1 BtuB in the cells incubated both on ice and at RT, the percentage of the short distance component increased and the long-distance component decreased (Figure 5.8). During the transport process, BtuB undergoes substrate induced conformational changes, but the current short distance represents a conformation which is also detected in the apo state. The big question was then, how does this conformational state of BtuB arise? And does this represent a BtuB conformational state that is pre-energized by the TonB complex instead of active transport after the substrate is bound with BtuB? Thus, we designed a few control experiments to explore the potential role of TonB on this new conformational state, by using (i) a transport deficient BtuB mutant (L8P BtuB), and (ii) using an uncoupling agent and an inhibitor of the pmf and electron transport chain (ETC) during the spin labeling process for V90R1-T188R1 BtuB.

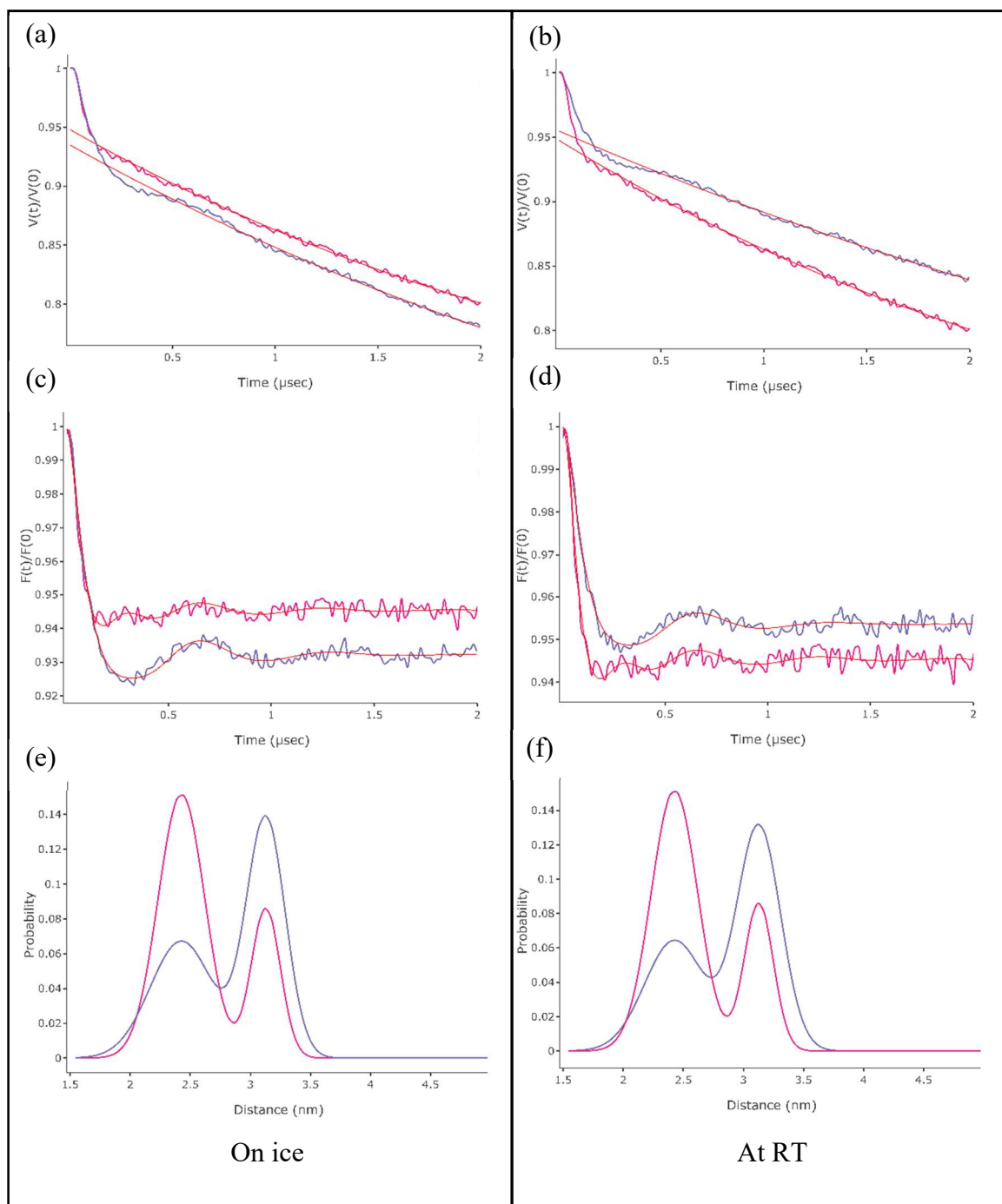


Figure 5.9. DEER signals collected from L8P-V90R1-T188R1 BtuB on ice and at RT with and without 100  $\mu\text{M}$  B<sub>12</sub>. (a) and (b) are the echo signals collected, (c) and (d) are the background subtracted form factors and (e) and (f) are the extracted distance distributions for apo and B<sub>12</sub> added L8P-V90R1-T188R1 BtuB.

L8P BtuB is a transport defective mutant, which has been previously studied using crosslinking experiments<sup>16</sup>. According to these studies, the transport defective mutant L8P, the interaction with TonB is less specific as the addition of a proline residue alters the specific orientation of the TonB box needed to contact TonB, however it could possibly allow for a single round of uptake<sup>16</sup>. Therefore, we used L8P-V90C-T188C BtuB expressed in *dsbA*<sup>-</sup> cells, to study the potential impact on the new BtuB conformational state by preventing the TonB box interaction with TonB. Based on the extracted distance distributions of samples, which were incubated on ice and at RT, both temperatures have similar levels of both the short and long distances in apo and B<sub>12</sub> bound states (Figure 5.9). Thus, this new conformational state is also present in the transport defective BtuB.

Even though L8P has a transport-defective phenotype, the exact mechanism by which it acts is unclear. The Ton box in the L8P mutant may fail to interact with TonB with high-affinity, or the substrate may be trapped in its binding site. Indeed, work in whole cells has shown that B<sub>12</sub> is irreversibly bound in L8P and other transport defective mutants<sup>17</sup>. In live cells, where all the necessary bacterial processes are occurring, the L8P mutant may still interact with TonB as indicated in the crosslinking experiments, and single molecule AFM show that the L8P mutant still has the ability to interact with TonB, albeit with reduced affinity<sup>18</sup>. Thus, we cannot completely eliminate the possibility that TonB energizes the core domain even in the presence of the L8P mutation.

As a second test to identify the pmf and TonB as the source of the shorter distance, CCCP and KCN were used to inhibit TonB from energizing the TonB box. CCCP is a protonophore which disrupts the pmf by diminishing the proton gradient across the IM while CN<sup>-</sup>, an ETC inhibitor, inhibiting cytochrome oxidase<sup>19, 20</sup>. So, addition of these inhibitors should prevent the

IM energy from coupling to BtuB and the eventual creation of new energized populations/states, however existing states/populations at the time of addition will be unaffected.

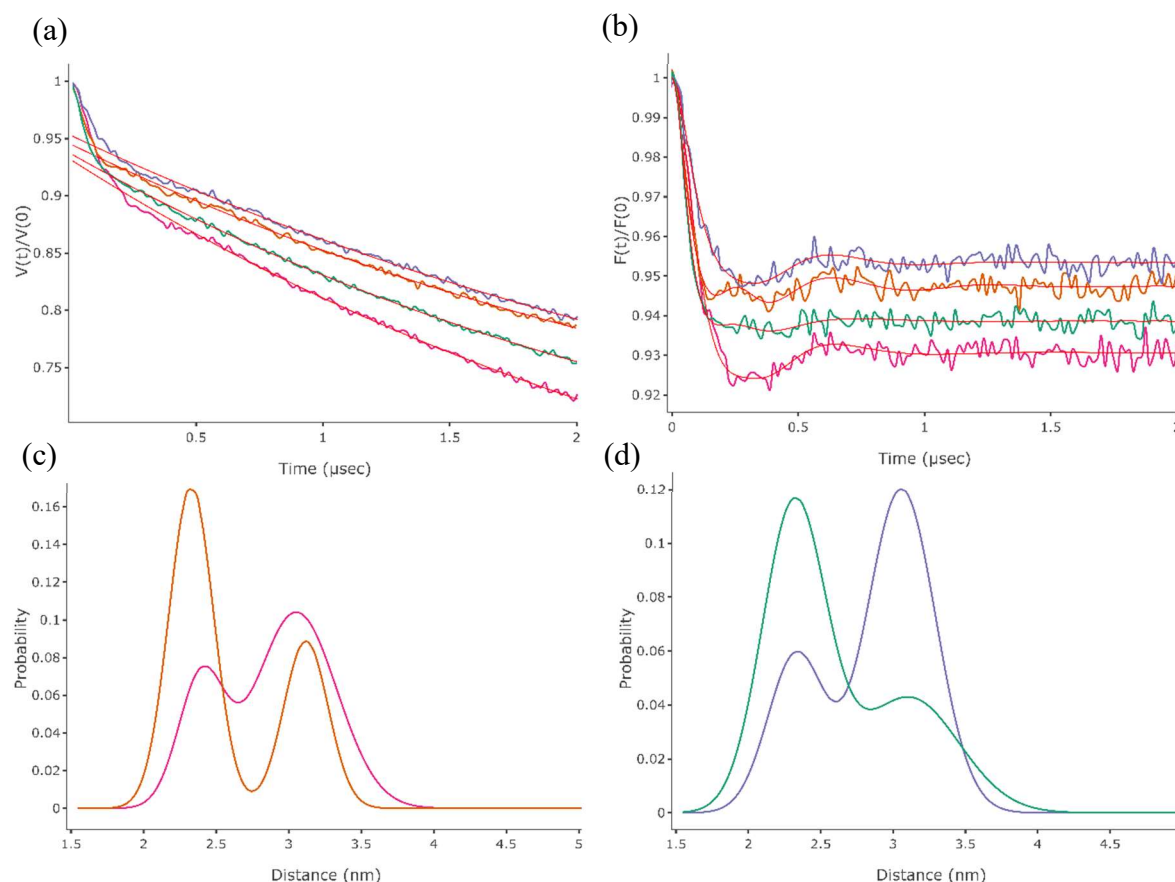


Figure 5.10. DEER signal of V90R1-T188R1 BtuB expressed in *dsbA*<sup>-</sup> cells treated with CCCP (50  $\mu\text{M}$ ) and KCN (1 mM). (a) the echo signals collected, (b) the background subtracted form factors of CCCP and KCN treated cells with spin labeled V90C-T188C BtuB. (c) The extracted distances for V90R1-T188R1 BtuB, **CCCP apo** and **CCCP with B<sub>12</sub>** and (d) the extracted distance distributions for V90R1-T188R1 BtuB, **KCN apo**, and **KCN with B<sub>12</sub>**.

The analyzed pulsed EPR spectra, for both CCCP and KCN treated cells with V90R1-T188R1 BtuB, show there are both short and long distances. Upon addition of B<sub>12</sub> the population

of short distance increases and the long-distance decreases (Figure 5.10). Based on our preliminary work, we cannot fully conclude that the new conformation is not an energized state as we need to further validate our results by modifying our conditions. For example, include of an energy source (glucose) with the cells to ensure that BtuB being energized throughout the experiment. However, this short distance represents a new BtuB conformational state involved in the substrate binding and transportation process, which may or may not be energized by TonB complex.

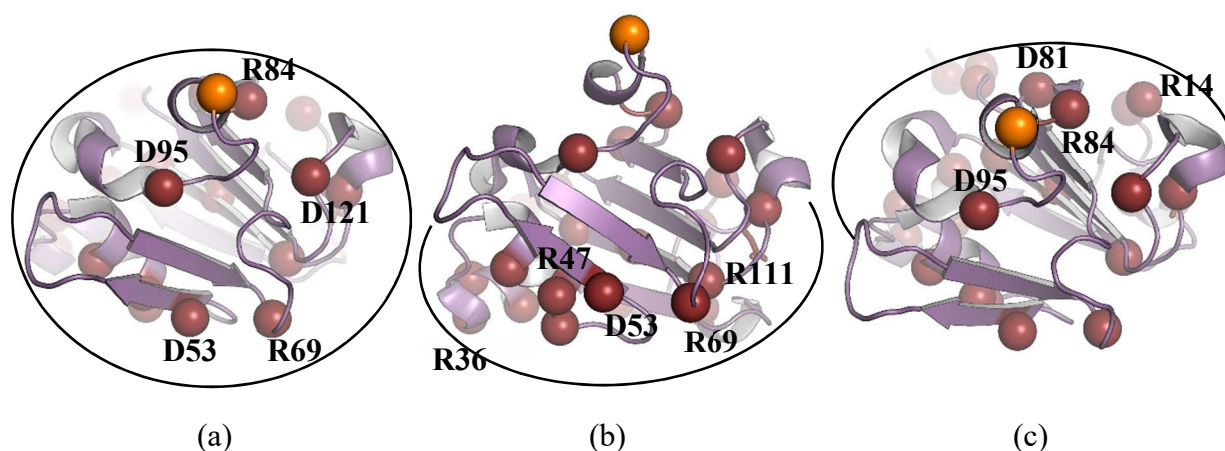


Figure 5.11. Top view of apo BtuB showing **aspartic acid**, **glutamic acid** and **arginine** residues in the **BtuB core** (1NQG). (a) **D53**, **R69**, **R84**, **V90**, **D95**, and **D121** residues are near the top surface of the core. (b) **R36**, **R47**, **D53**, **R69** and **R111** conserved residues are in the opposite side of **V90**. (c) **D81** and **R14** in the same side of **V90** form ion pairs with the barrel.

An examination of the core domain shows that there are 24 charged amino acid residues (aspartic acid, glutamic acid and arginine) which could form salt bridges between the core and the barrel or within the core itself. Salt bridges formed between the core and the barrel would stabilize the core within the barrel. Figure 5.10 shows all of these residues in the core along with

the V90 residue. Analysis of the core reveals that, from the top view, there are five residues (D53, R69, R84, D95, and D121) which can potentially form ion pairs (Figure 5.11a). Residues D53 and R69 form conserved ion pairs with residues R526 and E419 in the barrel. Residue R84 is  $\sim 4.3 - 4.5$  Å away from D121 while residue D95, located in the B<sub>12</sub> binding site and does not have any partnering residue within 4 Å to form a salt bridge. Residue R243 in the barrel is however about 5.5 - 6 Å away from D95 (Figure 5.11a).

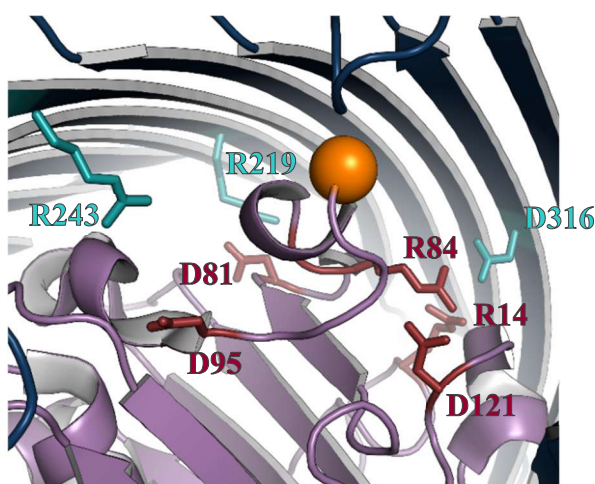


Figure 5.12. The apo BtuB showing potential ion pair forming residues near **V90**. The residues **D81** and **R14** form salt bridges with **R219** and D316, respectively, while **R84** and **D95** could potentially interact with **D121** and **R243** via small structural rearrangements.

As discussed in Chapter 6, ion pairs between the core and the barrel of BtuB play important roles in stabilizing the core within the barrel. They are also involved in propagating the substrate induced signal transduction and/or unfolding of the TonB box, which is required for an interaction with TonB, enabling substrate transport. In Chapter 6, five conserved ion pairs were studied (R36-D515, R47-D548, D53-R526, R69-E419, and R111-E465). Interestingly, all of these ion pairs are localized on the direct opposite side of the site V90 in core (Figure 5.11b). On

the same side as site V90, there are two residues, D81 and R14 in the core, which can form salt bridges with R219 and D316 within the barrel respectively in order to stabilize the core (Figure 5.11c). However, these ion pairs are not near the V90 core region.

Therefore, we hypothesize that two residues R84 and D95 near V90 undergo a structural change to form salt bridges with D121 and R243 (Figure 5.12), thereby moving the core region near V90 upward, which would stabilize BtuB in this state through ion pairs formed by these two residues. Whether such changes can actually bring site V90 about 7 Å nearer to loop two, will be tested in the future. As discussed in Chapter 6, the conserved ion pairs are required for the substrate binding and/or the transmembrane signal transduction processes. Mutating these residues into an alanine blocks the proper signal transduction process and inhibits the TonB box unfolding events. Then, upon substrate binding, substrate induced signal transduction propagates through these ion pairs, which are acting as conformational switches. Thus, one can speculate that the D53-R526 and R69-E419 conserved ion pairs are destabilizing which propagates the signal down to the periplasmic side of the protein. Therefore, in the presence of B<sub>12</sub> these conformational rearrangements could promote an increased short distance population resulting from the core region near V90 moving upward.

As the signal propagates downward, finally the conformational switch, R14-D316, will be destabilized. This ion pair has been previously studied using the two mutant cycle coupled to SDSL-EPR. According to the study, the R14-D316 conformational switch regulates the unfolding of the TonB box and its extension into the periplasm<sup>21</sup>. Thus, it is possible that this destabilization also aids the increasing short distance population even without the TonB interacting with the unfolded TonB box. On the other hand, as R14-D316 destabilizes, the

unfolded TonB box can interact with TonB which can energize the core, resulting in a significant BtuB population shift into the new state.

## 5.5 Future directions

Throughout these experiments, one thing we have learned is that we cannot plan too far ahead with experiments without knowing the recently performed or near future experiments' results. Our current knowledge of this novel conformational state is based on the DEER experiments which have been done up to this date. Thus, this chapter is only based on those experiments' results. However, this current knowledge will be subject to change as new data is collected. Below are the nearby future experiments which aim to deepen our understanding of this new conformational state.

All two cysteine mutants are studied in the Dsb mutant strains, as it was necessary to use these cell lines to spin label cysteine double mutants (Chapter 3). Even though many studies have been done in the isolated systems for BtuB, all these have used BtuB expressed in the RK5016 cell line. Since we have not done intact OM or reconstituted lipid bilayer studies for V90R1-T188R1 BtuB purified from *dsbA*<sup>-</sup> and *dsbB*<sup>-</sup>, it will be interesting to investigate if the short distance is present in these cases, in the apo, as well as, in the B<sub>12</sub> added state. These will allow us to understand the potential process(es) that result in and produce this conformation. If the outcome of these experiments is different than what we have observed in the RK5016 cell line it will be important to investigate why. Another future experiment set will be to compare the Dsb mutant cell lines we are using with respect to the RK5016 cells in order to understand any potential impact on the OM composition in these cell lines. Because, LptD-E inserts lipopolysaccharides into the outer leaflet of the OM and LptD is a substrate of DsbA and DsbC

in the disulfide bond formation pathway<sup>22</sup>. Moreover, it is also known that removing the functioning LptD-E complex results in insertion of phospholipids instead of LPS<sup>23</sup>. Therefore, assessing the lipid composition in the OM will allow us to understand any potential impacts of the OM in these cell lines and also whether or not the new conformational state is influenced by the lipid composition present.

In addition to above studies, it will be interesting to determine the role of the pairs: D53-R526, R69-E419, and R14-D316 with respect to short distance conformation of BtuB. Mutating these ion pair residues into alanine mutants in the V90C-T188C BtuB construct will enable us to explore the potential roles of each residue. If these ion pair mutants promote the short distance population in the apo state of V90R1-T188R1 BtuB, this will support the vital role of each residue with respect to new conformational state(s) associated with the substrate induced signal transduction process. In addition to the use of L8P-V90R1-T188R1 BtuB, CCCP and KCN treated V90R1-T188R1, cells can be grown with supplemented TonB box pentapeptide<sup>24</sup> in order to inhibit interactions with BtuB and to study its impact on each conformational state. This could be a very convenient strategy to see if TonB produces a pre-energized population of BtuB or not without affecting other cellular processes.

As the short distance component increases in the presence of B<sub>12</sub> while reducing the long-distance component, it is unclear if these two components return to their original populations over the course of time. Or will this allow us to isolate other new distance distributions? Thus, in addition to the above experiments, we will also explore the impact on both the short and long-distance populations in the presence of B<sub>12</sub> over time to extract information related to the functionality of each conformational state. Furthermore, we are planning to incorporate other cysteine double mutants, such as S74C-T188C, to better screen and to explore this

conformational state, as well as, to identify any more conformational states which might be involved in the resting and active states of apo BtuB and the substrate transport process.

## 5.6 References

- [1] Koebnik, R., Locher, K. P., and Van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell, *Mol Microbiol* 37, 239-253.
- [2] Nikaido, H. (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited, *Microbiology and Molecular Biology Reviews* 67, 593-656.
- [3] Konovalova, A., Kahne, D. E., and Silhavy, T. J. (2017) Outer Membrane Biogenesis, *Annu Rev Microbiol* 71, 539-556.
- [4] Chimento, D. P., Mohanty, A. K., Kadner, R. J., and Wiener, M. C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB, *Nature structural biology* 10, 394-401.
- [5] Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010) TonB-dependent transporters: regulation, structure, and function, *Annual review of microbiology* 64, 43-60.
- [6] Flores Jimenez, R. H., and Cafiso, D. S. (2012) The N-terminal domain of a TonB-dependent transporter undergoes a reversible stepwise denaturation, *Biochemistry* 51, 3642-3650.
- [7] Mokdad, A., Herrick, D. Z., Kahn, A. K., Andrews, E., Kim, M., and Cafiso, D. S. (2012) Ligand-induced structural changes in the Escherichia coli ferric citrate transporter reveal modes for regulating protein-protein interactions, *Journal of molecular biology* 423, 818-830.
- [8] Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex, *Science* 312, 1396-1399.
- [9] Fanucci, G. E., Coggs, K. A., Cadieux, N., Kim, M., Kadner, R. J., and Cafiso, D. S. (2003) Substrate-Induced Conformational Changes of the Periplasmic N-Terminus of an Outer-Membrane Transporter by Site-Directed Spin Labeling, *Biochemistry* 42, 1391-1400.
- [10] Cadieux, N., Barekzi, N., and Bradbeer, C. (2007) Observations on the calcium dependence and reversibility of cobalamin transport across the outer membrane of Escherichia coli, *The Journal of biological chemistry* 282, 34921-34928.
- [11] Stein, R. A., Beth, A. H., and Hustedt, E. J. (2015) A Straightforward Approach to the Analysis of Double Electron-Electron Resonance Data, *Methods in enzymology* 563, 531-567.
- [12] Kattinig, D. R., Reichenwallner, J., and Hinderberger, D. (2013) Modeling Excluded Volume Effects for the Faithful Description of the Background Signal in Double Electron-Electron Resonance, *The Journal of Physical Chemistry B* 117, 16542-16557.
- [13] Polyhach, Y., Bordignon, E., and Jeschke, G. (2011) Rotamer libraries of spin labelled cysteines for protein studies, *Physical chemistry chemical physics : PCCP* 13, 2356-2366.

- [14] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Research* 28, 235-242.
- [15] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.
- [16] Cadieux, N., Bradbeer, C., and Kadner, R. J. (2000) Sequence Changes in the Ton Box Region of BtuB Affect Its Transport Activities and Interaction with TonB Protein, *Journal of Bacteriology* 182, 5954-5961.
- [17] Nazir, B. (2009) Mutational Analysis of the Proteins BtuB and TonB: Their Roles in Cobalamin Transport, In *Department of Microbiology*, p 202, The University of Virginia, University of Virginia Library.
- [18] Hickman, S. J., Cooper, R. E. M., Bellucci, L., Paci, E., and Brockwell, D. J. (2017) Gating of TonB-dependent transporters by substrate-specific forced remodelling, *Nature Communications* 8, 14804.
- [19] Krulwich, T. A., Quirk, P. G., and Guffanti, A. A. (1990) Uncoupler-resistant mutants of bacteria, *Microbiological Reviews* 54, 52-65.
- [20] Cooper, C. E., and Brown, G. C. (2008) The inhibition of mitochondrial cytochrome oxidase by the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical mechanism and physiological significance, *Journal of Bioenergetics and Biomembranes* 40, 533.
- [21] Lukasik, S. M., David Ho, K. W., and Cafiso, D. S. (2007) Molecular Basis for Substrate-Dependent Transmembrane Signaling in an Outer-Membrane Transporter, *Journal of molecular biology* 370, 807-811.
- [22] Ren, G., Champion, M. M., and Huntley, J. F. (2014) Identification of disulfide bond isomerase substrates reveals bacterial virulence factors, *Mol Microbiol* 94, 926-944.
- [23] Rojas, E. R., Billings, G., Odermatt, P. D., Auer, G. K., Zhu, L., Miguel, A., Chang, F., Weibel, D. B., Theriot, J. A., and Huang, K. C. (2018) The outer membrane is an essential load-bearing element in Gram-negative bacteria, *Nature* 559, 617-621.
- [24] Tuckman, M., and Osburne, M. S. (1992) In vivo inhibition of TonB-dependent processes by a TonB box consensus pentapeptide, *Journal of Bacteriology* 174, 320-323.

## **Chapter 6: Exploring the potential roles of BtuB conserved ion pairs during the signal transduction and/or substrate transport process**

### 6.1 Abstract

*Escherichia coli* BtuB has many salt bridges which hold the core within the barrel. According to two mutant cycle analysis coupled to site directed spin labeling (SDSL) EPR studies, BtuB ion pair, R14-D316 acts as a conformational switch that signals the TonB box unfolding process during the substrate transport process. Here, five conserved ion pairs (R36-D515, R47-D548, D53-R526, R69-E419, and R111-E465) were studied using the same method in order to investigate their effect on the TonB box unfolding. In contrast to R14-D316, these ion pair mutants inhibit the TonB box unfolding even in the presence of B<sub>12</sub>. Interestingly, the R69, R111, E419, and E465 sites form interchangeable salt bridges. This sequence of salt bridge formation propagates the signal through E419-R69, R69-E465, and then E465-R111. Thus, elimination of both R111 and E419 can destabilize the core, which leads to the unfolding of the TonB box in the apo state.

As was performed previously, a two mutant cycle was used to identify interacting ion pair residues based on the conformational equilibrium constants for the TonB box folding and unfolding in the apo and B<sub>12</sub> bound states<sup>1</sup>. However, for some of the investigated ion pairs this technique did not reliably determine whether the residues were interacting, as these residues inhibit the TonB box unfolding. A subset of alanine mutants was also used to investigate spin labeled B<sub>12</sub> binding, to see if the TonB box unfolding was inhibited due to loss of substrate binding. Based on the substrate binding, R47A-D548A-V10C and R36A-D515A-V10C mutants have altered the substrate binding.

## 6.2 Introduction

A BtuB ortholog present in *Escherichia coli* (*E. coli*) OM)<sup>2, 3</sup>. which has been studied since 1970s after Professor Bradbeer, University of Virginia, reinitiated the investigations of vitamin B<sub>12</sub> uptake in bacteria which was frozen since the 1950s<sup>4</sup>. Upon substrate binding, the TonB box of BtuB undergoes substrate induced conformational changes. SDSL EPR studies of the TonB box region have previously revealed that the folded TonB box unfolds into the periplasm during substrate binding<sup>5, 6</sup>. During this unfolding event, substrate binding induces a signal transduction in the core starting from the extracellular binding site that propagates toward the TonB box on the periplasmic side. Conformational changes facilitate this signal transduction across the membrane through ion pairs within the protein which are known as molecular switches<sup>1</sup>. As discussed in Chapter 5.4, there are many ion pairs between the core and the barrel which are holding the core inside of the barrel<sup>7</sup>.

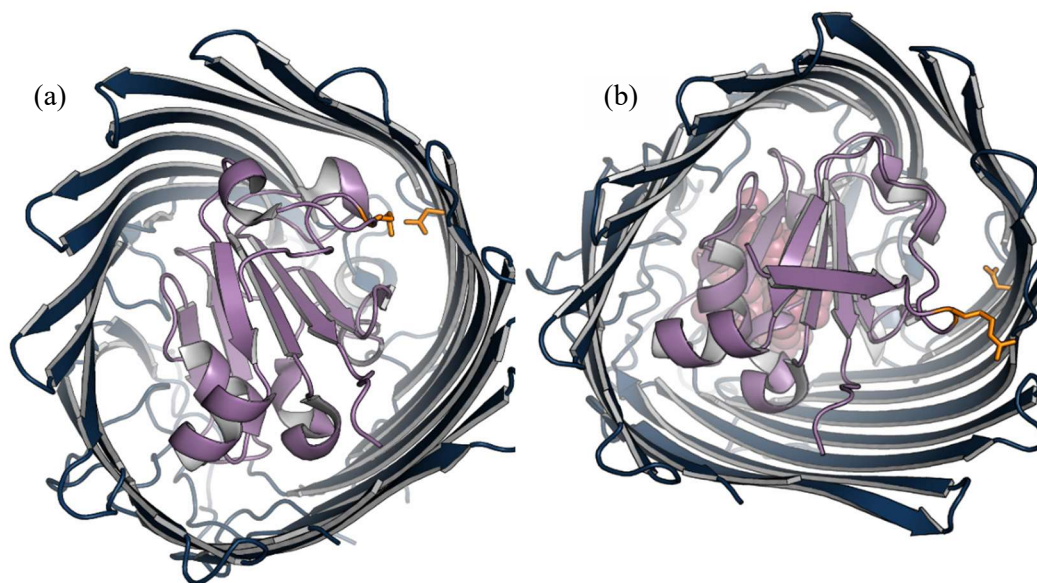


Figure 6.1 Structures of BtuB periplasmic view (a) apo (1NQG) and (b) B<sub>12</sub> bound (2GSK). **R14-D316** ion pair between the **core** and the **barrel** forming hydrogen bonds in the apo state that are weakened by substrate (**B<sub>12</sub>**) induced conformational changes.

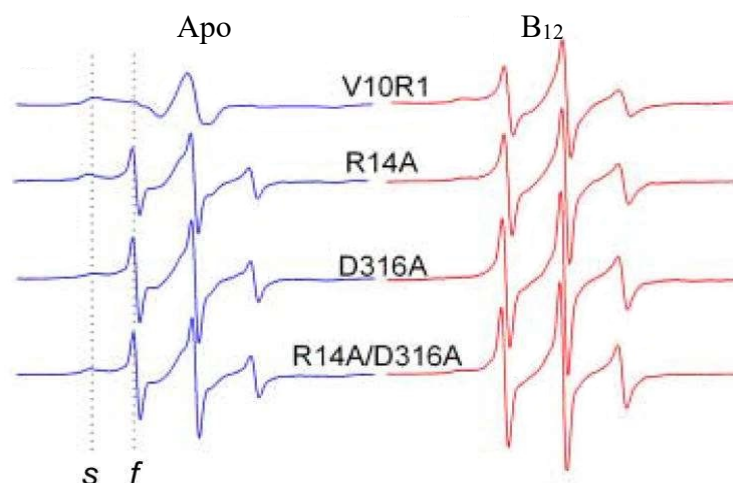


Figure 6.2. Normalized CW EPR spectra of **apo** and **substrate bound** BtuB. Mutating the R14-D316 ion pair, the single and double alanine mutants unfold V10R1 even without the substrate binding. Mobile (f) and immobile (s) components of the spectra are due to nitroxide spin dynamics of unfolded and folded V10R1 in the TonB box. Figure is from Lukasik, David Ho and Cafiso. (2007). *Journal of molecular biology*. **370**. 087-811.

Comparing the apo (1NQG) and B<sub>12</sub> bound (2GSK) BtuB crystal structures, there are various conformational changes associated with (potential) substrate transport across the BtuB. The ion pair R14-D316, between the core and the barrel, forms hydrogen bonds which hold the core within the barrel (Figure 6.1a). However, when B<sub>12</sub> is bound this interaction is broken which results in propagating the transduction signal for TonB box unfolding (Figure 6.1b). This ion pair was studied using two mutant cycle analysis coupled to SDSL EPR. Since the spin label dynamics are influenced by restriction in their environments and their secondary structure, the label at the TonB box (V10C) in the folded and unfolded states produces unique EPR spectra. These were used to study the role of the R14-D316 ion pair upon substrate binding with respect to the apo BtuB (Figure 6.2). According to the studied alanine mutants of the R14-D316 ion pair,

this pair forms ionic bonds in the apo BtuB, which are weakened and broken by substrate induced conformational changes, leading to TonB box unfolding<sup>1</sup>.

The crystal structures provide insight into stable conformations, but they lack information concerning intrinsic changes associated with the signal transduction and the substrate transportation. There are many other ion pairs within the barrel and the core that can also potentially mediate signal transduction processes as well. Based on sequence alignments with other BtuB orthologs, there are a few conserved ion pairs. These include: R36-D515, R47-D548, R69-E419, R111-E465, and D53-R526 of *E. coli* BtuB (Figure 6.3). The R69, R111, E419 and E465 residues are highly conserved among the BtuB orthologs<sup>2</sup>. Thus, these five ion pairs were chosen to be studied by two mutant analysis coupled SDSL-EPR.

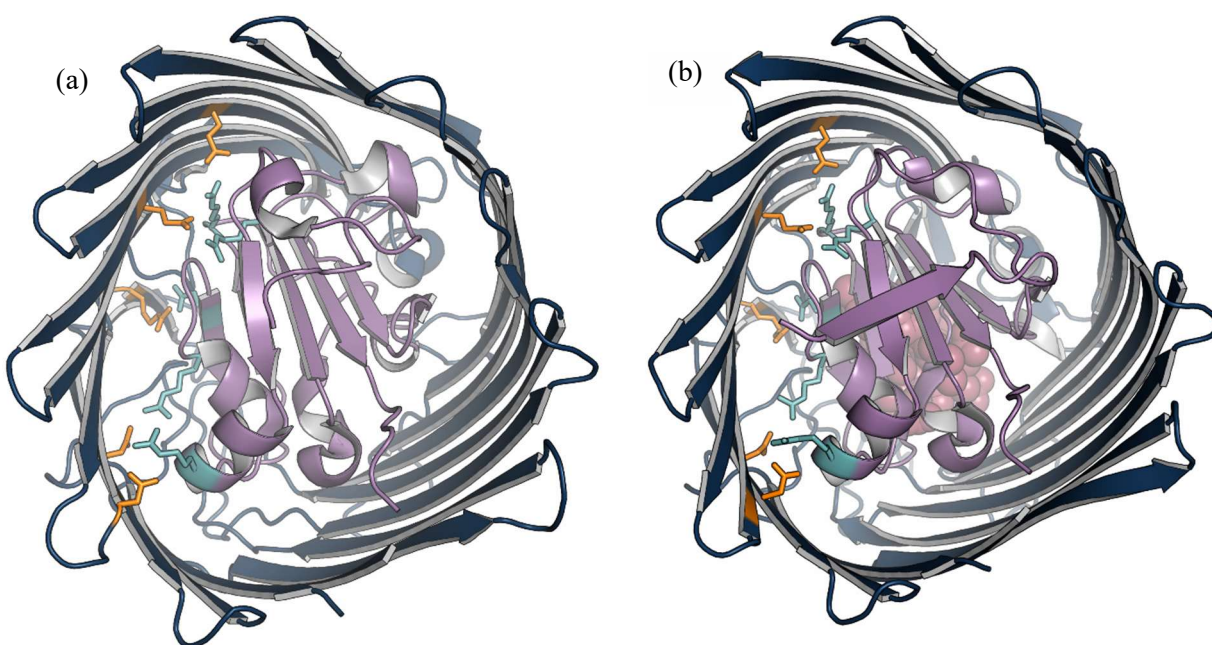


Figure 6.3 Conserved ion pairs in BtuB. Structures of BtuB periplasmic view of (a) apo and (b) **B<sub>12</sub>** bound. Residues **R36**, **R47**, **D53**, **R69** and **R111** on the **core** potentially interact with **D515**, **D548**, **R526**, **E419** and **E465** on the **barrel**.

## 6.3 Methods

### 6.3.1 Expression of BtuB mutants

Minimal media (100 mM phosphate buffer with 8 mM  $(\text{NH}_4)_2\text{SO}_4$  and 2 mM sodium citrate) supplemented with 100  $\mu\text{g/mL}$  ampicillin, 0.2 % w/v glucose, 150  $\mu\text{M}$  thiamine, 3 mM  $\text{MgSO}_4$ , 300  $\mu\text{M}$   $\text{CaCl}_2$ , 0.01 % w/v methionine and 0.01 % w/v arginine were used to grow RK5016 cells expressing BtuB mutants. Briefly, pre-cultures were directly inoculated with small aliquots of their corresponding glycerol stocks, incubated at 34° C overnight (O/N), and then used to inoculate main cultures at 34° C for about 8 hrs. The cell cultures were centrifuged at 6000 rpm for 10 min at 4° C using a Sorval SLA-3000 rotor and then harvested cells were either stored few days at -20° C or directly used to purify BtuB.

### 6.3.2 Purification and spin labelling of BtuB mutants

Cell pellets were resuspended in 20 mL of BtuB lysis buffer (10 mM HEPES pH 6.5 buffer) containing 10 nM of 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF). Cells were then lysed two to three times by French press and centrifuged at 12000 rpm for 20 min at 4° C using a Sorval SS-34 rotor in order to remove cellular debris. In order to isolate intact OM, 5 % sarkosyl was added to the supernatant which was then centrifuged at 32000 rpm (Beckman Coulter, Ti45 rotor) for 90 min at 4° C. Resultant OM pellet was resuspended in 5 mL of BtuB lysis buffer. The outer membranes (OMs) were solubilized in 0.5-0.7 g of OG containing 8 mM EDTA and 100 mM Tris pH 8.0. First, the OM suspension was incubated for 10 min at 37° C and then 2 hrs at room temperature (RT) followed by centrifugation at 25000 rpm (Beckman Coulter, Ti70 rotor) for 60 min at 4° C.

The supernatant was used to spin label solubilized V10C BtuB ion pair mutants within OG micelles. 200  $\mu\text{L}$  of a freshly prepared 12 mM MTSSL stock was added to the supernatant which

was then incubated for two to three hrs at RT in the dark. During the spin labeling reaction, the R1 side chain is linked to the sulfhydryl group of the cysteine (V10C) by disulfide bond formation. A subset of each BtuB mutant was purified without spin labeling in order to study spin labeled B<sub>12</sub> binding. At the end of the incubation, spin labeled BtuB was purified using 6 CV of BtuB wash buffer, WB (17 mM OG, 25 mM Tris pH 8.0), 12 CV of 0 -100 % gradient of BtuB elution buffer, EB (1 M NaCl, 17 mM OG, 25 mM Tris pH 8.0), and 6 CV of 100 % BtuB EB using a Q column. Sample fractions from the purification were analyzed by SDS-PAGE gel electrophoresis and the fractions containing BtuB were pooled and concentrated the down to about 30 mL.

#### 6.3.3 Reconstitution of BtuB into lipid vesicles

POPC (20 mg/mL) was solubilized in the presence of OG in the reconstitution buffer (150 mM NaCl, 100 nM EDTA, 10 mM HEPES pH 6.5) and then added to purified BtuB samples. BtuB was reconstituted into liposomes by dialysis over six buffer exchanges using reconstitution buffer and bio-beads (with minimum of six hrs dialysis per exchange). Reconstituted BtuB was pelleted by centrifugation at 14000 rpm for 40 min at 4° C using a Sorval SS-34 rotor and then resuspended in 200 µL of reconstituted buffer and further concentrated by using Beckman airfuge (20-25 psi for 30 min). The samples were frozen and stored at -80° C.

#### 6.3.4 (a) CW EPR analysis of substrate induced conformation changes of V10R1 BtuB

A 6 µL aliquot from the reconstituted V10R1 BtuB samples was loaded into glass capillaries (0.84 mm O.D. / 0.6 mm I.D.) and CW EPR spectra were collected using an X-band operating Bruker EMX spectrometer with an ER 4123D dielectric resonator. A 100 G sweep width, 1 G

modulation, and 2 mW of incident microwave power was used. In parallel, 100 mM B<sub>12</sub> was added to the reconstituted V10R1 BtuB ion pair mutants, three freeze-thaw cycles followed, then CW-EPR was run. All data was collected at RT.

#### 6.3.4 (b) TEMPO-B<sub>12</sub> binding with reconstituted ion pair mutant V10C BtuB

A 10  $\mu$ L aliquot from reconstituted V10C BtuB ion pair mutant samples were mixed resulting in different TEMPO-B<sub>12</sub> final concentrations (10, 5, and 2.5  $\mu$ M) which were then followed by three freeze-thaw cycles. The 6  $\mu$ L samples were then loaded into glass capillaries and CW EPR spectra were collected at RT.

#### 6.3.4 (c) TEMPO-B<sub>12</sub> binding with ion pair mutant V10C BtuB in intact cells

A subset of V10C BtuB ion pair mutants were grown in minimal media until OD<sub>600</sub> to around 0.6 and then 50 mL aliquots of cells were harvested and centrifuged at 4000 rpm for 10 min at 4° C. Harvested cells were resuspended in 100 mM HEPES, pH 7.0 and then centrifuged again. The washing step was repeated one more time then finally, cell pellets were stored on ice until CW EPR analysis. 10  $\mu$ M TEMPO-B<sub>12</sub> sample was added to the cells, which were incubated for 15 min followed by a washing step in order to remove excess unbound TEMPO-B<sub>12</sub>. At the end of the incubation, cells were loaded into a capillary tube and analyzed by CW EPR at RT.

#### 6.3.4 (d) CW EPR data analysis

All the collected CW-EPR spectra were normalized, phase corrected, and plotted using various in-house programs (DavePlot CW and CWPhase by David Nyenhuis). In order to obtain the TonB box folded and unfolded populations from the CW EPR spectra of V10R1 BtuB in the reconstituted system each spectrum was subtracted using FreeBee software (by Christian Altenbach). The percent populations were then used to calculate the free energies ( $\Delta^\circ G$ ) and then

to obtain the free energy difference between the mutants ( $\Delta\Delta G$ ) within the two mutant cycle analysis:

$$k = x/(100 - x)$$

$$\Delta G = -RT \ln k$$

Where x is the population percentage of the unfolded TonB box, (100-x) is then the folded population percentage, and k is conformational equilibrium constant.

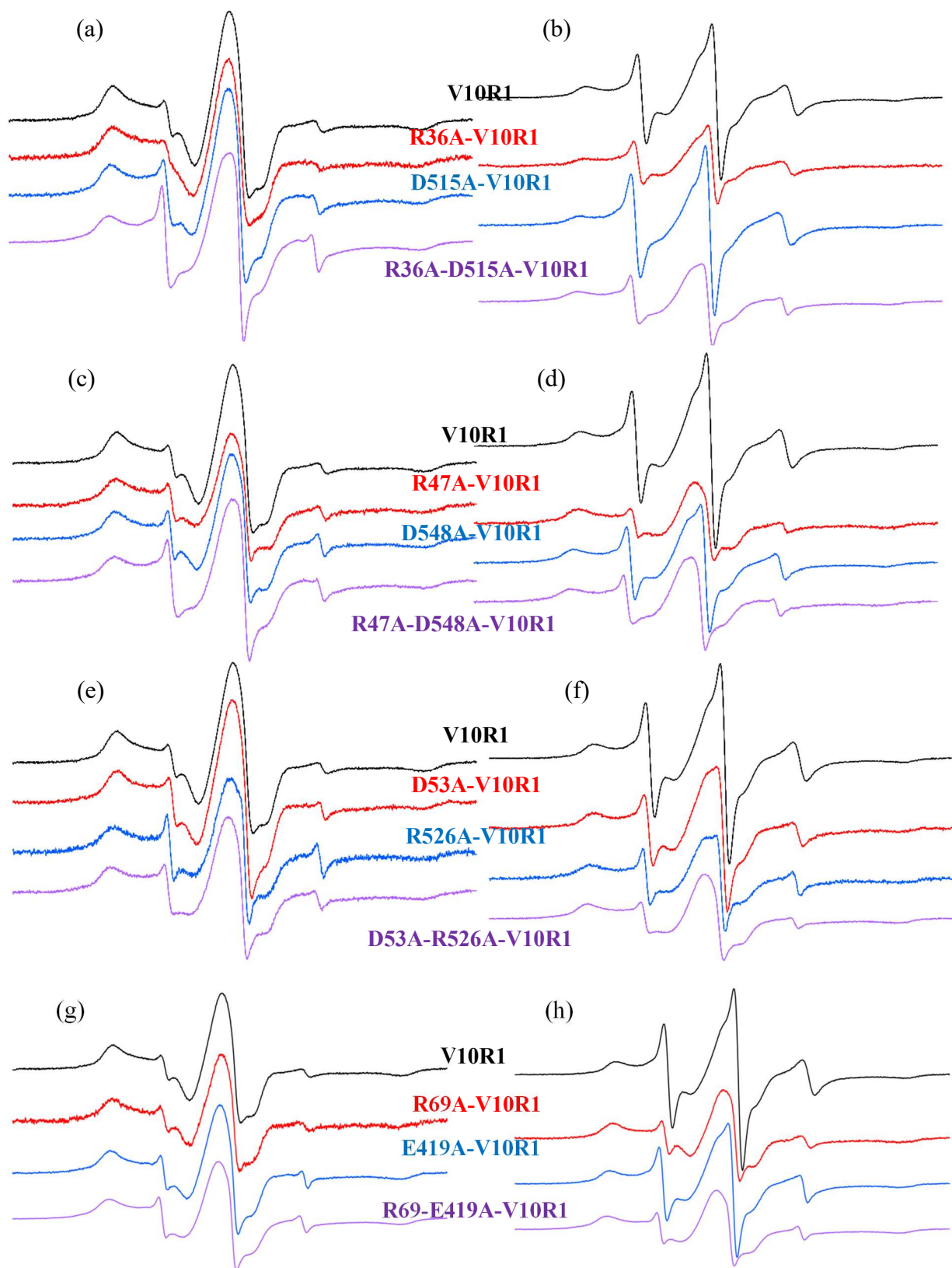
### 6.3.5 Generated BtuB images

Crystal structures of BtuB (1NQG and 2GSK)<sup>7, 8</sup> were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>9</sup> and the images were created using PyMol Molecular Graphics System version 1.8<sup>10</sup>.

## 6.4 Results and discussion

Upon the substrate binding, BtuB undergoes substrate induced conformational changes which unfold the TonB box and extends it into the periplasm. Previous studies using SDSL EPR have demonstrated that during these conformational changes, V10R1 of TonB box can be used to monitor these folding and unfolding events<sup>5</sup>. When V10R1 is in the folded state, the resultant EPR spectra is broad, indicating that V10R1 is in tertiary contact within the BtuB barrel and the core. In the unfolded state the V10R1 spectrum is narrow due to motional averaging of the nitroxide<sup>7</sup>.

Based on the normalized CW EPR spectra collected from each ion pair's single and double alanine mutants probing with V10R1 versus a mutant-less BtuB control (V10R1), CW EPR spectra consist of both a highly immobile population and a mobile population representing the TonB box's folded and unfolded states. In the apo state for V10R1 BtuB, the TonB box is almost fully folded resulting in the broad CW EPR spectra with low amplitude. This is observed for all alanine mutants, with the exception of E465A-V10R1 BtuB, where a larger population of the TonB box is unfolded than in V10R1-BtuB (Figure 6.4).



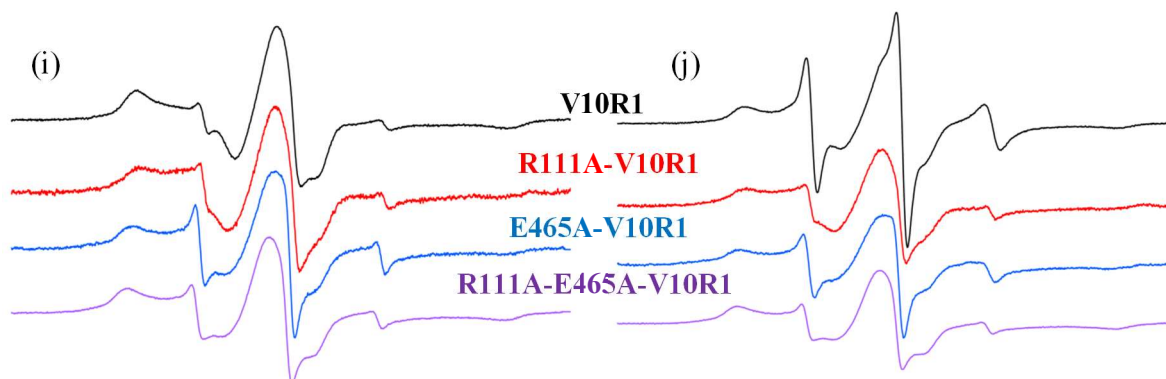


Figure 6.4. Normalized EPR spectra of alanine mutations coupled to V10R1 BtuB. R36A-D515A-V10R1 (a) apo and (b) B<sub>12</sub> bound, R47A-D548A-V10R1 (c) apo and (d) B<sub>12</sub> bound, D53A-R526A-V10R1 (e) apo and (f) B<sub>12</sub> bound, R69A-E419A-V10R1 (g) apo and (h) B<sub>12</sub> bound and R111A-E465A-V10R1 (i) apo and (j) B<sub>12</sub> bound.

Substrate binding induces conformational changes which result in TonB box unfolding. Thus, when probing with V10R1 in BtuB, the unfolded TonB box population increases, which results in an increase of the narrow and high amplitude components in the EPR spectrum. However, in all alanine mutants in this study, the percent of mobile population is inhibited. The degree of this inhibition differs from one mutant to the next, but overall each of these mutations could potentially be transport defective. Overall, D515A-V10R1 and E419A-V10R1 BtuB display a significant increase in the mobile population upon addition of B<sub>12</sub>, relative to other alanine mutants. Consequently, these findings demonstrate that each ion pair plays a crucial role in the substrate transport and/or during the signal transduction process. These results however are unique when compared to previously reported ion pair studies, as they do not purely leave the TonB box unfolded and unsensitive to B<sub>12</sub> binding/signal transduction. In the previous studies the R14-D316 ion pair, introducing alanine as either a single or double mutants leads to unfolding of TonB box even in the absence of B<sub>12</sub><sup>1</sup>. Thus, the R14-D316 pair is a molecular

switch which is directly involved in the transport process. However, the ion pairs roles in the present work may involve substrate binding, substrates transport, or signal transduction.

Further analysis of BtuB crystal structures shows that the R111-E465 and R69-E419 ion pairs are in close proximity meaning these two ion pairs can potentially interact with each other, coordinating their signal transduction process (Figure 6.5). Hence, we also examined the R111A-E419A-V10R1 and the R69A-E465A-V10R1 pairs in BtuB. The apo BtuB (1NQG) crystal structure displays distances between each arginine residue's nitrogen relative to each glutamic acid residue's oxygen, where E465 is about 2.4 Å away from R111 and 2.6 Å away from R69, while R69 is 3.3 Å away from E419 (Figure 6.5). Thus, these residues can interact and form salt bridges between the R69-E419, R69-E465 and R111-E465 residues in BtuB. The distance between the R111 and E419 residues is greater than 4 Å and will not form salt bridges.

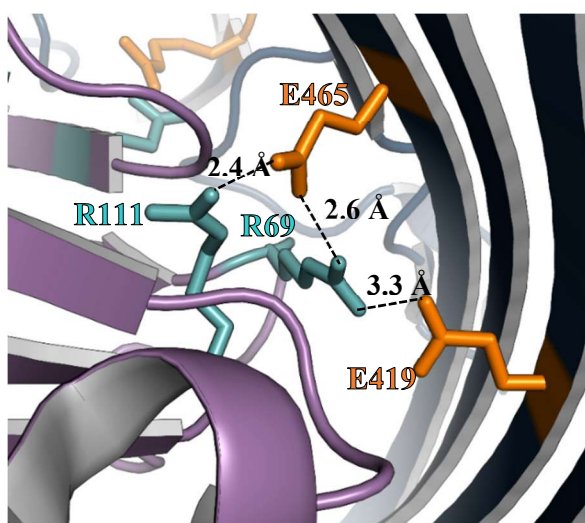


Figure 6.5. Structure of apo BtuB (1NQG) looking from the periplasmic side. The **R69** and **R111** residues in the **core domain** interact with the **E419** and **E465** residues in the **barrel domain**.

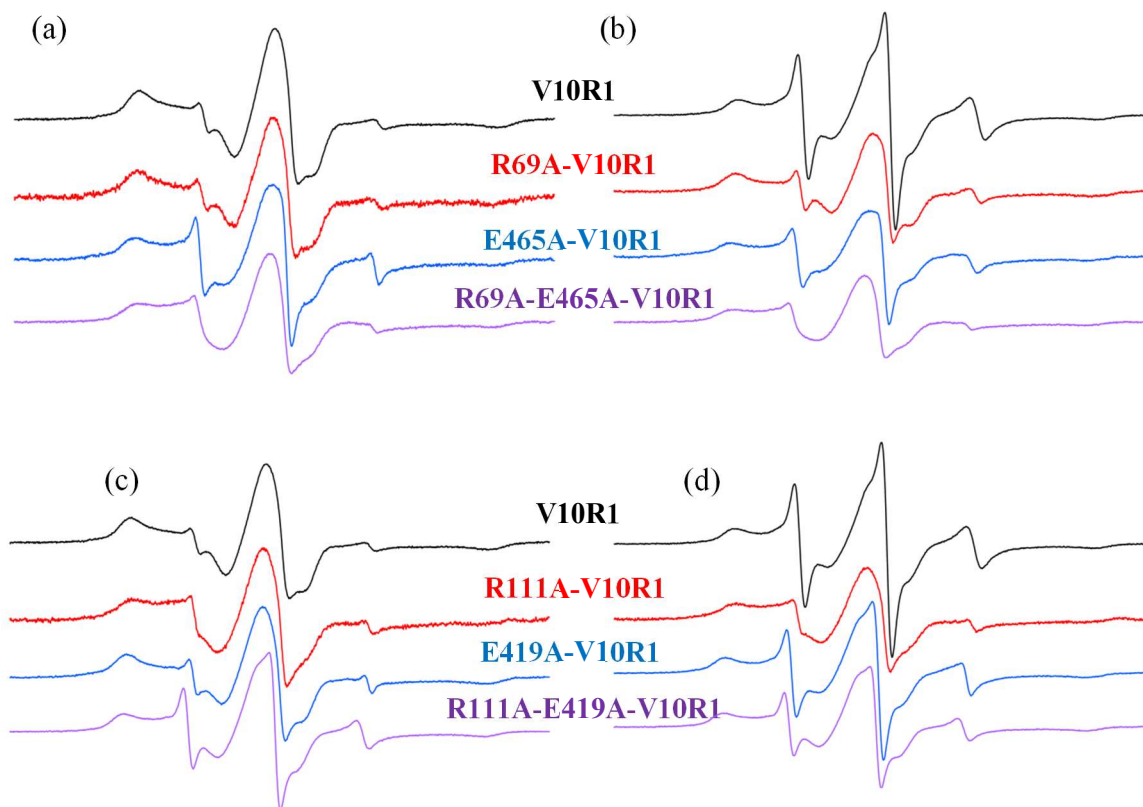


Figure 6.6. Normalized EPR spectra of the alanine mutations coupled to V10R1 BtuB. R69A-E465A-V10R1(a) apo and (b) B<sub>12</sub> bound, R111A-E419A-V10R1 (c) apo and (d) B<sub>12</sub> bound.

According to the normalized EPR spectra, R111A-E419A-V10R1 in BtuB displays a significant unfolded TonB box population in the apo state, confirming the significant role in core stability by R111 and E419, potentially through a network of salt bridges. This is the largest unfolded population observed throughout all of the alanine mutants (Figure 6.6). However, there are no significant changes upon addition of substrate and binding for both R111A-E419A-V10R1 and R69A-E465A-V10R1 BtuB. Thus, these four residues play vital roles in stabilizing the BtuB core within the barrel, as explained later.

The relative populations of the folded and the unfolded TonB box can be estimated by spectral subtraction of the mobile component from the multi-component spectrum using a LabVIEW based program called FreeBee<sup>1</sup>. Each of the collected EPR spectra is a composite of the mobile and immobile components of the nitroxide's spin dynamics. During the subtraction process, the mobile component is simulated then subtracted from the composite spectra until the immobile lineshape is obtained (Figure 6.7). To quantify the conformational equilibrium between the folded and unfolded states, the double integral of the initial spectrum, which is proportional to the total number of spins, was compared to the double integral of the mobile spectrum. The ratio of mobile to the total spins was then used to obtain the percent contributions of bound and unbound states to the total signal.

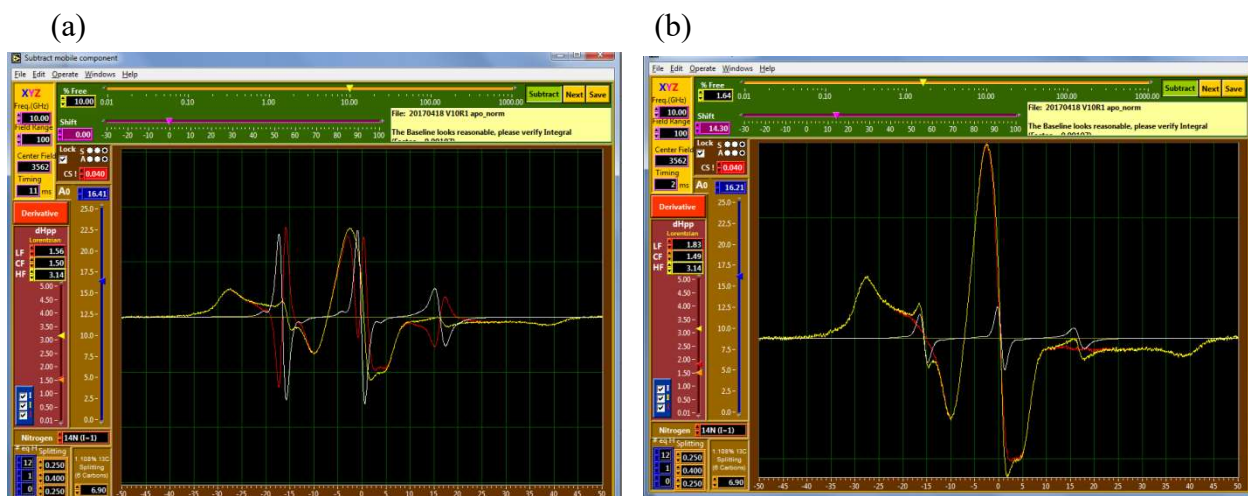


Figure 6.7. Spectral subtraction using FreeBee software. The composite spectrum (yellow solid-line) is a sum of both mobile (white solid-line) and immobile (red solid-line) components. During the subtraction, mobile component is simulated and subtracted until the immobile lineshape is remained. (a) Before and (b) after the spectral subtraction of the original spectrum.

Ion pair	Ion pair mutant with V10R1	x of apo (%)	x of B <sub>12</sub> bound (%)
R36-D515	R36A	1.64	7.67
	D515A	0.97	8.38
	R36A-D515A	0.91	1.48
R47-D548	R47A	1.64	2.38
	D548A	1.86	3.28
	R47A-D548A	1.79	1.86
D53-R526	D53A	0.92	5.88
	R526A	1.35	2.22
	D53A-R526	1.86	1.15
R69-E419	R69A	1.63	4.05
	E419A	0.86	**
	R69A-E419A	1.03	0.61
R69-E465	R69A-E465A	0.72	1.21
R111-E456	R111A	1.5	1.79
	E465A	2.23	5.07
	R111A-E465A	1.1	1.07
R111-E419	R111A-E419A	7.14	6.2
Control	V10R1	1.61	20.02

Table 6.1. The unfolded populations (x) in the apo and substrate bound states for each single and double alanine mutant obtained from the spectral subtraction of the mobile component from the normalized spectra. \*\* Unable to perform spectral subtraction due to the presence of three or more states in the B<sub>12</sub> bound E419A-V10R1 BtuB.

Based on the percentages of unfolded population (Table 6.1) in the apo state, with respect to the control (V10R1 BtuB), the majority of the alanine mutations have similar percentages of the unfolded state. However, in both the R69A-E465A-V10R1 and the E419A-V10R1 BtuB, the unfolded population of the TonB box is reduced about two-fold. Thus, each of these mutations cause increased TonB box folding within the BtuB relative to the control V10R1 BtuB. In contrast, in the R111A-E419A-V10R1 and the E465A-V10R1 BtuB, there are increased populations of unfolded TonB box in the apo state. Consequently, each of these mutations affect the salt bridge network between the barrel and the core domains which causes an increase in the unfolded TonB box population.

In the substrate bound state, the V10R1 BtuB unfolded population increases about 12-fold. Based on the unfolded populations for each alanine mutant, relative to the apo state, few single alanine mutations (R36A-V10R1, D515A-V10R1, D53A-V10R1, R69A-V10R1 and E465A-V10R1 BtuB show enhanced unfolded populations. The highest increase observed is in the D515A-V10R1 BtuB mutant, which is about 8-fold (Table 6.1). Unlike the other mutants, the unfolded population of the E419A-V10R1 BtuB mutant in the B<sub>12</sub> bound state was unable to be obtained via spectral subtraction because it had more than the two states (folded and unfolded). However, based on the EPR spectra it is evident that upon substrate binding a significant E419A-V10R1 BtuB population undergoes unfolding.

Based on the unfolded populations obtained from the spectral subtraction and the normalized EPR spectra, the R69, R111, E419 and E465 residues play interesting roles in the TonB box folding and unfolding. Between the two arginine (R69 and R111) residues and the two glutamic acid (E419 and E465) residues, R111 is vital. In the presence of substrate, mutating this arginine residue potentially produces transport defective BtuB, while the other three residues

comparatively (two-fold or more) enhance the unfolding of the TonB box (Table 6.1). However, for all four alanine double mutants (R69A-E419A-V10R1, R69A-E465A-V10R1, R111A-E465A-V10R1 and R111A-E419A) there was no significant increase in unfolding upon B<sub>12</sub> binding.

According to the crystal structures of BtuB (1NQG and 2GSK), both of the arginine residues, R69 and R111, are in close proximity to the glutamic acid residues, E419 and E465. Thus, R69, R111, E419 and E465 are likely forming salt bridges which can facilitate the signal transduction process in BtuB as similar to FhuA. Mutagenesis studies of FhuA reveal, residues R93, R133, E522, and E571 form salt bridges (R133-E571, E571-R93 and R93-E522) which also act as molecular switches<sup>11</sup>. Residues R69, R111, E419 and E465 in BtuB are analogous to residues R93, R133, E522 and E571 in FhuA. Therefore, similar to that of FhuA, BtuB residues R69 and E465 are forming salt bridges between both glutamic acids and arginine residues (Figure 6.5).

During the signal transduction and/or substrate transport process, the salt bridges formed between these four residues may stabilize and destabilize the BtuB core domain. This most likely propagates the signal from the bridge R69-E419 to R69-E465, and then to R111-E465. As a result, the corresponding alanine double mutations are unable to propagate the signal and inhibits unfolding of the TonB box. Here however, if any of the residues R69, E419, or E465 are absent, the remaining three residues can still potentially form salt bridges to stabilize BtuB. Thus, the R69A, E419A and E465A mutants could potentially still propagate the signal and/or substrate transport and ultimately unfold the TonB box in the presence of B<sub>12</sub>, but not the R111A mutant.

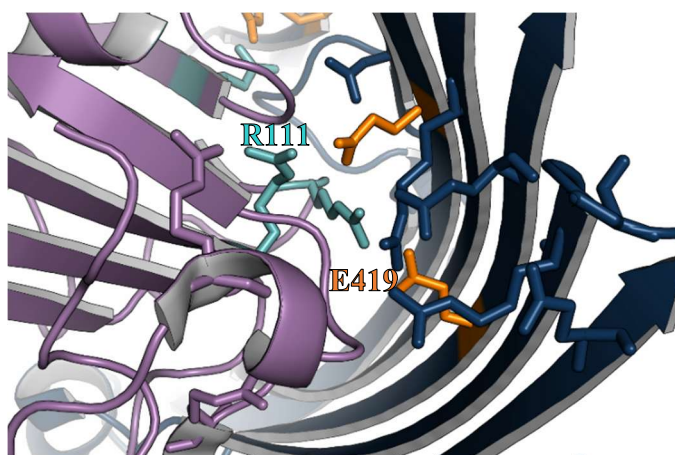
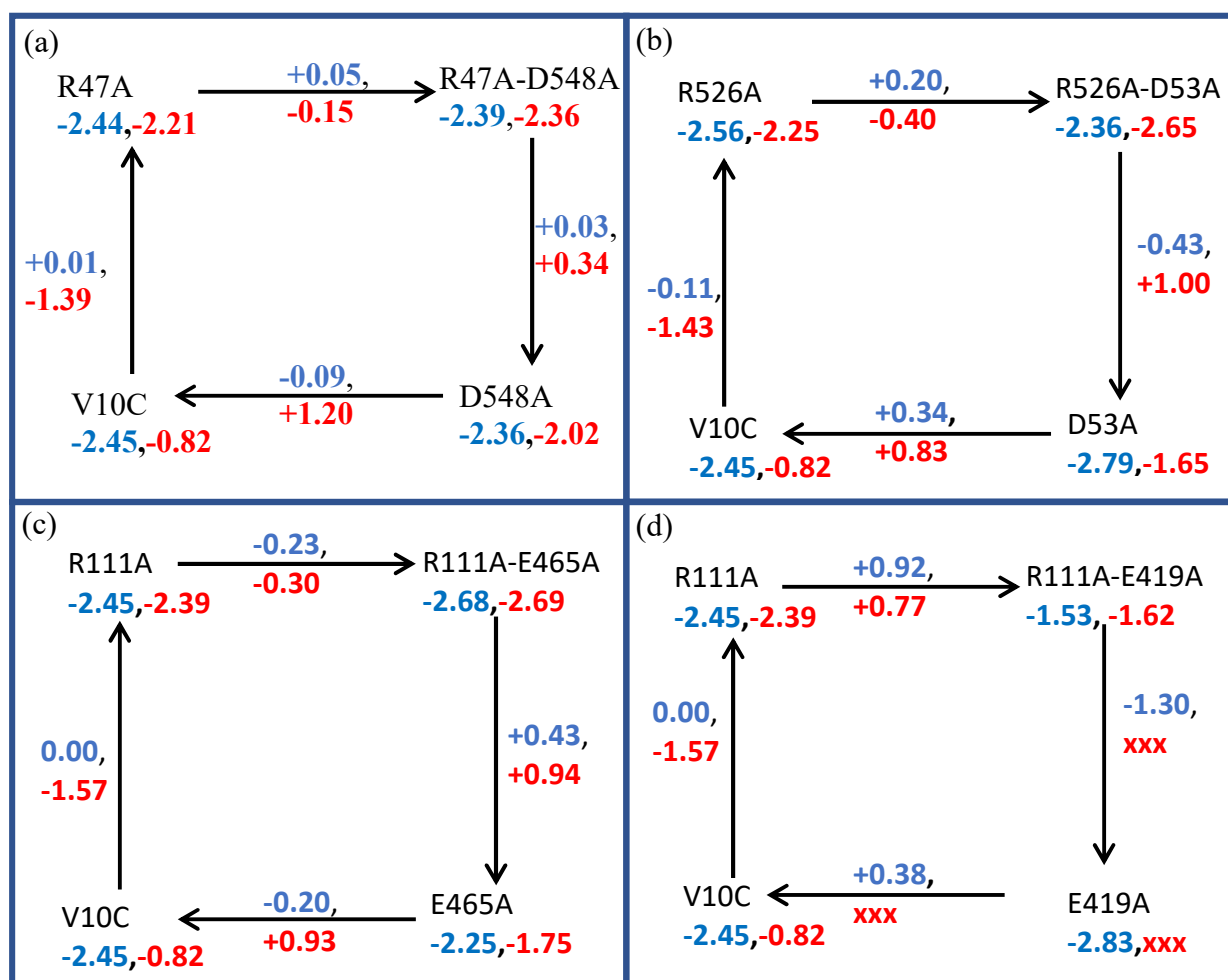


Figure 6.8. Structure of BtuB from periplasmic side showing the arginine, lysine, aspartic acid, and glutamic acid residues in the **core** and the **barrel** near sites **R111** and **E419**.

Moreover, the significant increase in unfolded population in the apo R111A-E419A-V10R1 BtuB could be the result of altering the salt bridge network between these four residues which destabilizes the BtuB core. In the R111A-E419A-V10C BtuB mutant, both the R111-E465 and the R69-E419 salt bridges are eliminated which leaves only the R69-E465 salt bridge. Looking at the periplasmic side of the BtuB apo crystal structure (1NQG), there are a few arginine, lysine, aspartic acid, and glutamic acid residues which could potentially form replacement salt bridges near the R111 and E419 sites, but all of these residues are more than 4 Å apart. Thus, we speculate that in the R111A-E419A-V10C BtuB mutant these residues will not form salt bridges which would aid BtuB core folding. Furthermore, in the structure, all of these residues, including R111 and E419, are localized within close proximity to the TonB box in the folded state (Figure 6.8). Consequently, the R111A-E419A-V10C destabilizes the folded core, especially near the TonB box area, which results in a large population of unfolded TonB box in the apo state and there is no increased unfolding in the presence of B<sub>12</sub>. Thus, this destabilization also hinders the signal propagation pathway.

Conformational equilibria were quantified and used to estimate the free energy differences ( $\Delta^\circ G$ ) between folded and unfolded states in the apo and the B<sub>12</sub> bound states according to the folded and unfolded populations, <sup>1</sup>. These free energy differences were then used to calculate the free energy ( $\Delta\Delta^\circ G$ ) associated with each mutant in the two mutant cycle analysis (Figure 6.9). The sum of  $\Delta\Delta^\circ G$  going around the two mutant cycle is about zero (kcal/mol) for all mutants with the exception of the B<sub>12</sub> bound R69A-E419A-V10R1 and the R111A-E419A-V10R1 BtuB, because the unfolded population was not obtained.



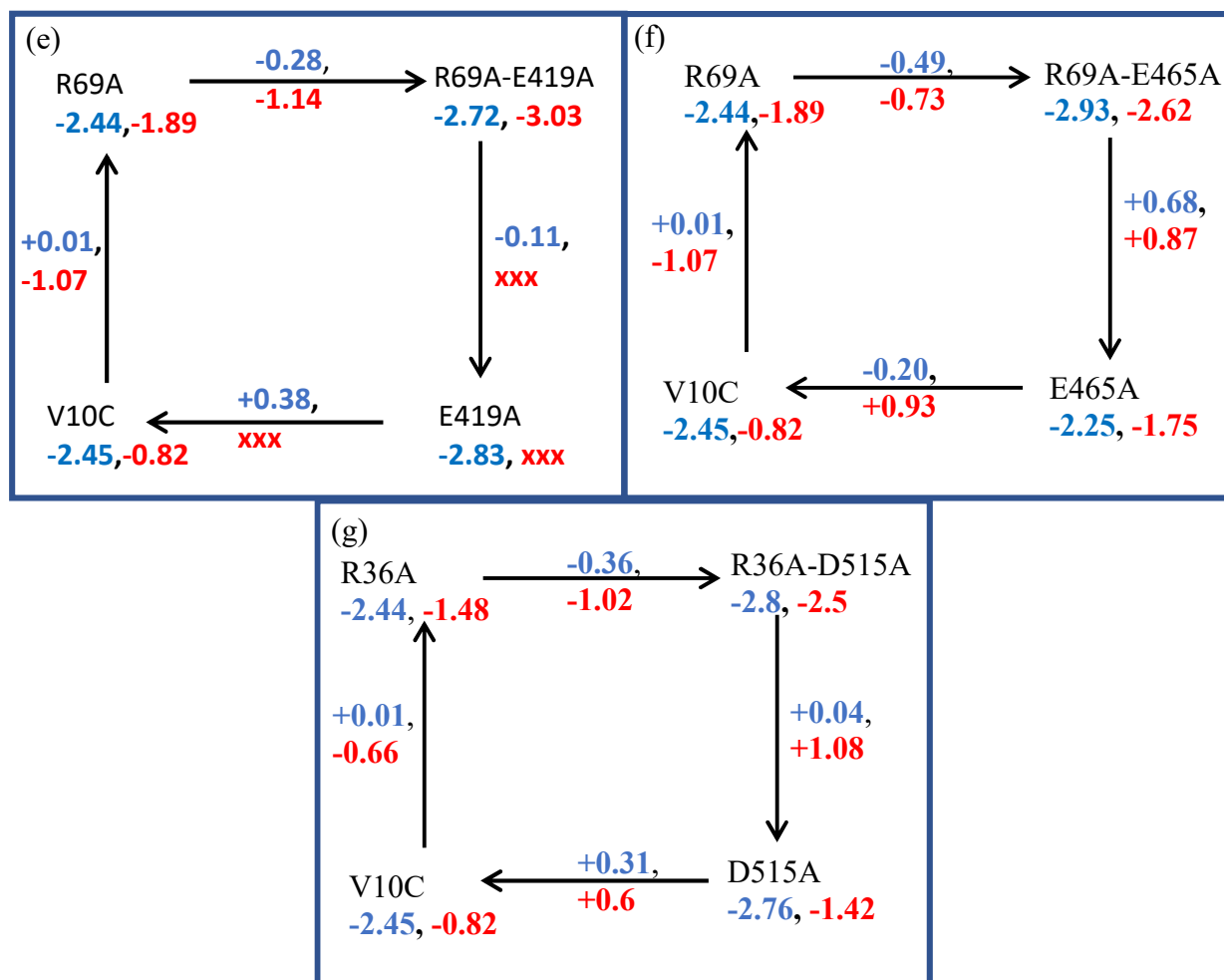


Figure 6.9. Two mutant cycle analysis of  $\Delta^\circ G$  calculated based on the conformational equilibria of the **apo** and **B<sub>12</sub>** bound states. The  $\Delta\Delta^\circ G$  associated with each mutant in the **apo** and **B<sub>12</sub>** bound states were calculated based on the  $\Delta^\circ G$  in the corresponding mutant cycle. (a) R47A-D548A-V10R1, (b) D53A-R526A-V10R1, (c) R111A-E465A-V10R1, (d) R111A-E419A-V10R1, (e) R69A-E419A-V10R1, (f) R69A-E465A-V10R1 and R36A-D515A-V10R1 BtuB.

The  $\Delta\Delta^\circ G$  values between each opposite side in the two mutant cycle; left and right ( $\Delta\Delta^\circ G_1$  and  $\Delta\Delta^\circ G_3$ ) or top and bottom ( $\Delta\Delta^\circ G_2$  and  $\Delta\Delta^\circ G_4$ ) can be compared to determine residue interactions. When there is an energy difference between the  $\Delta\Delta^\circ G_1$  and the  $\Delta\Delta^\circ G_3$ ; or the  $\Delta\Delta^\circ G_2$  and the  $\Delta\Delta^\circ G_4$ , this indicates that these residues must interact. Here, all calculations are based on the associated folding and unfolding populations of V10R1. Since the ion pair residues studied in this project mostly inhibit the TonB box unfolding, the calculated  $\Delta\Delta^\circ G$  values are not always significantly different from their counterpart  $\Delta\Delta^\circ G$ , such as the R36A-D515A-V10R1 mutant and the R111A-E465A-V10R1 mutant in their apo states. For the ion pairs with nonequivalent  $\Delta\Delta^\circ G$  values, R69A-E465A-V10R1 and D53A-R526A-V10R1 BtuB, these ion pair residues interact. However, for other ion pairs, we cannot use the TonB box unfolding with the two mutant cycle as a system to demonstrate interactions between these residues either at apo or B<sub>12</sub> bound states, as they do not directly involving in V10R1 unfolding event.

In addition to investigating the folding-unfolding of the V10R1 site on BtuB, a subset of unlabeled (V10C) ion pair alanine mutants were used to examine B<sub>12</sub> binding using either reconstituted BtuB or intact *E. coli* cells. These experiments were performed to determine if any of these mutations disrupt substrate binding, in order to understand the link between substrate binding and the promotion or inhibition of the TonB box unfolding in the presence of B<sub>12</sub>. In these studies, spin labeled B<sub>12</sub>, TEMPO-B<sub>12</sub> (Figure 6.10 a)<sup>12, 13</sup> binding to each alanine mutant was determined using CW-EPR. The free, unbound TEMPO-B<sub>12</sub> spectrum is narrow with a high amplitude because it is in fast, isotropic motion. When it binds to the V10C BtuB, the CW spectrum broadens producing lower amplitude due to reduced tumbling rate and motional restriction of the TEMPO-B<sub>12</sub>.

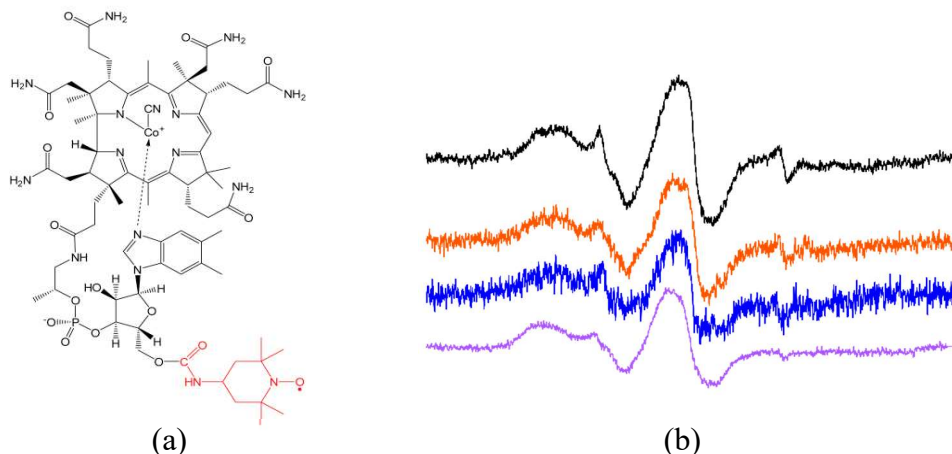


Figure 6.10. (a) Structure of TEMPO-B<sub>12</sub><sup>12</sup> (figure is from Joseph *et al.* 2015. *Angewandte Chemie (International ed. in English)*. **54**, 6196-6199) and (b) the normalized CW EPR spectra of V10C BtuB bound to TEMPO-B<sub>12</sub>. TEMPO-B<sub>12</sub> (**10 μM**, **5 μM** and **2.5 μM**) was incubated with the reconstituted BtuB and with the **intact live *E. coli* cells**.

Each spectrum (Figure 6.10 b) contains a both broad immobile and a narrow mobile component arising from mixed BtuB bound and unbound populations of the TEMPO-B<sub>12</sub>. The bound population in both the reconstituted and the intact cells is larger than the small fraction of unbound TEMPO-B<sub>12</sub>, freely diffusing. Looking at the D53A-R526A-V10C BtuB spectra, each alanine mutant binds with the TEMPO-B<sub>12</sub> (Figure 6.11 a and b). The BtuB in live cells, an anisotropic powder spectrum is observed, demonstrating binding. In the reconstituted BtuB a combination of bound and small percentage of unbound TEMPO-B<sub>12</sub>, due to leftover residual TEMPO-B<sub>12</sub> in the samples. In the intact cells, there are small differences in the hyperfine splitting observed for the alanine mutants (more prominently in the double mutant) potentially due to some differences in the spin dynamics and spin rotamers for the TEMPO-B<sub>12</sub> binding in these mutants.

For the R47A-D548A-V10C BtuB set, the single alanine mutants bind with the TEMPO-B<sub>12</sub> in both the intact cells and the reconstituted systems (Figure 6.11 c and d). However, the double mutant, R47A-D548A-V10C, did not bind as significantly with the BtuB as seen in the larger narrow, mobile spectra for both the reconstituted and the intact cells. Because this was observed in both the reconstituted and the intact cells, the R47A-D548A-V10C BtuB mutation impacts the substrate binding site and increases the dissociation of substrate due to destabilization in the BtuB core. Consequently, preventing the TonB box unfolding process (Figure 6.4 b). Similar to the D53A-R526A-V10C, there are changes in the hyperfine splitting observed in these single alanine mutants. The spectra of the R36A-D515A-V10C (Figure 6.11 e and f) in the intact cells and the reconstituted BtuB however differ. The D515A-V10C BtuB mutant in the cells displays a large unbound, mobile TEMPO-B<sub>12</sub> population, while in the reconstituted BtuB, a significant population of substrate is bound to the BtuB. These results could be due to low expression levels for the D515A-V10C BtuB mutant in the cells; thus, during the experiment the majority of the substrate was unbound and mobile. The spectrum for the R36A-D515-V10C BtuB mutant is also different from the control, V10C, BtuB in both the intact and the reconstituted systems. Thus, this double alanine mutation potentially affects the substrate binding in BtuB.

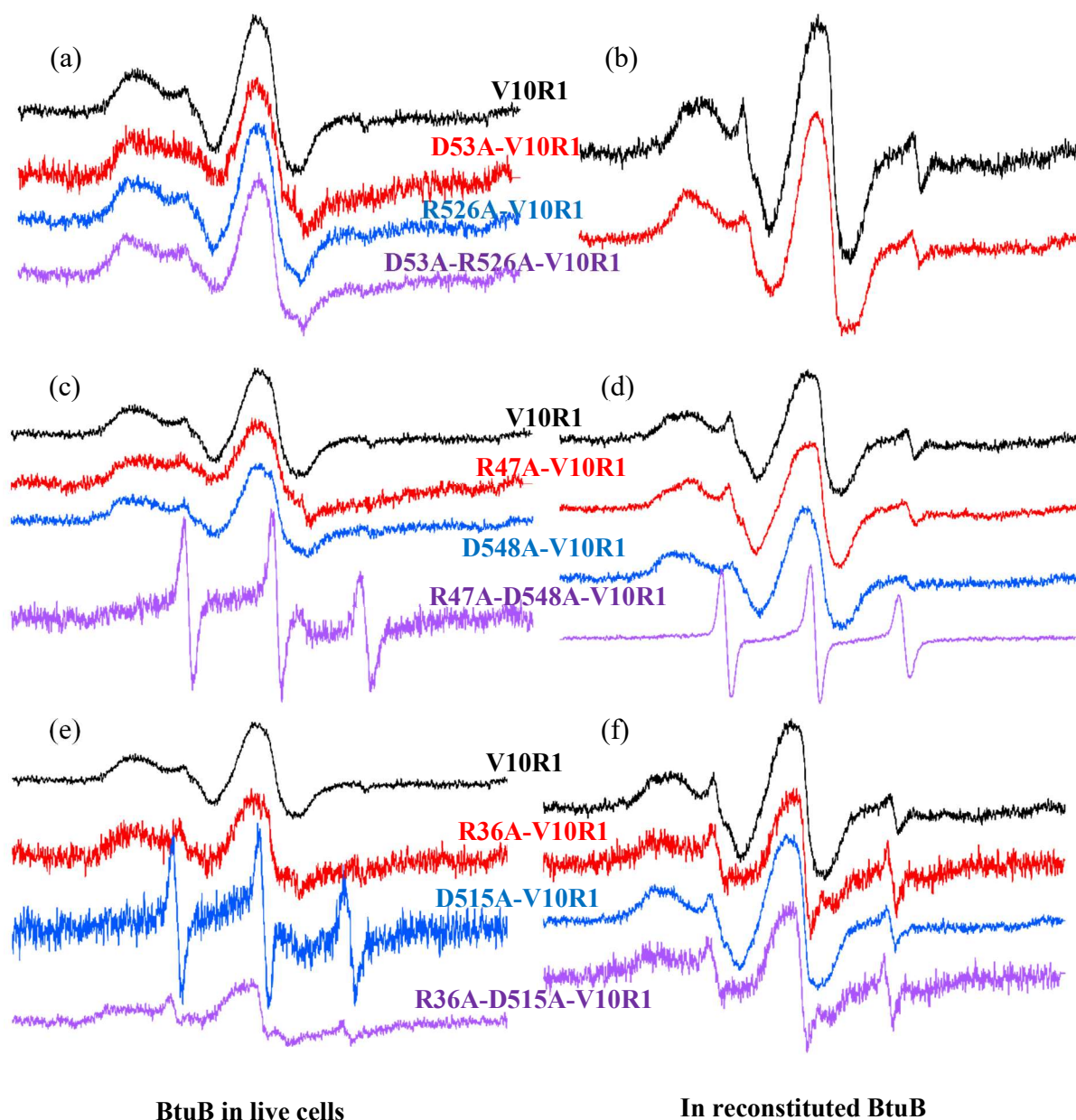


Figure 6.11. Normalized CW EPR spectra of TEMPO-B<sub>12</sub> binding to the ion pair mutants. D53A-R526A-V10C (a) in live cells and (b) in reconstituted BtuB, R47A-D548A-V10C (c) in live cells and (d) in reconstituted BtuB and R36A-D515A-V10C (e) in live cells and (f) in reconstituted BtuB bound with TEMPO-B<sub>12</sub> (V10C, single alanine mutant in the **core** and **barrel**, **double mutant**).

Even though these ion pairs are not directly connected with the folding and the unfolding events of the TonB box as reported previously, these ion pairs are involved in both the substrate binding and the signal transduction processes. Eliminating one or both ion pair residues inhibits the TonB box unfolding associated with the substrate binding as seen for V10R1 control sample. Furthermore, four residues (R69, R111, E419, and E465) from a salt bridge network in which the signal propagation is through E419-R69 to R69-E465 and then to E465-R111. Thus, eliminating both R111 and E419 led to destabilize the core and significantly unfold TonB box at the apo state and become B<sub>12</sub> insensitive by blocking the signal transduction pathway. Some of these studied ion pair mutations also impact substrate binding at the binding site of BtuB and affect the transport. Based on TEMPO-B<sub>12</sub> binding with the R47A-D548A-V10C BtuB mutant, elimination of R47 and D548 alters the binding of B<sub>12</sub> and increases the dissociation. As a result, in the presence of substrate, the double mutant is preventing the TonB box unfolding.

## 6.5 Future directions

There are many salt bridges between the core and barrel, which keeps the BtuB core within the barrel. These bridges also propagate a signal across the BtuB to coordinate the folding and the unfolding of the TonB box which is associated with substrate binding. The current work focused on five conserved ion pairs. In addition to these pairs, other ion pairs could play significant roles in these above processes and potentially play new roles associated with BtuB substrate (B<sub>12</sub>) transport. Additionally, these pairs could also be involved in colicin E1-9 binding as they have been reported to stabilize TonB box folding when the colicin E3 receptor domain binds with BtuB<sup>1</sup>. Hence, in the future, other ion pairs between the core and barrel can also be studied to identify potential ion pairs involved in signal transduction and substrate transportation

processes. Moreover, other sites in BtuB ion pair mutants could be spin labeled on the extracellular side of BtuB (S74, V90 and T188) and studied in addition to V10R1.

In addition to TEMPO-B<sub>12</sub> binding studies, transport assays using radiolabeled substrates could be used to determine if the substrate is being transported in each of the alanine mutants to see if the TEMPO-B<sub>12</sub> immobilized mutants are capable of substrate up-take, not just binding. Moreover, TEMPO-B<sub>12</sub> binding studies can be incorporated to determine the binding affinities for these mutants. It would also be interesting to determine substrate transport of the mutants where TEMPO-B<sub>12</sub> binds, but the unfolding of the TonB box is inhibited. This would investigate any alternative pathways for substrate transport that do not unfold the TonB box. This may occur by destabilization in the core of the BtuB.

## 6.6 References

- [1] Lukasik, S. M., David Ho, K. W., and Cafiso, D. S. (2007) Molecular Basis for Substrate-Dependent Transmembrane Signaling in an Outer-Membrane Transporter, *Journal of molecular biology* 370, 807-811.
- [2] Chimento, D. P., Kadner, R. J., and Wiener, M. C. (2003) The Escherichia coli Outer Membrane Cobalamin Transporter BtuB: Structural Analysis of Calcium and Substrate Binding, and Identification of Orthologous Transporters by Sequence/Structure Conservation, *Journal of molecular biology* 332, 999-1014.
- [3] Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010) TonB-dependent transporters: regulation, structure, and function, *Annual review of microbiology* 64, 43-60.
- [4] Di Girolamo, P. M., and Bradbeer, C. (1971) Transport of vitamin B 12 in Escherichia coli, *J Bacteriol* 106, 745-750.
- [5] Fanucci, G. E., Coggs, K. A., Cadieux, N., Kim, M., Kadner, R. J., and Cafiso, D. S. (2003) Substrate-Induced Conformational Changes of the Periplasmic N-Terminus of an Outer-Membrane Transporter by Site-Directed Spin Labeling, *Biochemistry* 42, 1391-1400.
- [6] Xu, Q., Ellena, J. F., Kim, M., and Cafiso, D. S. (2006) Substrate-Dependent Unfolding of the Energy Coupling Motif of a Membrane Transport Protein Determined by Double Electron-Electron Resonance, *Biochemistry* 45, 10847-10854.
- [7] Chimento, D. P., Mohanty, A. K., Kadner, R. J., and Wiener, M. C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB, *Nature structural biology* 10, 394-401.
- [8] Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex, *Science* 312, 1396-1399.
- [9] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Research* 28, 235-242.
- [10] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.
- [11] Endriß, F., Braun, M., Killmann, H., and Braun, V. (2003) Mutant Analysis of the *Escherichia coli* FhuA Protein Reveals Sites of FhuA Activity, *Journal of Bacteriology* 185, 4683-4692.
- [12] Sikora, A., Joseph, B., Matson, M., Staley, J. R., and Cafiso, D. S. (2016) Allosteric Signaling Is Bidirectional in an Outer-Membrane Transport Protein, *Biophysical journal* 111, 1908-1918.
- [13] Joseph, B., Sikora, A., Bordignon, E., Jeschke, G., Cafiso, D. S., and Prisner, T. F. (2015) Distance Measurement on an Endogenous Membrane Transporter in E. coli Cells and Native Membranes Using EPR Spectroscopy, *Angewandte Chemie (International ed. in English)* 54, 6196-6199.

## **Appendix 1. Preliminary biophysical characterization of Colicin E3 receptor domain binding to the primary receptor, BtuB**

### **Abstract**

Bacteriocins are antibiotics which are produced by bacteria when they are under stress. Colicin is one class of bacteriocins, which is produced by *Escherichia coli* (*E. coli*) against related *E. coli* strains. This class of bacteriocins is the most widely studied using molecular biological, biochemical, and biophysical techniques. However, current knowledge regarding the mechanisms and the associated intrinsic conformational changes that take place during colicin binding to the OM receptor and its uptake into the cell is limited.

In the present work, site directed spin labeling (SDSL) coupled to electron paramagnetic resonance (EPR) spectroscopy was used to investigate the conformational changes associated with receptor binding events of the colicin E3 receptor binding domain (colicin E3R) with its OM receptor, BtuB. Based on our preliminary work using the colicin E3R fragment (residues 343-417), highly mobile shorter R domain fragment reduces in mobility upon the binding with the wild type (WT) BtuB, both in live *E. coli* cells and in the reconstituted bilayer system. Furthermore, preliminary DEER data was used to predict the potential distance distributions between the T188R1 site on the unstructured BtuB loop 2 (in colicin E3R bound crystal structure, (PDB ID: 1UJW) and several spin labeled colicin mutants. Based on the distances, loop 2 is in close contact with colicin E3R.

## Introduction

The first bacteriocin was discovered by André Grutia in 1925, it was identified as a lethal compound produced by *Escherichia coli* (*E. coli*) V, which was toxic to *E. coli*  $\phi$  cells<sup>1</sup>. Since then, many other bacteriocins have been identified in both Gram-negative and Gram-positive bacteria (Chapter 1.6)<sup>2, 3</sup>. These antimicrobial agents provide an advantage by invading sensitive strains, as well as, providing protection over other invading bacterial species<sup>1</sup>. Additionally, it has been proposed that in Gram-positive bacteria bacteriocins mediate the quorum sensing process<sup>2</sup>.

As discussed in Chapter 1.6.4 and Chapter 1.6.6, during the invasion process, different colicins use different transporters in the outer membrane (OM) as their receptors and translocators across the OM and the inner membrane (IM). A group colicins are Tol-dependent, while B group colicins are TonB-dependent<sup>1, 4</sup>. Members of colicin E use BtuB as their primary receptor during the invasion process. Binding to BtuB facilitates binding between the above colicins and their corresponding translocator OM proteins<sup>1, 4</sup>. The colicin E3 crystal structure was first solved in 2001 by Soelaiman and *et. al*<sup>5</sup>(Chapter 1.6.3). Based on the colicin E3 crystal structure (PDB ID: 1JCH), colicin E3 is a Y-shaped molecule (Figure A1.1a)<sup>5</sup>.

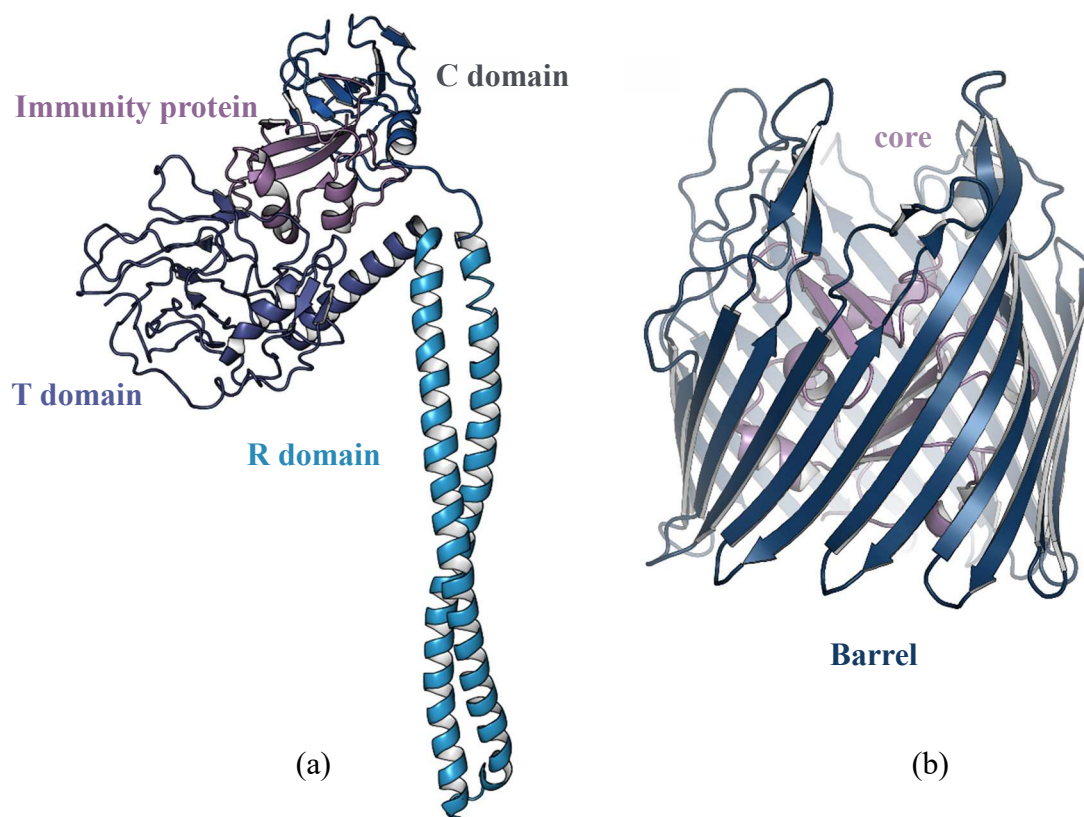


Figure A1.1. Structures of (a) colicin E3 (PDB ID: 1JCH), where the globular **T domain** and **C domain** are connected by antiparallel helical arms of **R domain**. The **immunity protein** is connected to both the T and C domains. (b) side view of apo BtuB with calcium (PDB ID: 1NQG) (calcium is not shown in the structure). The **core domain** occludes the pore of the **barrel**.

As discussed in Chapter 1.5.3, the BtuB forms a 22 anti-parallel  $\beta$  stranded barrel, in which the pore is occluded by the core domain (Figure A1.1b)<sup>6, 7</sup>. Kurisu and *et al.* were able to solve the crystal structure of colicin E3R bound to BtuB (PDB ID: 1UJW)<sup>8</sup>. During the co-crystallization, 135 amino acids of the R domain (R135) from Thr313 to Glu447 were bound to BtuB. Based on the structure, the binding of colicin E3R produces small conformational changes in the core domain, while some extracellular loops have significant changes in the barrel. The loops between the  $\beta$  strand 5-6 (loop 3) and 7-8 (loop 4) that are disordered in the BtuB apo

structure become ordered, while the extracellular loop two between the  $\beta$  strand 3-4 is partially ordered, and the position of loop ten between the strand 19-20 is changed (Figure A1.2). Thus, the binding of colicin has overall altered the extracellular loops. These changes however do not sufficiently disorder the core of the BtuB to allow passage of the colicin translocation domain<sup>8</sup>.

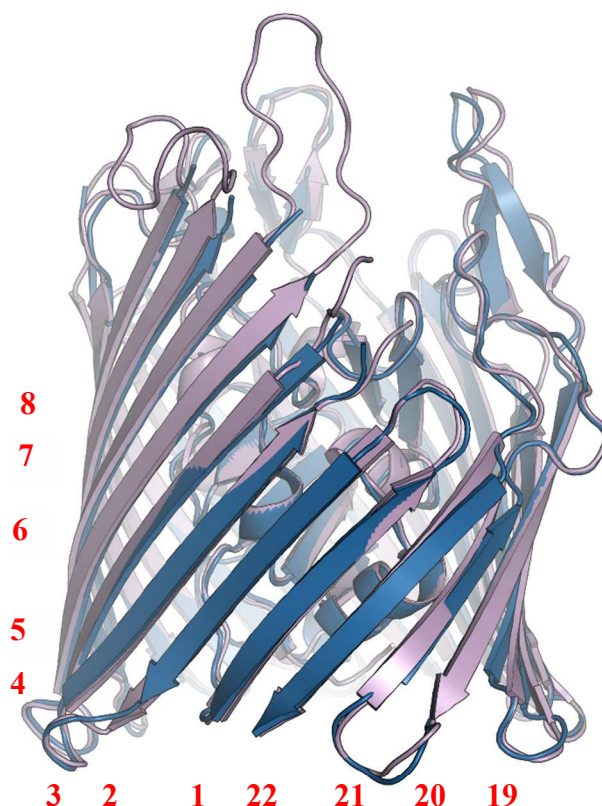


Figure A1.2. Structural alignment of **apo BtuB** (1NQE) and **colicin bound BtuB** (1UJW) with labeled  **$\beta$  strands**. Upon binding with BtuB, the extracellular loops 3 and 4 between  $\beta$  strand **5-6** and **7-8** are ordered and loop 2 between  $\beta$  strand **3-4** is partially ordered, while the position of loop 10 between  $\beta$  strand **19-20** altered.

Supporting the crystal structure observations, planer bilayer experiments have demonstrated that BtuB does not facilitate the translocation process<sup>9</sup>. Furthermore, these experiments suggest the translocation of the colicin E3 and colicin E1 occur through the OM porin (OmpF) and OM

exporter protein (TolC), respectively. OmpF, a trimeric porin, is formed by 16  $\beta$ -strands, where the extracellular loop L3 is bent into the pore, limiting the accessible pore size. In 2008, Yamashita and *et. al.* published a co-crystallization of the disordered N-terminus of the T domain (T83) with OmpF. Based on the structure, a seven residue long fragment was detected inside one of the OmpF pores (Figure A1.3). In combination with other biochemical and biophysical studies, it was concluded to be the N-terminus of the T83 fragment<sup>10</sup>.

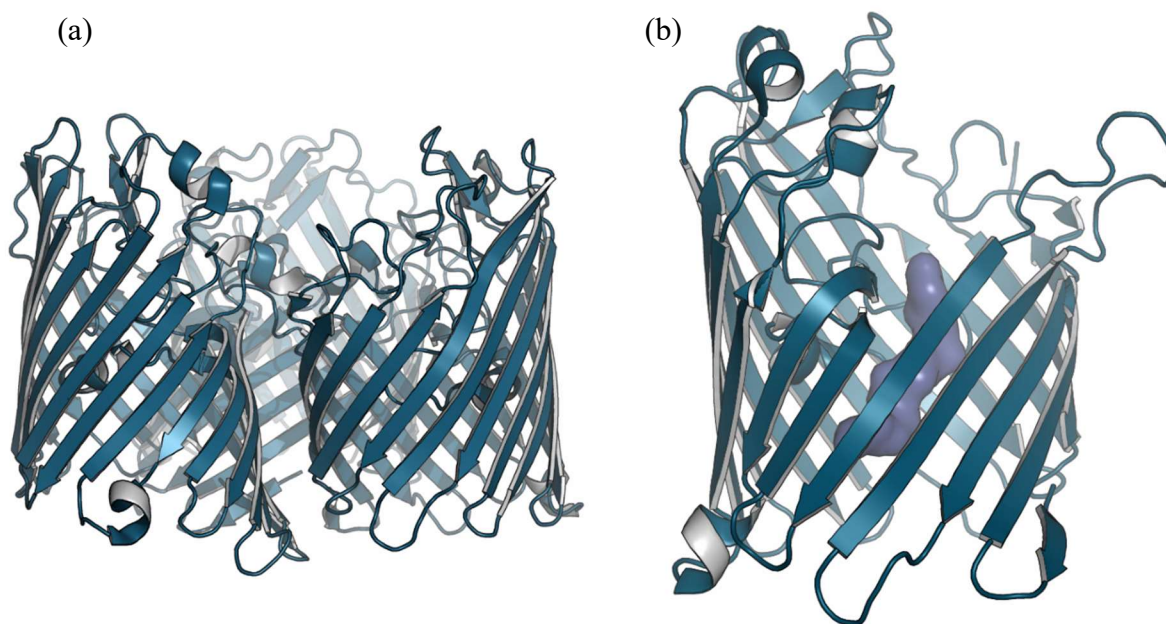


Figure A1.3. Side view of (a) trimeric **OmpF** (PDB ID: 3K19) and (b) **poly-glycine peptide** of colicin T83 fragment is occluding the pore of **OmpF** (PDB ID: 2ZLD).

According to the colicin R135 bound BtuB structure, the tip (M383) interacts with four residues within the BtuB core: T55, N57, L63, and S64 (Figure A1.4). Additionally, 27 amino acids of colicin interact with 29 amino acids on BtuB. A total of 19 residues from both the N- and C-terminus (313-322 and 439-447) of colicin E3R that are ordered also get disordered upon BtuB binding<sup>8</sup>. Thus, disordering both the N- and C- terminus could potentially enhance the dissociation of the T and C domains from the immunity protein prior to its translocation. This

could also facilitate scanning for the second OMP for translocation, using a proposed fishing pole mechanism for cellular entry.

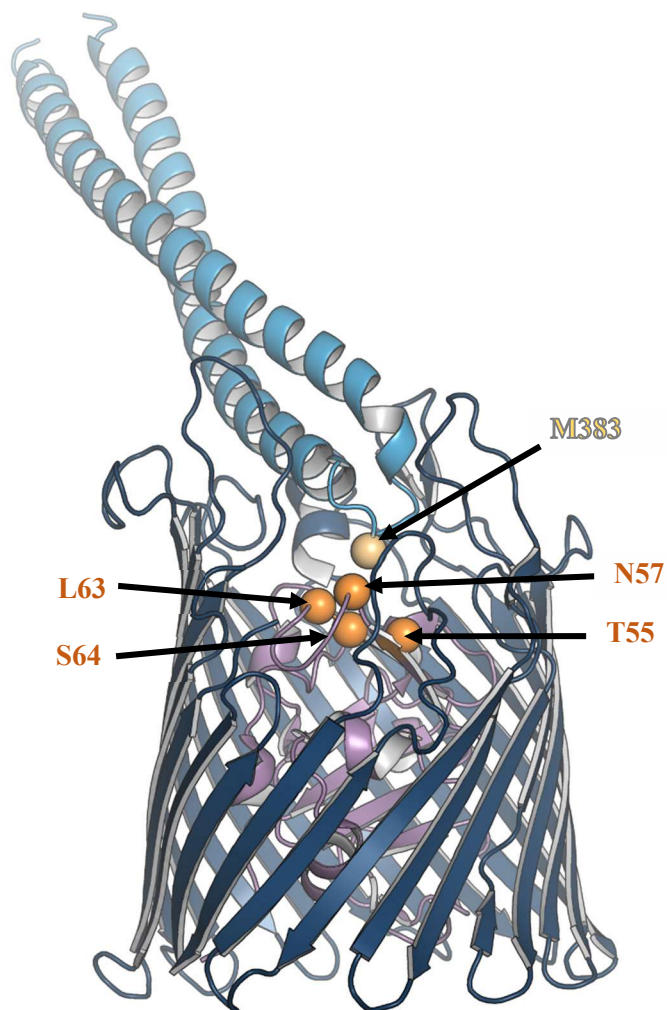


Figure A1.4. Side view of colicin **R135 fragment** bound to BtuB (1UJW). During binding colicin interacts with residues from both the **core** and the **barrel domain** of BtuB. The tip of colicin (**M383**) interacts with four residues (**T55**, **N57**, **L63** and **S64**) in the BtuB **core**.

These proposed receptor-induced conformational changes were studied using FRET analysis of dual-labeled colicin E3. Upon binding to BtuB, the C and T domains separate. These were originally interacting with the immunity protein (Figure A1.1a)<sup>11</sup>. Single particle tracking

experiments, using both fluorescently labeled antibodies and colicin E3, have been used to study the mobility of both BtuB and OmpF. Based on these findings, the BtuB is fast diffusing, while the trimeric OmpF is slow or immobile. However, the BtuB diffusion decreases in the wild type with respect to the TonB box mutant in BtuB (D6A-T7A-L8A-V9A). Eliminating the impact of TonB, the diffusion coefficient of BtuB is reduced ( $\sim 2$ -fold) upon the binding of colicin E3. However, the lateral diffusion of BtuB facilitates the search for OmpF on the OM<sup>12</sup>.

In addition to these studies of colicin E3, colicin E9 studies also aid in understanding the invasion process. In 2016, the complete E9 colicin structure was resolved by Klein and *et. al*<sup>13</sup>. Colicin E9 contains a TolB binding site (residues 32-47) between two OmpF binding sites (residues 2-18 and 54-63). All of these sites are in the intrinsically unstructured T domain (IUTD). Housden and *et. al.* was able to isolate and characterize the complex of the colicin E9-immunity protein-BtuB-OmpF-TolB. According to their studies, the IUTD of colicin E9 passed through two OmpF units in opposite directions and then recruited TolB via the TolB binding site, in order to initiate translocation<sup>14</sup>.

In contrast to crystal structures, other techniques can provide more functional information related to colicin invasion, such as: planar bilayer experiments, single particle tracking experiments, FRET analysis, *etc*<sup>9, 11, 12, 14</sup>. Many of these experiments are mostly focused on the isolated protein systems, where reconstituted OMPs are used to investigate the binding and translocation events of colicins E3<sup>9, 11</sup>. In contrast, in the single particle tracking experiments, bacterial cells were used for investigation BtuB mobility and screening for OmpF<sup>12</sup>. However, currently, the knowledge with respect to the intrinsic events associated with the colicin invasion process in live cells is limited. Moreover, conformational changes associated with the intact bacterial cell binding and translocation are lacking in the field.

Site directed spin labeling (SDSL) coupled to electron paramagnetic resonance (EPR) can be used to investigate these intrinsic but vital events associated with the colicin invasion mechanism. Here, the feasibility of detecting such events was the aim in this project, using both colicin and its primary receptor, BtuB, in both the isolated lipid bilayer system, as well as, intact cells with the future goal of mapping the T domain residues involved in the translocation process.

## Methods

### Expression, purification and spin labeling of colicin E3R

After sequence conformation, plasmids were transformed into the *E. coli* BL21 DE3 strain<sup>15</sup>. Luria-Bertani (LB) pre-cultures were inoculated with a single cell colony (SCC) and grown overnight (O/N) at 37 °C. Main LB cultures containing 100 ug/mL of ampicillin were inoculated with O/N grown pre-cultures and grown at 37 °C. Once the OD<sub>600</sub> reached 0.6-0.7, isopropyl β-D-1-thiogalactopyranoside, IPTG (0.25 mM) was added then cultures were further incubated for 4 hrs. Next, the cells were harvested by centrifugation at 6000 rpm for 10 min at 4 °C using a Sorval SLA-3000 rotor. The cell pellets were resuspended into colicin lysis buffer (300 mM NaCl and 25 mM HEPES, pH 8.0 buffer) containing 10 nM of 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF).

Resuspended cells were lysed using a French press then centrifuged using Sorval SS-34 rotor at 12000 rpm for 15 min at 4 °C to remove cell debris. The supernatant was incubated at 4 °C, O/N in Ni<sup>2+</sup> charged, immobilized metal affinity chromatography (IMAC) resin with gentle mixing. The column was run by first eluting the flow through (FT), then the column was washed with 2 column volumes (CV) of colicin lysis buffer. The bound proteins were eluted by 1 CV of 20% elution buffer (300 mM NaCl, 250 mM imidazole and 25 mM HEPES pH 8.0) followed by

3 CV of 100 % elution buffer. Samples from each fraction were analyzed by SDS-PAGE. The fractions containing colicin were combined, concentrated using 3 kDa MWCO concentrator, and stored at -80 °C.

During the spin labeling of cysteine mutants, on-column spin labeling was performed as followed. After washing any remnants from the cell lysate, 1 CV of colicin lysis buffer, containing 100 µl of freshly prepared 1mg/100 µl MTSSL stock, was added to the column then incubated 2 -3 hrs in the dark with gentle mixing at room temperature (RT). At the end of the incubation, the FT, containing free MTSSL spin label, was collected and the column was washed and the spin labeled colicin mutants were eluted as above.

#### Expression, purification and spin labeling of WT and cysteine mutant BtuB

*E. coli* RK5016 cells expression of WT and cysteine mutant BtuB (V90C, T188C, A288C and D492C) were grown in minimal media as described in Chapter 6. Harvested cells were lysed and centrifuged at 12000 rpm for 20 min at 4 °C to remove cellular debris. As explained in Chapter 6, the OM's were isolated and used to purify BtuB, then reconstituted into POPC lipid vesicles. Briefly, 5 % sarkosyl was added to the supernatant then centrifuged at 32000 rpm for 90 min at 4 °C. The resultant pellet was resuspended in BtuB lysis buffer (10 mM Hepes pH 6.5 buffer). Then the pelleted OM was solubilized into 0.5 g of OG, containing 8 mM EDTA and 100 mM Tris pH 8.0, by incubation at 37 °C for 10 min followed by 2 hrs at RT. The solubilized OM was next centrifuged at 25000 rpm for 60 min at 4 °C. Next the solubilized OM samples were used to purify the BtuB in OG, with or without spin labeling, at the desired cysteine mutation site in BtuB. During the spin labeling, BtuB cysteine mutants were incubated with 200 µL of a freshly prepared 12 mM MTSSL stock for 2 -3 hrs at RT in the dark.

BtuB was then purified using 6 CV of BtuB wash buffer (17 mM OG, 25 mM Tris pH 8.0), 12 CV of a 0 -100 % gradient of BtuB elution buffer, EB (1 M NaCl, 17 mM OG, 25 mM Tris pH 8.0), and 6 CV of 100 % BtuB EB on a Q column. Sample fractions were analyzed by SDS-PAGE and the fractions containing BtuB were pooled, then mixed with a sonicated POPC (20 mg/mL) with OG (100 mg/mL) lipid-detergent mixture, and reconstituted into POPC liposomes by detergent removal via six dialysis buffer exchanges. Reconstituted BtuB was pelleted by centrifugation at 14000 rpm for 40 min at 4° C then concentrated further using a Beckman airfuge. The samples were frozen and stored at -80° C until EPR sample preparation and analysis.

#### Colicin-BtuB binding assays

In order to analyze colicin binding with reconstituted BtuB, colicin E3R samples were first incubated with reconstituted WT or BtuB mutants (1:1 or 1:2 volume ratios). Samples were incubated for 10-15 min at RT followed by three freeze-thaw cycles. Next the samples were loaded into capillary tubes and analyzed by CW EPR and pulse EPR.

Colicin binding to BtuB was also studied using live *E. coli* cells. RK5016 cells expressing WT BtuB were grown in 250 mL minimal media cultures (as described in Chapter 3) until the OD<sub>600</sub> reached ~0.3. Aliquots of 50 mL were centrifuged at 4500 rpm for 10 min at 4° C, to remove any residual waste or nutrients from the culture media. The cell pellets were resuspended into 100 mM HEPES pH 7.0, incubated 5 to 10 min at RT, and then centrifuged at 4000 rpm for 6 min at 4° C. The supernatant was then discarded, the pellets were stored on ice during the EPR sample preparation, then analyzed by CW EPR.

### CW EPR analysis of colicin binding with BtuB

Spin labeled colicin samples were loaded into glass capillaries (0.84 mm O.D. / 0.6 mm I.D.) and CW EPR spectra were collected using an X-band operating Bruker EMX spectrometer with an ER 4123D dielectric resonator, a 100 G sweep width, 1 G modulation, and 2 mW of incident microwave power. Colicin samples incubated with BtuB (reconstituted or intact cells) were analyzed by CW-EPR. All the data was collected at RT. The collected CW EPR spectra were normalized, phase corrected, and plotted using in-house programs (DavePlot CW and CWPhase by David Nyenhuis).

### DEER analysis of colicin binding with reconstituted BtuB

Spin labeled colicin was mixed with reconstituted, spin labeled BtuB. The mixed samples were loaded into quartz capillary tubes and flash frozen using liquid nitrogen. DEER data was collected using a dead-time free, four pulsed DEER sequence on a Bruker ELEXSYS E580 at Q band with a dielectric resonator, EN5107D2 dielectric resonator and a 10 W Bruker AmpQ. During the data collection 32 ns and 36 ns lengths were used for the observe and pump  $\pi$  pulses, given with a 75 MHz frequency offset between the observe and pump pulse.

DEER data were analyzed using the DEER analysis 2015 software<sup>16</sup>. After removing the 2+1 pulse artifact by overlapping excitation bands of the pump and observe pluses, the background was subtracted from the echo signal to collect form factor. The distance distributions were generated from form factors using Tikhonov regularization. The predict DEER distributions for T188R1 BtuB with A370R1 or K363R1 colicin were generated using *in silico* labeling and simulations using MMM 2017.2 with the standard library for R1 label<sup>17</sup>. Data were plotted using

DavePlot (by David Nyenhuis) through Plotly and Dash frameworks developed by the Plotly corporation.

#### Generated BtuB images

Crystal structures of BtuB (1NQG, 2GSK 2ZLD, 3K19 and 1UJW)<sup>6-8, 10, 18</sup> were obtained from the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org)<sup>19</sup> and the images were created using PyMol Molecular Graphics System version 1.8<sup>20</sup>.

## Results and Discussion

The colicin fragment used in these studies is about 10 kDa in size. Thus, this small antiparallel helical fragment tumbles rapidly in solution, which averages the magnetic anisotropy of the label and yields narrow EPR lineshapes characteristic of labels with short rotational correlation times. This was clearly observed for four of the colicin mutants (K363R1, A370R1, T409R1, and K416R1) (Figure A1.5).

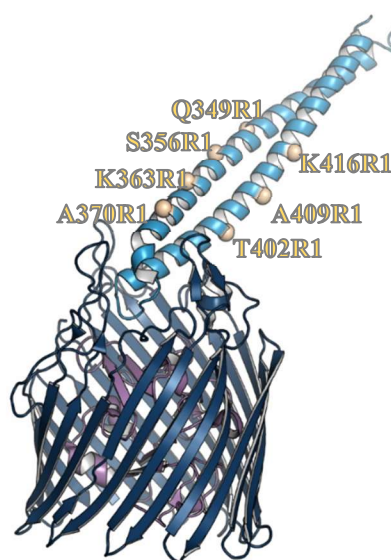


Figure A1.5. Structure of colicin R135 bound to BtuB. The residues **Q349R1**, **S356R1**, **K363R1** and **A370R1** on the first helical arm and residues **T402R1**, **A409R1** and **K416R1** are on the second helical arm.

Upon addition of reconstituted WT BtuB, the EPR lineshape of these mutants' changes. Significant changes were observed for the K363R1 and A370R1 colicin mutants. For these mutants, the fast-moving components were reduced due to limitations in the dynamics of the nitroxide spin labels on colicin the backbone (Figure A1.6a and b) when they bound to the BtuB. However, spin labels' dynamic restriction is based on the proximity of labeled site to the BtuB

binding site. Here, the barrel residues and the core residues interact with the colicin, as well as, the moving BtuB extracellular loops. Based on the crystal structure, both K363C and A370C are in the first helical arm and close to a few extracellular loops (loop 3, 8, 10). This would reduce the mobility of both the K363R1 and A370R1 spins upon binding.

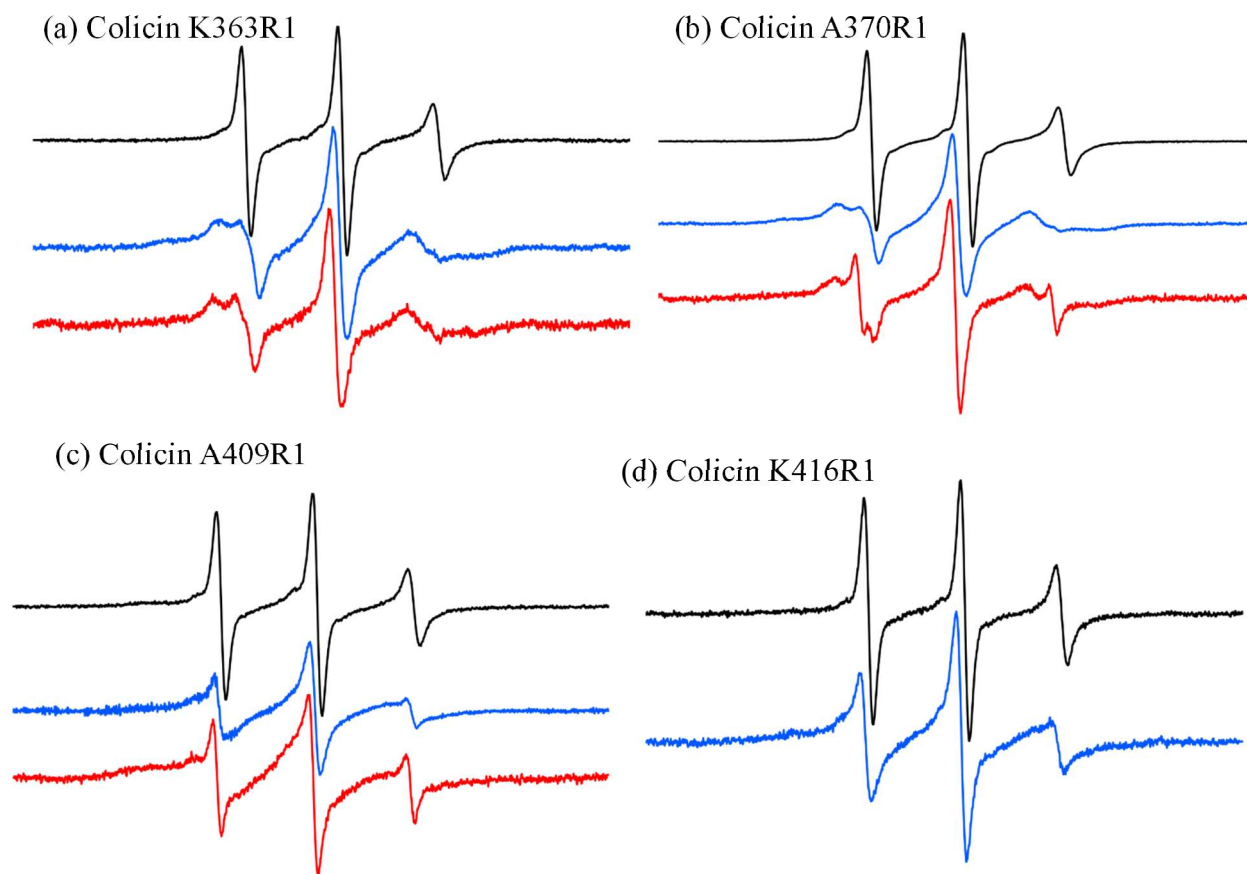


Figure A1.6 Normalized CW EPR spectra obtained from spin labeled colicin E3R mutants (a) K363R1, (b) A370R1, (c) A409R1, and (d) K416R1 in the **apo state**, **bound to reconstituted WT BtuB**, and to **intact *E. coli* expressing WT BtuB**.

However, for the K416R1 colicin mutant on helical arm two, the EPR lineshape change is minimal in the presence of reconstituted BtuB. In contrast to the colicin E3 R135 fragment, in

these studies, a shorter fragment (residues 343-417) was used with a C-terminal His<sub>6</sub> tag. Site K416R1 is near the C-terminus of the shorter colicin R fragment. Hence, when bound to BtuB, colicin's helical arm ends could be more flexible than the full R domain (Figure A1.5d). Additionally, the K416R1 site is just one residue away from the His<sub>6</sub> tag. Thus, upon binding with BtuB the nitroxide spin dynamics change minimally and contains a significant spin population in fast motion.

Consistent with the above results for BtuB (reconstituted), colicin mutants K363R1 and A370R1 bound, upon incubation, with live *E. coli* cells. Again, the corresponding EPR spectra showed reduced nitroxide spin dynamics at these two sites (Figure A1.6a and b). Based on EPR spectra of colicin A409R1 bound BtuB (Figure A1.6c), it has both mobile and immobile components. This could be due some mobility in this arm or due to the use of the shorter fragment of colicin or due to some leftover unbound colicin A409R1 in the sample.

In the colicin bound crystal structure (PDB ID: 1UJW), the extracellular loop 2 is not resolved; however, a structural alignment between colicin R135 bound to BtuB (1UJW) with Ca<sup>2+</sup> and apo BtuB (1NQG) (where the loop 2 is resolved) (Figure A1.7)<sup>7</sup> can be used to predict the distance distributions from site T188R1 on BtuB to colicin sites K363R1 A370R1 using MMM. The preliminary DEER experiments of reconstituted BtuB T188R1 with colicin K363R1 (Figure A1.8) and A370R1 sites (Figure A1.9), display short distance distributions. However, in the initial experiments, the modulation depths are small, in future work optimization should yield larger modulation depths.

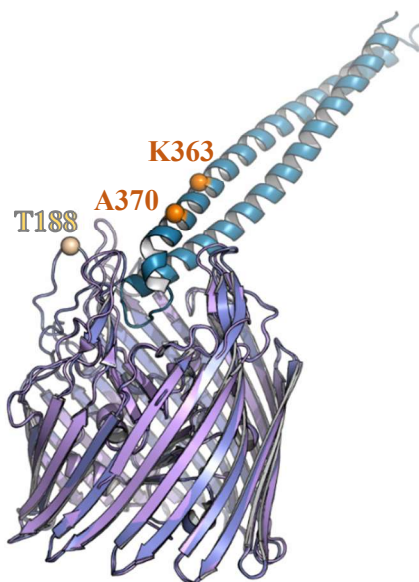


Figure A1.7. Structural alignment of **apo (Ca<sup>2+</sup> bound) BtuB** (PDB ID: 1NQG) and **colicin R135 bound BtuB** (PDB ID: 1UJW). **T188** is in loop two of BtuB and **K363** and **A370** are in close proximity.

The DEER data for colicin K363R1 bound to the reconstituted T188R1 BtuB show two main short distances around 29.8 Å and 22.7 Å. These normalized distance distributions are within the MMM rotamer distribution predictions based on the crystal structure. Based on DEER data, when colicin K363R1 is bound to BtuB, there are two main populations that are in close proximity to loop two of BtuB with about 7 Å difference and 29.8 Å being the predominant distance (Figure A1.8c). For colicin A370R1 bound with T188R1 BtuB, the normalized distance distribution shows a main distance around 28.2 Å (Figure A.1.9). With respect to the MMM simulations, the distance distribution should be significant near 2 nm. However, the distance distribution obtained based on DEER experiment is shifted with higher probability around 28.2 Å. Even though the loop 2 has not been resolved in the colicin bound BtuB crystal structure (1UJW), based on preliminary DEER data, the loop 2 is in close proximity to the bound colicin.

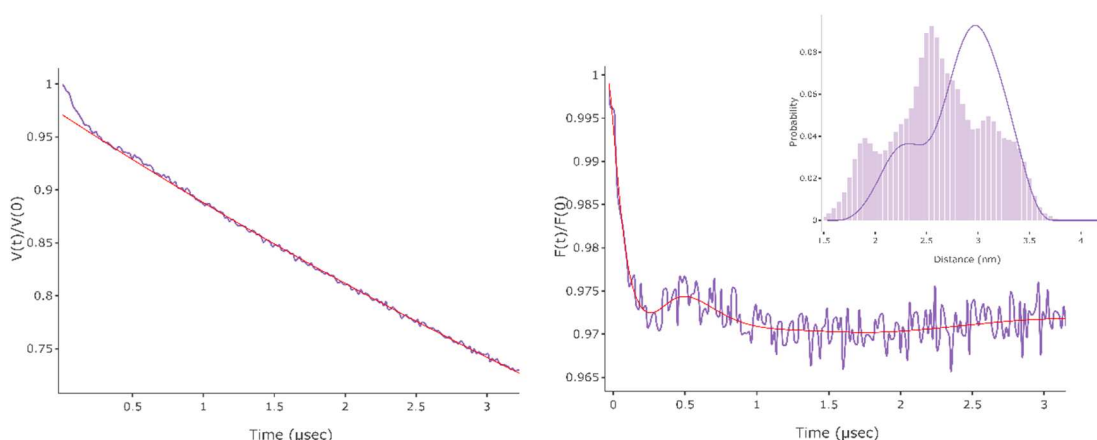


Figure A1.8. Echo signal obtained from pulsed EPR experiment for colicin K363R1 incubated with reconstituted T188R1 BtuB. (a) The **echo signal** collected from T188R1 BtuB bound to colicin K363R1 from the DEER experiment with the **fitted overlay**, (b) the **form factor** after subtracting the background using DEER analysis software with the **fitted overlay**. (c) The resultant **normalized probability distance distribution** and the histogram of distance distribution using MMM for the structural alignment between 1UJW and 1NQG for T188R1 BtuB bound colicin K363R1.

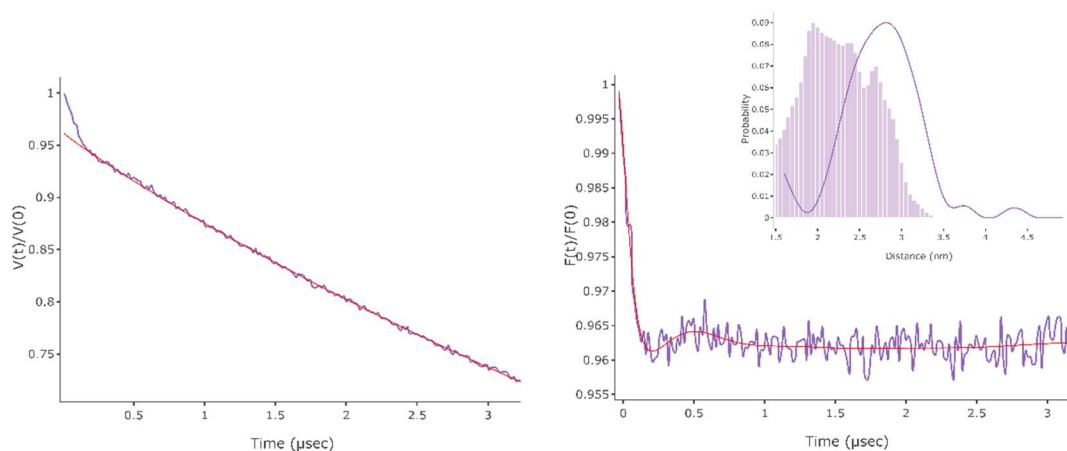


Figure A1.9. Echo signal obtained from pulsed EPR experiment for colicin A370R1 incubated with reconstituted T188R1 BtuB. (a) The **echo signal** collected from T188R1 BtuB bound colicin A370R1 from DEER experiment with the **fitting overlaid**, (b) the **form factor** after subtracting the background using DEER analysis software with the **fitted**

**overlaid** and (c) the resultant **normalized probability distance distribution** and the histogram of distance distribution using MMM stimulations for structure alignment between 1UJW and 1NQG for T188R1 BtuB bound colicin A370R1.

Once the colicin E binds with the primary receptor BtuB, it screens for the translocator in the OM. During these two main events, BtuB bound colicin undergoes conformational changes in the receptor domain and translocator domain. The distance distributions between A370 and K363 with respect to loop 2 broad. These intrinsic conformational changes are potentially enhanced by the R domain which moves the two arms apart. Furthermore, these events potentially direct the primary receptor bound colicin to unfold both ends of the R domain to allow the colicin to scan for the translocator, OmpF, via the fishing pole mechanism.

#### Future directions

Future work in the near future, will investigate the distance distributions associated with the binding events first using the short fragment with reconstituted BtuB. Based on the distances obtained using the crystal structure (Table A1.1) the two core domain mutants (S74 and V90) and the colicin mutants are within a measurable range (1.5 -8 nm) for DEER experiments. Thus, in addition to T188C BtuB these two core domain mutants can also be used to obtain distance distributions.

Colicin mutant	S74C BtuB (Å)	V90C BtuB (Å)
S349C	53.4	44.8
Q356C	44.7	37.0
K363C	36.0	29.8
A370C	29.0	25.8
T402C	29.9	21.8
A409C	39.4	31.8
K416C	49.4	42.3

Table A1.1. The distances between S74 and V90 BtuB in the core domain with colicin residues. The distances are based on the crystal structure 1UJW.

In addition, we will investigate these intrinsic events using different colicin constructs: the complete R domain (R135), the R and T domains, and the complete colicin E3. (i) According to the R135 bound BtuB crystal structure, the termini of the R domain become disordered (T313-asn322 and lys439-glu447)<sup>8</sup>. Using SDSL at these sites of colicin in R domain in the other constructs, we aim to study the disordering associated with binding events using whole cells. (ii) Based on FRET assays, binding enhances the intermolecular separation of the C and T domains<sup>11</sup>. Thus, the two helical arms potentially move apart from each other. So, SDSL near the binding site, center, and at the two ends (not the disordering regions) can map out these potential movements (Figure A1.11). This investigation can be further extended using labeled C and T domains to obtain distance distributions associated with T domain unfolding upon binding to BtuB.

Based on the findings by Housden and *et. al*, IUTD the colicin E9 passes through two OmpF pores and bind with TolB on the other side<sup>14</sup>. Based on the sequence analysis, the colicin E3 also

has two OmpF binding sites and a TolB binding site in between the OmpF binding sites (Figure A1.10) and N-terminus of T domain (residues 1-83) are disordered.

Col	E9	1	MSGGDGRGHNTGAHSTSG	NINGGPTGIGVSG	GASDGS	GSSEN	NPWG	GGSGSG	IHWGGGSG	RNG	65
			MSGGDGRGHNTGAHSTSG	NINGGPTG+GV	GGASDGS	GSSEN	NPWG	GGSGSG	IHWGGGSG	GNG	
Col	E3	1	MSGGDGRGHNTGAHSTSG	NINGGPTGLGVGG	GASDGS	GSSEN	NPWG	GGSGSG	IHWGGGSG	HNG	65

Figure A1.10. The sequence alignment of colicin E9 and E3 (residues 1-65). The TolB binding site is in between the two OmpF binding sites of colicin E9 which as identical sequence in the colicin E3 IUTD.

Thus, our whole cell method protocol can be modified to screen for colicin E3 IUTD binding events with the OmpF trimer. Selected single residue sites in the IUTD can be used with SDSL coupled to CW EPR to investigate the screening of the OmpF and translocation associated events (Figure A1.11). (iii) If the IUTD moves through the pore of one OmpF unit into the periplasm, then logically, the fast moving IUTD will undergo changes in spin mobility upon interacting with porin reducing the mobility. Then, if it enters the periplasm, the spin should be reduced. Thus, three potential, sequential events should be detectable using CW EPR with live *E. coli* cells. In this case it would be ideal to screen through the IUTD and T domain in order to map how far the full T domain can pass through the porin. (iv) Also, if the first OmpF binding site comes out of the second OmpF unit of the live cells, similar to the colicin E9 translocation complex detected by Housden and *et al*<sup>14</sup>, then we could potentially modify our whole cell experiment procedure to monitor re-entry to the first OmpF binding site back to extracellular side of the cell. Here, colicin mutants in/near the first OmpF binding site region should be bound to BtuB and any unbound colicin will be removed. If the first OmpF binding site region comes out from second porin unit, cysteine mutants in this region can be spin labeled to monitor these events in live cells on a real timescale.

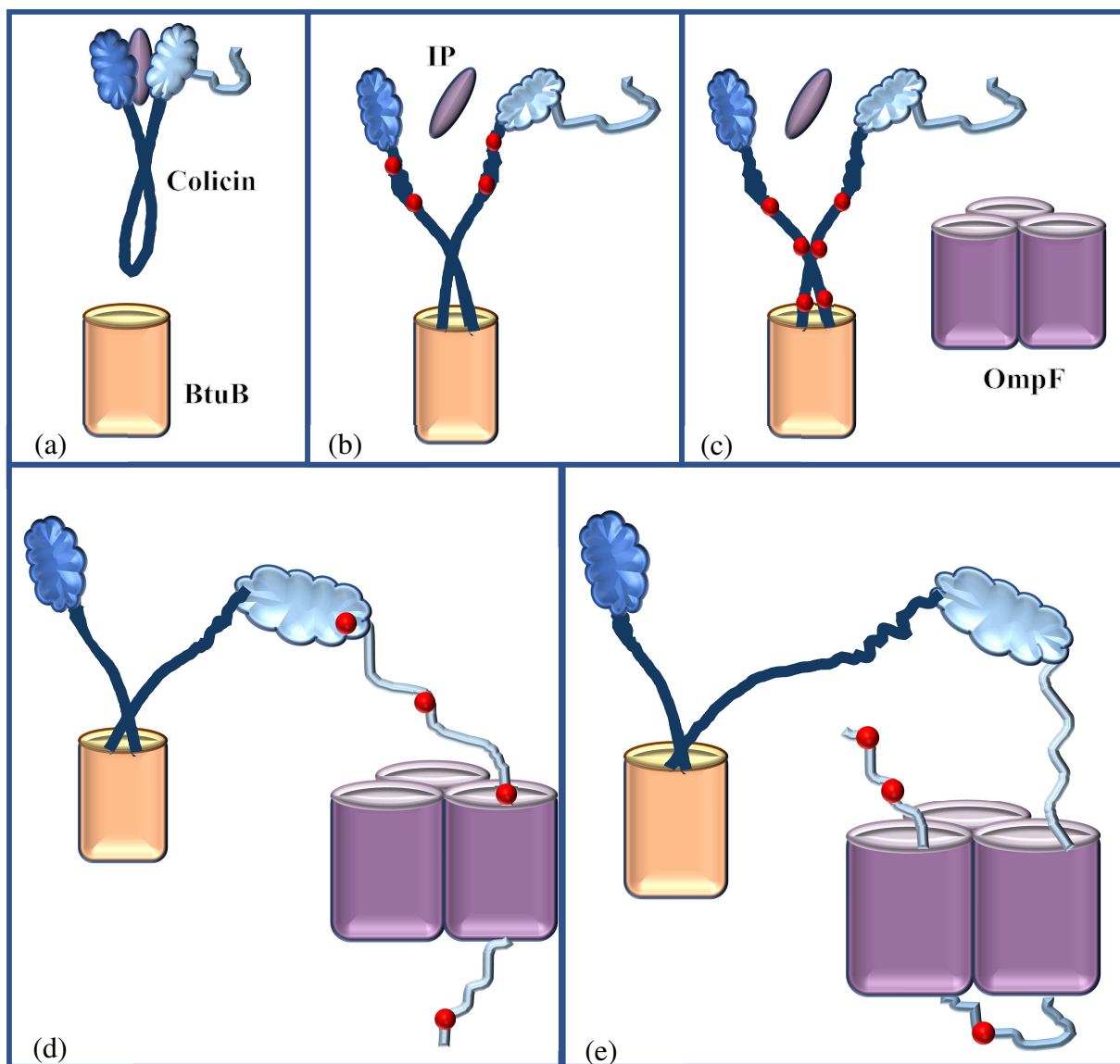


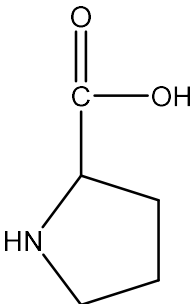
Figure A1.11. Schematic representation of colicin E3 binding and translocation events. (a) During the invasion, colicin E3R is with bound to **immunity proteins (IP)** that are released by host bacteria. (b) upon binding with **BtuB**, the receptor induces conformational changes which leads to dissociation of the **C** and **T** domains from the **IP**. (c) The **IUTD** of colicin E3 screens for **OmpF** then (f) passes though one **OmpF** unit and (e) potentially redirects a portion of the IUTD through the second **OmpF** unit. **SDSL of residues** can be used to screen for these events using EPR.

## References

- [1] Cascales, E., Buchanan, S. K., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S., and Cavard, D. (2007) Colicin biology, *Microbiology and molecular biology reviews* : *MMBR* 71, 158-229.
- [2] Rea, M. C., Ross, R. P., Cotter, P. D., and Hill, C. (2011) Classification of Bacteriocins from Gram-Positive Bacteria, In *Prokaryotic Antimicrobial Peptides: From Genes to Applications* (Dridger, D., and Rebuffat, S., Eds.), pp 29-53, Springer New York, New York, NY.
- [3] Ghequire, M. G. K., and De Mot, R. (2014) Ribosomally encoded antibacterial proteins and peptides from *Pseudomonas*, *FEMS Microbiology Reviews* 38, 523-568.
- [4] Jakes, K. S., and Cramer, W. A. (2012) Border crossings: colicins and transporters, *Annual review of genetics* 46, 209-231.
- [5] Soelaiman, S., Jakes, K., Wu, N., Li, C., and Shoham, M. (2001) Crystal structure of colicin E3: implications for cell entry and ribosome inactivation, *Mol Cell* 8, 1053-1062.
- [6] Chimento, D. P., Mohanty, A. K., Kadner, R. J., and Wiener, M. C. (2003) Substrate-induced transmembrane signaling in the cobalamin transporter BtuB, *Nature structural biology* 10, 394-401.
- [7] Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex, *Science* 312, 1396-1399.
- [8] Kurisu, G., Zakharov, S. D., Zhalnina, M. V., Bano, S., Eroukova, V. Y., Rokitskaya, T. I., Antonenko, Y. N., Wiener, M. C., and Cramer, W. A. (2003) The structure of BtuB with bound colicin E3 R-domain implies a translocon, *Nature structural biology* 10, 948-954.
- [9] Zakharov, S. D., Eroukova, V. Y., Rokitskaya, T. I., Zhalnina, M. V., Sharma, O., Loll, P. J., Zgurskaya, H. I., Antonenko, Y. N., and Cramer, W. A. (2004) Colicin occlusion of OmpF and TolC channels: outer membrane translocons for colicin import, *Biophysical journal* 87, 3901-3911.
- [10] Yamashita, E., Zhalnina, M. V., Zakharov, S. D., Sharma, O., and Cramer, W. A. (2008) Crystal structures of the OmpF porin: function in a colicin translocon, *Embo j* 27, 2171-2180.
- [11] Zakharov, S. D., Sharma, O., Zhalnina, M. V., and Cramer, W. A. (2008) Primary events in the colicin translocon: FRET analysis of colicin unfolding initiated by binding to BtuB and OmpF, *Biochemistry* 47, 12802-12809.
- [12] Spector, J., Zakharov, S., Lill, Y., Sharma, O., Cramer, W. A., and Ritchie, K. (2010) Mobility of BtuB and OmpF in the *Escherichia coli* outer membrane: implications for dynamic formation of a translocon complex, *Biophysical journal* 99, 3880-3886.
- [13] Klein, A., Wojdyla, J. A., Joshi, A., Josts, I., McCaughey, L. C., Housden, N. G., Kaminska, R., Byron, O., Walker, D., and Kleanthous, C. (2016) Structural and biophysical analysis of nuclease protein antibiotics, *The Biochemical journal* 473, 2799-2812.
- [14] Housden, N. G., Hopper, J. T. S., Lukyanova, N., Rodriguez-Larrea, D., Wojdyla, J. A., Klein, A., Kaminska, R., Bayley, H., Saibil, H. R., Robinson, C. V., and Kleanthous, C. (2013) Intrinsically Disordered Protein Threads Through the Bacterial Outer-Membrane Porin OmpF, *Science* 340, 1570-1574.
- [15] One Shot® TOP10 Competent Cells [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/oneshottop10\\_man.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/oneshottop10_man.pdf), Accessed on 02 August 2018.

- [16] Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data, *Applied Magnetic Resonance* 30, 473-498.
- [17] Polyhach, Y., Bordignon, E., and Jeschke, G. (2011) Rotamer libraries of spin labelled cysteines for protein studies, *Physical chemistry chemical physics : PCCP* 13, 2356-2366.
- [18] Kefala, G., Ahn, C., Krupa, M., Esquivies, L., Maslennikov, I., Kwiatkowski, W., and Choe, S. (2010) Structures of the OmpF porin crystallized in the presence of foscholine-12, *Protein Science* 19, 1117-1125.
- [19] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Research* 28, 235-242.
- [20] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.

## Appendix 2. Standard amino acids

<p>Glycine (Gly, G)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{H} \end{array}$	<p>Alanine (Ala, A)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{CH}_3 \end{array}$	<p>Valine (Val, V)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{CH}-\text{CH}_3 \\   \\ \text{CH}_3 \end{array}$
<p>Leucine (Leu, L)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{H}_2\text{C} \\   \\ \text{CH}-\text{CH}_3 \\   \\ \text{CH}_3 \end{array}$	<p>Isoleucine (Ile, I)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{CH}-\text{CH}_3 \\   \\ \text{CH}_2 \\   \\ \text{CH}_3 \end{array}$	<p>Proline (Pro, P)</p> 
<p>Methionine (Met, M)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{CH}_2 \\   \\ \text{CH}_2 \\   \\ \text{S} \\   \\ \text{CH}_3 \end{array}$	<p>Serine (Ser, S)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{CH}_2 \\   \\ \text{OH} \end{array}$	<p>Threonine (Thr, T)</p> $\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\   \\ \text{CH}-\text{OH} \\   \\ \text{CH}_3 \end{array}$

<p>Cysteine (Cys, C)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{SH}  \end{array}  $	<p>Phenylalanine (Phe, F)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{C}_6\text{H}_5  \end{array}  $	<p>Tyrosine (Tyr, Y)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{C}_6\text{H}_4 \\    \\  \text{OH}  \end{array}  $
<p>Tryptophan (Trp, W)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{Indole}  \end{array}  $	<p>Aspartic acid (Asp, D)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{C}=\text{O} \\    \\  \text{OH}  \end{array}  $	<p>Glutamic acid (Glu, E)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{C}=\text{O} \\    \\  \text{OH}  \end{array}  $
<p>Asparagine (Asn, N)</p>	<p>Glutamine (Gln, Q)</p>	<p>Histidine (His, H)</p>

$  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{C}=\text{O} \\    \\  \text{NH}_2  \end{array}  $	$  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{C}=\text{O} \\    \\  \text{NH}_2  \end{array}  $	$  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{C}_4\text{H}_3\text{N}  \end{array}  $
<p>Lysine</p> <p>(Lys, K)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{NH}_2  \end{array}  $	<p>Arginine</p> <p>(Arg, R)</p> $  \begin{array}{c}  \text{O} \\  \parallel \\  \text{H}_2\text{N}-\text{CH}-\text{C}-\text{O} \\    \qquad \qquad \qquad \text{H} \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  \text{NH} \\    \\  \text{C}=\text{NH} \\    \\  \text{NH}_2  \end{array}  $	

## Appendix 3

*E. coli* wild type BtuB amino acid sequence (UniProtKB id P06129)<sup>1</sup>

10	20	30	40	50
<b>MIKKASLLTA</b>	<b>CSVTAFSAWA</b>	QDTSPDTLVV	TANRFEQPRS	TVLAPTTVVT
60	70	80	90	100
RQDIDRWQST	SVNDVLRRLP	GVDITQNGGS	GQLSSIFIRG	TNASHVLVLI
110	120	130	140	150
DGVRLNLAGV	SGSADLSQFP	IALVQRVEYI	RGPRSAVYGS	DAIGGVVNII
160	170	180	190	200
TTRDEPGTEI	SAGWGSNSYQ	NYDVSTQQQL	GDKTRVTLLG	DYAHTHGYDV
210	220	230	240	250
VAYGNTGTQA	QTDNDGFLSK	TLYGALEHNF	TDAWSGFVRG	YGYDNRTNYD
260	270	280	290	300
AYYSPGSPLL	DTRKLYSQSW	DAGLRYNGEL	IKSQLITSYS	HSKDYNYPH
310	320	330	340	350
YGRYDSSATL	DEMKQYTVQW	ANNVIVGHGS	IGAGVDWQKQ	TTTPGTGYVE
360	370	380	390	400
DGYDQRNTGI	YLTGLQQVGD	FTFEGAARSD	DNSQFGRHGT	WQTSAGWEFI
410	420	430	440	450
EGYRFIASYG	TSYKAPNLGQ	LYGFYGNPNL	DPEKSKQWEG	AFEGLTAGVN
460	470	480	490	500
WRISGYRNDV	SDLIDYDDHT	LKYINEGKAR	IKGVEATANF	DTGPLTHTVS
510	520	530	540	550
YDYVDARNAI	TDTPLLRRAK	QQVKYQLDWQ	LYDFDWGITY	QYLGTRYDKD
560	570	580	590	600
YSSYPYQTVK	MGGVSLWDLA	VAYPVTSHLT	VRGKIANLFD	KDYETVYGYQ
610				
TAGREYTLTG	SYTF			

*E. coli* BtuB amino acid sequence contains N-terminus **translocation signal** (amino acids 1-20).

*E. coli* colicin E3 amino acid sequence (UniprotKB id P00646)<sup>1</sup>

10	20	30	40	50
MSGGDGRGHN	TGAHSTSGNI	NGGPTGLGVG	GGASDGSGWS	SENNPWGGGS
60	70	80	90	100
GSGIHWGGGS	GHGNGGGNGN	SGGGSGTGGN	LSAVAAPVAF	GFPALSTPGA
110	120	130	140	150
GGLAVSISAG	ALSAAIADIM	AALKGPFKFG	LWGVALYGVL	PSQIAKDDPN
160	170	180	190	200
MMSKIVTSLP	ADDITESPVS	SLPLDKATVN	VNVRVVDVVK	DERQNISSVS
210	220	230	240	250
GVPMSVPVVD	AKPTERPGVF	TASIPGAPVL	NISVNNSTPA	VQTLSPGVVN
260	270	280	290	300
NTDKDVRPAG	FTQGGNTRDA	VIRFPKDSGH	NAVYVSVSDV	LSPDQVKQRQ
310	320	330	340	350
DEENRRQQEW	DATHPVEAAE	RNYERARAEL	NQANEDVARN	QERQAKAVQV
360	370	380	390	400
YNSRKSELDA	ANKTLADAIA	EIKQFNRFAG	DPMAGGHRMW	QMAGLKAQRA
410	420	430	440	450
QTDVNNKQAA	FDAAAKEKSD	ADAALSSAME	SRKKKEDKKR	SAENNLNDEK
460	470	480	490	500
NKPRKGFKDY	GHDYHPAPKT	ENIKGLGDLK	PGIPKTPKQN	GGGKRKRWTG
510	520	530	540	550
DKGRKIYEW	SQHGELEGYR	ASDGQHLGSF	DPKTGNQLKG	PDPKRNIKKY

L

## Appendix 4

(a) *E. coli* BtuB mutations

Mutation	Protein	T <sub>m</sub> of the primers <sup>2</sup> (Reverse primer, Forward primer) (° C)	Annealing temperatures
R36A/V10C	BtuB	72.8, 71.7	58 - 68
D515A/V10C	BtuB	68.0, 64.0	60 - 73
R47A/V10C	BtuB	73.7, 74.4	62 - 72
D548A/V10C	BtuB	72.3, 72.8	60 - 73
R69A/V10C	BtuB	65.7, 64.3	58 - 68
E419A/V10C	BtuB	72.1, 73.4	60 - 73
R111A/V10C	BtuB	69.3, 70.6	58 - 68
E465A/V10C	BtuB	75.0, 72.2	60 - 73
R526A/V10C	BtuB	62.0, 62.0	60 - 73
D53A/V10C	BtuB	72.0, 69.4	66 - 71
T188C	BtuB	66.1, 67.3	63 - 69
V90C	BtuB	67.1, 67.1	60 - 68
A288C	BtuB	65.1, 65.5	62 - 68
D492C	BtuB	69.6, 67.7	63 - 69
S533C	BtuB	62.6, 62.9	60 - 65
S74C	BtuB	63.6, 63.6	60 - 68
D6C	BtuB	69.1, 67.3	60 - 69
G137C	BtuB	69.1, 65.3	62 - 71

N209C	BtuB	66.9, 67.1	62 - 71
I305C	BtuB	68.0, 64.7	62 - 71
L425C	BtuB	71.8, 71.1	62 -71
Q349C	Colicin E3R	60.4, 60.4	55 - 65
S356C	Colicin E3R	62.8, 62.8	55 -65
K363C	Colicin E3R	62.2, 63.8	52 - 65
A370C	Colicin E3R	54.2, 53.3	52 - 65
T402C	Colicin E3R	68.7, 62.8	55 - 65
A409C	Colicin E3R	62.3, 62.3	55 - 65
K416C	Colicin E3R	63.6, 63.6	55 - 65

## References.

- [1] The UniProt Consortium. (2017) UniProt: the universal protein knowledgebase, *Nucleic Acids Research* 45, D158-D169.
- [2] Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almagbrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntagart, N. O., Sailor, C. A., Dawson, R. B., and Peek, A. S. (2008) IDT SciTools: a suite for analysis and design of nucleic acid oligomers, *Nucleic Acids Research* 36, W163-W169.