Investigation of *Escherichia coli* Tat (twin arginine translocase) transport *in vitro*

Shee Chien Yong

Department of Biochemistry

and

Pembroke College

Trinity Term 2011



A thesis submitted to the Board of the Faculty of Medical Sciences of the University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

Investigation of *Escherichia coli* Tat transport *in vitro*
Shee Chien Yong, Pembroke College
DPhil Thesis, Trinity Term 2011

The Twin arginine translocase (Tat) system catalyzes movement of folded proteins across the cytoplasmic membrane of bacteria and the thylakoid membrane of plant chloroplasts. This transport process requires energy in the form of the transmembrane proton motive force (PMF). The Tat system can be studied *in vitro* using inner membrane vesicles (IMVs) from *E. coli* overproducing the Tat components, TatA, TatB and TatC. However, the transport efficiencies of current *in vitro* Tat transport assays are low.

In this work, current *in vitro* Tat transport assays were compared and parameters that affect transport efficiencies were identified. Mild French press treatment of IMVs resulted in larger IMVs with higher transport efficiencies. Chloride ions were shown to inhibit Tat transport *in vitro*. Generation of a PMF by the activity of ATP synthase gave higher transport efficiencies than generating a PMF by NADH respiration. This understanding was applied to develop an optimized *in vitro* Tat transport assay that showed a higher transport efficiency than currently published methods. Fluorescently labelled Tat substrates were developed to allow quantitative analysis of Tat transport.

The transport of the purified native Tat substrate, CueO into IMVs was characterized using the optimized *in vitro* Tat transport assay. It was shown that the proton concentration (ΔpH) component of the PMF was sufficient to support Tat transport *in vitro*. It was observed that transport of CueO ceased in a time-dependent manner in the *in vitro* Tat transport assays. This loss of transport efficiency could be due, at least in part, to the presence of a PMF since transport efficiency was reduced when IMVs were pre-energized.

Substrates for future *in vitro* single molecule fluorescence microscopy studies of the Tat transport were developed in this work. One of the substrates is fluorescently labelled CueO. The second substrate is the native Tat substrate alkaline phosphatase PhoX from *Vibrio fischeri* which was able to cleave the fluorogenic compound AttoPhos® and can thus be used as an enzymatic reporter of Tat transport.

The structure of a native Tat substrate from *Pseudomonas fluorescens*, PhoX, was solved by X-ray crystallography at a resolution of 1.4Å. PhoX is a monomeric six blade β propeller with two α-helical bundle subdomains. PhoX was shown to have optimum activity at pH8.0. PhoX has a novel catalytic site which requires two Fe^{3+} (including a Cys-coordinated Fe^{3+}) and three Ca^{2+} as cofactors. Mutagenesis studies showed that all the metal ions are required for the integrity of the active site. Co-crystallization of PhoX with vanadate, an inhibitor of PhoX which mimics the transition state, showed that hydrolysis of phosphomonoesters does not involve formation of a covalent phosphoenzyme intermediate. Instead, dephosphorylation of substrates is proposed to occur via a S_N2 reaction with OH^- as the attacking nucleophile.

Table of contents

Abstract i
Table of contents ii
Table of figures vi
Table of tables ix
General abbreviations x
Acknowledgements xii
Chapter 1:Introduction 1
1.1 Protein secretion and translocation 2
1.2 The Sec system 3
1.2.1 SecYEG translocase pore 3
1.2.2 Post-translational substrate targeting 6
1.2.3 Co-translational substrate targeting 6
1.3 The Tat system 8
1.3.1 Tat signal peptide 8
1.3.2 Tat substrate proteins 11
1.3.3 Bioenergetic requirements for Tat transport 16
1.3.4 Tat components 17
1.3.5 TatABC translocase 20
1.3.6 The Tat transport mechanism 23
1.3.7 Proofreading and quality control of Tat substrates 28
1.3.8 PspA and its potential role in Tat transport 29
1.3.9 Tat system in other microorganisms 30
1.3.9.1 Gram-positive bacteria 30
1.3.9.2 Archaea 33
1.3.10 Importance of Tat 34
1.4 Project Aims 35
Chapter 2: Materials and methods 36
2.1 Bacterial strains and plasmid constructs 37
2.2 Molecular biology techniques 43
2.3 Preparation of competent cells and cell transformation 44
2.4 Growth media and growth conditions 45
2.5 Preparation of inner inverted membrane vesicles (IMVs) 47
2.5.1 Müller protocol 47
2.5.2 Musser protocol 48
2.6 Protein purification 50

2.6.1 Preparation of cell lysates 50
2.6.2 Chromatography methods 51
2.6.2.1 GST (glutathione S-transferase) affinity chromatography 51
2.6.2.2 Ni^{2+} affinity chromatography 51
2.6.2.3 Size exclusion chromatography 52
2.6.2.4 Cationic exchange chromatography 52
2.6.3 Ultrafiltration of purified proteins 53
2.6.4 Protein concentration determination 54
2.7 Purification of α-CueO primary antibody 54
2.8 Fluorescent labelling of proteins 55
2.9 Protein analysis and characterization 56
2.9.1 SDS-PAGE 56
2.9.2 Immunoblotting 57
2.9.3 Immunodetection 58
2.10 *In vitro* Tat transport assay 59
2.11 Determination of ΔpH and ΔΨ gradients 60
2.12 ATP synthase activity assay 61
2.13 Characterization of PhoX 61
2.13.1 Phosphatase activity assay using paranitrophenol phosphate 61
2.13.2 Substrate specificity of PhoX using the malachite green phosphatase assay 62
Chapter 3: Optimization of the *in vitro* Tat transport system 63
3.1 Introduction 64
3.2 Purification of model Tat substrates for *in vitro* assays 68
3.2.1 Purification of SufI 68
3.2.2 Purification of CueO 68
3.3 Comparison of current IMV isolation protocols 73
3.4 Optimization of the protocol for IMV preparation 78
3.4.1 Selection of bacterial strain 78
3.4.2 Morphology of IMVs 80
3.4.3 French press procedure 83
3.4.4 Sucrose density purification of IMVs 85
3.4.5 Reconstitution of IMVs with the soluble fraction 90
3.4.6 Buffer components 93
3.4.7 Final optimized IMV preparation protocol 94
3.5 Optimization of the *in vitro* Tat transport assay 96
3.5.1 Energization of IMVs 96

3.5.2 Effect of Cl^- ions 101
3.5.3 Optimization of the IMV to substrate ratio 106
3.5.4 Improving the ratio of mature to full length substrate observed after transport 108
3.5.5 Effect of temperature 109
3.5.6 Final comparison between original and modified Musser *in vitro* Tat transport assays 110
3.6 Alternative *in vitro* Tat transport detection methods 111
3.6.1 Two-step immunoblotting using α-CueO 111
3.6.2 Detection of transport by fluorescence 113
3.6.3 Dependency of Tat transport efficiency on substrate concentration 117
3.7 *In vitro* Tat transport of other Tat substrates 120
3.7.1 Transport of MdoD 121
3.7.2 Transport of $MdoD_{C17I}$ 123
3.8 Discussion 127
Chapter 4: Characterization of *in vitro* Tat transport 129
4.1 Introduction 130
4.2 Bioenergetic studies of the Tat machinery 130
4.2.1 Effect of protonophores and ionophores on Tat transport 130
4.2.2 Sodium motif force (SMF) 136
4.3 Transport kinetics 138
4.3.1 Time trial assay of *in vitro* Tat transport 138
4.3.2 Effect of ADP 141
4.3.3 Role of ATP 143
4.3.4 Sustainability of PMF during *in vitro* transport 147
4.3.5 Transport competency of Tat substrates and IMVs 149
4.3.6 Order of substrate addition 149
4.4 Discussion of the bioenergetic and kinetic studies on Tat transport *in vitro* 151
4.5 Development of a fluorescence-based assay for fluorescence microscopy analysis of the Tat transport *in vitro* 154
4.5.1 TorA-PhoA 156
4.5.2 PhoX 157
4.5.2.1 *P. fluorescens* PhoX 157
4.5.2.2 *V. fischeri* PhoX 159
4.5.3 Discussion 165
Chapter 5: Biochemical and structural studies of *Pseudomonas fluorescens* PhoX 166
5.1 Introduction 167

5.2 Purification of *P. fluorescens* PhoX 169
5.3 Characterization of *P. fluorescens* PhoX 174
5.3.1 Determination of optimum pH for PhoX activity 174
5.3.2 Metal dependency of PhoX 175
5.3.3 Substrate specificity of PhoX 179
5.3.4 Characterization of *Vibrio fischeri* PhoX 183
5.3.5 Comparison of biochemical properties between PhoX and PhoA 185
5.4 Structural studies of *P. fluorescens* PhoX 186
5.4.1 Crystallization of *P. fluorescens* PhoX 186
5.4.2 Crystallographic data 187
5.4.3 Overall architecture of PhoX 190
5.4.3.1 Blade alignment 193
5.4.4 Sequence and structural comparison with close structural homologues of PhoX 194
5.4.5 Multiple sequence alignment of PhoX homologues 196
5.4.6 Metal binding sites 198
5.4.7 Site directed mutagenesis studies of active site residues 205
5.4.8 Oxidation state of the iron cofactors 207
5.4.9 Cys-Fe^{3+} coordination 211
5.4.10 PhoX active site 212
5.5 Discussion 216
5.5.1 Proposed mechanism of PhoX 217
5.5.2 PhoX and the Tat system 219
5.5.3 Regulation of PhoX in the cytoplasm 219
5.5.4 Comparison between PhoX and PhoA 220
Chapter 6: Conclusions and future directions 222
6.1 Insights from *in vitro* Tat transport assays 223
6.2 PhoX 226
6.3 Future directions 228
Chapter 7: Bibliography 230
Chapter 8: Appendix 244

Table of figures

Chapter 1

Figure 1- 1: Crystal structure of SecYEβ from *M. jannaschii* 5
Figure 1- 2: Schematic overview of protein translocation and membrane protein insertion by the Sec pathway in bacteria 7
Figure 1- 3: Schematic representation of Sec and Tat signal peptides 9
Figure 1- 4: Tat components that make up the Tat translocase 17
Figure 1- 5: Solution structure of *B. subtilis* $TatA_d$ 18
Figure 1- 6: 3D architecture of TatA complexes determined by electron microscopy 22
Figure 1- 7: 3D map of TatBC complex 23
Figure 1- 8: 3D maps of TatBC-$SufI_{His}$ complexes 24
Figure 1- 9: Proposed mechanism of Tat transport. 27

Chapter 3

Figure 3- 1: Illustration of the *in vitro* Tat transport assay 65
Figure 3- 2: Tat substrates used in Musser's *in vitro* transport assay 67
Figure 3- 3: Purification of full length $CueO_{His}$ by GSTrap affinity chromatography 69
Figure 3- 4: Purification of full length $CueO_{His}$ by size exclusion chromatography 70
Figure 3- 5: Degradation of $CueO_{His}$ 71
Figure 3- 6: Cationic exchange chromatography of CueO 72
Figure 3- 7: Representative immunoblot analysis of overproduction of TatABC in IMVs 75
Figure 3- 8: TatB accessibility assay to test orientation of IMVs 76
Figure 3- 9: Comparison of transport efficiency between Müller and Musser IMVs 77
Figure 3- 10: Selection of *E. coli* strain 79
Figure 3- 11: Negative stain electron microscopy images of IMVs 80
Figure 3- 12 Analysis of Musser IMVs by transmission EM 82
Figure 3- 13: Effect of French press treatment on Tat transport efficiency of IMVs 84
Figure 3- 14: Effect of French press pressure on Tat transport efficiency 85
Figure 3- 15: Comparison of Musser and Müller IMV purification methods 86
Figure 3- 16: Effect of different sucrose density separation methods on Tat transport efficiency. 87
Figure 3- 17: Negative stain EM analysis of IMVs isolated by sucrose cushion and sucrose gradient 89
Figure 3- 18: Schematic representation of preparation of IMVs 91
Figure 3- 19: Reconstitution of IMVs with soluble fraction 92
Figure 3- 20: Effect of buffer composition used for isolation of IMVs on Tat transport efficiency 94
Figure 3- 21: Final comparison of Tat transport competency of IMVs 95
Figure 3- 22: Effect of IMVs energization method on Tat transport efficiency 96
Figure 3- 23: PMF generation by NADH and ATP in *E. coli* 97
Figure 3- 24: Transport assay under the conditions of the energization tests 98
Figure 3- 25: ΔpH and ΔΨ traces of IMVs energized by NADH or ATP 100
Figure 3- 26: Effect of Cl^- on *in vitro* Tat transport efficiency 101

Figure 3- 27: Effect of Cl^-, Ac^- and SO_4^{2-} on generation of PMF by IMVs 103
Figure 3- 28: Illustration of the possible effects of ClC-ec1 transporters on ΔpH and ΔΨ of IMVs 105
Figure 3- 29: Immunoblot analysis of overproduction of TatABC in IMVs made from ΔClC mutants 105
Figure 3- 30: Optimization of IMVs to CueO ratio 107
Figure 3- 31: Effect of overexpression of *lepB* on *in vitro* Tat transport efficiency 109
Figure 3- 32: Effect of temperature on *in vitro* Tat transport assay 110
Figure 3- 33: Final comparison between original and modified Musser *in vitro* Tat transport assays 110
Figure 3- 34: One-step vs two-step immunoblot 111
Figure 3- 35: Detection of transport by α-CueO 113
Figure 3- 36: CueO and construction of CueO_GS3 115
Figure 3- 37: Transport competency of $CueO_GS3^{633}$ 116
Figure 3- 38: Representative standard curve used for quantitative analyses of fluorescence- based *in vitro* Tat transport assays 117
Figure 3- 39: Dependency of Tat transport efficiency on substrate concentration 119
Figure 3- 40: Characterization of MdoD 122
Figure 3- 41: *In vitro* Tat transport assay using MdoD 123
Figure 3- 42: *In vitro* Tat transport assay using $MdoD_{C17I}$ 124
Figure 3- 43: Model of MdoD 125
Figure 3- 44: Labelling and transport of $MdoD_{C17I}$ 126

Chapter 4

Figure 4- 1: Transport assay in the presence of ionophores and protonophores 132
Figure 4- 2: Effect of ionophores on ΔpH and ΔΨ of IMVs 133
Figure 4- 3: Effect of sodium motive force on *in vitro* Tat transport 137
Figure 4- 4: *In vitro* transport kinetics of $CueO_GS3^{633}$ 140
Figure 4- 5: Effect of ATP regenerating systems on Tat transport 142
Figure 4- 6: Effect of ADP on Tat transport efficiency 143
Figure 4- 7: Effect of addition of second dose of ATP on transport assay efficiency 144
Figure 4- 8: Measurement of ATP synthase activity 146
Figure 4- 9: Effect of active Tat transport of CueO on ΔpH and ΔΨ 148
Figure 4- 10: Effect of order of IMV energization and substrate addition 150
Figure 4- 11: Schematic representation of hypothesized conformational changes of TatABC upon substrate binding 153
Figure 4- 12: Schematic representation of TIRF setup for single molecule fluorescence microscopy to investigate Tat transport 155
Figure 4- 13: The reaction of AttoPhos® substrate with alkaline phosphatase 156
Figure 4- 14: Purification of *P. fluorescens* $PhoX_{His}$ by GSTrap affinity chromatography 158
Figure 4- 15: *In vitro* Tat transport of *P. fluorescens* PhoX 159
Figure 4- 16: Purification of *V. fischeri* $PhoX_{His}$ by GSTrap affinity chromatography 160
Figure 4- 17: *In vitro* transport assay of *V. fischeri* PhoX 161
Figure 4- 18: Purification of *V. fischeri* PhoX by Ni^{2+} affinity chromatography 162

Figure 4- 19:*In vitro* transport assay of *V. fischeri* PhoX purified via C-terminal His-tag 163

Chapter 5

Figure 5- 1: Purification of PhoX by Ni^{2+} affinity chromatography 170
Figure 5- 2: Size exclusion chromatography of PhoX 171
Figure 5- 3: PhoX is a coloured protein 172
Figure 5- 4: Sequence of *P. fluorescens* PhoX 173
Figure 5- 5: pH optimum of PhoX activity 175
Figure 5- 6: PhoX activity in the presence of EDTA 176
Figure 5- 7: Effect of EDTA on the visible absorption spectrum of PhoX 177
Figure 5- 8: Activity of EDTA-treated PhoX 178
Figure 5- 9: Cation requirement for enzymatic activity of PhoX 179
Figure 5- 10: Substrate specificity of PhoX 182
Figure 5- 11: Characterization of *V. fischeri* PhoX 184
Figure 5- 12: Crystal of PhoX 187
Figure 5- 13: Overall structure of PhoX 191
Figure 5- 14: The active site cavity of PhoX 192
Figure 5- 15: Electrostatic surface potential of PhoX 193
Figure 5- 16: Amino acid sequence alignment of the six β-propeller blades of PhoX from *P. fluorescens.* 194
Figure 5- 17: Sequence and structural alignment of PhoX with gluconolactonase 196
Figure 5- 18: PhoX amino acid conservation displayed on the *P. fluorescens* PhoX structure. 197
Figure 5- 19: MicroPIXE analysis of *P. fluorescens* PhoX 200
Figure 5- 20: The metal ions and their coordinating ligands in PhoX 201
Figure 5- 21: Metal binding sites in PhoX 202
Figure 5- 22: Sequence alignment of metal coordinating ligands in bacterial PhoX homologues 203
Figure 5- 23: CD spectra of PhoX variants 206
Figure 5- 24: EPR spectrum of EDTA treated PhoX 208
Figure 5- 25: Effect of dithionite on PhoX 210
Figure 5- 26: Active site of PhoX in the presence of phosphate or vanadate 214
Figure 5- 27: Vanadate and phosphate inhibition assays 215
Figure 5- 28: Schematic representation of proposed mechanism of PhoX 218
Figure 5- 29: X-ray crystal structure of PhoA 221

Chapter 6

Figure 6- 1: Schematic representation of issues with the *in vitro* Tat transport system 225

Chapter 8

Appendix 1: Multiple sequence alignment of PhoX from 50 different species 25845

Table of tables

Chapter 1

Table 1- 1: List of known and predicted *E. coli* Tat substrates 13

Table 1- 2: Predicted number of Tat substrates and Tat components numbers in prokaryotes 31

Chapter 2

Table 2- 1: List of *E. coli* strains used 37

Table 2- 2: List of plasmids used 38

Table 2- 3: List of primers used 41

Table 2- 4: Buffers for making competent cells 45

Table 2- 5: Growth media 46

Table 2- 6: List of antibiotics 46

Table 2- 7: Buffers for Musser IMV protocol 49

Table 2- 8: List of proteins purified 50

Table 2- 9: Buffers for affinity chromatography and gel filtration chromatography 53

Table 2- 10: Buffers for affinity purification of α-CueO 55

Table 2- 11: Buffers for SDS-PAGE 57

Table 2- 12: Buffers for immunoblotting and immunodetection 58

Table 2- 13: Buffers for translocation assay 60

Chapter 3

Table 3- 1: Comparison of Müller and Musser IMV isolation protocols 74

Table 3- 2: Buffers for storing IMVs 94

Chapter 4

Table 4- 1: Reagents used to dissipate ΔpH and/or ΔΨ in *in vitro* Tat transport assays 131

Table 4- 2: Activity assay of *V. fischeri* PhoX purified via different protocols 163

Chapter 5

Table 5- 1: List of substrate for PhoX activity 181

Table 5- 2: Kinetic properties of PhoX and PhoA 185

Table 5- 3: PhoX X-ray data collection and crystallographic statistics 189

Table 5- 4: Bond distances between metal ions and coordinating ligands 204

Table 5- 5: Activity and stability of PhoX variants 206

Table 5- 6: Effect of dithionite on stability of PhoX 211

Table 5- 7: Sequence alignment of PhoX around C179 212

Table 5- 8: Specific activity of PhoX_C179D 212

General abbreviations

A	Absorbance
ATP	adenosine-5'-triphosphate
ACMA	9-Amino-6-chloro-2-methoxyacridine
Amp	ampicillin
BSA	bovine serum albumin
CD	circular dichroism
C-terminal	carboxyl- terminal
Δ pH	proton gradient
Δ Ψ	electric field gradient
DMSO	dimethyl sulfoxied
DSF	Differential scanning fluorimetry
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
EPR	electron paramagnetic resonance
FeS	iron-sulfur
GFP	green fluorescent protein
GST	glutathione-S-transferase
GTP	guanosine-5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horse radish peroxidase
IMVs	inner (inverted) membrane vesicles
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
λ	wavelength
LB	Luria broth
MES	2-(N-morpholino) ethanesulfonic acid
microPIXE	proton induced X-ray emission
MOPS	3-(N-morpholino) propanesulfonic acid
MWCO	molecular weight cut off
NAD+	nicotinamide adenine dinucleotide (oxidized)
NADH	nicotinamide adenine dinucleotide (reduced)
NMR	nuclear magnetic resonance
N-terminal	amino-terminal
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBST	Phosphate buffered saline
PEG	Polyethylene glycol
P_i	inorganic phosphate
PK	proteinase K
PMF	proton motive force
PMSF	phenylmethylsulfonyl fluoride

pNPP	p-nitrophenyl phosphate
SDS	sodium dodecyl sulphate
Sec	secretory
SRP	signal recognition particle
Tat	twin arginine translocase
TCA	trichloroacetic acid
TCEP	tris-(2-carboxyethyl)phosphine
TMAO	trimethylamine-N-oxide
Tris	tris(hydroxymethyl)aminomethane

Acknowledgements

I would like to begin by thanking my supervisor, Dr. Ben Berks, for this opportunity to undertake DPhil studies in his laboratory. I am extremely grateful for his patience and constant guidance. His support and encouragement helped me through the good and bad times over the past three years.

I would also like to thank every member from the Berks group, past and present, for their companionship and assistance during and outside of lab hours. Special mention to Dr. Nicholas Greene for picking me up from the bus station early on New Year's day when I first arrived and helping me settle down in Oxford.

I am also indebted to Prof. Susan Lea and Dr. Pietro Roversi for their help in X-ray crystallography, especially when I have no prior experience in structural biology. I would also like to acknowledge James Lillington, Oliver Zeldin, Prof. Elspeth Garmen and Dr. Mike Shaw for helping me with the biophysical experiments.

Special thanks to Mary Gregoriou for her help in organizing my pre-departure documents and making my transition to the UK an easy one. Thank you to the staff of Pembroke College for helping out with admissions related issues.

I am also extremely grateful for the financial support provided by Clarendon Fund and Pembroke College during my DPhil years.

Thanks are also due to all my friends in Oxford for making my social life an enjoyable one and for being there for me when times are tough.

And finally, a huge thank you to my family, especially my parents. None of these would have been possible without their love and constant support.

Chapter 1 :

Introduction

1.1 Protein secretion and translocation

Proteins are essential parts of organisms and participate in virtually every process within cells such as enzymatic catalysis and cell signaling. Protein synthesis occurs in the cytoplasm in all organisms, eukaryotes and prokaryotes alike. Some proteins remain in the cytoplasm while others are targeted out of the cytoplasm. Gram-negative bacteria, such as *Escherichia coli* have both an inner (cytoplasmic) membrane and an outer membrane with an aqueous space in between the membranes known as the periplasm. In these organisms, proteins can be targeted to the inner or outer membrane, periplasm or the extracellular milieu.

Proteins can be transported in their folded or unfolded state depending on the translocation system. Due to the overall complexity of protein translocation where different substrates have to be targeted to various destinations via distinct pathways, these processes need to be under tight regulation so that only the correct pathways are targeted. One of the mechanisms employed to ensure proper targeting of substrates is to tag substrates with a signal peptide at either terminus that is recognized by their cognate translocation system. Energy is likely to be required for the translocation process and can be in the form of nucleotide-5'-triphosphate (NTP) hydrolysis or the proton motive force (PMF) (Holland 2004).

In *E. coli*, proteins are transported across the cytoplasmic membrane by two main transport systems: the general secretory (Sec) pathway and the twin-arginine translocase (Tat) pathway. Approximately 20% (~900) of *E. coli* proteins are transported via the Sec pathway (Pugsley 1993) while only 30 are known or predicted to be native *E. coli* Tat substrates (Berks *et. al.* 2000; Hatzixanthis *et. al.* 2003). Proteins targeted to the Sec and Tat pathways can be transported across the cytoplasmic membranes to the periplasm or inserted in the cytoplasmic membrane.

1.2 The Sec system

The Sec system is the major route of transport of unfolded polypeptide chains across the inner membrane to the periplasm. Peptide substrates targeted for the Sec pathway contain an N-terminal Sec signal sequence. The Sec signal sequence has a tripartite structure with positively charged amino acid residues at the N-terminus region (n-region), a hydrophobic core (h-region) and a polar C-terminus carrying the signal peptidase cleavage site (c-region) (Driessen and Nouwen 2008). The cleavage site is preceded by the 'AxA' motif, which comprises of small and neutral residues at the -1 and -3 positions before the cleavage position. This motif is recognized and cleaved by the signal peptidase LepB after transport into the periplasm (Von Heijne 1983, 1985). Sec substrates can either be targeted post-translationally or co-translationally.

1.2.1 SecYEG translocase pore

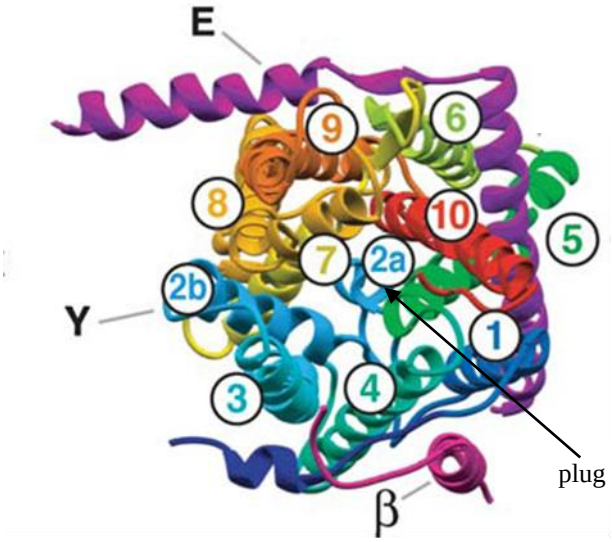
Translocation of both soluble proteins and membrane proteins by the Sec pathway involves a protein conducting channel. The Sec protein conducting channel in *E. coli* is made up of SecY, SecE and SecG subunits. The structure of the detergent solubilised SecYEβ (where Secβ is the archaeal SecG orthologue) complex from the archaeon *Methanococcus jannaschii* has been solved at a resolution of 3.2Å (Figure 1- 1A) (Van Den Berg *et. al.* 2004). This structure most likely reflects the structure of SecYEG in *E. coli* as the transmembrane segments of the *E. coli* SecYEG complex seen in low resolution 2D electron microscopy structure (Breyton *et. al.* 2002) are superimposable onto the transmembrane segments of the *M. jannaschii* SecYEβ complex.

The crystal structure shows a monomeric SecYEβ complex containing a central pore that acts as a translocation channel (Van Den Berg *et. al.* 2004). The SecY subunit forms the core of the translocase. It contains ten transmembrane helices that are split into two domains. The first domain consists of transmembrane helices 1-5 (N-terminus) while the second

domain consists of transmembrane helices 6-10 (C-terminus). The two domains are arranged into a pseudo-symmetrical structure with the two remaining smaller subunits, SecE and Secβ, at the periphery. It is hypothesized that a large, funnel-like cavity which lies at the cytoplasmic side of SecY serves as the channel entrance. Another funnel-like cavity lies at the periplasmic side of SecY. Both funnels meet about halfway across the membrane at a constriction site lined by a 'pore ring' made from hydrophobic residues. This 'pore ring' could serve as a seal to maintain the membrane barrier during transport. In the closed state, the translocon channel is blocked by a 'plug' formed by part of the transmembrane helix 2 of SecY. Van den Berg *et. al.* (2004) propose that once the signal sequence binds, the channel is opened by displacement of the plug. This would result in an hourglass shaped channel with aqueous funnels at both ends and a hydrophobic constriction site in the middle. The translocon also has a lateral gate formed by relatively short segments of transmembrane helices 2b, 3, 7 and 8 which allows transmembrane segments of nascent membrane proteins targeted to the Sec system to exit laterally from the channel into the lipid phase.

Biochemical studies using cross-linking agents suggest that SecYEG complex is functional as a monomer (Osborne and Rapoport 2007). Although dimers of SecYEG complexes have also been observed by electron microscopy (Breyton *et. al.* 2002; Mitra *et. al.* 2005), the role of SecYEG oligomerization has yet to be determined as SecYEG is active as a monomer.

A)



B)

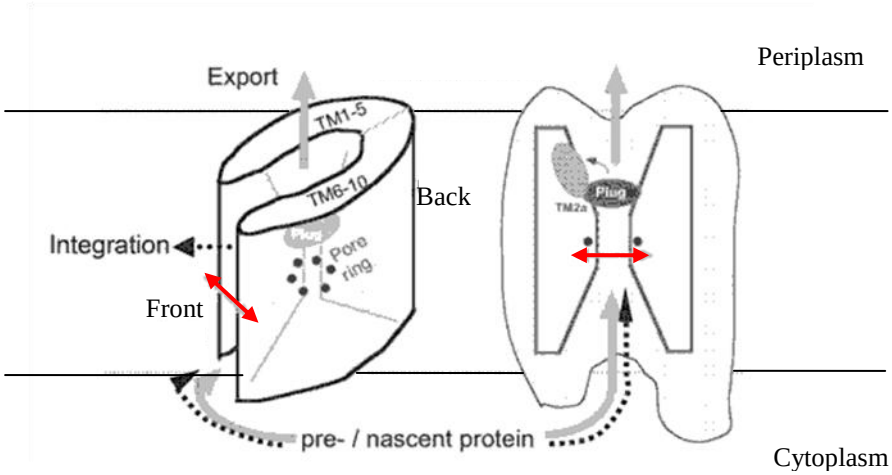


Figure 1- 1: Crystal structure of SecYEβ from *M. jannaschii*

(A) Ribbon representation of a cytoplasmic view of the *M. janaschii* SecYEβ complex (PDBID 1RHZ) with individual helices coloured and labelled. SecE (in purple) and Secβ (orthologus of SecG in *E. coli*) (in pink) are located at the periphery of SecY. Transmembrane helix 2a forms the plug which seals the channel in the closed state while transmembrane helices 2b, 3, 7 and 8 form the lateral pore. Figure adapted from van den Berg *et. al.* (2004). (B) Diagrammatic sketch of the SecYEβ translocon in the closed state. The left figure shows the clam shell-like pseudo-symmetrical arrangement, with a lateral gate on the front side. The figure on the right shows the hourglass shaped channel with the two funnel like cavities, the plug and the pore ring made up of hydrophobic residues at the constriction site. The channel opening movements are indicated by the red arrow. The grey arrows and dotted black arrows indicate the route that secreted proteins and membrane proteins may follow, respectively. Figure adapted from (Ito and Mori 2009)

1.2.2 Post-translational substrate targeting

Most secretory proteins in bacteria are targeted to the Sec translocase in a post-translational manner. The substrate protein is completely synthesized and released by the ribosome before transport. Many proteins secreted by this pathway utilize the molecular chaperone SecB. SecB binds to the mature domain of the newly synthesized precursor proteins and stabilizes them in an unfolded state (Kumamoto 1989; Liu *et. al.* 1989).

The SecB-substrate complex then interacts with SecA (Figure 1- 2), which is a soluble protein that functions as a receptor for the secretory proteins. SecA interacts with the precursor protein, SecB, SecY as well as acidic phospholipids. SecA is also an ATPase. Together with a proton motive force (PMF), SecA provides the driving force for the translocation of substrate through the SecYEG pore (Hartl *et. al.* 1990; Bauer and Rapoport 2009). The oligomeric state of SecA remains controversial. The crystal structure of *Bacillus subtilis* SecA has been solved in both the dimeric (Hunt *et. al.* 2002) and monomeric (Osborne *et. al.* 2004) states. It has been shown that SecA dimers are in rapid equilibrium with monomers in solution (Woodbury *et. al.* 2002).

1.2.3 Co-translational substrate targeting

Co-translational protein targeting is mediated by the signal recognition particle (SRP) pathway. Most of the substrates targeted to this pathway are membrane bound proteins (Valent *et. al.* 1997). In *E. coli*, the SRP pathway consists of a 4.5S RNA, a 48kDa guanosine-5'-triphosphate (GTP) -binding protein known as Fifty-four homologue (Ffh) and FtsY, which is a peripheral membrane receptor protein that also binds GTP (Bernstein *et. al.* 1989; Luirink and Sinning 2004). Proteins bound for the SRP pathway are recognized by the hydrophobic region of their signal sequence or by a transmembrane segment emerging from the ribosome. The nascent protein-ribosome-mRNA-SRP complex is subsequently targeted to FtsY. The polypeptide chain is transferred to the SecYEG translocase (Figure 1- 2). Release

of membrane protein from SRP and FtsY depends on GTP hydrolysis (De Keyzer *et. al.* 2003; Driessen and Nouwen 2008).

For nascent inner membrane proteins, the transmembrane segments exit the SecYEG complex by the lateral pore into the lipid bilayer. Insertion of inner membrane proteins into the lipid bilayer can be assisted by the membrane insertase, YidC. For example, cross-linking studies show that YidC interacts with the nascent transmembrane helix of FtsQ, a Sec dependent membrane protein, and with SecYEG after co-translational targeting by the SRP pathway (Scotti *et. al.* 2000). This suggests that YidC can be involved in lateral transfer of transmembrane segments from the Sec translocon into the lipid bilayer.

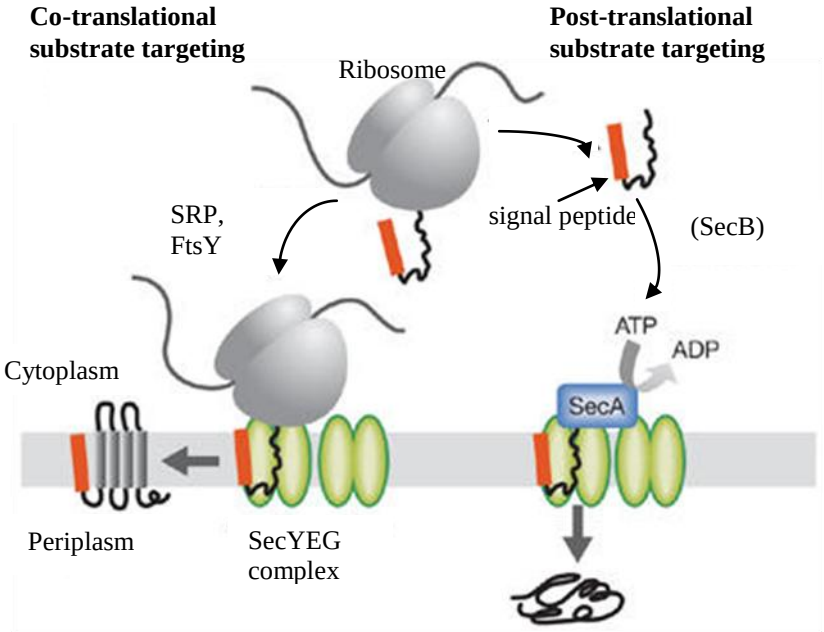


Figure 1- 2: Schematic overview of protein translocation and membrane protein insertion by the Sec pathway in bacteria

Nascent chains with a signal sequence (red) emerge from the ribosomes (grey) and are targeted to the SecYEG complex (green) co-translationally by the signal recognition particle (SRP) and FtsY (SRP receptor). Alternatively, they can be presented post-translationally through SecA with the aid of SecB (right). Figure adapted from (Robson and Collinson 2006)

1.3 The Tat system

The Tat system, like the Sec system, transports proteins across the inner membrane into the periplasm. However, the Tat system mainly transports folded proteins while the Sec system is specific for unfolded polypeptides. The bacterial Tat translocase system is both structurally and mechanistically closely related to the thylakoid ΔpH/Tat pathway (cpTat) (Dalbey and Robinson 1999). Homologues of the essential bacterial Tat components, TatA, TatB and TatC, can be found in the thylakoids as Tha4/Tha9, Hcf106 and cpTatC respectively. Signal sequences seem to be interchangeable between the Tat pathway of bacteria and thylakoids (Mori and Cline 1998; Wexler *et. al.* 1998). To date, most work on the bacterial Tat system has been focused on the Tat system found in *E. coli*.

1.3.1 Tat signal peptide

An N-terminal signal peptide is required for Tat transport. The Tat signal peptide has a tripartite structure consisting of 3 distinct regions as in the Sec signal peptides: positively charged N-terminus (n-region); hydrophobic region (h-region) and the c-region where the cleavage site is usually found. The major determinant of the cleavage site is the presence of amino acids with small, neutral side chains at the -1 and -3 positions relative to the site of cleavage (Von Heijne 1990). However, the Tat signal sequences have an extended n-region which results in a longer signal sequence compared to the Sec signal sequences. The h-region of Tat signal peptides are also, on average, less hydrophobic than Sec signal peptides. It has been shown that by increasing the hydrophobicity of the h-region it is possible to direct a Tat substrate to the Sec pathway. In addition, the c-region of Tat signal peptides can be more basic than Sec signal peptides as basic residues in this region in Sec signal peptides interfere with Sec transport (Figure 1- 3) (Cristóbal *et. al.* 1999).

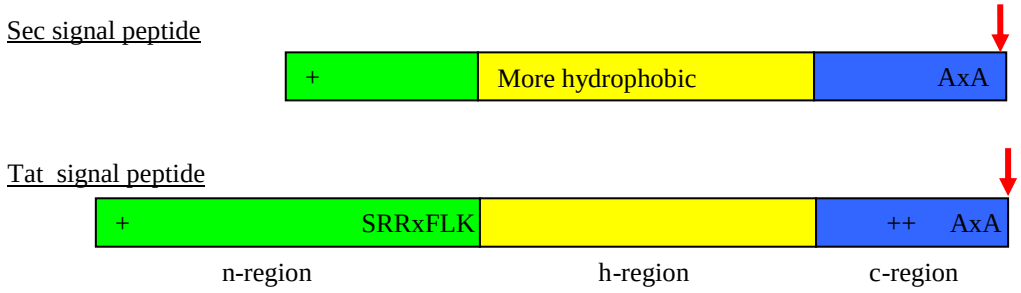


Figure 1- 3: Schematic representation of Sec and Tat signal peptides.
The key difference between the Tat and Sec signal peptides is the presence of the SRRxFLK motif at the n-region/h-region boundary in Tat signal peptides. Tat signal peptides are longer than the Sec signal peptides. The h-region of Tat signal peptides are less hydrophobic than the Sec signal peptides. The c-region in Tat signal peptides are more positively charged compared to the Sec signal peptides. The signal peptidase, LepB, cleavage site is represented by red arrows. This cleavage site is preceded by the 'AxA' motif, which comprises of small and neutral residues at the -1 and -3 positions before the cleavage position.

The Tat signal peptide in bacteria is characterized by the presence of the consensus motif Ser/Thr-Arg-Arg-X-Phe-Leu-Lys (where X can be any polar amino acid) at the n-region/h- region boundary. The twin arginine dipeptide in this consensus sequence gave rise to the name twin-arginine translocase (Tat) system. This twin arginine consensus dipeptide is almost invariant in all Tat signal peptides while the frequency of occurrence of the other amino acids exceeds 50% (Berks 1996). The importance of the twin arginine (RR) consensus dipeptide was first observed in glucose-fructose oxidoreductase (GFOR) in *Zymomonas mobilis* (Halbig *et. al.* 1999). The authors showed that substitution of the twin arginines for lysines (RR→KK) as well as a single substitution of the second arginine with either lysine (RR→RK) or glutamine (RR→RQ) abolished transport. On the other hand, substitution of the first arginine for lysine (RR→KR) resulted in lower levels of transport compared to the wild type signal peptide. Studies by Stanley *et. al.* (2000) further support the importance of the twin arginine residues in the Tat signal peptide. This work showed that mutation of both arginines to lysines (RR→KK) in the signal peptide of the native *E. coli* Tat substrates, SufI and CueO, resulted in loss of transport. Substitutions of either of the arginine residues show

varying degrees of Tat transport inhibition. Single arginine to lysine mutations (RR→RK or RR→KR) in SufI still allow transport of SufI by the Tat translocase though the transit half-time time through the translocase is greatly increased. On the other hand, substitution of the first arginine to lysine (RR→KR) completely abolished transport of CueO. Other studies show that the substitution of both arginines to lysines (RR→KK) or the substitution of the second arginine to lysine (RR→RK) in the *E. coli* trimethylamine-N-oxide reductase (TorA) blocked transport while substitution of the first arginine to lysine (RR→KR) is tolerated (Buchanan *et. al.* 2001). Although rare, deviations from the arginine pair in the consensus motif have been reported in nature. Examples include the tetrathionate reductase (TtrB) from *Salmonella entericia* where the first arginine is replaced by lysine (i.e. KR) (Hinsley *et. al.* 2001), pre-pro-penicillin amidase from *E. coli* where the two arginine residues are separated by an asparagine (i.e. RNR) (Ignatova *et. al.* 2002) and a flavin adenine dinucleotide (FAD)-containing oxidoreductase from *Mycobacterium tuberculosis* with a arginine-glutamine (RQ) pair at the twin arginine position (Widdick *et. al.* 2008).

Structural studies on the Tat signal peptide using solution nuclear magnetic resonance (NMR) have shown that the Tat signal peptide of the high potential iron-sulfur protein (HiPIP) from *Allochromatium vinosur* is largely unstructured in solution (Kipping *et. al.* 2003). Circular dichroism (CD) experiments done on the Tat signal peptide of *E. coli* SufI showed that signal peptides can adopt a structure containing up to ~40% of α-helical structure when incubated in membrane-mimetic environments such as detergent micelles. The α-helical region was predicted to begin close to the RR consensus motif (San Miguel *et. al.* 2003). The authors propose a model where the Tat signal peptide adopts the α-helical conformation either upon interaction with the membrane prior to interacting with the Tat translocase or upon direct interaction with the Tat translocase.

1.3.2 Tat substrate proteins

The Tat pathway in bacteria was originally identified by its ability to transport co-factor bound proteins that are assembled in the cytoplasm prior to transport (Berks 1996). In *E. coli*, there are at least 30 putative Tat substrates that contain the twin arginine motif, characteristic of Tat transport, in the signal sequence (Tullman-Ercek *et. al.* 2007). Some of the Tat substrate proteins could be integral membrane proteins (Hatzixanthis *et. al.* 2003). Table 1- 1shows the list of known and predicted Tat substrates.

There are several lines of evidences that support the idea that the Tat pathway transports folded proteins. It has been shown that Tat precursor proteins that accumulate in the cytoplasm of *tat* mutant strains contain bound cofactors, implying that these precursor proteins are folded (Sargent *et. al.* 1998). In addition, if cytoplasmic insertion of cofactors into precursor Tat proteins is blocked, the precursor protein is not transported (Santini *et. al.* 1998). A further example is the transport of fluorescent green fluorescent protein (GFP) to the periplasm by the Tat system. GFP is not fluorescent when translocated in an unfolded fashion to the periplasm of *E. coli* by the Sec system (Feilmeier *et. al.* 2000). However, GFP fused to a Tat signal peptide can be transported via the Tat pathway and was shown to fluoresce in the periplasmic space, indicating that it must have been folded in the cytoplasm prior to transport (Santini *et. al.* 2001). Another piece of evidence is provided by studies done on the transport of the *E. coli* alkaline phosphatase, PhoA. PhoA is a native Sec substrate as it requires formation of disulfides bonds during folding which cannot be achieved in the reducing environment in the cytoplasm. However, PhoA fused with a Tat signal peptide at the N-terminus was transported to the periplasm by the Tat system when PhoA was expressed in mutant *E. coli* strains with an oxidizing cytoplasm, which allows formation of disulfide bonds (Delisa *et. al.* 2003). On the other hand, transport of TorA-PhoA by the Tat system was not observed in wild type *E. coli* strains as TorA-PhoA was unable to fold in the cytoplasm.

Protein	Physiological Role	Cofactors	Co-exported partner?	Sequence of signal peptide
HyaA	Hydrogen oxidation	3 x Fe-S clusters	HyaB	MNNEETFYQAMRRQGVTRRSFLKYCSLAATSLGLGAGMAPKIAWA
HybO	Hydrogen oxidation	3 x Fe-S clusters	HybC	MTGDNTLIHSHGINRRDFMKLCAALAATMGLSSKAAA
HybA	Hydrogen oxidation	4 x Fe-S clusters*	unknown	MNRRNFIKAASCGALLTGALPSVSHA
NapG	Nitrate reduction	4 x Fe-S clusters*	unknown	MSRSAKPQNGRRRFLRDVVRTAGGLAAVGVALGLQQQTARA
NrfC	Nitrite reduction	4 x Fe-S clusters*	unknown	MTWSRRQFLTGVGVLAAVSGTAGRVVA
YagT	Aldehyde oxidoreductase	2 x Fe-S clusters	YagR ,YagS	MSNQGEYPEDNRVGKHEPHDLSLTRRDLIKVSAATAVVYPHSTLAASVPA
YdhX	Component of aldehyde ferredoxin oxidoreductase?	4 x Fe-S clusters*	YdhV? (tungsten)	MSWIGWTVAATALGDNQMSFTRRKFVLGMGTVIFFTGSASSLLA
TorA	TMAO reduction	MGD	none	MNNNDLFQASRRRFLAQLGGLTVAGMLGPSLLTPRRATAAQA
TorZ	TMAO reduction	MGD	none	MTLTRREFIKHSGIAAGALVVTSAAPLPAWA
NapA	Nitrate reduction	MGD, 1 x Fe-S cluster	none	MKLSRRSFMKANAVAAAAAAAGLSVPGVA
DmsA	DMSO reduction	MGD, 1 x Fe-S cluster	DmsB	MKTKIPDAVLAAEVSRRGLVKTTAIGGLAMASSALTLPFSRIAHA
YnfE	selenate reductase	MGD, 1 x Fe-S cluster*	YnfG*	MSKNERMVGISRRTLVKSTAIGSLALAAGGFSLPFTLRNAAA
YnfF	DMSO reduction	MGD, 1 x Fe-S cluster*	YnfG*	MMKIHTTEALMKAEISRRSLMKTSALGSLALASSAFTLPFSQMVRA
FdnG	Formate oxidation	MGD, 1 x Fe-S cluster	FdnH	MDVSRRQFFKICAGGMAGTTVAALGFAPKQALA
FdoG	Formate oxidation	MGD, 1 x Fe-S cluster*	FdoH*	MQVSRRQFFKICAGGMAGTTAAALGFAPSVALA
YedY	TMAO/DMSO reduction?	MPT	none	MKRRQVLKALGISATALSLPHAAHA
CueO	Copper homeostasis	4 x Cu ions	none	MQRRDFLKYSVALGVASALPLWSRAVFA
PcoA	Copper resistance	4 x Cu ions*	none	MLLKTSRRTFLKGLTLSGVAGSLGVWSFNARSSLSLPVAA
SufI	Cell division	none	unknown	MSLSRRQFIQASGIALCAGAVPLKASA
YahJ	Putative deaminase	1 x Fe ion*	unknown	MKESNSRREFLSQSGKMVTAAALFGTSVPLAHA
WcaM	Colanic acid biosynthesis	unknown	unknown	MPFKKLSRRTFLTASSALAFLHTPFARA
MdoD	Glucans biosynthesis	unknown	unknown	MDRRRFIKGSMAMAAVCGTSGIASLFSQAAFA
EfeB	Peroxidase	heme	none	MQYKDENGVNEPSRRRLLKVIGALALAGSCPVAHA

Protein	Physiological Role	Cofactors	Co-exported partner?	Sequence of signal peptide
YaeI	Possible phosphodiesterase	unknown	unknown	MISRRRFLQATAATIATSSGFGYMHYC
AmiA	Cell wall amidase	none	unknown	MSTFKPLKTLTSRRQVLKAGLAALTLSGMSQAIA
AmiC	Cell wall amidase	none	unknown	MSGSNTAISRRRLLQGAGAMWLLSVSQVSLA
FhuD	Ferrichrome binding	none	unknown	MSGLPLISRRRLLTAMALSPLLWQMNTAHA
YcbK	unknown – peptidase M15 superfamily	unknown	unknown	MDKFDANRRKLLALGGVALGAAILPTPAFA
Pac	Penicillin amidase	Ca^{2+}	none	MKNRNRMIVNCVTASLMYYWSLPALA
C3736	Posible diene lactone hydrolase	unknown	unknown	MPRLTAKDFPQELLDYYDYYAHGKISKREFLNLAAKYAVGGMTALA

Table 1- 1: List of known and predicted *E. coli* Tat substrates.

The export of CueO (formerly YacK), SufI (Sargent *et. al.* 1999), MdoD (Lequette *et. al.* 2004), AmiA, AmiC (Ize *et. al.* 2003), FhuD (Ize *et. al.* 2004), Pac (Ignatova *et. al.* 2002), C3736 (Ize and Palmer, unpublished) and EfeB (formerly known as YcdB) (Sturm *et. al.* 2006)have been demonstrated experimentally to require the Tat pathway. The signal peptides of YahJ, WcaM and YcbK have been shown to engage in the Tat pathway when fused to a Tat-specific reporter protein (Tullman-Ercek *et. al.* 2007). Some of the leader peptidase cleavage sites have been determined experimentally (Berks 1996), others predicted using SignalP (http://www.cbs.dtu.dk/services/SignalP/). * predicted by homology, genetic linkage or sequence analysis. Table from (Palmer *et. al.* 2010)

The majority of identified Tat substrates in *E. coli* require redox active cofactors to be inserted in the cytoplasm before export. Examples include molybdoenzymes trimethylamine N-oxide (TMAO) reductase (TorA), dimethyl sulfoxide (DMSO) reductase (DmsA) or the Rieske iron-sulfur cluster containing proteins (Berks 1996). These proteins are transported by the Tat system after acquisition of the appropriate cofactors in the cytoplasm because periplasmic insertion of cofactors is limited by the oxidizing environment in the periplasm and the lack of nucleotide triphosphates in the periplasm to drive this process (Santini *et. al.* 1998).

Proteins that require association with protein partners in heteroligomeric complexes are also transported by the Tat pathway. In most cases, only one of the partner proteins carries the Tat signal peptide. An example of this is the [NiFe] hydrogenase-2 protein complex which is a heterodimer made up of HybO and HybC subunits. The Tat signal sequence is only found on the smaller subunit, HybO. However, HybO cannot be transported across the membrane without HybC (Rodrigue *et. al.* 1999).

Integral membrane proteins can also be targeted to the Tat pathway. Membrane tethering can be via an uncleaved signal peptide or a C-terminal transmembrane helix. It is believed that the Tat signal peptide is unstructured in aqueous solution but helical in hydrophobic environment. Therefore, it could act as an anchor for membrane proteins exported by the Tat system (San Miguel *et. al.* 2003). For example, the Tat signal peptide of the Rieske Fe-S protein subunit of cytochrome bc_1 in *Paracoccus denitrificans* lacks the AxA cleavage site at the c-region. This signal peptide remains uncleaved after transport and instead anchors the protein in the membrane (Bachmann *et. al.* 2006). Sequence analysis has also indicated that at least five known *E. coli* Tat substrates are integral membrane proteins with C-terminal transmembrane anchors. These proteins are FdnH (formate dehydrogenase-N

β-subunit); FdoH, (formate dehydrogenase-O β-subunit); HyaA ([NiFe] hydrogenase-1 β-subunit); HybO ([NiFe] hydrogenase-2 β-subunit) and HybA ('16Fe' ferredoxin from hydrogenase-2 complex) (Sargent *et. al.* 2002). The exact mechanism of transmembrane insertion remains unknown. However, it has been shown that YidC, a protein essential for proper insertion of membrane proteins in other pathways, is not essential for insertion of Tat-dependent integral membrane proteins (Hatzixanthis *et. al.* 2003).

There are also examples of cofactor-less and monomeric Tat substrates such as the cell division protein SufI (Tarry *et. al.* 2009a). It is still unclear why these proteins are targeted to the Tat system and not the Sec system. One hypothesis is that these proteins fold very rapidly in the cytoplasm and are therefore incompatible to the Sec pathway. Another possibility is that the proteins would fold too slowly in the periplasm, which would have resulted in degradation by the periplasmic heat shock protein, DegP (Groll *et. al.* 2005). Some proteins might require cytoplasmic chaperones to fold. Thus these proteins need to fold in the cytoplasm before transport.

Since folded proteins vary in shape and size, the Tat system has to accommodate transport of proteins of various shapes and sizes while at the same time maintaining the membrane integrity. The sizes of known Tat substrates transported by the *E. coli* Tat system range from the 9kDa high-potential iron-sulphur protein (HiPIP) from *Allochromatium vinosum* with a minimal diameter of around 20Å (Carter *et. al.* 1974) to the 142kDa formate dehydrogenase-N from *E. coli* with a minimal diameter of around 70Å (Berg *et. al.* 1991).

However, there are also studies which demonstrated the ability of the Tat system to transport unstructured proteins. Richter *et. al.* (2007) showed that naturally unfolded FG repeats from the yeast Nsp1p nuclear pore protein can be transported by the Tat system when fused with a Tat signal peptide. However, transport was blocked when hydrophobic surface

patches were introduced. These observations suggest that unfolded substrates can be transported by the Tat substrate provided these proteins are sufficiently hydrophilic on their surfaces (Richter *et. al.* 2007). This observation was also reported in plant thylakoids where the thylakoidal Tat system was able to transport largely unstructured hydrophilic peptide substrates (Cline and Mccaffery 2007).

1.3.3 Bioenergetic requirements for Tat transport

When the Tat system was first identified in plant thylakoids it was observed that only the proton motive force (PMF), and not ATP hydrolysis, was required for transport. The PMF is made up of two components, the transmembrane potential gradient ($\Delta\Psi$) and the proton gradient (ΔpH). *In vitro* methods using intact thylakoids showed that thylakoids only require ΔpH for Tat transport (Cline *et. al.* 1992). However, this finding was contradicted by *in vivo* studies on the thylakoidal Tat system (cpTat) which showed that cpTat transport is actually ΔpH insensitive (Finazzi *et. al.* 2003).

In the bacterial Tat system, a similar dependence on the PMF and independence of ATP hydrolysis is observed (Yahr and Wickner 2001; Alami *et. al.* 2002). Further studies on the bacterial Tat system showed that at least two $\Delta\Psi$ steps are necessary for *in vitro* Tat transport. The current model proposes that an early transport step requires a substantial but short duration $\Delta\Psi$ followed by a small but long duration $\Delta\Psi$ to drive transport in the later steps (Bageshwar and Musser 2007).

There is a relatively high energetic cost required for Tat transport. In chloroplasts, it has been estimated that, on average, 8×10^4 protons are required for the translocation of a single Tat substrate, which is equivalent to the hydrolysis of 10^4 ATP molecules (Alder and Theg 2003).

1.3.4 Tat components

Three main components, Tat A, TatB and TatC, make up the Tat translocase (Figure 1- 4). Genes involved in the Tat transport pathway (*tatA*, *tatB* and *tatC*) lie within an operon at 86min on the *E. coli* chromosome. Another Tat gene, *tatD*, which does not appear to be involved in Tat transport, is also located in this operon (Wexler *et. al.* 2000). The gene of a TatA homologue, *tatE*, is mapped to an unlinked, monocistronic gene at 14min on the *E. coli* chromosome (Sargent *et. al.* 1998). Figure 1- 4 shows the predicted topologies and secondary structures of the three main Tat components, TatA, TatB and TatC.

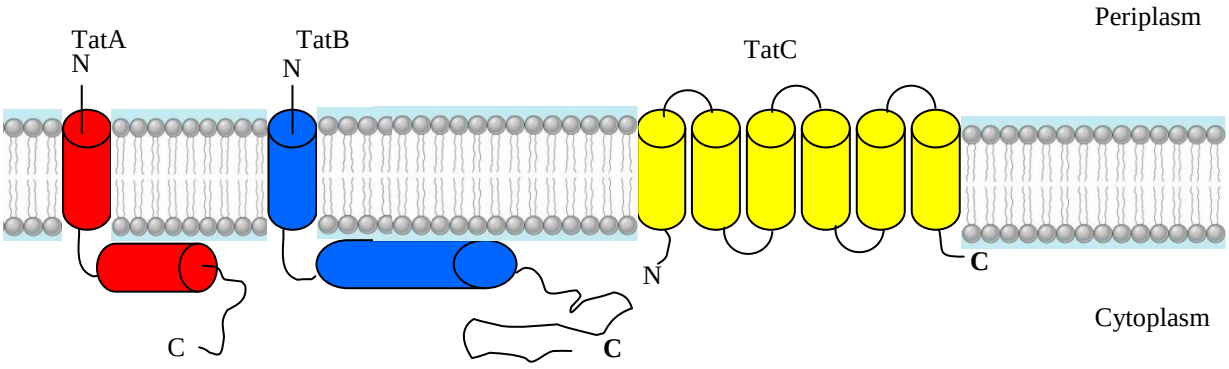


Figure 1- 4: Tat components that make up the Tat translocase
Illustration of TatA, TatB and TatC and their respective predicted topology and secondary structure.

The TatA protein is the most abundant of the Tat components. It is present in 20 fold excess over TatB and TatC and is believed to be synthesized at a higher rate than other Tat components (Jack *et. al.* 2001; Sargent *et. al.* 2001). TatA is made up of a single N-terminal transmembrane helix followed by a hinge region, an amphipathic α-helix and an unstructured tail region. TatA is transported and inserted into the inner membrane without involving the Sec system and the membrane insertase, YidC (Fincher *et. al.* 2003). It is believed that the C-terminus of TatA lies in the cytoplasm while the N-terminus is in the periplasm. This is based on the positive inside rule and a series of protease sensitivity experiments (Settles *et. al.* 1997;

Porcelli *et. al.* 2002). Truncation analysis shows that only the N-terminal transmembrane helix, 'hinge' region and amphipathic helix, but not the unstructured tail on the C-terminus, are essential for Tat transport (Berks *et. al.* 2000; Lee *et. al.* 2002; Barrett *et. al.* 2003; Hicks *et. al.* 2003). Further work has implied that an acidic motif after the amphipathic helix is also important for TatA function (Warren *et. al.* 2009).

Recently, the structure of $TatA_d$ from *Bacillus subtilis* was solved by solution NMR. $TatA_d$ was revealed to be a monomer in dodecylphosphocholine (DPC) micelles and has a L-shaped structure formed by the amphipathic helix and the transmembrane helix, while the C-terminal tail is mainly unstructured (Figure 1- 5). In this structure, the N-terminus was modelled to lie in the periplasm. The L-shaped structure is stabilized by interaction amongst residues near the hinge region (Hu *et. al.* 2010). These data are complemented by solid state NMR experiments, which also revealed that the amphipathic helix does not lie flat on the bilayer but is instead aligned at a steep tilt angle to the membrane (Walther *et. al.* 2010).

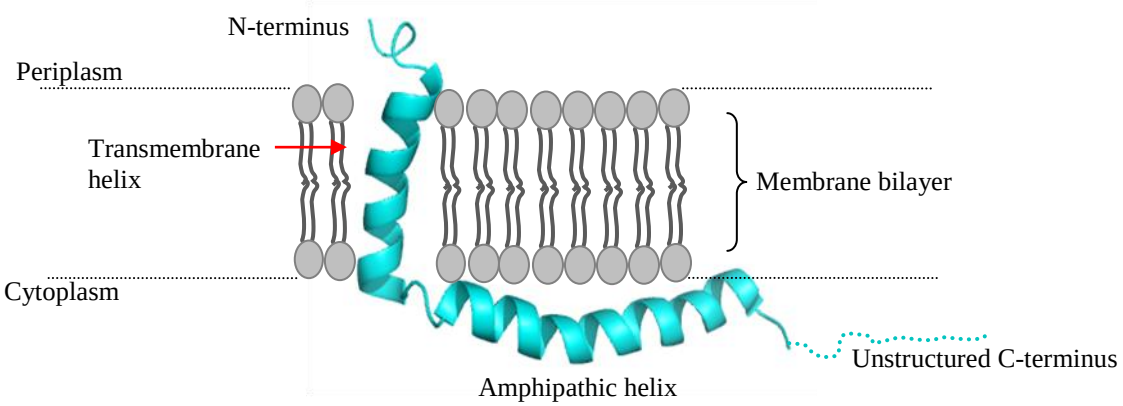


Figure 1- 5: Solution structure of *B. subtilis* $TatA_d$
Ribbon representation of solution NMR structure of *B. subtilis* $TatA_d$. (Adapted from Hu *et. al.* (2010), PDBID 2L16) The N-terminus is modeled to be in the periplasm while the C-terminus faces the cytoplasm.

TatE shares 60% amino acid sequence similarity with TatA. *tatE* probably arose from gene duplication of *tatA.* The expression level of *tatE* is between 50 and 200 fold less than the expression of *tatA.* (Jack *et. al.* 2001). Studies by Sargent *et. al.* (1998) showed that TatA plays a more important role than TatE as Tat transport in a Δ*tatE* strain is significantly less compromised than the Δ*tatA* strain for most Tat substrates tested in this work. However, neither strain is totally defective in the Tat transport pathway. Therefore, it is hypothesized that TatA and TatE have overlapping functions and can partially substitute for each other. However, Tat transport was completely abolished in a Δ*tatA*Δ*tatE* double mutant (Sargent *et. al.* 1998).

TatB has a similar structure to TatA but has a longer unstructured cytoplasmic tail at its C-terminus. TatB is considered to be an essential component for the Tat complex in *E. coli* since Tat transport in a Δ*tatB* mutant strain was blocked (Sargent *et. al.* 1999). *E. coli* TatB was shown to act as a mediator between TatA and TatC (Alami *et. al.* 2003). Although TatB shares ~25% sequence identity with TatA, TatB plays a distinct role in the Tat system. Complementation analysis showed that *ΔtatB* mutants are not complemented by *tatA.* (Sargent *et. al.* 1999). However, there have been reports that bifunctional TatA mutants with mutations at the extreme N-terminus allow transport in the absence of TatB though the transport is less efficient (Blaudeck *et. al.* 2005). Although TatB is essential for Tat transport in *E. coli* and most species, some micro-organisms, such as *Bacillus subtilis*, only has TatA and TatC and can still transport Tat substrates (Jongbloed *et. al.* 2004).

TatC is the largest subunit of the Tat system. TatC has 258 amino acids (~28.9kDa) while TatA only has 89 amino acids (~9.6kDa) and TatB has 171 amino acids (~18.4kDa). Early studies using a *tatC* deletion strain showed that TatC is essential for Tat transport in *E. coli* (Bogsch *et. al.* 1998). It has been shown experimentally to have six-transmembrane

helices with both the N- and C- termini in the cytoplasm (Punginelli *et. al.* 2007). It is also the most hydrophobic and the most highly conserved element in the Tat machinery (Bogsch *et. al.* 1998).

Translational fusion of TatB and TatC was shown to support Tat transport (Bolhuis *et. al.* 2001) suggesting that TatC forms a complex with TatB. TatC, together with TatB, serves as the initial recognition site for Tat signal peptides prior to transport (Alami *et. al.* 2003). The precise binding site is not known but point mutational studies show that the entire N-terminal half of TatC is involved, with the first cytosolic loop of TatC being of particular importance (Holzapfel *et. al.* 2007; Kreutzenbeck *et. al.* 2007; Strauch and Georgiou 2007).

The gene *tatD* is also found in the same operon coding for *tatABC*. TatD is characterized as a cytoplasmic protein with DNase activities. However, TatD is not required for Tat transport as the Tat transport activity of *ΔtatD* strains is similar to that of wild type strains (Wexler *et. al.* 2000).

1.3.5 TatABC translocase

The Tat translocase is made up of TatA, TatB and TatC. Several studies have been carried out to investigate the organization of the Tat components in the Tat translocase. Attempts to purify Tat components resulted in characterisation of two distinct high molecular mass complexes, suggesting that there are at least two Tat complexes in the cell. One of the two high molecular mass complexes purified from cells overexpressing *tatABCDE* is composed of mainly TatA and small amounts of TatB. This is known as the TatA(B) complex. This purified TatA(B) complex has an apparent molecular mass of 600 kDa as measured by gel permeation chromatography (Sargent *et. al.* 2001).

TatA can also be purified as homo-oligomers from cells in the absence of other Tat components (Porcelli *et. al.* 2002; Gohlke *et. al.* 2005). Electron paramagnetic resonance

(EPR) spectroscopy studies showed that Ile12 and Val14 within the transmembrane helices of TatA subunits are important for TatA oligomerization. Exchange between TatA subunits in TatA complexes was also observed (White *et. al.* 2010).

TatA complex migrates in a blue native-PAGE gel as a ladder of bands with sizes ranging from under 100kDa to over 500kDa. These bands differed by about 34kDa, suggesting that the TatA complex is built up from modules of 3 to 4 TatA monomers (Oates *et. al.* 2005). Visualization of yellow fluorescent protein (YFP) labelled TatA by fluorescence microscopy was also used to determine the oligomeric state of TatA that was expressed at native levels. This study supports the hypothesis that TatA complexes are assembled from tetramers and that ~25 TatA monomers, on average, are required to form a TatA complex in living cells. The PMF has no affect on TatA stoichiometry (Leake *et. al.* 2008). When Tat components were overproduced, it was observed that TatA clusters together to from tubular-like structures *in vivo* (Berthelmann *et. al.* 2008).

The TatA(B) complex was revealed to form ring shaped particles when visualized by negative stain electron microscopy. This observation suggests that the TatA(B) complex could form a protein transport channel (Sargent *et. al.* 2001). This observation was further affirmed by 3D images of TatA complexes produced using random conical tilt electron microscopy of negatively stained specimens. In this work, TatA complexes of different sizes were observed. The diameter of the channel was shown to vary between 30Å-70Å. This range matches the diameter of known Tat substrates. It is believed that varying amounts of TatA subunits in the complex contributes to the different sizes observed. This would require dynamic interaction between protomers to facilitate addition or removal of protomers from the complex to allow transport of substrates with varied shapes and sizes. Membrane integrity

is speculated to be maintained by a lid region at the cytosolic side of the membrane (Figure 1-6) (Gohlke *et. al.* 2005).

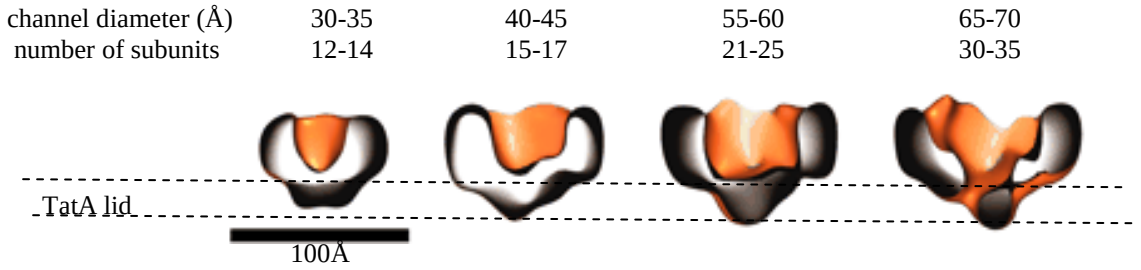


Figure 1- 6: 3D architecture of TatA complexes determined by electron microscopy
Cross-section view of four classes of TatA complexes as observed by low resolution EM. Sizes of the channel diameter of TatA complexes vary between 30Å-70Å. The lid section is proposed to be facing the cytoplasm. The predicted number of subunits that make up individual TatA class is annotated above the figure. Figure adapted from Gohlke *et. al.* (2005).

The second high molecular weight complex isolated from *E. coli* overexpressing *tatABCDE* is the Tat(A)BC complex, which is primarily composed of TatB and TatC with small amount of co-purified TatA (Bolhuis *et. al.* 2001; De Leeuw *et. al.* 2002). Co-purification of TatA could be an artefact of overexpression of *tatABC* as TatA and TatBC form distinct complexes at native levels of expression of *tatABC* (Mcdevitt *et. al.* 2006). TatB and TatC form a high molecular weight complex containing multiple copies of each subunit. Early studies showed that TatB and TatC are present in a 1:1 ratio in the complex with varying amounts of TatA in small quantities (Bolhuis *et. al.* 2001). Single particle electron microscopy of Tat(A)BC complexes shows oval shaped particle with lobes of density surrounding a core. Each lobe was estimated to contain seven or eight transmembrane helices, which could correspond to a single TatBC subunit. It was estimated that each complex contains between five and seven TatBC subunits (Oates *et. al.* 2003). 3D maps of TatBC

complex constructed from electron microscopy images revealed the roughly hemispherical shape of the TatBC complex (Tarry *et. al.* 2009b).

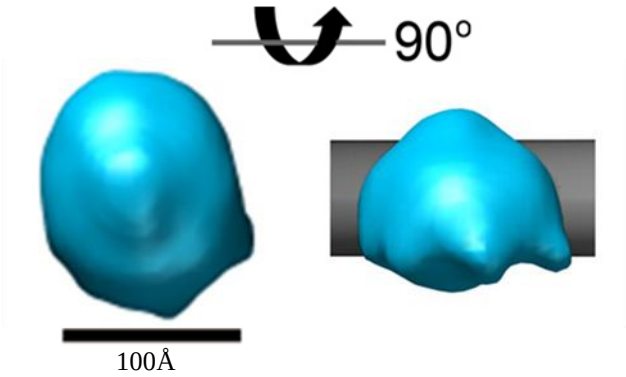


Figure 1- 7: 3D map of TatBC complex
3D reconstruction of TatBC by random conical tilt negative stain electron microscopy. Figure adapted from Tarry *et. al.* (2009)

1.3.6 The Tat transport mechanism

The Tat translocase is responsible for the transport of folded proteins of varying shapes and sizes. This has to be achieved without compromising membrane integrity. Successful Tat transport depends on several factors, including the presence of PMF, the folding state of the substrate, the Tat signal peptide, the composition of the Tat translocase and the nature of the passenger domain.

The Tat transport cycle involves a dynamic system where assembly and disassembly of a TatABC complex is required (Figure 1- 9). The N-terminal Tat signal sequence of Tat subtrates is recognized by the TatBC complex. It is hypothesized that TatC serves as the primary binding site for the consensus twin arginine (RR) motif, in which the entire N-terminus of TatC is involved (Alami *et. al.* 2003). It has been shown that TatA is not required at this stage (Cline and Mori 2001). *In vitro* studies have also shown that the precursor protein not only interacts with the Tat translocon but also interacts with the membrane lipids

via the signal peptide (Bageshwar *et. al.* 2009). Therefore, it is proposed that the early step of transport involves interaction of the signal peptide with the cytoplasmic membrane. This interaction could be due to electrostatic forces between the positively charged n-region and negatively charged phospholipid heads or via hydrophobic interactions, possibly by insertion of the h-region of the signal peptide into the hydrophobic core of membrane (Hou *et. al.* 2006; Shanmugham *et. al.* 2006). It has been demonstrated that the lipid- and translocon-bound forms of the precursor are interconvertible (Bageshwar *et. al.* 2009). It has also been reported that the N-terminus of the passenger protein participates in the membrane binding stage (Schlesier and Klösgen 2010).

Since the TatBC complex is made up of multiple TatC monomers, multiple Tat substrates can potentially bind to the TatBC complex. EM images of TatBC with bound SufI showed that there can be up to two substrates binding to a TatBC complex (Figure 1- 8). Whenever two substrates are observed, they are bound at adjacent sites. Nonetheless, binding of only one substrate is needed to induce a structural reorganization of TatBC, resulting in a decrease in the diameter of TatBC, presumably due to loss of subunits from the TatBC complex (Tarry *et. al.* 2009b). It has also been shown that TatB can interact with surface exposed residues of the mature domain of the substrate (Maurer *et. al.* 2010). Substrate binding does not require energy (Alami *et. al.* 2003; Holzapfel *et. al.* 2007).

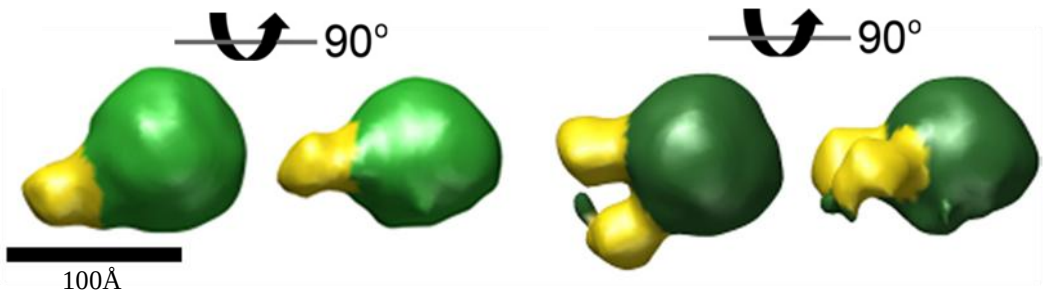


Figure 1- 8: 3D maps of TatBC-SufI$_{His}$ complexes
TatBC (green) with 1 SufI$_{His}$ (yellow) bound. (*D*) TatBC (green) with 2 SufI$_{His}$ (yellow) bound. Each map is shown in 2 orientations related by a 90° rotation about the horizontal axis.

After binding to the TatBC complex, the Tat substrate has to be translocated across the membrane. To do so, TatBC recruits the TatA complex. Cross-linking experiments done on the chloroplast Tat system showed that when substrate is bound to the thylakoid homologue of TatBC, this complex can now crosslink to Tha4, the thylakoidal homologue of TatA. This step was shown to require a PMF (Mori and Cline 2002). Cross-linking studies showed that Tha4 then undergoes conformational changes and undergoes oligomerization via its C-terminal tail domain (Dabney-Smith *et. al.* 2006; Dabney-Smith and Cline 2009). Further work using fluorescent microscopy to visualize TatA-YFP complexes expressed at native levels showed that the TatA complexes only formed large foci in the presence of TatB and TatC (Leake *et. al.* 2008). These studies suggest that protomers of TatA form a translocation channel. Moreover, 3D images of TatA complexes constructed from electron microscopy images shows the formation of a ring structure (Gohlke *et. al.* 2005).

There have been speculations that the C-terminus of TatA has a dual topology. It is believed that TatA assumes its L-shaped single transmembrane topology during the resting phase with hydrophobic residues facing the interior of the cavity. During active transport, the C-terminal amphipathic helix is proposed to undergo structural rearrangements and inserts into the inner membrane to form a double transmembrane orientation with hydrophilic residues lining the pore interior. This phenomenon is dependent on the PMF (Gouffi *et. al.* 2004; Chan *et. al.* 2007). Such rearrangements could be required to provide a hydrophilic lining to the translocase pore during protein transport.

The exact mechanism of transport remains unknown. It is unclear whether a push or pull force is required to drive the translocation of substrate. No known proteins on either side of the membranes have been found which provides the driving force. It has also been

speculated that the force could be applied within the membrane domain or the Tat translocase itself (Robinson and Bolhuis 2004).

Processing of the signal peptide occurs after the translocation step. In most cases, the signal peptide is cleaved after the AxA motif, which is found in the c-region of the Tat signal peptide, by the signal peptidase, LepB. This was demonstrated by Yahr and Wickner (2001) when they showed that processing of SufI in *E. coli* is sensitive to the β-lactam inhibitor, which irreversibly inhibits LepB activity. LepB has a soluble globular domain and is anchored to the periplasmic side of the inner membrane. LepB is postulated to cleave signal peptides only during the later stages of the Tat transport cycle. Since the serine protease active site of LepB is believed to be located at the interface of the headgroups of the bilayer and its globular domain, substrates in the cytoplasm or free floating in periplasm will not be processed (Sargent *et. al.* 2006; Lüke *et. al.* 2009). This is also true for the LepB homologue, TPP (thylakoidal-processing protease), in thylakoids. Studies in thylakoids showed that precursor substrates that have been cross-linked to cpTatC via the Tat signal peptide were efficiently translocated. However, the signal peptide was uncleaved, implying that the signal peptide has to be transported across the membrane for TPP to access it (Gerard and Cline 2006). Cleavage of lipoproteins bearing the Tat signal peptides containing the [L/I/G/A][A/G/S]C 'lipo box' motif by 'type-II' signal peptidases also occurs after transport (Sargent 2007). Nonetheless, the Tat transport process is independent of the cleavage process (Frielingsdorf and Klösgen 2007) and not all signal peptides are cleaved after transport (Molik *et. al.* 2001). For some Tat substrates, namely integral membrane proteins such as the Rieske Fe/S protein, the signal peptide is left uncleaved after transport to be used as a transmembrane anchor (De Buck *et. al.* 2007).

The TatBC-substrate-TatA supercomplex is transient. TatABC dissociates to TatA and TatBC after translocation and resets the cycle in preparation for subsequent transport (Mori and Cline 2002). The requirements for resetting the Tat transport cycle remain elusive.

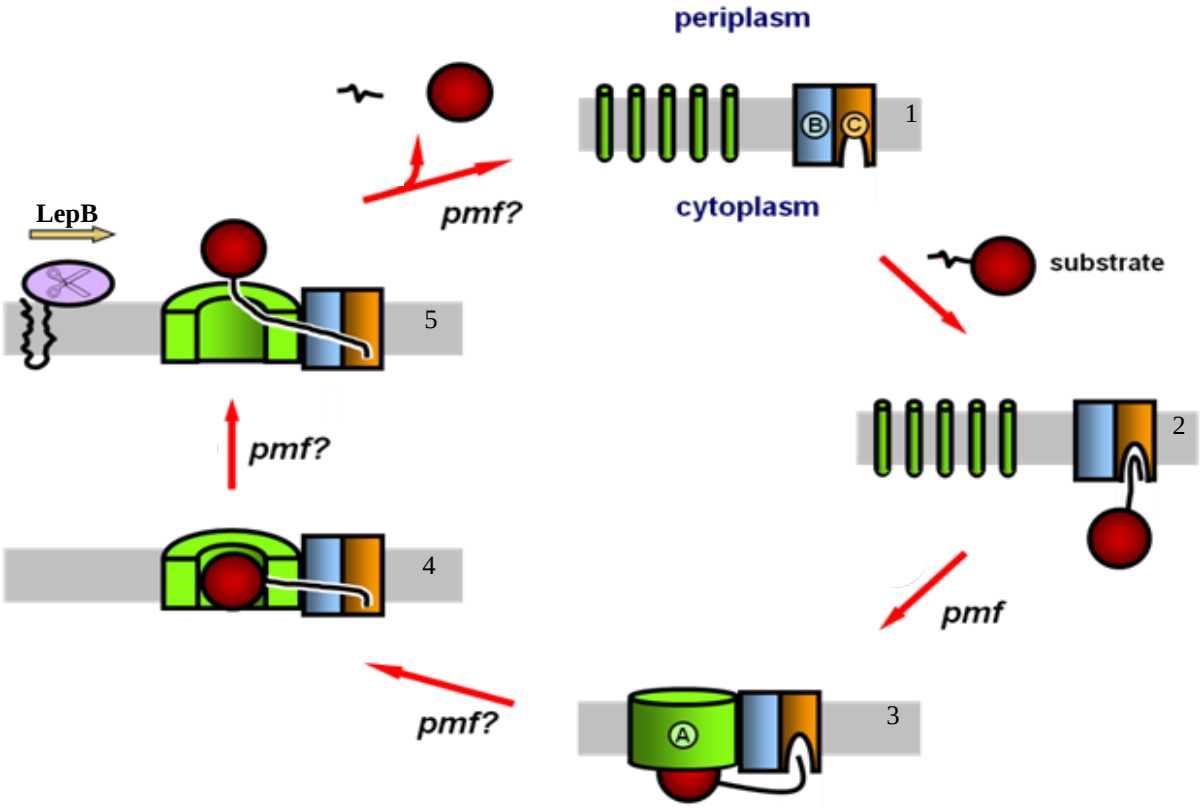


Figure 1- 9: Proposed mechanism of Tat transport.
1) In the resting state, TatA and TatBC are present in the membrane as separate complexes. **2)** A Tat precursor binds to the TatBC complex via interactions of its signal peptide with TatC. **3)** Following binding, and in the presence of the proton motive force (PMF), a TatA complex is formed and associates with TatBC complexes to form the active translocase. **4)** The substrate is transported through the membrane, possibly via a channel formed by the TatA complex. The energy requirement for this step is unclear. **5)** After translocation across the membrane, the signal peptide is cleaved by the action of the signal peptidase, LepB. The mature substrate is released from the translocase and the TatA and TatBC complexes dissociate.

1.3.7 Proofreading and quality control of Tat substrates

The main distinguishing feature of the Tat system is its ability to transport folded proteins, some of which are bound to cofactors. The Tat system has to ensure that only proteins that are properly folded with appropriately bound cofactors or partner proteins are transported.

The concept of the Tat translocase having an intrinsic quality-control mechanism that was able to 'sense' regions of unfolded proteins and only allowed transport of folded proteins was first suggested by DeLisa *et. al.* (2003). However, other studies have shown binding of unfolded proteins to TatB and TatC (Richter and Brüser 2005).

Transport of Tat substrates can also be regulated at an earlier stage on the Tat pathway to prevent premature targeting of Tat substrates and to ensure that only fully assembled substrates are presented to the Tat translocase. It has been proposed that chaperones are involved in this 'proofreading' process. General chaperones such as GroEL, DnaK and SlyD have been shown to be involved in the folding of Tat substrates (Graubner *et. al.* 2007; Pérez-Rodríguez *et. al.* 2007). More importantly, there are chaperones that are specific for Tat substrates. These chaperones are collectively known as redox enzyme maturation proteins (REMP) and they are specific for individual redox enzymes, which make up a large proportion of the Tat substrates. The first example of a chaperone identified that is involved in proofreading is DmsD. DmsD was isolated through its interaction with the precursor form of DmsA, a native Tat substrate. However, DmsD does not interact with the mature form of DmsA suggesting that DmsD binds to the signal peptide of DmsA (Oresnik *et. al.* 2001). Following this, the chaperones HyaE and HybE was identified using a bacterial two-hybrid system. HyaE and HybE are required for the assembly of [NiFe] hydrogenase-1 and [NiFe] hydrogenase-2 respectively (Dubini and Sargent 2003). HyaE binds to the

precursor form of HyaA of the [NiFe] hydrogenase-1 complex, which bears the signal peptide. In the case of [NiFe] hydrogenase-2, HybE only interacts with the small HybO subunit and this interaction was shown to be dependent on the signal peptide. Another chaperone that has been identified is TorD, which binds to TorA (Pommier *et. al.* 1998; Jack *et. al.* 2004). Binding of TorD was shown to be specific for the Tat signal peptide of TorA regardless of the passenger protein.

The identification of signal sequence specific chaperones led to the proposal that the Tat signal peptides can be divided into two classes (Sargent *et. al.* 2006). 'Class 1' Tat signal peptides interact with enzyme specific chaperones before transport as well as the Tat translocase. 'Class 2' Tat signal peptides do not interact with chaperones and its sole function is to interact with the Tat translocase and trigger Tat transport.

1.3.8 PspA and its potential role in Tat transport

It was observed in *in vivo* Tat transport assays that the export of Tat substrates is blocked when the expression of Tat substrates is elevated (Chanal *et. al.* 2003). This could most likely be due to saturation of the Tat translocase. It has been shown that co-expression of *pspA*, which codes for a phage shock protein, with Tat substrates was able to relieve saturation of the Tat translocase caused by excess Tat substrates (Delisa *et. al.* 2004). *pspA* is part of the *psp* operon which usually plays a role in survival of *E. coli* in late stationary phase at alkaline pH and protects the cell against dissipation of PMF when subjected to stress (Model *et. al.* 1997). However, expression of *pspA* was only induced when the Tat translocase is saturated or when the *tat* genes are inactivated. It is hypothesized that PspA aids the Tat transport by maintaining the PMF during translocation (Delisa *et. al.* 2004).

1.3.9 Tat system in other microorganisms

Putative Tat substrates have been identified in at least 84 prokaryotic proteomes using TATFIND (Table 1- 2) (Dilks *et. al.* 2003). Although much of the studies have been done on the chloroplast and *E. coli* Tat system, there have been attempts to understand the Tat system in other microorganisms.

1.3.9.1 Gram-positive bacteria

The Tat system has been quite extensively studied in Gram-positive bacteria, especially *Bacillus subtilis*. *B. subtilis* has a highly selective Tat pathway that only requires TatA and TatC. Unlike the *E. coli* Tat system, there are at least two minimal Tat translocases in *B. subtilis*, each composed of a specific TatA and a specific TatC. These complexes are also substrate specific in *B. subtilis*. For example, the $TatA_dC_d$ complex transports the phosphodiesterase PhoD but not YwbN while the $TatA_yC_y$ complex transports YwbN but not PhoD (Jongbloed *et. al.* 2002; Jongbloed *et. al.* 2004). Studies on *B. subtilis* are likely representative of all Gram-positive bacteria except *Streptomyces*, which have TatABC translocases (Schaerlaekens *et. al.* 2001).

When the *B. subtilis* Tat complexes were expressed in *E. coli*, they were able to transport some, but not all, *E. coli* Tat substrates. This shows that the *B. subtilis* and *E. coli* Tat complexes are able to recognize similar targeting determinants within the Tat signal peptide (Barnett *et. al.* 2009). Nonetheless, the *B. subtilis* Tat complex is different from the *E. coli* Tat complex. Estimation of the sizes of the complexes $TatA_dC_d$ (~350kDa) and $TatA_yC_y$ (~200kDa) by gel filtration shows that these complexes are much smaller than the TatABC complex (~600kDa) in *E. coli*. In addition, the membrane bound *B. subtilis* $TatA_d$ and $TatA_y$ appear to form homogenous complexes and are still able to transport substrates of various sizes. This observation gave rise to speculations as to whether *E. coli* TatA heterogeneity is

Table 1- 2: Predicted number of Tat substrates and Tat components numbers in prokaryotes

Prediction based on TatFIND (Rose *et. al.* 2002). Table adapted from (Dilks *et. al.* 2003). bac: eubacteria

Organism	Domain	Phylum	Putative Tat substrates	Tat Components		
				Tat A/E	Tat B	Tat C
Streptomyces coelicolor A3(2)	bac	Actinobacteria	230	1	-	1
Mesorhizobium loti	bac	α - Proteobacteria	95	1	-	1
Sinorhizobium meliloti	bac	α - Proteobacteria	94	1	1	1
Caulobacter crescentus	bac	α - Proteobacteria	88	1	1	1
Ralstonia solanacearum	bac	β - Proteobacteria	71	1	1	1
Halobacterium sp. NRC-1	archaea	Euryarchaeota	68	1	-	2
Pseudomonas aeruginosa	bac	γ - Proteobacteria	57	1	1	1
Xanthomonas campestris 3391	bac	γ - Proteobacteria	55	1	1	1
Agrobacterium tumefaciens	bac	α - Proteobacteria	51	1	1	1
Xanthomonas axonopodis 306	bac	γ - Proteobacteria	50	1	1	1
Escherichia coli K12	bac	γ - Proteobacteria	34	2	1	1
Escherichia coli O157-H7 EDL933	bac	γ - Proteobacteria	33	2	1	1
Salmonella typhimurium LT2	bac	γ - Proteobacteria	33	2	1	1
Escherichia coli O157-H7	bac	γ - Proteobacteria	32	2	1	1
Mycobacterium tuberculosis H37Rv	bac	Actinobacteria	31	1	-	1
Nostoc sp.PCC7120	bac	Cyanobacteria	31	2	-	1
Mycobacterium tuberculosis CDC1551	bac	Actinobacteria	29	1	-	1
Salmonella enterica Typhi	bac	γ - Proteobacteria	28	2	1	1
Deinococcus radiodurans	bac	Deinococcus-Thermus	22	2	-	1
Synechocystis sp. PPC6803	bac	Cyanobacteria	21	2	-	1
Yersinia pestis	bac	γ - Proteobacteria	19	2	1	1
Brucella melitensis	bac	α - Proteobacteria	19	1	-	1
Xylella fastidiosa 9a5c	bac	γ - Proteobacteria	17	1	1	1
Aquifex aeolicus	bac	Aquificae	15	2	-	1
Corynebacterium glutamicum	bac	Actinobacteria	15	2	-	1
Pyrobaculum aerophilum	archaea	Euryarchaeota	14	1	-	1
Neisseria meningitidis Z2491	bac	β - Proteobacteria	12	1	1	1
Pasteurella multocida	bac	γ - Proteobacteria	12	1	1	1
Campylobacter jejuni	bac	ε - Proteobacteria	11	1	1	1
Neisseria meningitidis MC58	bac	β - Proteobacteria	11	1	1	1
Haemophilus influenza	bac	γ - Proteobacteria	9	1	1	1
Archaeoglobus fulgidus	archaea	Euryarchaeota	9	2	-	2
Mycobacterium leprae	bac	Actinobacteria	9	1	-	1
Vibrio cholerae	bac	γ - Proteobacteria	7	2	1	1
Bacillus subtilis	bac	Firmicutes	7	3	-	2
Aeropyrum pernix	archaea	Crenarchaeota	7	2	-	1
Methanosarcina mazei Goe1	archaea	Euryarchaeota	6	2	-	2
Treponema pallidum	bac	Spirochaetes	6	-	-	-
Bacillus halodurans	bac	Firmicutes	5	2	-	2
Methanosarcina acetivorans str.C2A	archaea	Euryarchaeota	5	2	-	2
Sulfolobus solfataricus	archaea	Crenarchaeota	5	3	-	2
Chlorobium tepidum TLS	bac	Chlorobi	5	2	-	1
Pyrococcus horikoshii	archaea	Euryarchaeota	5	-	-	-
Sulfolobus tokodaii	archaea	Crenarchaeota	4	2	-	1
Helicobacter pylorii 26695	bac	ε -Proteobacteria	3	1	1	1
Helicobacter pylorii J99	bac	ε - Proteobacteria	3	1	1	1
Clostridium perfringens	bac	Firmicutes	3	-	-	-
Pyrococcus furiosus DSM3638	archaea	Euryarchaeota	3	-	-	-
Thermotoga maritima	bac	Thermotogae	3	-	-	-
Listeria innocua	bac	Firmicutes	2	1	-	1
Staphylococcus aureus Mu50	bac	Firmicutes	2	1	-	1
Staphylococcus aureus MW2	bac	Firmicutes	2	1	-	1
Staphylococcus aureus N315	bac	Firmicutes	2	1	-	1
Thermoplasma acidphilum	archaea	Euryarchaeota	2	1	-	1
Thermoplasma volcanium	archaea	Euryarchaeota	2	1	-	1
Chlamydia trachomatis	bac	Chlamydiae	2	-	-	-
Chlamydophila pneumoniae AR39	bac	Chlamydiae	2	-	-	-
Chlamydophila pneumoniae CWL029	bac	Chlamydiae	2	-	-	-

Organism	Domain	Phylum	Putative Tat substrates	Tat Components		
				Tat A/E	Tat B	Tat C
Pyrococcus abyssi	archaea	Euryarchaeota	2	-	-	-
Listeria monocytogenes EGD-e	bac	Firmicutes	1	1	-	1
Rickettsia conorii	bac	α - Proteobacteria	1	1	-	1
Rickettsia prowazekii	bac	α - Proteobacteria	1	1	-	1
Chlamydia muridarum	bac	Chlamydiae	1	-	-	-
Chlamydophila pneumoniae J138	bac	Chlamydiae	1	-	-	-
Clostridium acetobutylicum	bac	Firmicutes	1	-	-	-
Lactococcus lactis subsp. lactis	bac	Firmicutes	1	-	-	-
Methanothermobacter thermautotrophicus	archaea	Euryarchaeota	1	-	-	-
Streptococcus pneumoniae R6	bac	Firmicutes	1	-	-	-
Streptococcus pyogenes	bac	Firmicutes	1	-	-	-
Streptococcus pyogenes MGAS8232	bac	Firmicutes	1	-	-	-
Thermoanaerobacter tengcongensis	archaea	Euryarchaeota	1	-	-	-
Buchnera aphidicola	bac	γ - Proteobacteria	-	-	-	-
Buchnera sp. APS	bac	γ - Proteobacteria	-	-	-	-
Fusobacterium nucleatum 25586	bac	Fusobacteria	-	-	-	-
Methanococcus janaschii	archaea	Euryarchaeota	-	-	-	-
Methanopyrus kandleri AV19	archaea	Euryarchaeota	-	1	-	-
Mycoplasma genitalium	bac	Firmicutes	-	-	-	-
Mycoplasma pneumoniae	bac	Firmicutes	-	-	-	-
Mycoplasma pulmonis	bac	Firmicutes	-	-	-	-
Streptococcus pneumoniae TIGR4	bac	Firmicutes	-	-	-	-
Ureaplasma urealyticum	bac	Firmicutes	-	-	-	-
Borrelia burgdorferi	bac	Spirochaetes	-	-	-	-

necessarily relevant to substrate size (Barnett *et. al.* 2008; Barnett *et. al.* 2009). Cytosolic $TatA_d$ and $TatA_y$ are also present in *B. subtilis* (Pop *et. al.* 2003; Barnett *et. al.* 2008).

1.3.9.2 Archaea

Many archaea are found in extreme conditions such as in very low or high pH, temperature or salinity. The presence of the Tat system varies among the different species of archaea. Some do not appear to have a functional Tat pathway such as the thermophilic methanogen *Methanopyrus kandleri* while others have a Tat pathway that are more similar to the bacterial Tat system in that they only have a limited number of Tat substrates and most are cofactor-binding proteins. By contrast, some archaea, such as the haloarchaea *Halobacterium sp. NRC-1* have a high number of Tat substrates. This is believed to be due to the high intracellular and extracellular salt concentrations, which means that proteins would need to fold rapidly to prevent aggregation. In addition, proteins require chaperones for proper folding under high salt conditions and these chaperones cannot be found outside the cytoplasm. Therefore, these proteins are Sec-incompatible as they will be folded before reaching the membrane. This is likely the case for all halophilic archea living in hypersaline environments (Bolhuis 2002; Rose *et. al.* 2002; Hutcheon and Bolhuis 2003). Furthermore, lipoproteins can also be targeted to the Tat pathway in haloarchaea such as *Haloferax volcanii* (Giménez *et. al.* 2007).

The complexity of Tat complexes varies among archaea species. Most have multiple copies of Tat complexes though this does not correlate to the number of Tat substrates the organism has. TatA and TatC homologues can be found in archea but not TatB (Table 1- 2). In *H. volcanii*, there are two TatA proteins, TatAo and TatAt, and two TatC proteins, TatCo and TatCt (Dilks *et. al.* 2005). Similar to $TatA_d$ and $TatA_y$ in the Gram-positive *B. subtilitis*, both TatA paralogs of *H. volcanii* can be found in the membrane as well as the cytoplasm. On

the other hand, *Halobacterium sp. NRC-1* has only one TatA but two TatC proteins, TatC1 and TatC2 (Bolhuis 2002; Thomas and Bolhuis 2006). Table 1- 2 shows the number of Tat components that some archea are known to have.

The bioenergetic requirements for Tat transport appear to differ between the bacterial and the haloarchaeal Tat system. The bacterial Tat system relies on the presence of a PMF for transport. On the other hand, studies carried out on the haloarchaeaon *Haloarcula hispanica* demonstrated the sole dependence of the transport of the Tat substrate, AmyH, on the sodium motif force (Kwan *et. al.* 2008). Nonetheless, the archaeal Tat system still requires the twin arginine motif for transport (Kwan and Bolhuis 2010).

1.3.10 Importance of Tat

The Tat system of *E. coli* is constitutively expressed (Jack *et. al.* 2001). Mutations in the Tat system has a pleiotropic effect as the Tat system is involved in a whole range of cell function such as the respiratory process, the biogenesis of the cell envelope, heavy-metal resistance, biofilm formation, pathogenicity and antibiotic resistance (Ize *et. al.* 2003; Pradel *et. al.* 2003; Ize *et. al.* 2004).

The Tat system is coupled to virulence in a wide range of human, animal and plant pathogenic bacteria. This includes species of *Pseudomonas*, *Legionella*, *Salmonella*, *Agrobacterium*, *Yersinia* and *Mycobacterium* (De Buck *et. al.* 2008). Mutations of the Tat system of pathogen usually results in loss of virulence and pathogenicity. Since the Tat pathway is relatively well conserved among pathogens and is absent from mammalian cells, it has been proposed as a target for novel antimicrobial compounds (Cline and Mccaffery 2007).

The Tat system also has biotechnological applications. Secretion of recombinant proteins into the medium of growing cultures is advantageous for biotechnological productions as it simplifies the purification process. Although the Sec system is currently

used to secrete recombinant proteins, some recombinant proteins are misfolded or degraded after transport. This is because these proteins need to be folded in the cytoplasm with the aid of chaperones. Hence, the discovery of the Tat system could allow for biotechnological production of recombinant proteins that needs to be folded in the cytoplasm as long as the protein of interest is fused to a Tat signal peptide at the N-terminus. A Gram-positive bacterium with a Tat system that accepts a wide range of Tat substrates is preferred as the host to produce recombinant proteins as proteins are secreted directly to the medium and this simplifies the protein purification process. Export of Tat amenable recombinant proteins can be further improved by optimization of the Tat signal sequence, the abundance of the Tat translocase, chaperones or co-factors if necessary (Brüser 2007).

1.4 Project Aims

The initial aim of this project is to optimize the transport efficiency of *in vitro* Tat transport assays. Although *in vitro* Tat transport assays have been used to address several issues pertaining the Tat transport mechanism, the achieved transport efficiency was low and hardly achieved the transport efficiency of Tat transport *in vivo*. Following this, the optimized *in vitro* Tat transport assay will be used to gain a further understanding of the Tat transport mechanism.

The second part of the project is to determine the structure and biochemical properties of the alkaline phosphatase PhoX, a native Tat substrate in *Pseudomonas fluorescens*. PhoX will be used as a substrate for *in vitro* studies of the Tat transport. Therefore, characterization of PhoX would provide an insight into the relationship between PhoX and the Tat system.

Chapter 2:

Materials and methods

2.1 Bacterial strains and plasmid constructs

Table 2- 1: List of *E. coli* strains used

***E. coli* strain**	**Genotype**	**Reference**
DH5α	F$^-$ *end*A1 *gln*V44 *thi*-1 *rec*A1 *rel*A1 *gyr*A96 *deo*R *nup*G Φ80d*lacZ*ΔM15 Δ(*lacZYA*-argF)U169, *hsd*R17(r_K^- m_K^+), λ–	Promega
XL1-Blue	*end*A1 *gyr*A96(nal^R) *thi*-1 *rec*A1 *rel*A1 *lac gln*V44 F'[::Tn10 *pro*AB$^+$ *lac*I^q Δ(*lac*Z)M15] *hsd*R17(r_K^- m_K^+): Tet^R	Stratagene
BL21 (DE3)	F$^-$ *omp*T *gal dcm lon hsd*S_B(r_B^- m_B^-) λ(DE3)	(Studier and Moffatt 1986)
BL21 (DE3) Δ*tat*	BL21 (DE3) Δ*tatABC* :: Apr	Palmer, T
C43 (DE3)	BL21 (DE3) but with two uncharacterised mutations.	(Miroux and Walker 1996)
C43 (DE3) *Δtat*	C43(DE3), Δ*tatABC* :: Apr	Palmer, T
MC4100	F$^-$ *ara*D139 Δ(*arg*F-l*ac*)U169 *rsp*L150 *rel*A1 *flbB*5301 *fru*A25 *deo*C1 *pts*F25 *e*14-	(Casadaban 1976)
DADE	MC4100, Δ*tatABCD* Δ*tatE*	(Wexler *et. al.* 2000)
MG1655	F$^-$ λ$^-$ *ilvG- rfb*-50 *rph*-1	(Blattner *et. al.* 1997)
MG1655 Δ2	F$^-$ λ$^-$ *ilvG- rfb*-50 *rph*-1, Δ*eriC* Δ*mriT*	(Iyer *et. al.* 2002)
JM109	*endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB$^+$ Δ(lac-proAB) e14- [F' traD36 proAB$^+$ lacI^q lacZΔM15] hsdR17($r_K^-m_K^+$)*	(Yanisch-Perron *et. al.* 1985)

Table 2- 2: List of plasmids used

Plasmid	Vector	Inserts	Cloning method (if done in this work)	Description/ Use	Resistance	Source
pREP4	pACYC184	*lacI*^{q}	-	Repression of genes under control of *lac* promoter	Kan	Qiagen
pLysS	pACYC184		-	Encodes T7 lysozyme which inhibits T7 RNA polymerase in lysogenic hosts	Cm	(Studier 1991)
p8737	pET22b(+)	*tatABCD*	-	For overproduction of TatABCD	Amp	(Alami *et. al.* 2002)
pUniTat	pQE60	*tatABC*	-	For overproduction of TatABC	Amp	(Lee *et. al.* 2006)
pSC17	pQE80L	*CueO*$_{His}$ (mature domain)	-	For overproduction of mature CueO with 6xHis at the C-terminus	Amp	Chen, S
pSC22	pGEX-6P-2	*CueO*$_{His}$	-	For overproduction of CueO with glutathione S-transferase (GST) at the N-terminus and 6xHis at the C-terminus	Amp	Chen, S
pET28a-lepB	pET28a	*lepB*$_{His}$	-	For overproduction of LepB with 6xHis at the C-terminus	Kan	Greene, N
pGEX-CueO(nostop)	pGEX-6P-2	*cueO* with stop codon removed	1569bp *Bam*HI-*Xho*I fragment encoding for *cueO* generated by PCR using primers SC037 and CY002 with pSC-22 as template	For overproduction of CueO with GST at the N-terminus and stop codon removed.	Amp	This work

Plasmid	**Vector**	**Inserts**	**Cloning method (if done in this work)**	**Description/ Use**	**Resistance**	**Source**
pPCR-GS3	pPCR-Script			Codes for [GGGGS]x3 (GS3) linker	Amp	(Cline and Mccaffery 2007)
pGEX-CueO-$GS3_{His}$	pGEX	*cueO_GS3*$_{His}$	75bp *Xho*I –*Hind*III fragment encoding for *GS3* linker generated using primersCY005 and CY006 with pPCR-GS3 as template and cloned into pGEX-CueO(nostop)	For overproduction of CueO with GST at the N-terminus and GS3 linker followed by Cys and 6xHis at the C-terminus	Amp	This work
pGEX-3C-protease	pGEX-6P-2	*3C protease*		For overproduction of 3C-protease with GST at N-terminus	Amp	Greene, N.
pGEX-$MdoD_{His}$	pGEX-6P-2	*mdoD*$_{His}$	1431bp *Bam*HI-*Xho*I fragment encoding *mdoD* generated by PCR using primers CY020 and CY021 with *Escherichia coli* (MC4100) genomic DNA as template	For overproduction of MdoD with GST at N-terminus and 6xHis at C-terminus	Amp	This work
pGEX-PF_$PhoX_{His}$	pGEX-6P-2	*P. fluorescens phoX*$_{His}$	1939bp *Bgl*II-*Xho*I fragment encoding *phoX* generated by PCR using primers CY022 and CY023 with *Pseudomonas fluorescens Pf0-1*genomic DNA as template	For overproduction of PhoX from *Pseudomonas fluorescens Pf0-1* with GST at N-terminus and 6xHis at C-terminus	Amp	This work
pQE60-PF_$PhoX_{His}$	pQE60	*P. fluorescens phoX*$_{His}$	1939bp *Bgl*II-*Hind*III fragment encoding *phoX* generated by PCR using primers CY022 and CY030 with *Pseudomonas fluorescens Pf0-1*genomic DNA as template	For overproduction of PhoX from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work
pQE60-PF_PhoXC179S$_{His}$	pQE60	*P. fluorescens phoX(C179S)*$_{His}$	Quickchange mutagenesis using primers CY055 and CY056 and pQE60-PF_$PhoX_{His}$ as template	For overproduction of PhoX(C179S) from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work

Plasmid	Vector	Inserts	Cloning method (if done in this work)	Description/ Use	Resistance	Source
pQE60-PF_PhoXD494A$_{His}$	pQE60	*P. fluorescens phoX(D494A)$_{His}$*	Quickchange mutagenesis using primers CY053 and CY054 and pQE60-PF_PhoX$_{His}$ as template	For overproduction of PhoX(D494A) from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work
pQE60-PF_PhoXE194A$_{His}$	pQE60	*P. fluorescens phoX(E194A)$_{His}$*	Quickchange mutagenesis using primers CY066 and CY067 and pQE60-PF_PhoX$_{His}$ as template	For overproduction of PhoX(E194A) from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work
pQE60-PF_PhoXE532A$_{His}$	pQE60	*P. fluorescens phoX(E532A)$_{His}$*	Quickchange mutagenesis using primers CY070 and CY071 and pQE60-PF_PhoX$_{His}$ as template	For overproduction of PhoX(E532A) from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work
pQE60-PF_PhoXD292A$_{His}$	pQE60	*P. fluorescens phoX(D292A)$_{His}$*	Quickchange mutagenesis using primers CY068 and CY069 and pQE60-PF_PhoX$_{His}$ as template	For overproduction of PhoX(D292A) from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work
pQE60-PF_PhoXC179D$_{His}$	pQE60	*P. fluorescens phoX(C179D)$_{His}$*	Quickchange mutagenesis using primers CY078 and CY079 and pQE60-PF_PhoX$_{His}$ as template	For overproduction of PhoX(C179D) from *Pseudomonas fluorescens Pf0-1* with 6xHis at C-terminus	Amp	This work
pGEX-VF_PhoX$_{His}$	pGEX-6P-2	*V. fischeri PhoX$_{His}$*	1904bp *Bgl*II-*Xho*I fragment encoding *phoX* generated by PCR using primers CY024 and CY025 with *Vibrio fischeri* (ES114) genomic DNA as template	For overproduction of PhoX from *Vibrio fischeri* ES114 with glutathione S-transferase at N-terminus and 6xHis at C-terminus	Amp	This work

Table 2- 3: List of primers used

Name	Description	Sequence (restriction sites underlined)	Restriction sites
SC037	Forward primer for amplification of *cueO*	5'-GCGCGGATCCATGCAACGTCGTGATTTCTTAAAAT-3'	*Bam*HI
CY002	Reverse primer for amplification of *cueO* without the stop codon	5'-GCGCCTCGAGCGTTTAAAGCTTTACCGTAAACCCTAACATC-3'	*Hind*III, *Xho*I
CY005	Forward primer for amplification of *[GGGGS]x3* linker	5'- GCGCAAGCTTGGCGGCGGCGGCTCTGG-3'	*Hind*III
CY006	Reverse primer for amplification of *[GGGGS]x3* linker followed by *CHHHHHH*	5'- GCGCCTCGAGTTAGTGATGGTGATGGTGATGGCAGCTC CCGCCACCCCCTGAGC-3'	*Xho*I
CY020	Forward primer for amplification of *mdoD*	5'-GCGCGGATCCATGGATCGTAGACGATTTATTAAAGG-3'	*Bam*HI
CY021	Reverse primer for amplification of *mdoD*	5' -GCGCCTCGAGTTAGTGATGGTGATGGTGATGAGATCTAC TCATCACGCGGTCGTCAAC-3'	*Bgl*II, *Xho*I
CY022	Forward primer for amplification of *P. fluorescens phoX*	5' -GCGCAGATCTATGAGCCTATTAGAAGAAAACC-3'	*BglII*
CY023	Reverse primer for amplification of *P. fluorescens phoX* into pGEX-6P-2	5' - GCGCCTCGAGTTAGTGATGGTGATGGTGATGGGCGCCGACGAT CCC-3'	*XhoI*
CY024	Forward primer for amplification of *V. fischeri phoX*	5'-GCGCAGATCTATGAATAAATTTGTAGAGAAAG-3'	*Bgl*II
CY025	Reverse primer for amplification of *V. fischeri phoX*	5'- GCGCCTCGAGTTAGTGATGGTGATGGTGATGACTGCCGATCACG CCGC-3'	*Xho*I
CY030	Reverse primer for amplification of *P. fluorescens phoX* into pQE60	5'-GCGCAAGCTTTTAGTGATGGTGATGGTGATGG-3'	*Hind*III
CY053	Sense primer for *P. fluorescens phoX* D494A mutation	5'- GTGGATCCTCACCGCGGGAGATTCGAGCA -3'	None

Name	Description	Sequence (restriction sites underlined)	Restriction sites
CY054	Antisense primer for *P. fluorescens phoX* D494A mutation	5'- TGCTCGAATCTCCCGCGGTGAGGATCCAC -3'	None
CY055	Sense primer for *P. fluorescens phoX* C179S mutation	5'-GCACCTTCCAGAACAGCGCCAACGGCAAA-3'	None
CY056	Antisense primer for *P. fluorescens phoX* C179S mutation	5'-TTTGCCGTTGGCGCTGTTCTGGAAGGTGC-3'	None
CY066	Sense primer for *P. fluorescens phoX* E194S mutation	5'-TATCTGACCTGCGAAGCGAACTTCACCGACTGC-3'	None
CY067	Antisense primer for *P. fluorescens phoX* E194S mutation	5'-GCAGTCGGTGAAGTTCGCTTCGCAGGTCAGATA-3'	None
CY068	Sense primer for *P. fluorescens phoX* D292A mutation	5'-CATGGGCGACGCCGAACGCGGCG-3'	None
CY069	Antisense primer for *P. fluorescens phoX* D292A mutation	5'-CGCCGCGTTCGGCGTCGCCCATG-3'	None
CY070	Sense primer for *P. fluorescens phoX* E532A mutation	5'-CCGATCGGCTGTGCAGTCACCGGGATC-3'	None
CY071	Antisense primer for *P. fluorescens phoX* E532A mutation	5'-GATCCCGGTGACTGCACAGCCGATCGG-3'	None
CY078	Sense primer for *P. fluorescens phoX* C179D mutation	5'-CGGCACCTTCCAGAACGACGCCAACGGCAAAACG-3'	None
CY079	Antisense primer for *P. fluorescens phoX* C179D mutation	5'- CGTTTTGCCGTTGGCATCGTTCTGGAAGGTGCCG -3'	None

2.2 Molecular biology techniques

DNA of inserts were amplified using polymerase chain reaction (PCR). Primers were designed against sequences complementary to the target region with restriction sites introduced at both the 5' and 3' end of the target gene. Templates used were either genomic DNA or plasmids containing the target sequence. All PCR reactions were carried out in the presence of the heat-stable, high fidelity polymerase KOD (Novagen) or KOD Hot Start (Novagen) according to the manufacturer's protocol. Quickchange site directed mutagenesis was also done using KOD Hot Start polymerase. PCR products were run on 1% agarose gels (Roche agarose dissolved in TAE buffer: 40mM tris(hydroxymethyl)aminomethane (Tris), 20mM acetic acid, 1mM EDTA, pH8.0) in TAE buffer. PCR products were purified from the agarose gels using the QIAquick gel extraction kit (Qiagen).

Digests of vectors and inserts were carried out using restriction digests from New England Biolabs (NEB). Digests were usually left overnight under conditions recommended by NEB (single digests: http://www.neb.com/nebecomm/EnzymeFinder.asp); double digests: http://www.neb.com/nebecomm/DoubleDigestCalculator.asp) to ensure complete digestion. Digested vectors were treated with antartic alkaline phosphatase (NEB) for 1h at 37°C. This step removes 5' phosphate groups from the digested vector to prevent self ligation. Antartic alkaline phosphatase was inactivated by a 5min incubation at 65°C. Digested inserts and vectors were purified from 1% agarose gels using QIAquick gel extraction kit (Qiagen). Inserts were ligated into vectors with compatible cohesive ends in the presence of T4 DNA ligase (NEB) and left overnight at room temperature. Vectors containing inserts were transformed into either competent DH5α or XLI-blue cells (see 2.3) and plated on 1.5% agar in LB broth supplemented with the appropriate antibiotic.

All plasmids were purified using the Qiagen® Miniprep kit system or the Invitrogen Midiprep kit according to the manufacturer's protocol. Plasmids were sent for sequencing at Gene Service in Oxford.

2.3 Preparation of competent cells and cell transformation

Plasmids were maintained in either DH5α or XL1-Blue cells. Cells were made competent using the rubidium chloride method. A single colony was used to inoculate 5ml LB supplemented with appropriate antibiotics and grown overnight at 37°C. The overnight culture was diluted 1:100 into 50ml LB and grown at 37°C to OD_{600nm} ~0.4. The culture was cooled on ice for 30min and harvested at 2,250xg for 10min at 4°C. The cell pellet was resuspended in 30ml TFBI (Table 2- 4) and incubated on ice for 15min. Cells were pelleted at 2,250xg for 10min at 4°C and resuspended in 2ml TFBII (Table 2- 4). The cells were then incubated on ice for at least 1h and then used immediately or flash frozen with liquid nitrogen and stored at -80°C. Competent cells were transformed with plasmids via the heat shock method. 100µl competent cells were incubated with 1µl plasmid on ice for 30min and then heat shocked at 42°C for 1min before transferring the cells to ice for 1min. 900µl of LB was added and cells were incubated at 37°C for at least 1h before spreading onto 1.5% agar (Merck) in LB broth supplemented with antibiotics for selection of transformed cells and incubated overnight at 37°C.

Table 2- 4: Buffers for making competent cells

Solution	Components
TFBI	30mM potassium acetate 100mM $RbCl_2$ 10mM $CaCl_2$ 50mM $MnCl_2$ 3mM hexamine cobalt chloride 15% v/v glycerol, pH5.8
TFBII	10mM MOPS 75mM $CaCl_2$ 10mM $RbCl_2$ 15% v/v glycerol, pH 6.5

2.4 Growth media and growth conditions

A single colony of the appropriate strain was used to inoculate into 50ml of appropriate medium (Table 2- 5) and grown at 37°C overnight while shaking at 200rpm. If required, growth medium was supplemented with antibiotics (Table 2- 6). A 1:100 dilution of this overnight culture was used to inoculate 750ml of appropriate media. Cells were cultured at 37°C while shaking at 200rpm till OD_{600nm}~0.4. Expression of gene products was then induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) and further cultured for 4-5h. Alternatively, cells were grown at 25°C while shaking till OD_{600nm}~0.4. Expression of gene products was induced by addition of IPTG and cells were left to grow overnight. Cells that do not contain plasmids and therefore do not require induction for protein expression were grown for 5h after inoculation into 750ml medium before harvesting. The optical density of cultures (OD_{600nm}) was measured on an Ultraspec 2100 Prospectrophotometer (GE Healthcare). Cells were harvested by centrifugation at 5,000xg for 12min at 4°C.

Table 2- 5: Growth media

Media	Components	Final concentration
Luria-Broth (LB) (Melford)	Tryptone Yeast Extract NaCl	10g/l 5g/l 10g/l
Low salt LB (Bageshwar and Musser 2007)	Tryptone Yeast extract NaCl Glycerol	10g/l 5g/l 2.5g/l 5% v/v
Terrific broth (Tartoff and Hobbs 1987)	Tryptone Yeast extract Glycerol KH_2PO_4 K_2HPO_4	24g/l 12g/l 0.4% v/v 17mM 72mM
YE/Tryp broth	Tryptone Yeast extract Glucose KH_2PO_4 K_2HPO_4	13.22g/l 13.22g/l 1% w/v 40mM 40mM

Table 2- 6: List of antibiotics

Antibiotic	Final concentration
Ampicillin (Amp)	100µg/ml
Kanamycin (Kan)	50µg/ml
Chloramphenicol (Cm)	25µg/ml
Apramycin (Apr)	50µg/ml
Tetracycline (Tet)	10µg/ml

2.5 Preparation of inner inverted membrane vesicles (IMVs)

2.5.1 Müller protocol

IMVs were prepared from BL21 (DE3)-pLysS cells transformed with p8737. Cells were grown in YE/Tryp broth (Table 2- 5) supplemented with ampicillin and chloramphenicol. Growth conditions are as described in section 2.4. Expression of *tatABCD* was induced by addition of 1mM IPTG. IMVs were isolated based on the method developed by Alami et al (2002). Harvested cells were resuspended in Buffer A (50mM triethanolamine acetate, 1mM EDTA, 250mM sucrose, 1mM DTT, pH7.5) with 2 complete mini EDTA free protease inhibitor tablets (Roche). Cells were centrifuged at 5,500xg for 15min at 4°C and resuspended in 1mg/ml Buffer A. Cells were homogenized in Buffer A in the presence of a few crystals of DNaseI (Sigma). Cells were disrupted by passage through a French pressure cell three times at 8,000 psi. The supernatant was centrifuged at 185,000xg for 1h at 4°C. The membrane pellet was resuspended in Buffer A. The resulting crude membrane extract was loaded on a 3 step sucrose gradient of 11ml 0.77M sucrose/ 11ml 1.44M sucrose/ 9ml 2.02M sucrose made up in sucrose free Buffer A and centrifuged at 83,000xg for 16h at 4°C. Fraction A was collected from the 0.77M/1.44M sucrose interface while fraction B was collected from the 1.44M/2.02M sucrose interface. Collected fractions were diluted four-fold with 50mM triethanolamine acetate, pH7.5 and centrifuged at 185,000xg for 1h at 4°C. The pellet was resuspended and homogenized in EDTA free Buffer A. Samples were stored at -80°C.

2.5.2 Musser protocol

Musser IMVs were isolated according to Bageshwar and Musser (2007). IMVs were prepared from BL21 (DE3)-pLysS cells transformed with p8737. Cells were grown in low salt LB (Table 2- 5) in the presence of ampicillin and chloramphenicol. Growth conditions are as described in section 2.4. Expression of *tatABCD* was induced by addition of 1mM IPTG. Harvested cells were resuspended and homogenized in 50ml ice-cold Buffer A (Table 2- 7) with 0.4mg/ml lysozyme (Sigma), 0.5mM EDTA, and 2 complete mini EDTA free protease inhibitor tablets (Roche). After a 20min incubation on ice, the resulting spheroplasts were sedimented by centrifugation at 4,200xg for 12min at 4°C. The pellet was resuspended in Buffer A and passed through a French pressure cell once at 6,000 psi to produce IMVs. Cell debris was removed by centrifugation at 4,200xg for 12min at 4°C. The crude extract (10ml each) was layered over a 10ml 2.2M sucrose cushion in Buffer B (Table 2- 7) and centrifuged at 150,000xg for 75min at 4°C. IMVs were collected from the interface between the sucrose cushion and crude extract. The IMVs were diluted four-fold with Buffer B and centrifuged at 150,000xg for 45min at 4°C. The sample was resuspended in storage buffer (Table 2- 7) and stored at -80°C.

Table 2- 7: Buffers for Musser IMV protocol

Solution	Components
Buffer A	1mM $MgSO_4$ 0.5% polyvinylpyrrolidine 450mM mannitol 2mM DTT 50µg/ml DNase I 10µg/ml RNAase 1mM potassium acetate 100mM Tricine, pH 7.5
Buffer B	1mM potassium acetate 1mM $MgSO_4$ 2mM DTT 10mM HEPES, pH 7.0
Storage buffer	5mM magnesium acetate 50mM potassium acetate 200mM sucrose 57µg/ml BSA 25mM MOPS 25mM MES 1mM DTT 50% v/v glycerol, pH7.0

2.6 Protein purification

Table 2- 8: List of proteins purified

Protein	Plasmid	Host strain	Initial Purification step
CueO	pSC22	C43 (DE3) *Δtat*	GST affinity chromatography (see 2.6.2.1)
CueO (mature)	pSC17	C43 (DE3) *Δtat*	Ni^{2+} affinity chromatography (see 2.6.2.2)
CueO_GS3	pGEX-CueO-$GS3_{His}$	C43 (DE3) *Δtat*	GST affinity chromatography (see 2.6.2.1)
MdoD	pGEX-$MdoD_{His}$	C43 (DE3) *Δtat*	GST affinity chromatography (see 2.6.2.1)
3C protease	pGEX-3C-protease	Bl21 (DE3) pLysS	GST affinity chromatography (see 2.6.2.1)
P. fluorescens PhoX (incl. mutants)	pQE60-PF_$PhoX_{His}$	C43 (DE3) *Δtat* pREP4	Ni^{2+} affinity column (see 2.6.2.2)
V. fischeri PhoX	pGEX-VF_$PhoX_{His}$	C43 (DE3) *Δtat*	GST affinity chromatography (see 2.6.2.1)

2.6.1 Preparation of cell lysates

Host strains that had been transformed with the required plasmids (Table 2- 8) were grown in LB (Table 2- 5). Growth conditions are as described in section 2.4. Gene expression was induced by addition of 0.1mM IPTG. Harvested cells were resuspended in resuspension buffer (Table 2- 9) supplemented with protease inhibitor tablets (Roche) and a few crystals of lysozyme (Sigma) and DNase (Sigma). Cells were passed through a French pressure cell three times at 8,000psi. The supernatant was centrifuged for 20min at 25,000xg at 4°C and subjected to further centrifugation at 185,000xg for 1h at 4°C. The lysate was then filtered through a 0.45µm syringe filter and DTT was added to a final concentration of 1mM.

2.6.2 Chromatography methods

2.6.2.1 GST (glutathione S-transferase) affinity chromatography

GST fusion proteins were purified from cell lysates by affinity chromatography using an immobilized glutathione matrix where only GST-tagged proteins bind to glutathione. In this work, cell lysate was loaded at 1ml/min onto a 5ml GSTrap FF column (GE Healthcare), that had been equilibrated with binding buffer (Table 2- 9). GST fusion proteins were eluted in 50% elution buffer (Table 2- 9)/ 50% binding buffer. Glutathione was removed from the sample using 30kDa molecular weight cut off (MWCO) Amicon Ultra-4 centrifugal filter devices. The GST domain was cleaved from the protein of interest by addition of 1:20 molar ratio of GST tagged protein of interest to GST tagged 3C protease in resuspension buffer (Table 2- 9) overnight at 4°C. Post cleavage sample was loaded at 0.2ml/min onto a 1ml GSTrap FF column (GE Healthcare) that had been pre-equilibrated with binding buffer. Target proteins which had their GST tag cleaved would not be able to bind to glutathione and were collected in the flow through. Free GST and GST tagged 3C protease would remain bound on the matrix. Fractions were analysed on SDS-PAGE and proteins were flash frozen in binding buffer for storage at -80°C (Table 2- 9).

Purification of GST tagged 3C-protease was carried out in HEPES buffers (Table 2- 9). Peak fractions were collected and glutathione removed by dialysis using a dialysis membrane with MWCO of 6-8kDa at 4°C overnight. Purified GST tagged 3C-protease was flash frozen and stored in 50% v/v glycerol/50% HEPES resuspension buffer (Table 2- 9) at -80°C.

2.6.2.2 Ni^{2+} affinity chromatography

His-tag proteins (6x His at either the N- or C- terminus) were purified by Ni^{2+} affinity chromatography. Histidine residues will chelate the Ni^{2+} ion via their imidazole ring side chain. Cell lysate was loaded at 1ml/min onto a 5ml HisTrap column (GE healthcare), that

had been equilibrated with resuspension buffer (Table 2- 9) containing 30mM imidazole. The low concentration of imidazole was added to prevent binding of weakly chelating contaminating proteins while allowing His-tagged proteins to bind strongly onto the column. Cell lysates were added to the column and the column was washed with resuspension buffer containing 30mM imidazole to remove unbound samples. His-tagged proteins were then eluted at a linear gradient from 30mM to 1M imidazole in resuspension buffer over 100ml at 1min/ml. Peak A_{280nm} fractions were collected and analyzed on SDS-PAGE. Imidazole was removed by dialysis using a dialysis membrane with MWCO of 6-8kDa at 4°C overnight. Purified proteins were flash frozen and stored in resuspension buffer at -80°C (Table 2- 9).

2.6.2.3 Size exclusion chromatography

If the protein of interest was still contaminated by other proteins after affinity purification, the sample was loaded onto a Superdex 75 HR 10/30 gel filtration column (GE Healthcare) that had been pre-equilibrated with resuspension buffer (Table 2- 9). Proteins were eluted in resuspension buffer according to their size where larger proteins will be eluted first. Peak A_{280nm} fractions were analyzed by SDS-PAGE and the fractions containing the protein of interest are collected.

2.6.2.4 Cationic exchange chromatography

Sample proteins that were stored in resuspension buffer were diluted 40x into 50mM sodium acetate, pH5.5, and loaded at 1ml/min onto a 5ml SP column (GE Healthcare). When a stable A_{280nm} baseline was achieved, the sample was eluted at 1ml/min with a linear gradient of 0M to 1M NaCl. 1ml fractions were collected and analyzed.

Table 2- 9: Buffers for affinity chromatography and gel filtration chromatography

Solution	Components
Binding buffer	20mM MOPS 200mM NaCl 1mM DTT, pH7.5
Elution buffer	20mM MOPS 200mM NaCl 1mM DTT 20mM glutathione, pH7.5
Resuspension buffer	20mM MOPS 200mM NaCl, pH7.5
HEPES binding buffer	200mM NaCl 40mM HEPES 1m DTT, pH7.0
HEPES elution buffer	200mM NaCl 40mM HEPES 1m DTT 20mM glutathione, pH7.0
HEPES resuspension buffer	200mM NaCl 40mM HEPES, pH7.0

2.6.3 Ultrafiltration of purified proteins

Amicon Ultra-4 centrifugal filter device was used for rapid buffer exchange, desalting of samples and concentrating of purified protein samples. The process of ultrafiltration uses semi-permeable membranes and pressure to separate molecular species on the basis of size and shape. Purified protein samples were loaded into a filter unit that was then put into the collection tube. Typical molecular weight cut off (MWCO) values used in this work are 10kD and 30kDa, depending on the size of the protein. The filter device was centrifuged at 2,250xg and the retentate containing purified protein in resuspension buffer (Table 2- 9) was collected. Proteins were flash frozen and stored at -80^0C.

2.6.4 Protein concentration determination

Protein concentration was determined using the BioRad DC Protein Assay. This colorometric assay is based on the Lowry assay. 20µl of an appropriately diluted sample was mixed with 100µl of Biorad reagent A followed by 800µl of Biorad reagent B. The reaction was left for 10min at room temperature before taking the absorbance value at 750nm. The protein concentration was determined with reference to a calibration curve based on known concentration of BSA in the same buffer. Readings were typically taken in duplicate.

2.7 Purification of α-CueO primary antibody

α-CueO was raised in rabbits against the mature domain of CueO (i.e without signal peptide but with C-terminal His-tag) by Davids Biotechnology (Germany). α-CueO primary antibody was further purified by affinity chromatography. Mature CueO was coupled to a 1ml HiTrap NHS-activated HP column (GE Healthcare) as follows. All flow rates were at 1ml/min unless otherwise stated. The column was washed with 3x2ml of 1mM ice-cold hydrochloric acid before loading with 1ml of mature CueO (10mg/ml). The column was then left at 4°C overnight. Excess active groups that have not been coupled to mature CueO were deactivated by washing the column with 3x2ml deactivation buffer (Table 2- 10). The column was then washed with 3x2ml of wash buffer (Table 2- 10) to remove non-specifically bound ligands. The deactivation step was repeated again but the column was left at room temperature for 30min before the wash step. The deactivation and wash steps were repeated for a third time before injecting 2ml of resuspension buffer (Table 2- 9) into the column. Before purification, the column was washed with 3ml of binding buffer (Table 2- 10), followed by 3ml of elution buffer (Table 2- 10). The column was equilibrated with 10ml of binding buffer before loading on 1ml of rabbit anti-serum that had been filtered through a 0.45µm syringe filter at 0.2ml/min. The column was sealed and left at 4°C overnight. The

column was then washed with 10ml of binding buffer before α-CueO was eluted with 4ml of elution buffer. Neutralization buffer (Table 2- 10) was added until the α-CueO sample was at ~pH7.0. This purified α-CueO was stored at 4°C and used for all experiments in this work.

Table 2- 10: Buffers for affinity purification of α-CueO

Solution	Components
Deactivation buffer	0.5M ethanolamine, 0.5M NaCl, pH8.3
Wash buffer	0.1M acetic acid, 0.5M NaCl, pH4.0
Binding buffer	75mM Tris-HCl, pH8.0
Elution buffer	0.2M glycine, pH2.7
Neutralization buffer	1M Tris-HCl, pH9.0

2.8 Fluorescent labelling of proteins

Cysteine containing substrates (e.g. CueO_GS3) were labelled with AlexaFluor® 633 C_5-maleimide dye. Substrates (50μM) were initially reduced with 2mM tris-(2-carboxyethyl)phosphine (TCEP) solution, pH7.5 (5mg TCEP in 160μl 1M Tris-HCl, pH7.5) at room temperature for 45min. Alexa Fluor® 633 C_5-maleimide dye solution was added drop wise to the protein to give ~20 moles of reagent for each mole of protein. The sample was covered in foil to protect the sample from light and gently mixed overnight at 4ºC on a windmill. The dye-conjugated protein was separated from free dye on a PD-10 desalting column (GE healthcare) equilibrated in resuspension buffer (Table 2- 9).

SDS-PAGE gels were scanned on a Fuji FLA-3000 (Emission λ: 633nm, Excitation λ: 675nm) with resolution set at 50 and sensitivity set at F1000. Quantitative analyses were carried out using the manufacturer's ImageReader program.

2.9 Protein analysis and characterization

2.9.1 SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is widely used to separate proteins according to their molecular weight. The presence of SDS disrupts non-covalent bonds in the proteins. Hence denaturing them and causing proteins to return to their primary (linearized) form. Moreover, anions of SDS coat the proteins, resulting in a negative charge on the protein that is proportional to the mass of the protein.

The Mighty Small™ apparatus (GE Healthcare) was routinely used in this study for SDS-PAGE analysis. The resolving gel comprised 0.325 M Tris-HCl pH 8.8, 0.1% SDS, 0.01% v/v NNN'N'tetramethylenediamine (TEMED) (Sigma), 0.1% ammonium persulfate (APS) (Sigma) and acrylamide (National Diagnostics), which was added accordingly to give 8%, 10%, 12.5% or 15% SDS-PAGE gels. The resolving gel was poured to approximately 2 cm below the top of the gel plates. After polymerization, the stacking gel comprising 3% acrylamide, 0.125 M Tris-HCl pH 6.8, 0.1% SDS, 0.01% v/v TEMED, 0.1% APS was layered on top of the resolving gel and a comb was inserted. Once polymerized, the gels were placed in the gel running apparatus and 1x SDS-PAGE running buffer (Table 2- 11) was added. Samples that had been prepared in 5x SDS-PAGE sample buffer (Table 2- 11) were loaded. SDS-PAGE gels were either stained with Coomassie blue stain (Table 2- 11) followed by destaining in water or prepared for immunobloting (see 2.9.2 and 2.9.3) for further analysis.

Table 2- 11: Buffers for SDS-PAGE

Solution	Components
Coomassie blue stain	50% v/v ethanol 50% v/v glacial acetic acid 0.2% Coomassie Brilliant Blue R250
1x SDS-PAGE running buffer	25 mM Tris 250 mM glycine 0.1% SDS
5x SDS-PAGE sample buffer	100m M Tris-HCl 4% SDS 20% v/v glycerol 0.2% bromophenol blue 200mM DTT, pH 6.8

2.9.2 Immunoblotting

To enable protein detection by immunoblotting, proteins have to be transferred from SDS-PAGE gel onto nitrocellulose membranes. This was performed using a Trans-Blot SD semi-dry electrophoretic cell (BioRad). A piece of Hybond-ECL nitrocellulose membrane (GE Healthcare) was placed on top of four pieces of 3MM blotting paper (Whatman). The SDS-PAGE gel was then placed on top of the nitrocellulose membrane followed by another four pieces of blotting paper. All blotting papers, nitrocellulose membranes and SDS-PAGE gels were pre-soaked in transfer buffer (Table 2- 12) for at least 10min. Transfer was usually performed for 1h at a current of 100mA. After transfer, the membrane was stained in Ponceau S (Table 2- 12) to visualize molecular weight markers. Ponceau S was then rinsed off with PBST (Table 2- 12) before immunodetection.

In some cases, the iBlot® Dry Blotting (Invitrogen) was used. The iBlot® device is a self-contained unit, which uses disposable blotting stacks with integrated nitrocellulose membranes. Each iBlot™ Gel Transfer Stack contains a copper electrode and appropriate cathode and anode buffers in the gel matrix. SDS-PAGE gels were placed between the top and bottom stack and transfer was carried out for 7min30sec at 20V. As with semi-dry gel transfer, the nitrocellulose membrane was stained with Ponceau S and the stain was rinsed off before immunodetection.

Table 2- 12: Buffers for immunoblotting and immunodetection

Solution	Components
Transfer buffer (Semi-dry blotting)	48mM Tris-HCl 39mM glycine 0.037% SDS 20% methanol, pH 8.3
Phosphate buffered saline (PBST)	140mM NaCl 2.7mM KCl 10mM Na_2HPO_4 anhydrous 1.8mM KH_2PO4 anhydrous 0.1% v/v Tween-20, pH7.4
Ponceau S stain	0.1% Ponceau S 5% v/v glacial acetic acid

2.9.3 Immunodetection

Nitrocellulose membranes were blocked in blocking buffer (PBST containing 5% skimmed milk powder) for 1h at room temperature or overnight at 4°C to prevent non-specific interaction of antibodies with the membrane. Primary antibodies were diluted to the appropriate concentration in blocking buffer and added to the nitrocellulose membranes for 1h at room temperature. Unbound antibodies were removed by washing the membrane 6x10min with PBST (Table 2- 12). If the primary antibody used was not conjugated to

horseradish peroxidise (HRP), secondary antibodies conjugated with HRP were added to the membrane at the appropriate dilution in blocking buffer for 1h at room temperature. The membrane was then washed 6x10min with PBST. Chemi-luminescence detection was performed using the ECL Western blotting kit (GE Healthcare), which is based on the emission of light resulting from the HRP-catalyzed oxidation of luminol. A piece of autoradiography film (Hyperfilm ECL, GE Healthcare) was placed on top of membranes and exposed according to signal strength of the protein bands of interest. Films were developed using a XOgraph imaging systems Compact X4 developer.

2.10 *In vitro* Tat transport assay

Transport assays were based on the Bageshwar and Musser method (2007). Standard *in vitro* translocation assays used a 35-µl reaction volume containing 100nM substrate protein and 4mM ATP in translocation buffer (Table 2- 13). IMVs were added to preheated reaction mixture (37°C) to a final concentration of A_{280nm} = 10. The reaction was carried out at 37°C for 30min and the reaction was quenched by placing it on ice for 2min. The samples were then treated with 0.73mg/ml proteinase K (PK) for 40min at room temperature to digest untranslocated substrates left in the reaction mixture. Samples were precipitated with 5% trichloroacetic acid (TCA) to inactivate PK and incubated at 56°C for 10min before being placed on ice for at least 1 h. Samples were then pelleted at 16,000xg for 5min and the TCA precipitate was redissolved in 30µl loading buffer (Table 2- 13) and incubated at 50°C for 10min. Samples were then heated to 95°C for 5min before being resolved by 10% SDS-PAGE and detected by immunoblotting (see 2.9.2, 2.9.3) or in-gel fluorescence scan (see 2.8).

Table 2- 13: Buffers for translocation assay

Solution	Components
Translocation buffer	5mM magnesium acetate 50mM potassium acetate 200mM sucrose 57µg/ml BSA 25mM MOPS 25mM MES, pH 7.0
Loading buffer	132mM Tris-HCl, pH 6.8 1M EDTA 3% SDS 12% v/v glycerol 0.1M DTT 0.015% bromophenol blue

2.11 Determination of ΔpH and ΔΨ gradients

The presence of ΔpH and ΔΨ gradients across IMV membranes was determined by fluorescence spectroscopy (Perkin Ellman fluorimeter LS-55) using 0.1µM 9-Amino-6-chloro-2-methoxyacridine (ACMA) in from a 100µM stock in 10% v/v ethanol (Excitation λ: 419nm; Emission λ: 483nm; slits: 5nm each) and 0.2µM oxonol V in from a 200µM stock in 20% v/v ethanol (Excitation λ: 625nm; Emission λ: 645nm; slits: 10nm each) respectively. IMVs (A_{280nm}=10) were preincubated with the dyes in translocation buffer (Table 2- 13) at 37°C for 5min before the addition of ATP or NADH to a final concentration of 4mM in the presence or absence of protonophores or ionophores (Final total reaction volume = 105µl). Reactions were carried out in a 100µl quartz microcuvette.

2.12 ATP synthase activity assay

ATP hydrolysis rate by IMVs was measured via NADH consumption in ATP regenerating system. IMVs were diluted to 0.02mg/ml in assay buffer (20mM HEPES, 5mM $MgCl_2$, 100mM KCl, 5mM potassium cyanide, 2.5mM phosphoenolpyruvate (PEP), 0.1mg/ml pyruvate kinase (PyK), 0.1mg/ml LDH, pH7.5). A blank reading was obtained before 0.7mM NADH was added. Reaction was initiated by addition of 1mM ATP. The reaction was followed by the decrease in A_{340nm} due to consumption of NADH using a spectrophotometer (Cary 50; Version 1). When NADH was oxidized (A_{340nm} <0.1), another 0.7mM NADH was added to continue the measurement. Enzymatic activity was defined as the amount of enzyme required to convert 1µmol of substrate per minute per ml (U/ml). Extinction coefficient for NADH is $6220M^{-1}cm^{-1}$.

2.13 Characterization of PhoX

2.13.1 Phosphatase activity assay using paranitrophenol phosphate

PhoX activity was assayed by measuring the release of paranitrophenol from paranitrophenol phosphate (pNPP). Assays were performed in 1M Tris-HCl, pH8.0, 0.04% pNPP. 5µg of PhoX was added to 1ml reaction mixture and NaOH was added to a final concentration of 1M to stop the reaction. Measurements were taken at A_{405nm}. Specific activity was defined as 1 µmol of inorganic phosphate (P_i) released min^{-1} $(mg\ enzyme)^{-1}$. Extinction coefficient for paranitrophenol is 18,000 M^{-1} cm^{-1}. All experiments were repeated at least 3 times.

Substrate specificity of PhoX using the malachite green phosphatase assay

PhoX was tested for substrate specificity using phosphorylated compounds. This assay was based on the malachite green phosphatase assay where free phosphates react with molybdate in the reaction mix to form phosphomolybdate. Phosphomolybdate reacts with the malachite green reagent to form a chromogenic compound that absorbs at 620nm. Assays were performed in 96 well microplates using reaction mixtures (80µl) containing 50mM Tris-HCl, pH8.0, 0.5mM $CaCl_2$, 0.02mM substrate and 5µg PhoX. After 10min incubation at 37°C, the reaction was terminated by addition of 20µl malachite green reagent (AMS biotechnology) and left to incubate at room temperature for 30min. Production of P_i was measured at 620nm and quantified against a P_i standard curve.

Chapter 3:

Optimization of the *in vitro* Tat transport system

3.1 Introduction

The development of *in vitro* translocation assays is essential for detailed analysis of the mechanism of the Tat pathway. In general, *in vitro* transport assays rely on the transport of Tat substrates into energized IMVs (Figure 3- 1). Substrate proteins are added to IMVs prepared from *E. coli* strains overproducing TatABC. The generation of a proton motive force (PMF) across the IMVs is required to initiate Tat transport. The Tat translocase recognizes the signal peptide at the N-terminus of the full length substrate protein and transports proteins into IMVs. Subsequently, proteinase K (PK) is added to the reaction to remove non-transported substrate proteins. Transported substrates in the IMVs will be protected from PK digestion. In addition, translocated Tat substrates will be susceptible to cleavage by the signal peptidase, LepB, after transport into the IMVs (Yahr and Wickner 2001; Lüke *et. al.* 2009). LepB cleaves the N-terminal Tat signal peptide of transported substrates, to give only the mature domain of the substrate. Assessment of successful transport is based on the resistance of Tat substrate to PK digestion and production of mature protein.

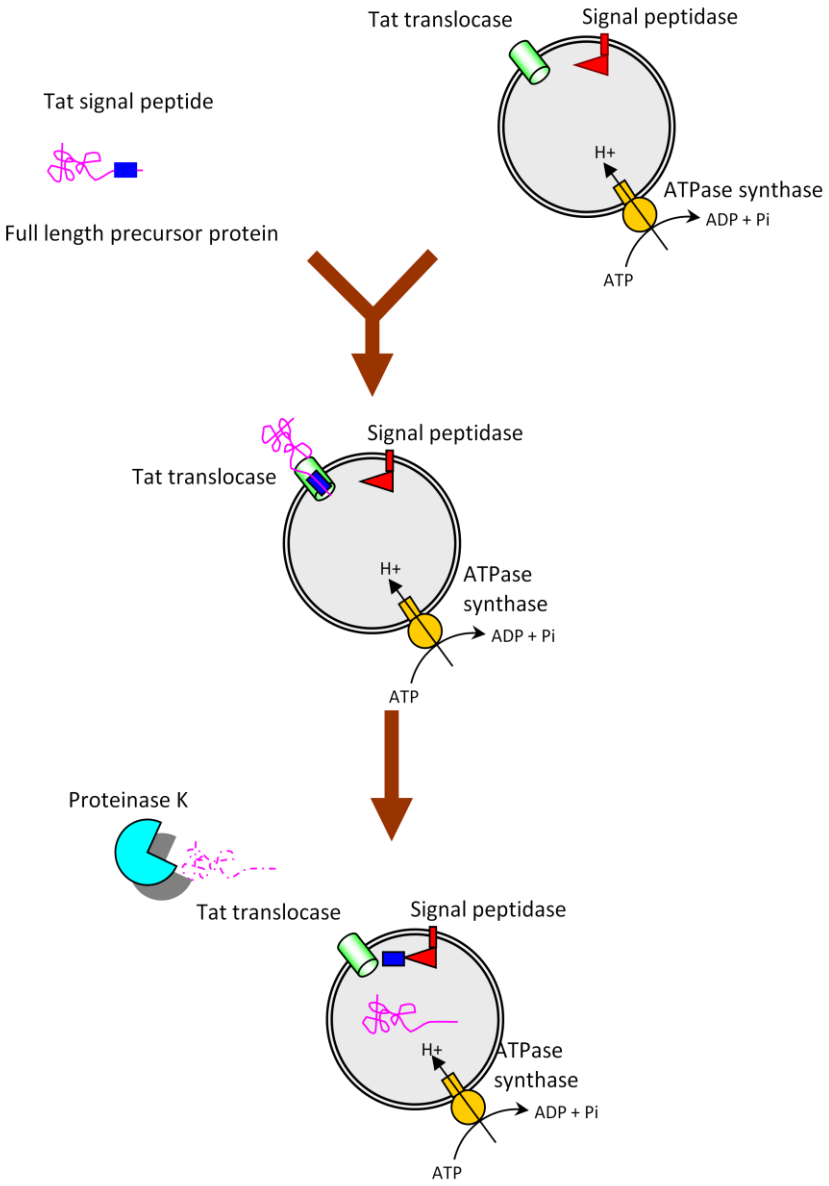


Figure 3- 1: Illustration of the *in vitro* Tat transport assay

Purified substrate proteins are added to IMVs from strains overproducing TatABC. The Tat translocase recognizes the Tat signal peptide and transports Tat substrates into energized IMVs. In this illustration, ATP was used to generate a PMF via the ATP synthase. The signal peptides of transported proteins are cleaved by signal peptidases on the interior face of the IMVs, yielding mature proteins. Subsequently, PK is added to the reaction to remove non-transported substrate proteins. Transported substrates in the IMVs will be protected from PK digestion. Assessment of successful transport is based on the resistance of the Tat substrate to PK digestion and production of mature protein.

The first generation of Tat *in vitro* assays involved import of radiolabelled substrate into inverted *E. coli* inner membrane vesicles (IMVs). The substrate used in these assays was SufI, a native *E. coli* Tat substrate, which was synthesized using an *in vitro* coupled translational-transcriptional system. This initial attempt showed that overproduction of all three main Tat components, TatA, TatB and TatC, was required for detectable transport of Tat substrates. However, the transport efficiency was reported to be only ~ 0.4% (Yahr and Wickner 2001).

As more *in vitro* assays were done, it was observed that the later the IMVs were added after the synthesis of SufI, the lower the transport efficiency. This gave rise to the second generation of *in vitro* Tat transport assays in which IMVs were added during the synthesis of radiolabelled SufI. This resulted in an increase of transport efficiency to ~20% (Alami *et. al.* 2002). As this method originated from Matthias Müller's laboratory, the IMVs and protocol used to make the IMVs for this *in vitro* Tat transport assay will be known as "Müller IMVs" and the "Müller protocol", respectively, in this work.

The transport efficiencies of these early assays are clearly not comparable to the Tat pathway *in vivo*. This constrains the types of mechanistic experiments that can be performed and may result in the assays failing to faithfully recapitulate all features of the physiological Tat mechanism. Recently an *in vitro* Tat transport assay with improved Tat transport efficiency has been reported. Similar to its predecessors, IMVs were derived from TatABC overproducing strains, albeit in a different manner than Müller IMVs. This difference in IMV preparative method will be investigated in detail in this chapter. The substrate used was full length SufI with a C-terminal His-tag that had been purified in biochemical amounts from cells overexpressing s*ufI*. Detection relied on immunoblotting against the C-terminal His-tag of SufI. Using this new protocol, transport efficiencies as high as ~70% was reported with SufI. However, the transport efficiency was reduced to ~15% when green fluorescent protein (GFP) fused with the signal peptide of a native Tat substrate, TorA, was used as the substrate.

Moreover, the authors failed to detect any transport when GFP was fused with the signal peptide of SufI (Figure 3- 2) (Bageshwar and Musser 2007). Since this method was published by Siegfried Musser's laboratory, the IMVs and protocol used to make IMVs from this publication will be known as "Musser IMVs" and the "Musser protocol", respectively, in this work.

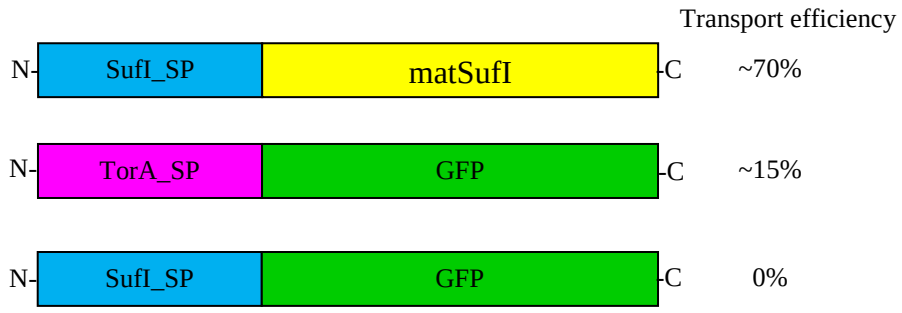


Figure 3- 2: Tat substrates used in Musser's *in vitro* transport assay
Diagrammatic representation of the constructs used for Musser's *in vitro* transport assay. The respective transport efficiency is noted on the right. SP: signal peptide, GFP: green fluorescent protein, matSufI: mature domain of SufI.

This project aimed to further optimize the efficiency of the assay to produce a more reliable *in vitro* Tat transport system. To do so, I first attempted to identify the experimental parameters responsible for the difference in transport efficiency between the Müller and Musser protocols. Identification of these parameters could provide useful mechanistic insights into the Tat transport system.

3.2 Purification of model Tat substrates for *in vitro* assays

In vitro Tat transport assays require full length Tat substrates. To avoid the use of radiolabelled substrates, full length Tat substrates have to be purified in biochemical amounts. Suitable candidates would be native *E. coli* Tat substrates since native substrates are more likely to be transported and are more physiologically relevant. Most native *E. coli* Tat substrates contain complex multiatom cofactors and are involved in redox reactions. Some of these proteins are also membrane bound which makes them difficult to purify. Another criterion for substrates to be used in *in vitro* assays is that the transport of these substrates must not depend on partner proteins. Hence, a good candidate would be a monomeric native Tat substrate that can be easily purified with an intact Tat signal peptide (*i.e.* full length protein) and remains soluble and folded after purification.

3.2.1 Purification of SufI

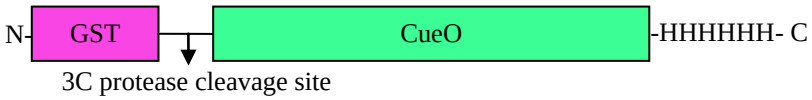
The substrate used for most of the published *in vitro* Tat transport experiments is SufI, a co-factorless Tat substrate involved in cell division (Tarry *et. al.* 2009b). However, I was unable to purify full length SufI as this protein is both highly insoluble and rapidly degraded either during or after purification (data not shown).

3.2.2 Purification of CueO

A possible alternative transport substrate is CueO, another native *E. coli* Tat substrate. CueO is a 56.5kDa multi-copper oxidase that binds 4 copper ions per molecule after it is transported to the periplasm (Grass and Rensing 2001). To prevent cleavage of the N-terminal Tat signal peptide, CueO was synthesized with a N-terminal glutathione-S-transferase (GST) domain with a 3C protease cleavage site separating the GST domain from CueO (Figure 3- 3A). A C-terminal hexa-histidine tag was introduced for immunoblotting purposes. CueO protein purified using the GST domain will contain the Tat signal peptide. The GST domain can then be cleaved off by addition of GST tagged 3C protease (GST-3C

protease) (Figure 3- 3B). The GST cleaved full length CueO was purified from free GST and GST-3C protease by a second GSTrap affinity chromatography step followed by gel filtration chromatography. This led to the elution of a single species of size ~60kDa as determined by gel filtration chromatography and SDS-PAGE analysis (Figure 3- 4). The molecular weight of $CueO_{His}$ could not be verified by mass spectrometry as no peak was observed in the mass spectrum (data not shown).

A)



B)

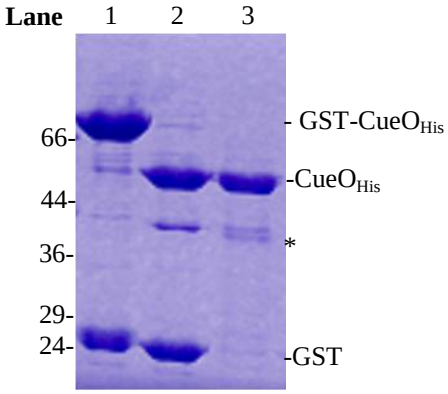
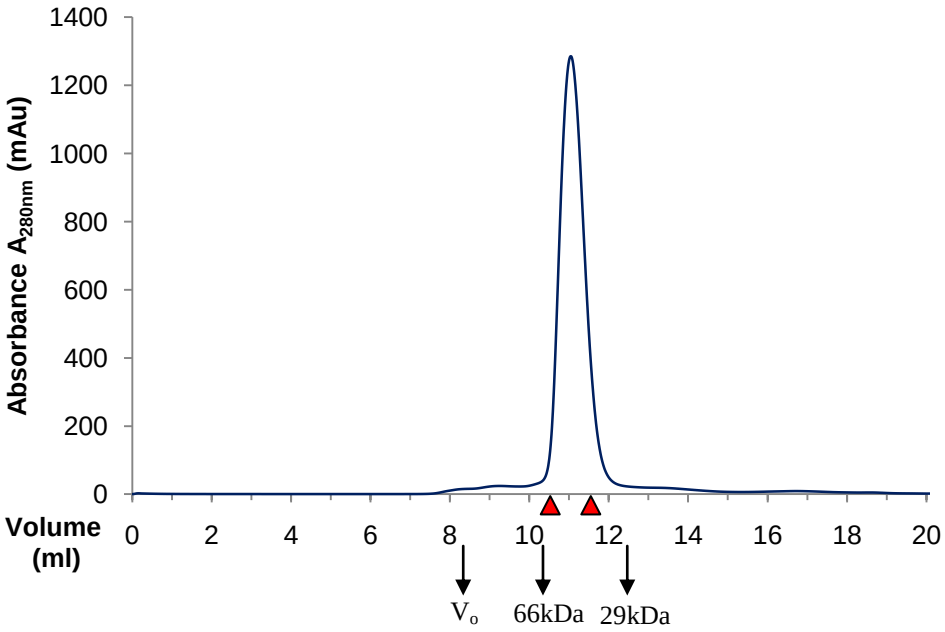


Figure 3- 3: Purification of full length $CueO_{His}$ by GSTrap affinity chromatography

(A) Diagrammatic representation of the GST-$CueO_{His}$ construct. (B) A cleared cell lysate was prepared from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 over producing GST-$CueO_{His}$ from the plasmid pSC22. The lysate was applied to a GSTrap affinity column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5). Bound GST-$CueO_{His}$ was eluted with 10mM glutathione. Glutathione was removed by ultrafiltration using a 30kDa molecular weight cut off (MWCO) membrane. GST was cleaved from GST-$CueO_{His}$ by GST-3C protease at 4°C overnight. The cleaved sample was applied to a GSTrap affinity column equilibrated in resuspension buffer. $CueO_{His}$ will elute in the flow through while free GST and GST-3C protease remained on the column. Fractions collected from both GSTrap affinity columns were analysed by 12.5% SDS-PAGE and proteins visualised by Coomassie Brilliant Blue staining. Lane1: Eluate from first GSTrap affinity column purification; Lane2: After 3C protease cleavage but before the second GSTrap affinity column purification; Lane 3: Flow through of second GSTrap affinity column purification. The molecular masses (kDa) of standard proteins are given to the left of the gel. The $CueO_{His}$ degradation product is indicated by an asterisk (*).

A)



B)

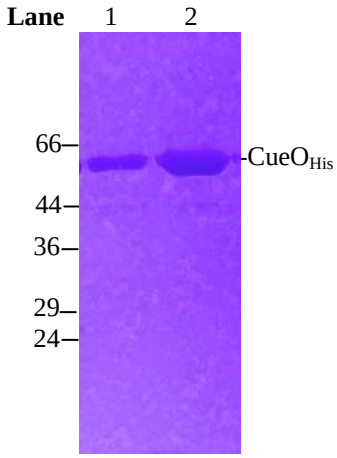


Figure 3- 4: Purification of full length $CueO_{His}$ by size exclusion chromatography

(A) The flow through from the second GSTrap affinity column purification was concentrated to 500µl by ultrafiltration using a 30kDa MWCO membrane and loaded onto a Superdex-75 column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5). The column was run at a flow rate of 0.4ml/min at room temperature and 1ml fractions were collected. The elution positions of water soluble standard proteins are shown below the column profile and labelled with the protein mass (kDa). Standards used were alcohol dehydrogenase (V_o), bovine serum albumin (66 kDa) and carbonic anhydrase (29 kDa). (B) Peak fractions from the size exclusion chromatography column were analysed by 12.5% SDS-PAGE and proteins visualised by Coomassie Brilliant Blue staining. Lane 1: fraction collected at 10; Lane 2: fraction collected at 11ml (both are marked as red triangles on A). The molecular masses (kDa) of standard proteins are given to the left of the gel.

It was observed that full length CueO was relatively well behaved when freshly purified but degraded upon storage. Degradation products became more prominent after two weeks of storage at -80°C (Figure 3- 5). The degradation products could not be separated from full length CueO by size exclusion chromatography alone as the resolution of the column was not high enough.

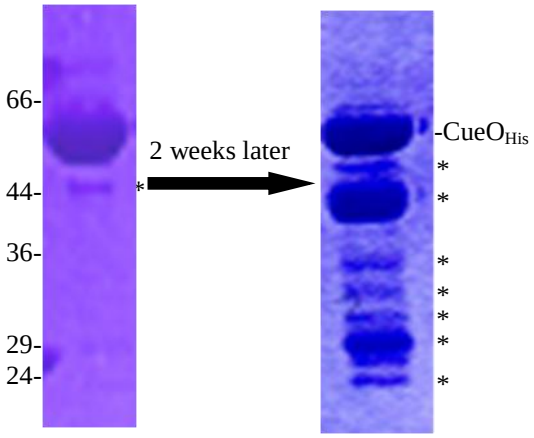
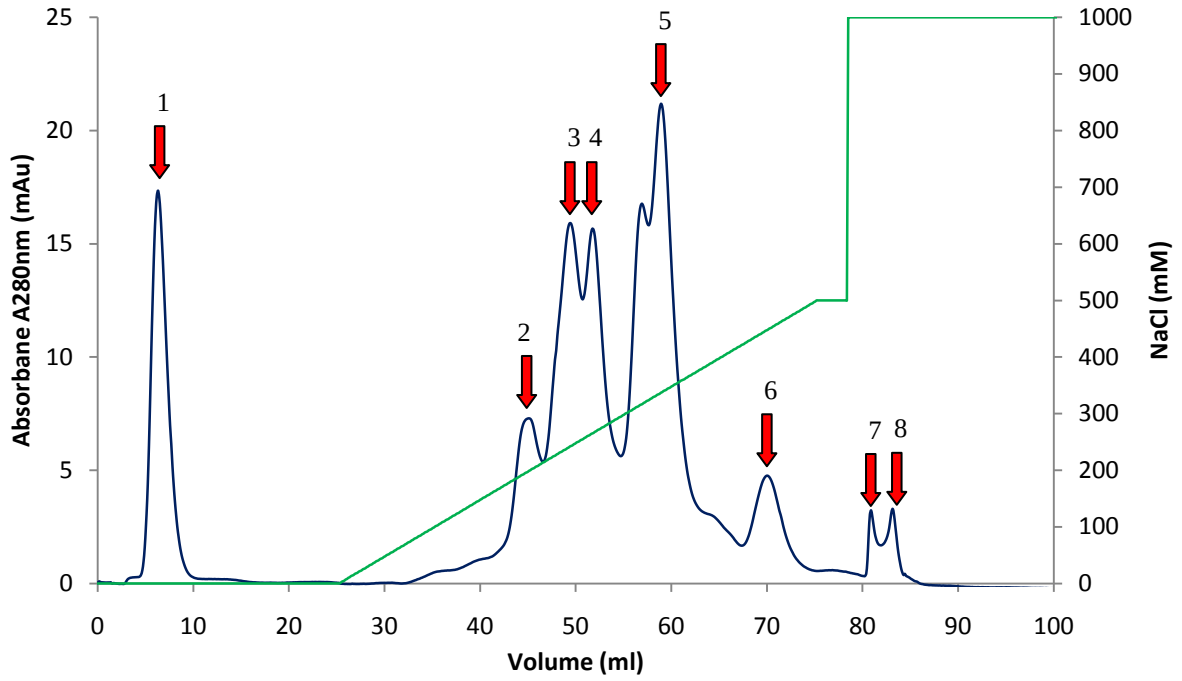


Figure 3- 5: Degradation of CueO$_{His}$
Freshly purified CueO$_{His}$ was flash frozen in liquid nitrogen and stored at -80°C. The sample was thawed two weeks later. Samples were analysed by 12.5% SDS-PAGE and proteins visualised by Coomassie Brilliant Blue staining. The molecular masses (kDa) of standard proteins are given to the left of the gel. CueO$_{His}$ degradation products are indicated by asterisks (*).

As an alternative to size exclusion chromatography, I tried to separate the degradation products using a cationic exchange column. CueO was loaded onto a cationic exchange column and eluted using a linear salt gradient (Figure 3- 6A). Peak fractions were concentrated 20-fold and loaded onto a 10% SDS-PAGE gel for analysis (Figure 3- 6B). Some of the degradation products were removed in the wash fraction (Fraction 1). Due to the effect of concentrating this wash fraction, the residual amount of GST-CueO$_{His}$ that remained in the sample became detectable. Multiple peaks were observed when the salt gradient was applied. Proteins in all of these fractions (Fractions 2-8) migrated to ~55kDa on SDS-PAGE and were assumed to be CueO. It is undetermined if this CueO is full length or mature since it is not possible to resolve the two species on a 10% SDS-PAGE gel. It is hypothesized that

these peaks represent different CueO species that contain different numbers of copper ions. As it is unclear whether the Tat system can transport CueO regardless of the composition of the copper ions bound, it is possible that only a proportion of CueO added into the reaction is Tat transport compatible.

A)



B)

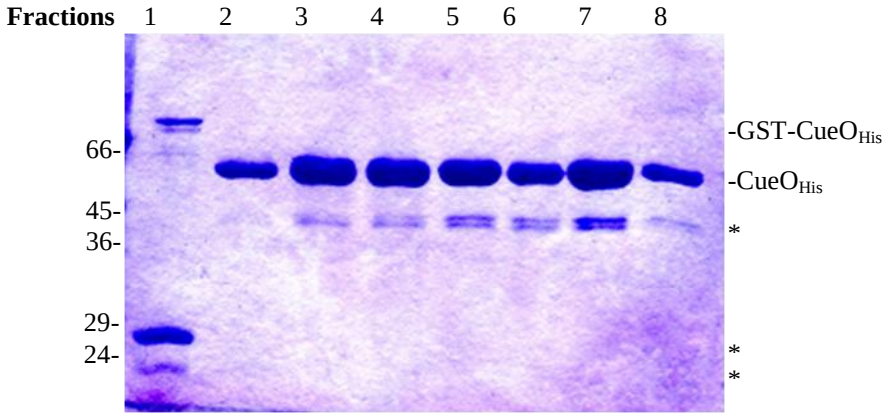


Figure 3- 6: Cationic exchange chromatography of CueO

(A) 100µl of CueO$_{His}$ (4mg/ml) was loaded onto a 5ml HiTrap SP HP column equilibrated in buffer A (50mM sodium acetate, pH5.5). CueO$_{His}$ was eluted with a linear NaCl gradient from 0mM to 500mM NaCl. Chromatography was done at room temperature. 1ml fractions were collected. (B) Peak fractions from the NaCl gradient elution were concentrated to 50µl. Samples were analyzed by 10% SDS-PAGE and visualized by Coomassie Brilliant Blue staining. Lanes 1-8 correspond to peak fractions from the NaCl elution gradient as indicated in (A). The molecular masses (kDa) of standard proteins are given to the left, and the position of CueO$_{His}$ to the right, of the gel.

In summary, I had successfully purified full length CueO. However, CueO degraded over time. This meant that the Tat transport competency of CueO would decrease upon storage. Nonetheless, fresh preparations of CueO (i.e. <2 weeks after purification) mainly consisted of full length CueO that can be used for *in vitro* transport. It was also observed that the CueO purified was made up of a heterogeneous population, possibly comprising of different copper binding forms of CueO. This could influence the efficiency of the transport assay if only a sub-population of the purified CueO is Tat transport competent.

3.3 Comparison of current IMV isolation protocols

In vitro Tat transport assays require the use of IMVs. IMVs made have to be in the inside-out orientation (lumen is equivalent to the periplasmic space; reaction mix is equivalent to the cytoplasm) so that the recognition site of the Tat signal peptide on the Tat complex receptor is exposed on the outside of the vesicles. In this orientation, Tat substrates that are added to the reaction mix would be transported into the lumen of the IMVs.

The two most widely used IMV isolation protocols are based on the protocols from the Müller laboratory (Alami *et. al.* 2002) or the Musser laboratory (Bageshwar and Musser 2007). The major differences between the two protocols are the buffers used, the French press procedure and the method of vesicle purification. In addition, Müller IMVs are made from whole cells while Musser IMVs are made from cells that have been spheroplasted to remove the cell wall. Table 3- 1 provides a more detailed comparison between the Müller and Musser protocols.

Table 3- 1: Comparison of Müller and Musser IMV isolation protocols

Differences in preparation of IMVs		Müller IMVs	Musser IMVs
French press procedure		3 x 8,000psi	1 x 6,000psi
Buffers	Lysis/ spheroplasting buffer	50mM triethanolamine acetate 250mM sucrose 1mM EDTA 1mM DTT Protease inhibitors, pH7.5	1mM magnesium sulfate (Mg_2SO_4) 0.5% w/v polyvinylpyrrolidine 450mM mannitol 1mM potassium chloride (KCl) 100mM tricine 2mM DTT 0.4mg/ml lysozyme 0.5mM EDTA 50µg/ml DNaseI 10µg/ml RNaseI Protease inhibitors, pH7.5
	IMV isolation/ dilution buffer	50mM triethanolamine acetate 1mM EDTA 1mM DTT, pH7.5	1mM KCl 1mM Mg_2SO_4 2mM DTT 20mM Hepes, pH 7.0
	Resuspension buffer	50mM triethanolamine acetate 250mM sucrose 1mM DTT, pH7.5	1mM KCl 1mM Mg_2SO_4 2mM DTT 50% v/v glycerol 20mM Hepes, pH 7.0
IMV isolation		ultracentrifugation of crude extract ⇩ 16h ultracentrifugation on a sucrose gradient (0.77M/1.44M/2.02M)	crude extract loaded onto a 2.2M sucrose cushion and ultracentrifuged for 1h

Although a higher transport efficiency was reported for the Musser protocol than the Müller protocol, to date, no one has directly compared the transport efficiency of Müller IMVs and Musser IMVs under identical transport conditions. To investigate whether the difference in IMV isolation protocol is responsible for the difference in transport efficiencies, I purified IMVs using either the Müller or the Musser protocol. All Müller and Musser IMVs used for these translocation assays contained overproduced TatABC (Figure 3- 7).

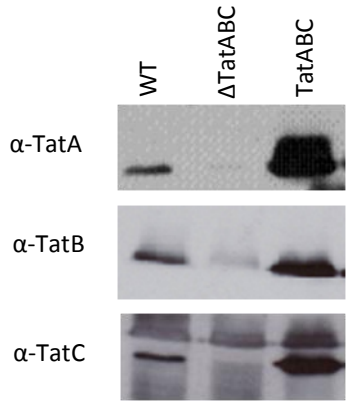


Figure 3- 7: Representative immunoblot analysis of overproduction of TatABC in IMVs
WT: IMVs isolated from *E. coli* strain MC4100 with wild type expression of *tatABC*; ΔTatABC: IMVs isolated from *E. coli ΔtatABCDΔtatE* strain DADE; TatABC: IMVs from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. Each lane contains 20µg of total protein. Proteins were detected using the indicated anti-sera (Dilution factors: 1° antibodies- 1/7,000; 2° antibody α-rabbit: 1/5,000). IMVs tested in this blot were Musser IMVs.

To ensure that the IMVs were in the desired inside-out orientation suitable for *in vitro* transport assays, IMVs were treated with PK to test for the accessibility of the C-terminal domain of TatB. TatB has a single transmembrane domain near the N-terminus and a C-terminal cytoplasmic domain. The cytoplasmic domain of TatB in IMVs of the inside-out orientation will be exposed to the reaction mixture and will be susceptible to PK digestion. Hence, the proportion of IMVs in the inside-out orientation can be determined by using anti-serum against the cytoplasmic domain of TatB. IMVs used in this work had less than 20% of the cytoplasmic domain of TatB remaining after PK digestion (Figure 3- 8, lane 7). This

means that more than 80% of the IMVs used were in the inside-out orientation. A control experiment done in the presence of Triton X-100, a detergent added to permeabilize membranes, confirmed that PK was able to completely digest the cytoplasmic tail of TatB (Figure 3- 8, lane 8).

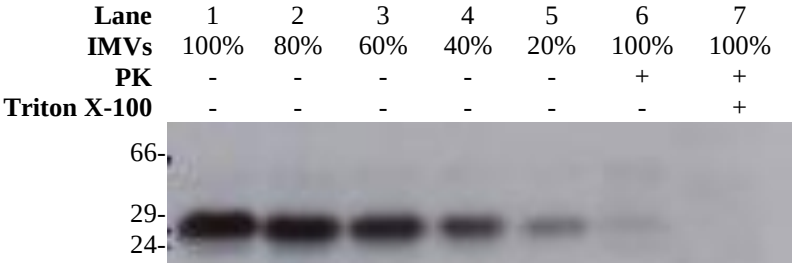


Figure 3- 8: TatB accessibility assay to test orientation of IMVs
IMVs (A_{280nm}=10) made from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 were subjected to PK treatment. The α-TatB antibody (1/7,000) used recognizes the cytoplasmic domain of TatB. Lanes 1-5 show serial dilutions of the stock IMV. The molecular masses (kDa) of standard proteins are given to the left of the blot.

I then compared the transport efficiency between Müller and Musser IMVs using the same transport assay conditions. In this assay, both types of IMVs were energized by ATP and purified full length CueO was used as the Tat substrate. Transport was done at 37°C for 30min and PK was added to digest untransported CueO. CueO could be completely digested by PK (Figure 3- 9; lane 2). Therefore, transported CueO that is protease resistant can be differentiated from non-transported CueO. Transport was detected by immunoblotting of the C-terminal His-tagged CueO. As observed in Figure 3- 9, Musser IMVs were significantly more transport competent than the Müller IMVs. Nonetheless, transport of CueO by Musser IMVs was not complete. Although no transport by Müller IMVs was detected by immunoblotting, transport was shown using radiolabelled *in vitro*-synthesised substrate (S. Chen, unpublished). Detection by radioactivity is more sensitive than immunoblotting.

To ensure that the protease resistant bands observed were due to real transport events, Triton X-100, a detergent which destabilizes the IMVs, was added after the transport reaction.

This treatment resulted in barely detectable bands (Figure 3- 9, lane 8). Addition of carbonyl cyanide *p*-(trifluoromethoxy) phenylhydrazone (FCCP), which dissipates the PMF required for functional Tat transport also resulted in almost complete digestion of CueO by PK (Figure 3- 9, lane 3). These observations mean that any CueO observed after PK digestion represented transported CueO.

Detection of successful transport can also be based on the appearance of mature CueO. Signal peptides of transported full length CueO would be exposed to the signal peptidase, LepB, on the interior face of the IMVs and cleaved to form the smaller sized mature CueO. However, two distinct protease-protected bands were observed in the immunoblots for the Tat transport assays. The band which ran faster corresponded to mature CueO. The other band represented full length CueO. Since this full length CueO was only observed in the presence of a PMF and was susceptible to protease digestion only after Triton X-100 treatment, this likely represented transported but uncleaved CueO. This observation could be due to saturation of LepB.

Lanes	1	2	3	4	5	6	8
IMV type	Muss	-	Muss	Muss	Muss	Mull	Muss
A_{280nm} IMV	10	-	10	20	10	20	10
CueO (nM)	100	100	100	200	100	200	100
FCCP	-	-	+	-	-	-	-
Triton X-100	-	-	-	-	-	-	+
PK	-	+	+	+	+	+	+

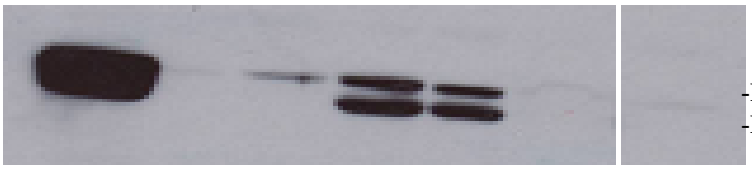


Figure 3- 9: Comparison of transport efficiency between Müller and Musser IMVs

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 according to the Müller (Mull) or Musser (Muss) protocol. IMVs added were standardized according to final A_{280nm} value. IMVs were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. 0.2% Triton X-100 was added following transport as required. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

By directly comparing the Müller IMVs and Musser IMVs in parallel under identical conditions, it is clear that the Musser IMVs have a higher transport efficiency than the Müller IMVs.

3.4 Optimization of the protocol for IMV preparation

3.4.1 Selection of bacterial strain

Different *E. coli* strains have subtle differences in their membrane components. As a result, the *in vitro* Tat transport efficiency could be affected by the strain of *E. coli* from which the IMVs were purified. The IMVs used in the first generation *in vitro* transport assay of Yahr and Wickner (2001) were made from *E. coli* strain MC4100. MC4100 is also used frequently for *tat* genetics. Müller IMVs, which were used for the second generation of *in vitro* assay, were made from BL21, an *E. coli* B-strain (Alami *et. al.* 2002). More recently, Musser IMVs were made from JM109, which is an *E. coli* K-12 strain (Bageshwar and Musser 2007).

To test the effect of *E. coli* strain on the Tat transport competency of IMVs, I overexpressed *tatABC* in four different *E. coli* strains – BL21, JM109, MC4100 and MG1655 (parental strain of MC4100), to select for the strain that supports the highest Tat transport efficiency. All IMVs were isolated in parallel in a similar manner. *tatABC* overexpression tests showed that all three Tat components were overproduced in all strains (Figure 3- 10A). However, BL21 strains produced less TatA than the other strains.

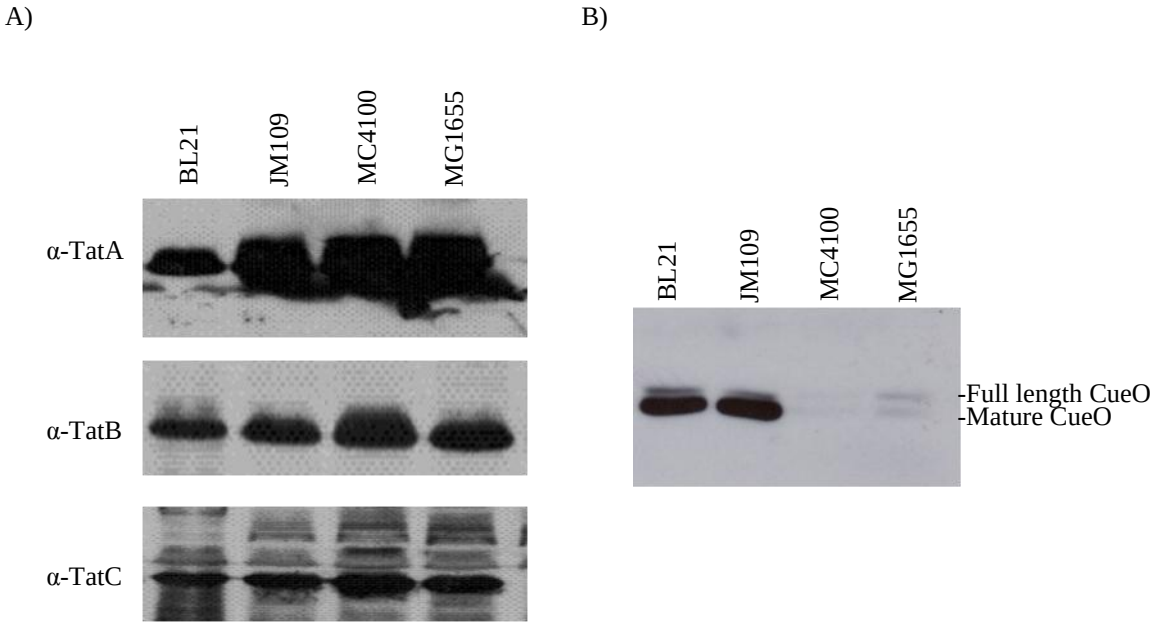


Figure 3- 10: Selection of *E. coli* strain

(A) Immunoblot of TatABC from different *E. coli* strains overproducing TatABC. TatABC were produced from plasmid p8737 for BL21(DE3)-pLysS. Strains JM109, MC4100 and MG1655 contained the accessory plasmid pREP4 and TatABC were produced from plasmid pUniTat. Strains in each lane are as indicated above the blots. Each lane contained 20µg of total protein. Proteins were detected using the indicated anti-sera. (Dilution factors: 1° antibodies: 1/7,000; 2° antibody (α-rabbit): 1/5,000) (B) Tat transport efficiencies of various strains overexpressing *tatABC*. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

Although all four strains had overproduced TatABC, transport was only observed for IMVs purified from the JM109 and BL21 strains even though BL21 IMVs had the least amount of TatA. This result suggests that overexpression of *tatABC* alone does not result in efficient *in vitro* Tat transport and that transport is not directly proportional to TatABC levels. This could be because the ratio of TatA:TatB:TatC is more important than the absolute amount of each individual component. This would make sense because a functional Tat translocase requires all three Tat components. Hence, in the case of JM109, TatB and TatC could be the limiting factor for the formation of functional Tat translocases. In addition, the membrane composition of each cell strain could have an effect. For example, differential

expression of ion channels or membrane proteins in different strains could affect the generation of a PMF, which is required for Tat transport.

Based on this result, I chose to use BL21 as my expression strain in future experiments.

3.4.2 Morphology of IMVs

To determine if Musser IMVs had a different morphology to Müller IMVs, both types of IMVs were viewed under electron microscopy (EM). EM images were taken by Dr. Mike Shaw at the pathology department. EM images revealed that the morphologies of the Müller IMVs and Musser IMVs were distinct from one another (Figure 3- 11). The IMV sample prepared using the Musser protocol had a more heterogeneous population of vesicles. Musser IMVs also appeared to be larger on average than Müller IMVs. This could be because Müller IMVs were made under harsher conditions where whole cells were passed through the French pressure cell three times at 8,000psi. By contrast, Musser IMVs were made from spheroplasts that were disrupted by French press only once at 6,000psi.

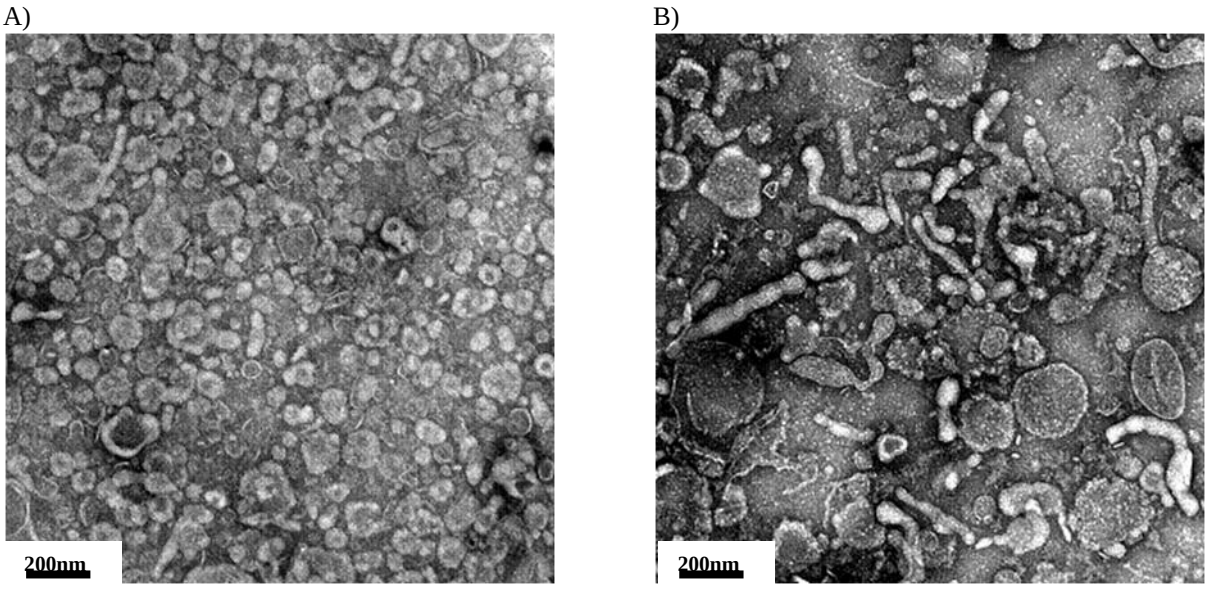


Figure 3- 11: Negative stain electron microscopy images of IMVs
IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 by either (A) the Müller protocol or (B) the Musser protocol. Samples were stained using phosphotungstic acid and viewed under a scanning electron microscope. Scale bar: 200nm. Images were taken by Dr. Mike Shaw.

From these observations, it can be hypothesized that the difference in the morphology of the Müller and Musser IMVs contributes to the difference observed in the transport efficiency. Tat transport can only occur when all three Tat components, TatA, TatB and TatC, are present in the correct stoichiometry to form a complete Tat translocase complex. The larger surface area of Musser IMVs would increase the chances of all required Tat components being present to form a transport compatible complex, thereby increasing transport efficiency.

Since Musser IMVs showed a higher transport efficiency, a more detailed analysis of Musser IMVs was carried out. To reveal further details on the morphology of the Musser IMVs, Musser IMVs were fixed and sectioned before visualization by EM. As seen in Figure 3- 12, the morphologies of the Musser IMVs differed from each other even within one sample. The vesicles generally appeared to be multilamellar. Interestingly, tubular structures were also observed (indicated by red arrows in Figure 3- 12D). Similar tubular structures were also reported by Berthelmann *et. al.* (2008) in *E. coli* overproducing TatABC (Figure 3- 12E). Due to the diversity in IMVs morphologies in any given sample, it is unclear which of these vesicles or tubular structures were Tat transport competent.

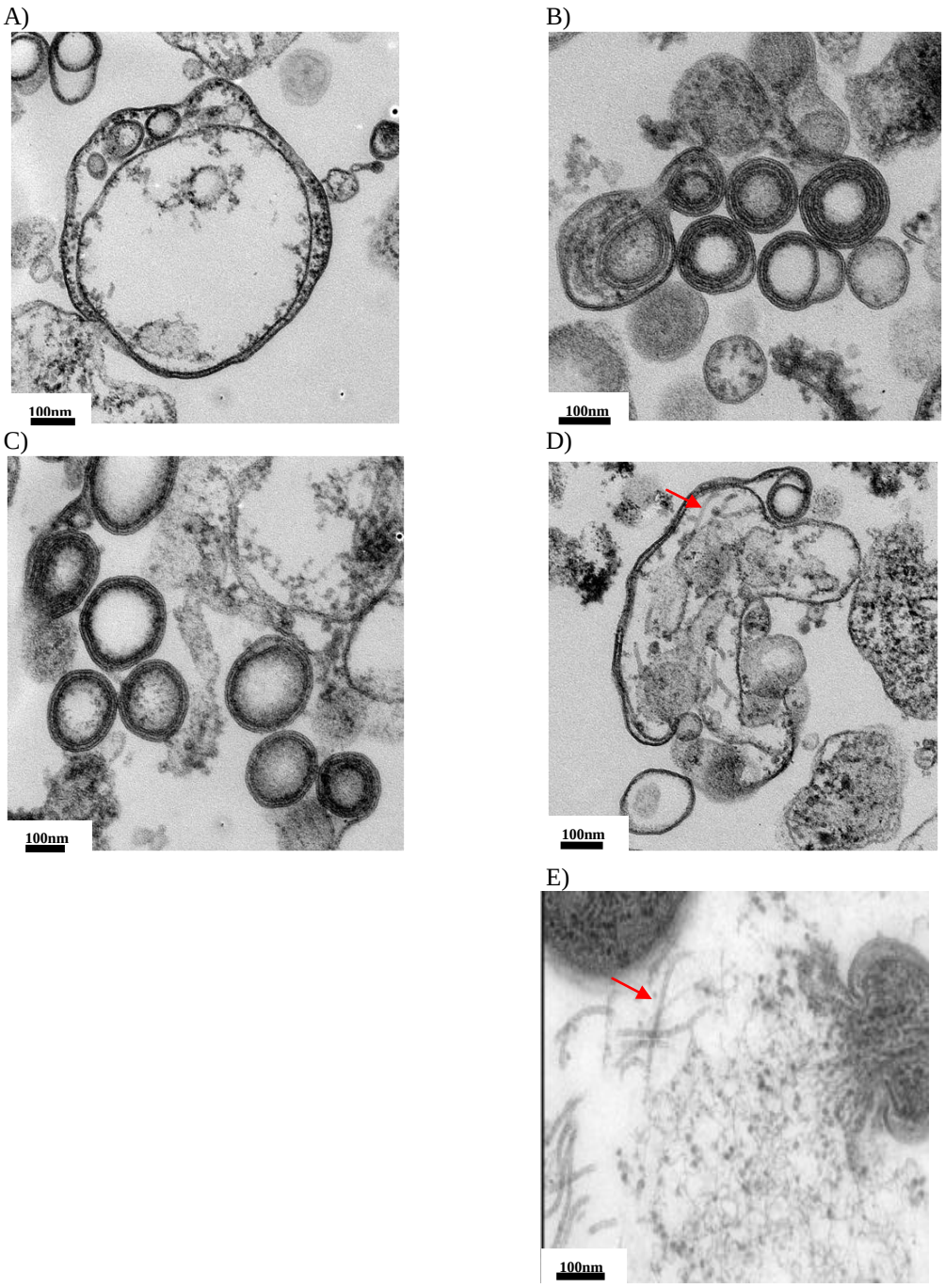


Figure 3- 12 Analysis of Musser IMVs by transmission EM

Musser IMVs were fixed with 2.5% glutaraldehyde and osmium tetroxide. Images A-D are representative images. Red arrows in D and E highlight the tubular structure observed. Image E was adapted from Berthelmann *et. al.* (2008) which shows the tubular structure in broken cells overproducing TatABC. Scale: 100nm. Images A-D were taken by Dr. Mike Shaw.

3.4.3 French press procedure

Since the Fresh press step appears to have an effect on the morphology of the IMVs, I proceeded to test if different French press treatment affected Tat transport efficiency. BL21 strain overexpressing *tatABC* was grown and split into two batches. Both batches were spheroplasted and treated according to the Musser protocol except that one batch of the spheroplasts was French pressed three times at 8,000psi while the other batch was French pressed once at 6,000psi. As a comparison, IMVs were also made according to the Müller protocol where whole cells were either French pressed three times at 8,000psi or once at 6,000psi. Figure 3- 13 shows that different French press treatments had an effect on Tat transport efficiency for both the Müller IMVs and the Musser IMVs. Cells and spheroplasts that were given the milder French press treatment (1x6,000psi) were more transport competent than those that had undergone harsher French press treatment (3x8,000psi). One possible explanation for this difference in transport efficiency could be that the mechanical stress caused by French press treatment had damaged the membranes. Moreover, harsh French press treatment would result in formation of smaller IMVs as seen in Figure 3- 11. These small IMVs might not have sufficient Tat components to form functional Tat translocases.

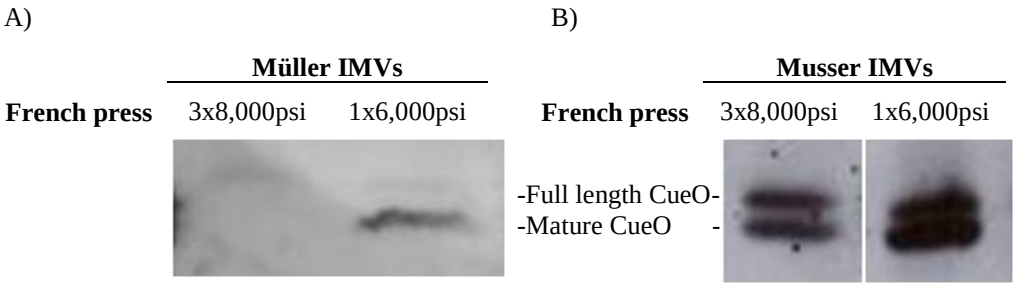


Figure 3- 13: Effect of French press treatment on Tat transport efficiency of IMVs

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 using (A) the Müller protocol or (B) the Musser protocol except for changes in the French press step. IMVs were either produced by mild (1x6,000psi) French press treatment or harsh (3x8,000psi) French press treatment. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

Since mild French press treatment resulted in more efficient transport, I investigated whether further reduction of the French press pressure would increase Tat transport efficiency of Musser IMVs. To achieve this, I split cells that were grown at the same time into several batches. Each batch of spheroplasts was French pressed at a different pressure, while the rest of the protocol remained unchanged. Figure 3- 14 showed that IMVs made from spheroplasts that were French pressed at lower pressures gave slightly better transport efficiencies. Spheroplasts that were not French pressed were also transport competent. This could be due to formation of some IMVs during the spheroplasting step prior to French pressing. Although lower French press pressures resulted in slightly higher transport efficiency, less IMVs were collected in the end as most membranes remained as spheroplasts and would have been removed during the process of isolating the IMVs. Hence, I decided to maintain the French press pressure at 6,000psi as this pressure gave relatively good transport efficiencies without compromising the yield of IMVs produced.

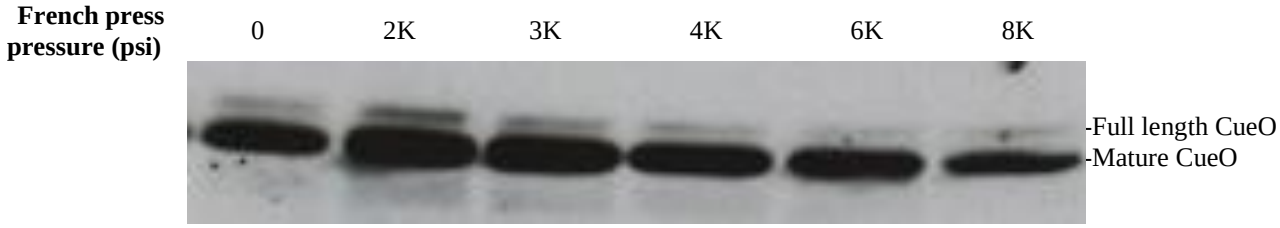


Figure 3- 14: Effect of French press pressure on Tat transport efficiency

IMVs used for the assay were prepared from spheroplasts from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 according to the Musser protocol and passed through the French press once at the indicated pressures. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that still remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

3.4.4 Sucrose density purification of IMVs

IMVs in both the Musser and Müller protocols were purified from the crude extract obtained after the French press step using density separation methods (Figure 3- 15). Musser IMVs were isolated by loading the crude extract directly onto a sucrose cushion followed by ultracentrifugation for 1-hour. By contrast, the Müller protocol for isolating IMVs involved ultracentrifugation of the crude extract. The resuspended pellet of crude extract was then subjected to a 16-hour ultracentrifugation on a 3-step sucrose gradient. The sucrose gradient used in the Müller protocol should provide better separation of IMVs from other cellular fractions, including the outer membrane, than the sucrose cushion used in the Musser protocol.

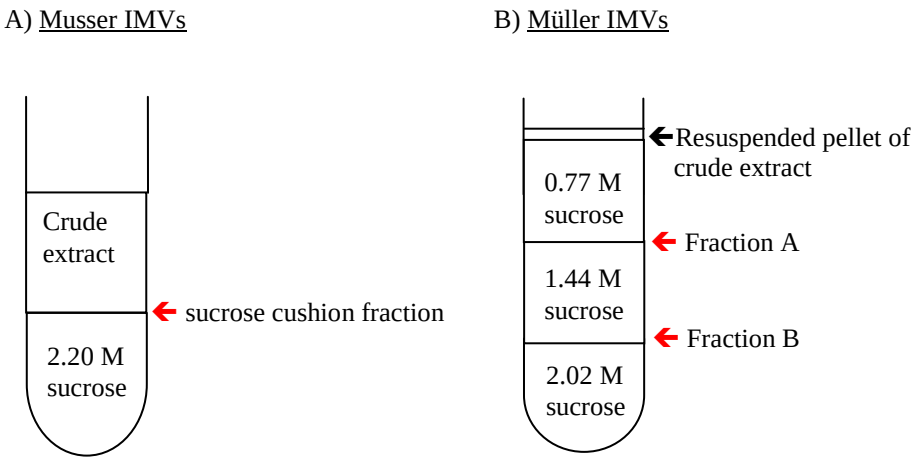


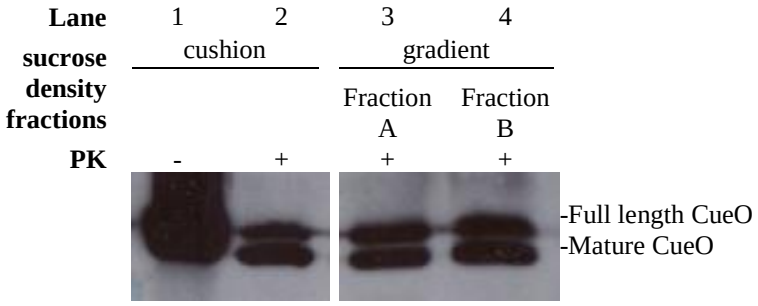
Figure 3- 15: Comparison of Musser and Müller IMV purification methods
Illustration of the different experimental setup in centrifuge tubes for (A) Musser's sucrose cushion and (B) Müller's 3-step sucrose gradient method. Red arrows indicate position of recovered fractions after centrifugation.

To assess the effect of these different purification methods on the Tat transport competency of IMVs, I compared both density separation methods (i.e. Müller's 3-step sucrose gradient vs Musser's sucrose cushion). To allow a direct comparison between the two density separation methods, I omitted the initial ultracentrifugation step in the Müller protocol and loaded the crude extract onto both the Müller sucrose gradient and the Musser sucrose cushion. In the sucrose cushion method, all types of membrane vesicles would accumulate at the sucrose-sample interface resulting in a heterogeneous population of membrane vesicles. In the Müller sucrose gradient, fraction A (interface between 0.77M/1.44M sucrose) should contain the IMVs while fraction B (interface between 1.44M/2.02M sucrose) contains the outer membrane vesicles.

Tat transport activity was observed for the fraction recovered from the sucrose cushion as well as both fractions A and B from the sucrose density method (Figure 3- 16A). All three fractions contained overproduced TatABC including fraction B (Figure 3- 16B, lane 6), which was supposed to contain only outer membranes vesicles. One possible explanation for this would be that there is a heterogeneous population of IMVs in the crude extract and

some of the IMVs might have sedimented with Fraction B. It was also observed that fraction A from the sucrose gradient contained more TatA than the sucrose cushion fraction or fraction B from the sucrose gradient (see Figure 3- 16B, lane 5). However, there was no significant increase in Tat transport efficiencies of IMVs in fraction A. This reinforces the observation made previously in section 3.4.1 where a greater amount of TatA, while maintaining the same amount of TatB and TatC, did not result in higher transport efficiency.

A)



B)

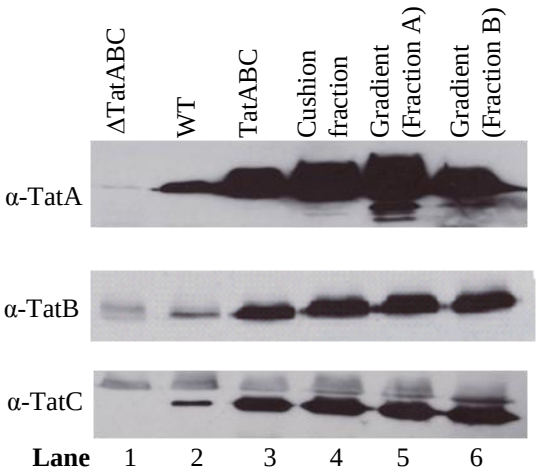


Figure 3- 16: Effect of different sucrose density separation methods on Tat transport efficiency.
(A) IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. Transport of CueO was compared between IMVs purified using a sucrose cushion (lane 2), and fractions A and B from the sucrose gradient as depicted in Figure 3- 15 (lanes 3 and 4 respectively). IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. (B) Immunoblot to test overproduction of TatA, TatB and TatC in IMVs used for this assay. Proteins were detected by the antisera indicated on the left (1° antibodies: 1/7,000; 2° antibody (α-rabbit): 1/5,000). WT: IMVs isolated from *E. coli* strain MC4100 with wild type expression of *tatABC*; ΔTatABC: IMVs isolated from *E. coli ΔtatABCDΔtatE* strain DADE; TatABC: IMVs from BL21 (DE3)-pLysS with overproduced TatABC from plasmid p8737. Each lane contained 20μg of total protein.

To determine if the sucrose gradient method had separated morphologically different population of IMVs, the two fractions recovered from the sucrose gradient were analyzed by EM. The EM images show that the fractions contained membrane vesicles of different morphology. Fraction A from the sucrose gradient contained a mixture of small vesicles and tubular structures (Figure 3- 17A) while fraction B from the sucrose gradient contained mainly large vesicles (Figure 3- 17B). However, EM analysis could not reveal which vesicles were made from the inner membrane and which were from the outer membrane. As a comparison, I also analyzed the fraction collected from the sucrose cushion. This fraction appeared similar to the fraction collected from fraction A (Figure 3- 17A and C). At this stage, it remains unclear as to which of the structures observed were responsible for Tat transport since all three fractions showed similar *in vitro* Tat transport efficiencies (Figure 3- 16).

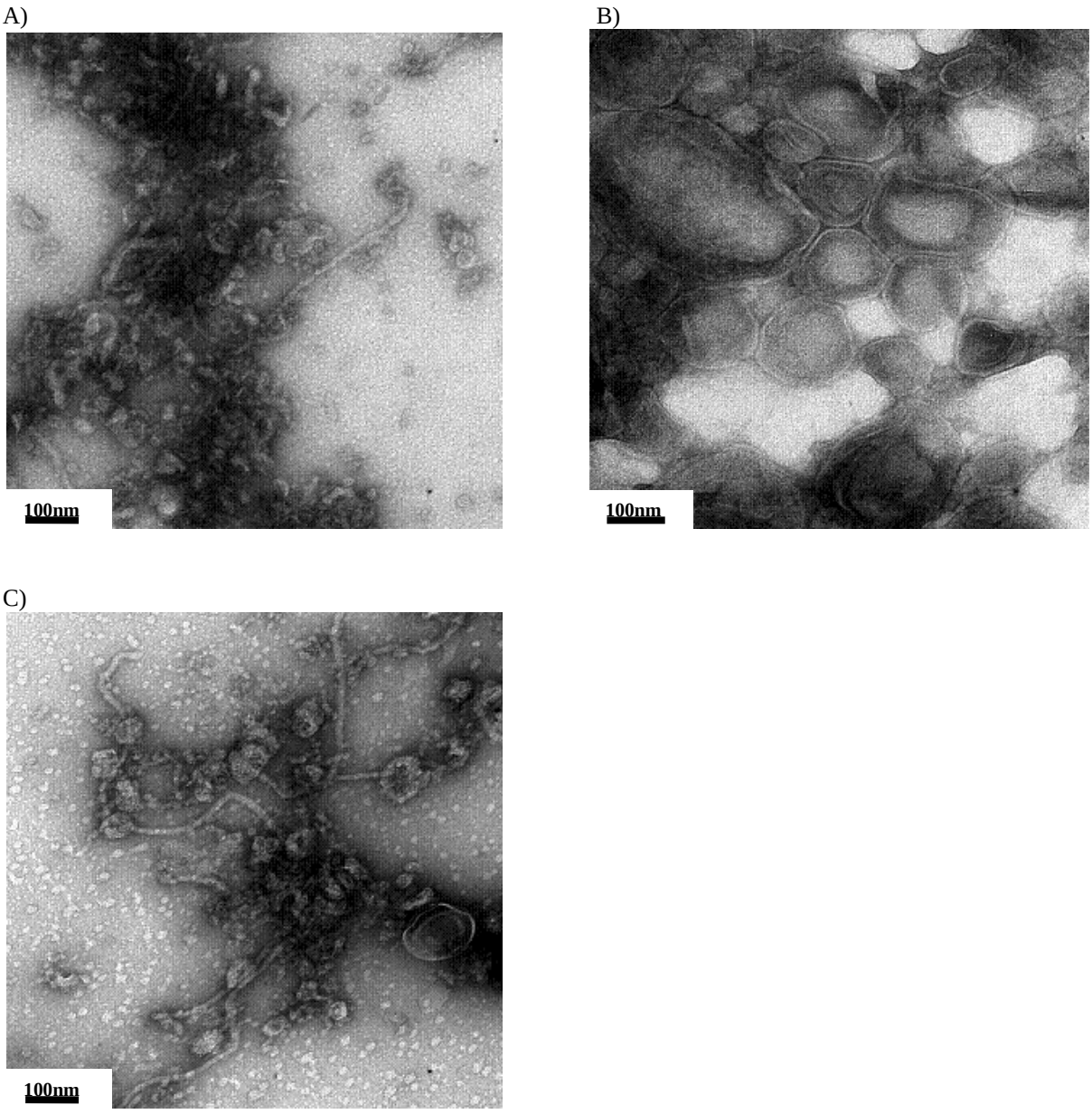


Figure 3- 17: Negative stain EM analysis of IMVs isolated by sucrose cushion and sucrose gradient

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 and isolated using either a sucrose cushion or a 3-step sucrose gradient. Isolated IMVs were analyzed by negative stain EM. Samples were stained using phosphotungstic acid. (A) Fraction A from the sucrose gradient; (B) fraction B from the sucrose gradient; (C) membrane fraction from the sucrose cushion. Scale bar: 100nm. Images were taken by Dr. Mike Shaw.

Overall, the above results indicate that the different density separation method used did not affect Tat transport efficiency. According to the EM images, both the sucrose cushion and sucrose gradient methods produced several populations of morphologically different membrane vesicles. Since the 16-hour sucrose gradient method was more time consuming than the 1-hour sucrose cushion method, I used the sucrose cushion protocol to purify IMVs for all IMV preparations in this thesis.

3.4.5 Reconstitution of IMVs with the soluble fraction

I wanted to test if there were any components in the soluble fraction that could improve Tat transport efficiency *in vitro*. In the standard Musser IMV protocol, crude extract was loaded directly onto the sucrose cushion. IMVs were collected from the sucrose-sample interface and concentrated by ultracentrifugation (Muss IMVs). To obtain the soluble fraction (SF), crude extract was ultracentrifuged before loading onto the sucrose cushion. The supernatant containing the soluble components was collected. The IMVs that were collected were either used directly in Tat transport assays (IMVs_U1) or were concentrated by another ultracentrifugation step (IMVs_U1+U2) before adding to the Tat transport assays. Figure 3-18 shows a diagrammatic representation of the different IMVs purification processes used in this experiment.

As shown in Figure 3- 19A, all three types of IMVs were Tat transport competent. However, IMVs_U1+U2 had the lowest transport efficiency, suggesting that additional ultracentrifugation steps and resuspension steps affect the transport competency of IMVs, possibly due to additional mechanical stress on the IMVs.

Figure 3- 19B showed that reconstitution of IMVs with the soluble fraction not only did not improve Tat transport efficiency but resulted in reduced or loss of Tat transport efficiency. One possible explanation would be that the supernatant contained Tat inhibitory factors or proteases that could cleave the Tat signal peptide from the Tat substrate. While this

work was still in progress, Frielingsdorf *et. al.* (2008) reported findings of a stromal pool of soluble TatA in chloroplasts which improved thylakoid Tat transport efficiency when added to the thylakoid Tat transport assay. Figure 3- 19B showed that the soluble fraction collected from purification of IMVs_U1 and IMVs_U1+U2 contained significant amounts of TatA and slight amounts of TatB and TatC. It can therefore be argued that there is a pool of TatA in the cytoplasm. However, the effect of this pool of soluble TatA on the transport efficiency could not be determined from this experiment due to the presence of other components in the soluble fraction.

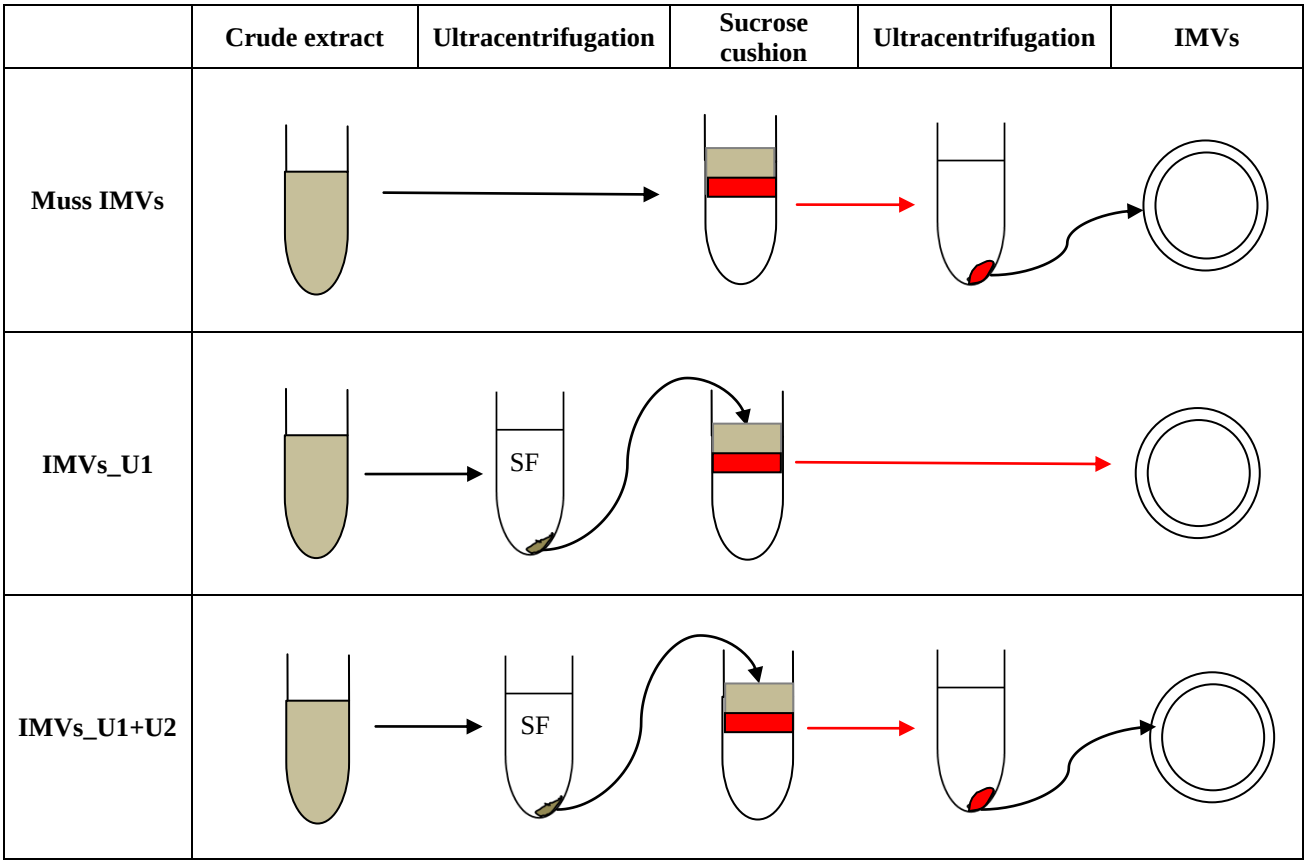
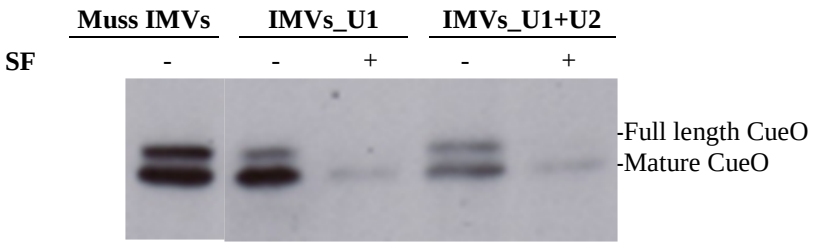


Figure 3- 18: Schematic representation of preparation of IMVs

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. Several types of IMVs were prepared according to different ultracentrifugation isolation methods. Ultracentrifugations were done for 1h at 150,000xg. IMVs were recovered from the surface-sample interfaces of the sucrose cushion (marked as red boxes in the figure). Muss IMVs: Standard Musser IMVs; IMVs_U1: IMVs were ultracentrifuged before loading onto the sucrose cushion, IMVs_U1+U2: IMVs were ultracentrifuged before and after loading onto the sucrose cushion, SF: Soluble fraction

A)



B)

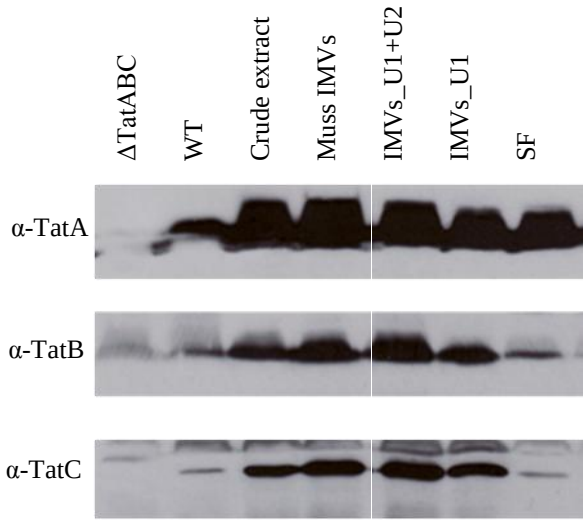


Figure 3- 19: Reconstitution of IMVs with soluble fraction

Several types of IMVs were prepared according to the isolation methods as described in Figure 3- 18. Muss IMVs: Standard Musser IMVs; IMVs_U1: IMVs that were ultracentrifuged before loading on a sucrose cushion; IMVs_U1+U2: IMVs ultracentrifuged before and after the sucrose cushion step; SF: soluble fraction. (A) Transport assays using the different IMV types reconstituted with the soluble fraction (10µg). IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. (B) Immunoblot for TatABC in all types of IMVs and the soluble fraction. ΔTatABC: IMVs isolated from *E. coli ΔtatABCDΔtatE* strain DADE WT: IMVs isolated from *E. coli* strain MC4100 with wild type expression of *tatABC*. 20µg of total protein was loaded for each lane and blotted against the indicated anti-sera. (Dilution factors: 1° antibodies: 1/7,000; 2° antibody (α-rabbit): 1/5,000).

Since the standard Musser IMVs had the best *in vitro* Tat transport efficiency and reconstitution of soluble fraction did not improve Tat transport efficiency, IMVs for further experiments were purified according to the Musser protocol.

3.4.6 Buffer components

Next, I investigated the components in the buffers used to isolate IMVs which could have an effect on *in vitro* Tat transport efficiency. Musser's original IMV protocol used chloride containing buffers. Our laboratory had made the preliminary observation that Cl^- ions have an inhibitory effect on Tat transport. Since acetate (Ac^-) was used as the counterion in the Müller protocol (Alami *et. al.* 2002), I used Ac^- to replace Cl^- as the counterion in all the buffers used in the Musser protocol. These buffers include the spheroplasting buffer, IMV isolation buffer and resuspension buffer (Table 3- 1). However, there was no significant difference in transport efficiency between IMVs made in Cl^- (i.e. Cl^--IMVs) and Ac^- (i.e. Ac^--IMVs) containing buffers (Figure 3- 20, lanes 1 and 2). This could be because the amount of Cl^- in the buffer (i.e. 1mM) was not enough to affect Tat transport. Investigation of the effect of Cl^- on Tat transport will be discussed in section 3.5.2.

In the Musser protocol, the IMVs were isolated and stored in resuspension buffer (RB) while the translocation assay was done in translocation buffer (TB) (Table 3- 2). This mean that if different amounts of IMVs were added to different reaction assays, the proportion of RB to TB would be different for each reaction assay. This could potentially cause variation in results. Therefore, I tried to eliminate this buffer variation by storing IMVs in storage buffer (SB) that is made up of all the components in TB supplemented with 1mM DTT and 50% v/v glycerol. Comparison of transport efficiencies between IMVs stored in RB and SB showed that IMVs could be stored in SB without loss of Tat transport competency (Figure 3- 20, lanes 2 and 4). However, the glycerol in SB is essential to maintain Tat transport competency

(Figure 3- 20, lanes 3 and 4). This could be because glycerol provides protection against ice crystal formation and stabilizes proteins. Hence, I have stored all IMVs in SB.

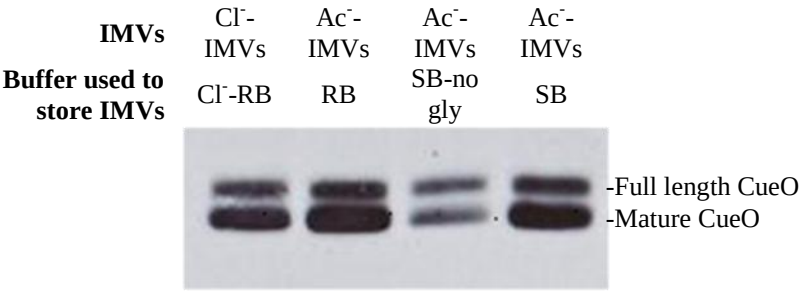


Figure 3- 20: Effect of buffer composition used for isolation of IMVs on Tat transport efficiency
IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 and stored in buffers as described in Table 3- 2. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. Cl^--IMVs : IMVs made in Cl^--containing buffer; Ac^--IMVs : IMVs made in acetate-containing buffers; Cl^--RB: Cl^- containing resuspension buffer; RB: Resuspension buffer; SB-no gly: Storage buffer without glycerol; SB: Storage buffer.

Table 3- 2: Buffers for storing IMVs

Resuspension buffer (RB)	Translocation buffer (TB)	Storage buffer (SB)
1mM potassium acetate 1mM $MgSO_4$ 2mM DTT 10mM HEPES, 50% v/v glycerol, pH 7.0	5mM magnesium acetate 50mM potassium acetate 200mM sucrose 57µg/ml BSA 2mM MOPS 25mM MES, pH7.0	5mM magnesium acetate 50mM potassium acetate 200mM sucrose 57µg/ml BSA 2mM MOPS 25mM MES 1mM DTT 50% v/v glycerol , pH7.0

3.4.7 Final optimized IMV preparation protocol

Comparison of Müller's and Musser's IMV preparation protocol showed that the Musser protocol produced IMVs of higher Tat transport efficiency (Figure 3- 21, lanes 1 and 3). While experimenting on various methods to improve the transport efficiency of the IMVs made, I was able to ascertain that IMVs were sensitive to mechanistic disruption, such as French pressing procedures. I have also shown that the transport efficiency of IMVs were dependent on the *E. coli* strain it was purified from. In addition, I have introduced slight

modifications to reduce variability in the buffers that would otherwise be carried over to *in vitro* Tat transport assays. These include storing the IMVs in SB and replacing Cl^- with Ac^- in all buffers used during the purification process. These additional modifications did not affect Tat transport efficiency (Figure 3- 21, lanes 3 and 4).

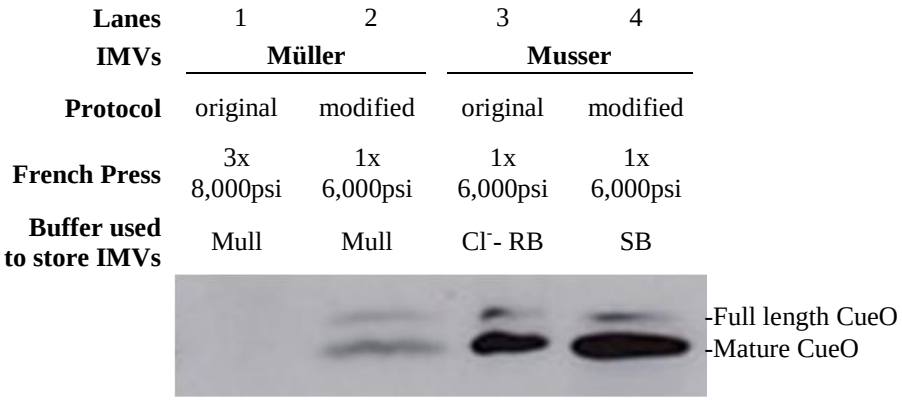


Figure 3- 21: Final comparison of Tat transport competency of IMVs

Tat transport efficiencies of IMVs made from original or modified Müller and Musser protocols were compared under identical *in vitro* Tat transport assay conditions. All IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. Buffer used to store IMVs: Mull: Müller's IMV storage buffer (50mM triethanolamine acetate, 250mM sucrose, 1mM DTT); Cl^--RB: Cl^--containing resuspension buffer (Table 3- 2); SB: storage buffer (Table 3- 2).

Therefore, all the IMVs used in *in vitro* Tat transport assays described in the remainder of this thesis were made using the modified Musser protocol as below:

Cell growth
(*E. coli* strain BL21(DE3)-pLysS overproducing TatABC)
↓
Spheroplast
↓
French press protocol (1x6,000psi)
↓
Sucrose cushion (2.2M sucrose)
↓
Ultracentrifugation (150,000xg for 1h)
↓
Resuspend IMVs in storage buffer and store at -80°C

3.5 Optimization of the *in vitro* Tat transport assay

Having optimized the protocol for isolating IMVs used for Tat transport, I then focused on optimizing the conditions for *in vitro* Tat transport assays.

3.5.1 Energization of IMVs

Tat transport is an active process which requires a proton motive force (PMF) to drive the transport of substrates across membranes. Hence, IMVs used for *in vitro* Tat transport assays need to be energized to produce a PMF across membranes. Methods used for PMF generation include energization by NADH and ATP. NADH was used in the Musser protocol (Bageshwar and Musser 2007) while ATP was used in the Müller protocol (Alami *et. al.* 2002). Here, I compared both energization methods and observed that ATP energized IMVs were more transport competent than NADH energized IMVs (Figure 3- 22).



Figure 3- 22: Effect of IMVs energization method on Tat transport efficiency
IMVs were prepared from E. coli strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP or 4mM NADH and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

This difference in transport efficiency observed could be due to how NADH and ATP generate a PMF (Figure 3- 23). NADH generates a PMF by utilizing the NADH dehydrogenase I complex and cytochrome oxidase-bo_3 in the electron transport chain (ETC) in the inner membrane to pump protons across the membrane. This process relies on the presence of an intact ETC. On the other hand, generation of PMF by ATP involves an ATP synthase coupling ATP hydrolysis to proton movement across the membrane. In addition,

energization by NADH is limited by the presence of oxygen, which acts as the final electron acceptor in the ETC. ATP is considerably more soluble in solution than oxygen.

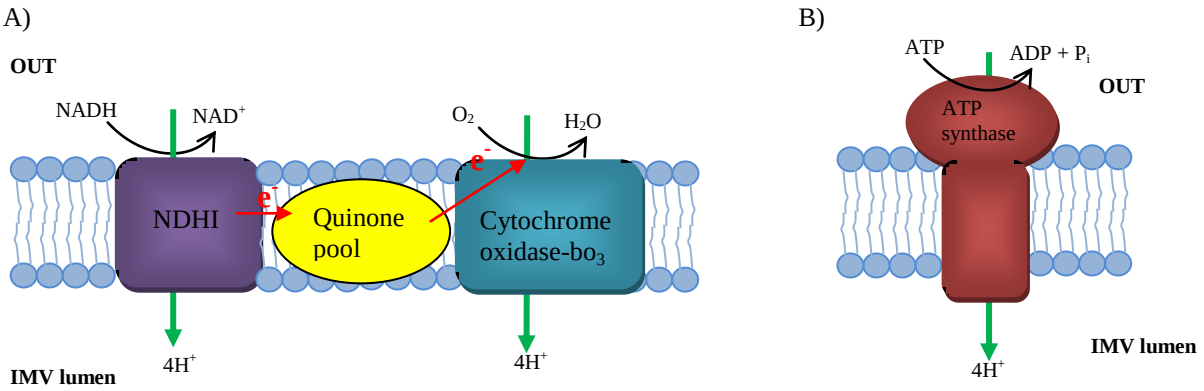


Figure 3- 23: PMF generation by NADH and ATP in *E. coli*

(A) NADH requires an intact ETC to transfer its electron to O_2, the final electron acceptor, for generation of a PMF. Only the proton pumps NDHI and cytochrome oxidase-bo_3 are depicted in this figure. NDHI= NADH dehydrogenase I. (B) ATP only requires an active ATP synthase to generate a PMF.

I compared the ability of NADH and ATP to generate a PMF across IMVs. Both the transmembrane pH gradient (ΔpH) and transmembrane electrical potential difference (ΔΨ) components, which contribute to the PMF, were monitored. The presence of ΔpH in IMVs was assessed by using the pH-sensitive fluorescent dye ACMA (9-Amino-6-chloro-2-methoxyacridine). Upon generation of ΔpH, ACMA binds onto the membrane. Aggregation of ACMA on the membrane surface reduces the amount of free ACMA dyes that are fluorescent. Hence, fluorescence quenching is observed (Huang *et. al.* 1983). The fluorescent dye oxonol V was used to assess the presence of ΔΨ. Oxonol dyes are lipophilic anions and are attracted to positive charged surfaces (Cooper *et. al.* 1990). Aggregation of oxonol dyes on positive surfaces results in a decrease in fluorescence intensity (Ahmed and Krishnamoorthy 1994). However, ACMA and oxonol V quenching measurements are only semi-quantitative and do not provide the exact magnitude of the ΔpH and ΔΨ generated (Rottenberg 1989).

To ensure that IMVs remained transport competent in the presence of the fluorescence dyes, I tested the transport efficiencies of IMVs in the presence of ACMA and oxonol V. Figure 3- 24 shows that addition of ACMA and oxonol V dyes did not affect Tat transport *in vitro*. Therefore, IMVs were capable of active Tat transport under the conditions in which energization tests were done.

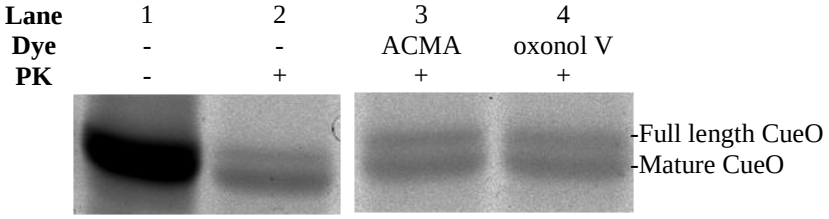


Figure 3- 24: Transport assay under the conditions of the energization tests

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. *In vitro* Tat transport assay was carried out under conditions identical to that during energization tests where the IMVs (A_{280nm}=10) were pre-incubated with 0.1µM ACMA or 0.2µM oxonol V in translocation buffer containing 100nM CueO prior to addition of 4mM ATP. The reaction was left for 30min at 37°C before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Analysis was based on in gel-fluorescence scan of fluorescent dye labelled CueO (CueO_GS3^{633}) (Excitation λ: 633nm, Emission λ: 675nm) (see 3.6.2).

ACMA was quenched upon addition of ATP (Figure 3- 25), suggesting that a ΔpH was generated upon energization by ATP. A control experiment was done where nigericin, an ionophore which dissipates ΔpH, was added after IMVs were energized with ATP. In this control experiment, dissipation of ΔpH by nigericin completely abolished the quenching of the ACMA dye and fluorescence returned to baseline. In NADH energized IMVs, ACMA fluorescence was quenched upon addition of NADH but the quenching recovered almost immediately and remained stable throughout the experimental duration. On the other hand, the recovery of ACMA fluorescence for IMVs energized by ATP was more gradual. However, the ACMA fluorescence for both NADH and ATP energized IMVs did not return to the baseline measurement during the duration of the experiment suggesting there could be residual ΔpH that remained throughout the experiment.

A greater reduction in the fluorescence of oxonol V was observed when IMVs were energized by NADH compared to ATP. However, fluorescence quenching by NADH was generated slightly more gradually than the quenching generated by ATP (Figure 3- 25B). This decrease in oxonol V fluorescence was maintained throughout the experimental duration. As a control experiment, thiocynate (SCN^-) was added after IMVs were energized by ATP. SCN^- is a membrane permeable anion and is therefore capable of dissipating $\Delta\Psi$. Oxonol V quenching was abolished by addition of SCN^-. This confirms that the stable decrease in oxonol V quenching observed for ATP or NADH energized IMVs was due to the presence of a $\Delta\Psi$.

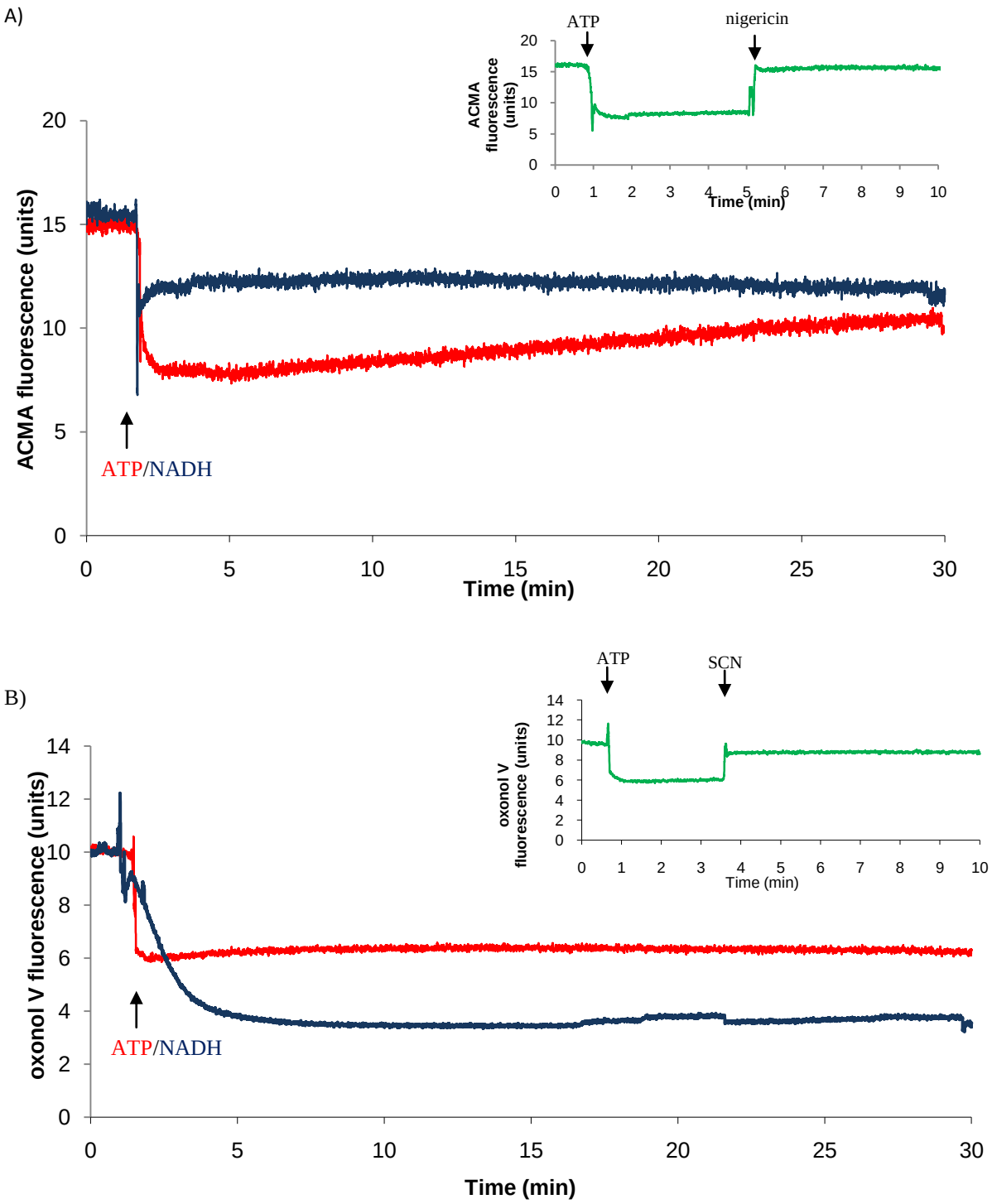


Figure 3- 25: ΔpH and ΔΨ traces of IMVs energized by NADH or ATP

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were incubated at 37°C in 100µl translocation buffer before addition of fluorescent dyes and the absorbance was followed. When a stable fluorescent baseline was obtained, 4mM of ATP or NADH was added. (A) ACMA dye (0.1µM) was used to monitor ΔpH across membranes (Excitation λ: 419nm, Emission λ: 483nm). Insert shows a control experiment where 10µM of the ionophore nigericin was added after IMV energization by ATP to abolish ΔpH (B) Oxonol V dye (0.2µM) was used to monitor ΔΨ (Excitation λ: 625nm, Emission λ: 645nm). The insert shows a control experiment where 25mM SCN^- was added after IMV energization by ATP to dissipate ΔΨ. The experiment was done in triplicate and the average plotted. **Blue**: energization of IMVs by NADH; **Red**: energization of IMVs by ATP.

In summary, NADH and ATP were both able to energize IMVs for *in vitro* Tat transport. IMVs that were energized by ATP resulted in higher Tat transport efficiencies.

3.5.2 Effect of Cl^- ions

Our laboratory has previously observed that Cl^- ions have an inhibitory effect on transport efficiency. The original Musser transport assay had at least 50mM Cl^- ions in the translocation buffer. To investigate the effect of Cl^- ions on transport efficiency, Cl^- ions were replaced with other anions, namely acetate (Ac^-), which was used in the Müller protocol and SO_4^{2-}. These results showed that the transport efficiency of transport assays carried out in Cl^- containing buffers were significantly less than the translocation efficiency of transport assays carried out in either Ac^- or SO_4^{2-} buffers (Figure 3- 26A). A Cl^- ion titration assay showed that concentrations of $Cl^- \geq 10$mM had an inhibitory effect on Tat transport (Figure 3- 26B).

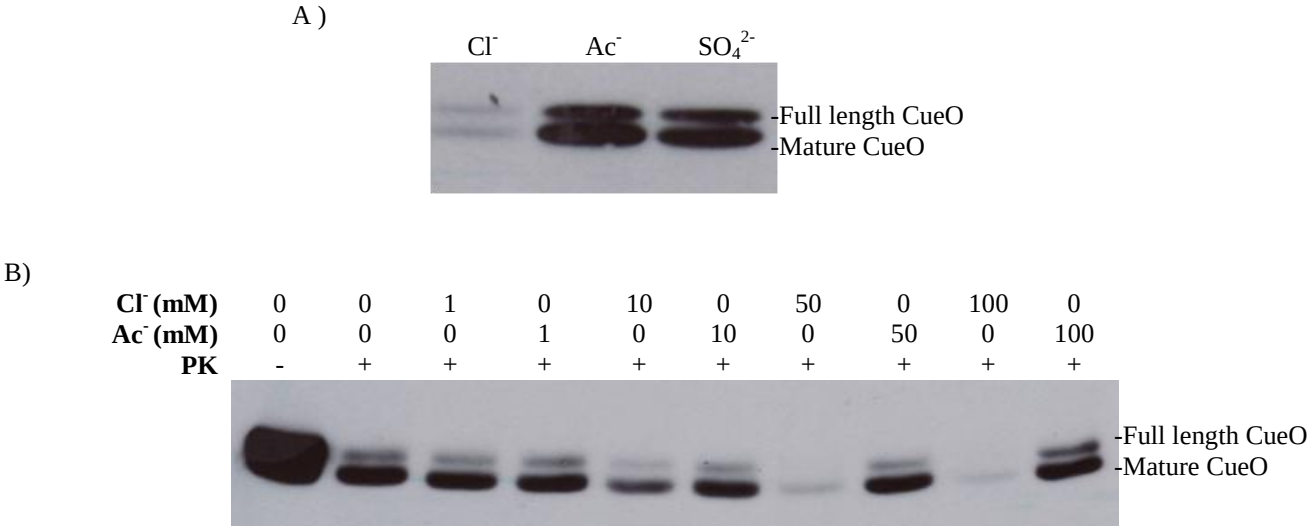
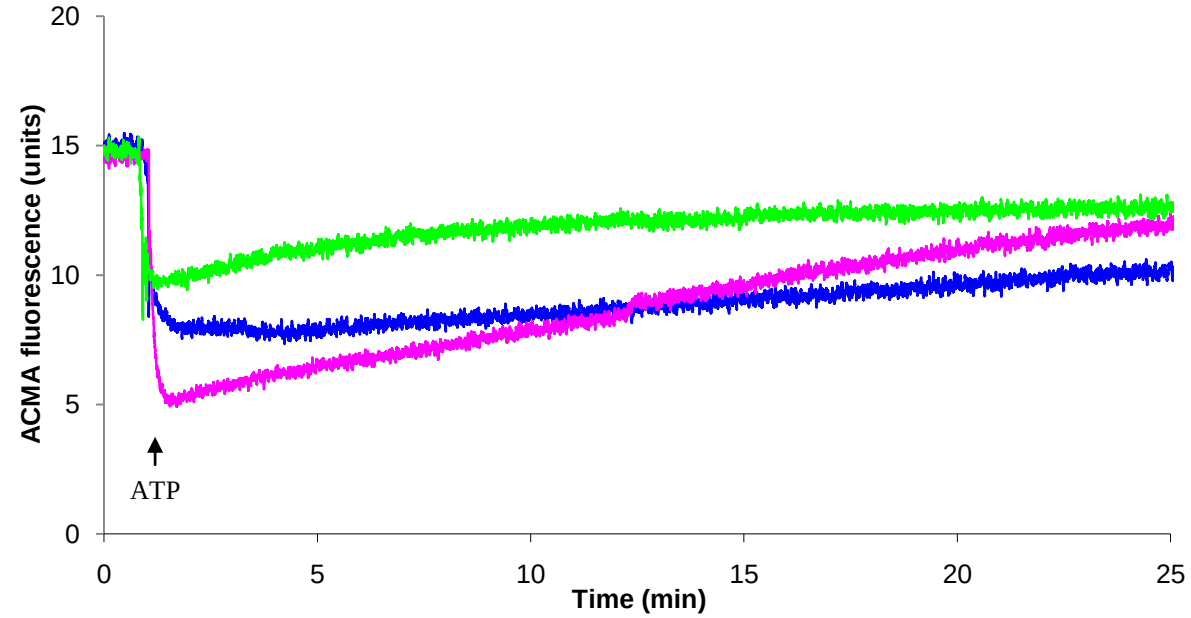


Figure 3- 26: Effect of Cl^- on *in vitro* Tat transport efficiency

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. (A) Cl^- ions (50mM) in the original Musser translocation buffer was replaced with Ac^- (50mM) or SO_4^{2-} (25mM). (B) Transport assay was carried out in increasing amounts of KCl or KAc in the translocation buffer.

I tested if the inhibitory effect of Cl^- ions was due to an effect on the PMF. To do this, I carried out energization tests using ACMA and oxonol V, as described previously. Figure 3-27 shows that the presence of Cl^-, Ac^- and SO_4^{2-} did not prevent generation of PMF. Since the magnitude of quenching of ACMA and oxonol V dyes have been shown to also be influenced by experimental conditions including the presence of different ions (Rottenberg 1989), it is not possible from this method to directly correlate the magnitude of quenching of the dyes with the magnitude of PMF the generated.

A)



B)

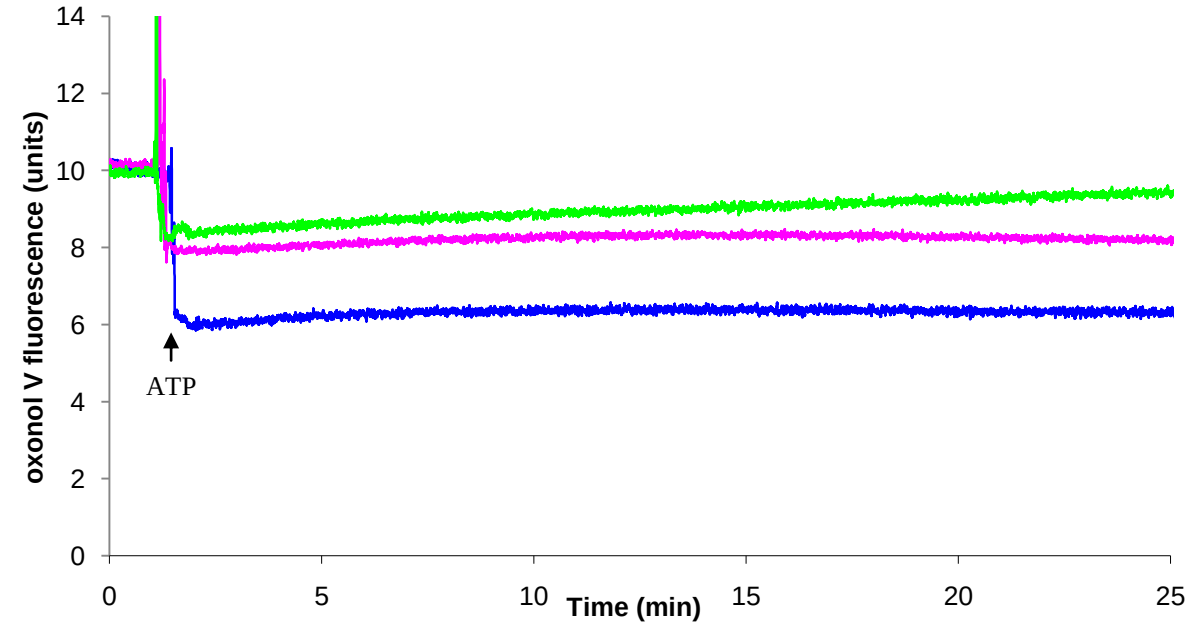


Figure 3- 27: Effect of Cl^-, Ac^- and SO_4^{2-} on generation of PMF by IMVs

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were incubated at 37°C in 100µl translocation buffer (TB) before addition of fluorescent dyes and the absorbance was followed. When a stable fluorescent baseline was obtained, 4mM of ATP was added. (A) ACMA dye (0.1µM) was used to monitor ΔpH across membranes (Excitation λ: 419nm, Emission λ: 483nm). (B) Oxonol V dye (0.2µM) was used to monitor ΔΨ (Excitation λ: 625nm, Emission λ: 645nm). The experiment was done in triplicates and the average plotted. **Blue**: TB containing acetate; **pink**: TB containing Cl^-; green: TB containing SO_4^{2-} .

Although the ACMA and oxonol V data indicated that PMF was generated across IMVs in buffers containing Cl^- ions, it was not possible to determine the direction of the ΔpH or ΔΨ generated from these experiments. There may be Cl^- transporters or channels in the IMVs that can facilitate movement of Cl^- ions into the membrane under high Cl^- concentrations in translocation buffer. This could potentially result in the reversal of direction of ΔpH and ΔΨ generated. Such Cl^- transporters include the Cl^-/H^+ antiporter, ClC-ec1, which exchanges 2 Cl^- for 1 H^+ (Accardi and Miller 2004). A high concentration of Cl^- outside IMVs would result in transport of two Cl^- ions into the IMVs via ClC-ec1 in exchange for one H^+ (Figure 3- 28). This could explain why ΔpH cannot be maintained and the direction of the ΔpH might be reversed in the presence of Cl^-. Theoretically, the import of two negative Cl^- coupled with export of one positive H^+ would render the lumen of the IMVs more negative. Therefore, the influx of Cl^- meant that ΔΨ might also potentially be reversed. As a result, the PMF generated in the presence of Cl^- could be in the opposite direction (*i.e.* low $[H^+]$; negative inside) to when H^+ is transported into the lumen of the IMVs by the ATP synthase when energized by ATP (*i.e.* high $[H^+]$; positive inside). Since the IMVs also have other ion channels and/or transporters in the inner membrane that affect the membrane permeability towards different ions, the contributing factors of these proteins towards the PMF should not be discounted.

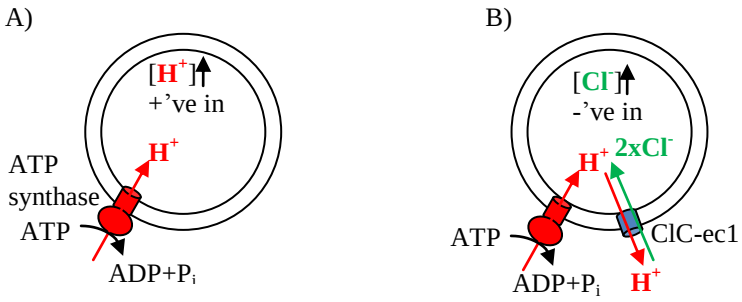


Figure 3- 28: Illustration of the possible effects of ClC-ec1 transporters on ΔpH and ΔΨ of IMVs
During *in vitro* Tat transport assays, IMVs were energized with ATP. Protons (H^+) would be pumped across the membrane into the lumen by ATP synthase. (A) Hydrolysis of ATP by ATP synthase increases the H^+ concentration in the lumen of the IMVs. This would also result in the lumen being more positive than the exterior buffer. (B) If ClC-ec1 is active when there is Cl^- in the reaction mixture, Cl^- will be transported into the lumen in exchange for H^+. If the concentration of Cl^- in the reaction mixture is sufficiently high, this could result in reversal of direction of PMF generated.

In order to determine the role of ClC-ec1 in Tat transport, I tried to carry out a transport assay using a ClC-ec1 mutant, MG1655 Δ2, where the genes *eriC* and *mriT*, coding for ClC-ec1 chloride channel homologues were deleted. However, this mutant failed to overexpress *tatABC* (Figure 3- 29). Hence, it could not be tested whether ClC-ec1 has a contributing factor to the inhibitory effect of Cl^- on *in vitro* Tat transport.

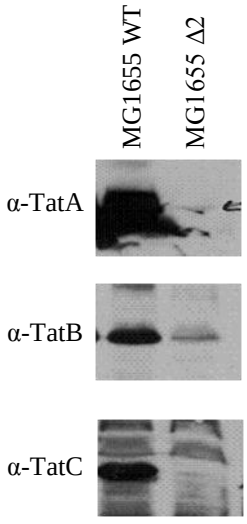


Figure 3- 29: Immunoblot analysis of overproduction of TatABC in IMVs made from ΔClC mutants
IMVs were prepared from *E. coli* strains MG1655-pREP4 and MG1655 Δ2-pREP4 overproducing TatABC from plasmid pUniTat. Expression of Tat proteins was assessed by immunoblotting. Each lane contains 20μg of total protein. Proteins were detected using the indicated anti-sera (1° antibodies: 1/7,000; 2° antibody α-rabbit: 1/5,000). MG1655 Δ2: Cl^- channels mutant

Since Tat transport was affected by the presence of Cl^- ions, Cl^- in the original translocation buffer (Cl^--TB) was substituted with Ac^- (Ac^--TB) to improve transport efficiency.

3.5.3 Optimization of the IMV to substrate ratio

Next, I tested the effect of the IMV to substrate ratio on Tat transport efficiency. To do this, I increased the amount of IMVs added to the *in vitro* assay. It was observed that the amount of mature protein at the end of the assay decreased as the amount of IMVs added was increased. By contrast, the band corresponding to the full length protein was more intense (Figure 3- 30A, lanes 1-3 and B, lanes 6-8). However, it has to be noted that the IMVs were stored in 50% v/v glycerol. This means that by adding more IMVs, more glycerol would inevitable be added into the reaction mix. Therefore, the effect that was observed could be due to the additional glycerol. To test this possibility, I added glycerol back into the reaction while keeping the amount of IMVs in the assay the same. The assay with the standard amount of IMVs (*i.e.* A_{280nm}=10) had 5% v/v glycerol.

Figure 3- 30 showed that addition of glycerol had the same effect as adding more IMVs to give the same final concentration of glycerol in the assay. This suggests that the effects observed when increased amounts of IMVs were added to the assay were due to glycerol. Control experiments were done to determine which bands reflect true transport events. As shown in lane 3 of Figure 3- 30B, addition of glycerol partially protected the full length CueO from PK digestion. This is supported by the observation of a small amount of full length CueO when FCCP was added (lane 5). Addition of Triton X-100 (lane 9) resulted in complete digestion of the mature band but the full length band was still observed in the presence of glycerol. This means that only the mature band represented transported CueO. As the intensity of the mature band decreased with increasing amounts of glycerol, these results suggest that addition of glycerol inhibits Tat transport.

A)

Lane	1	2	3	4	5	6
IMV A_{280nm}	10	20	40	10	10	10
Final glycerol (v/v)	5%	10%	20%	10%	20%	5%
Triton X-100	-	-	-	-	-	+



B)

Lane	1	2	3	4	5	6	7	8	9
Final glycerol (v/v)	5%	5%	25%	5%	25%	5%	25%	45%	25%
ATP	-	-	-	+	+	+	+	+	+
FCCP	-	-	-	+	+	-	-	-	-
Triton X-100	-	-	-	-	-	-	-	-	+
PK	-	+	+	+	+	+	+	+	+



Figure 3- 30: Optimization of IMVs to CueO ratio

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs were added to the translocation buffer containing 4mM ATP, 100nM CueO and glycerol as indicated. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added at the end of the assay if required. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. (A) Increasing amounts of IMVs were added in lanes 1-3 (A_{280nm} as indicated above the gel). The amount of glycerol added in control experiments (lanes 4 and 5) mimic the extra amount of glycerol that is inevitably added when increasing amounts of IMVs are used (i.e. 5% extra glycerol is added when IMVs A_{280nm} =20 and 15% extra glycerol is added when IMVs A_{280nm} =40). (B) This assay was done using only the standardized amount of IMVs (A_{280nm}=10) in each lane. This amount of IMVs would introduce 5% glycerol into the assay. Additional glycerol was added in lanes 3, 5, 7, 8 and 9.

From these experimental setups, it is not possible to determine if increasing the amount of IMVs in the transport assay would improve transport efficiency as the amount of glycerol that would be inevitably added affected the results. Given the problematic effect of glycerol, caution has to be taken when considering whether to include both the full length and mature CueO in assessing the transport efficiency. To account for this, control experiments in the absence of PMF or in the presence of Triton X-100 had to be done concurrently while doing the experiment. Since the results of this experiment to determine the optimum IMV: substrate ratio is inconclusive, the IMVs: substrate ratio used for following experiments is maintained at A_{280nm} =10 (IMVs): 100nM (CueO).

3.5.4 Improving the ratio of mature to full length substrate observed after transport

Full length Tat substrates that have been transported by the Tat translocase will have their Tat signal peptides cleaved by LepB, a membrane bound signal peptidase (Yahr and Wickner 2001; Lüke *et. al.* 2009). However, transported full length substrate was always observed together with the mature substrate in *in vitro* Tat transport assays (Figure 3- 9). One possible reason that the signal peptide had not been cleaved was that there was insufficient LepB in some IMVs to cleave the Tat signal peptide.

To address this issue, IMVs were isolated from cells that co-overexpressed *lepB* with *tatABC*. Figure 3- 31A showed that IMVs enriched with both LepB and TatABC resulted in less full length CueO but transport efficiencies were lower. This lower transport efficiency could be because production of TatABC was compromised when LepB was overproduced (Figure 3- 31B). One reason less full length CueO was observed could be because the presence of overproduced LepB was able to deal with this reduced amount of transported CueO and cleave the signal peptide of most CueO transported. Since the transport efficiency of IMVs made from cells overexpressing *lepB* and *tatABC* was low, IMVs for this thesis were made from cells overexpressing *tatABC* only.

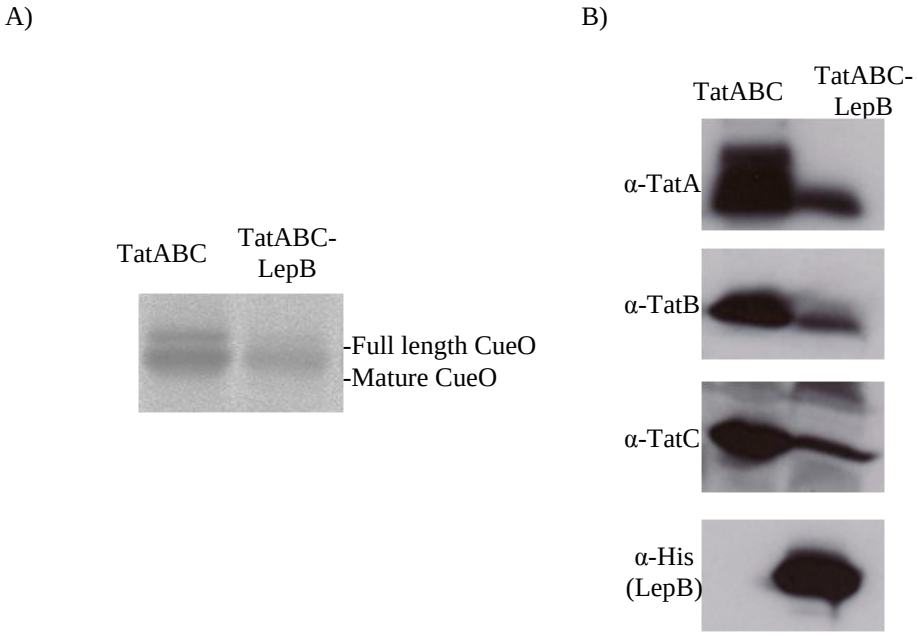


Figure 3- 31: Effect of overexpression of *lepB* on *in vitro* Tat transport efficiency
IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overexpressing *tatABC* from plasmid p8737 (TatABC) or from cells co-overexpressing *tatABC* from plasmid p8738 and *lepB* from plasmid pET28a-LepB (TatABC-LepB). (A) *In vitro* transport assay to compare transport efficiency between IMVs overproducing either TatABC only or both TatABC and LepB. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Analysis was based on in gel-fluorescence scan of fluorescent dye labelled CueO (CueO_GS3^{633}) (Excitation λ: 633nm, Emission λ: 675nm) (see 3.6.2). (B) Overexpression blots of Tat components. Proteins were blotted against the indicated antisera (1° antibodies: 1/7,000; 2° antibody (α-rabbit): 1/5,000) and LepB (α-His: 1/5,000).

3.5.5 Effect of temperature

The original assay was done at 37°C. I carried out the *in vitro* transport assay at different temperatures to determine if 37°C is the optimum temperature for transport assays. No transport was observed when the assay was done at room temperature (19°C) (Figure 3-32). Less transport was observed at 30°C than at 37°C while the transport efficiency at 42°C was comparable to the transport efficiency of the assay that was done at 37°C. Therefore, the temperature in which transport assays were carried out was maintained at 37°C.

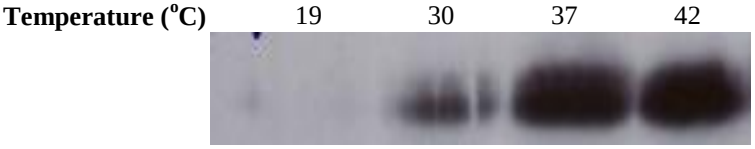


Figure 3- 32: Effect of temperature on *in vitro* Tat transport assay

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. The *in vitro* Tat transport assay was carried out at different temperatures. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at indicated temperatures for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

3.5.6 Final comparison between original and modified Musser *in vitro* Tat transport assays

Comparison of the *in vitro* Tat transport assay between the original Musser protocol and the modified protocol showed that the modified protocol has a higher Tat transport efficiency (Figure 3- 33).

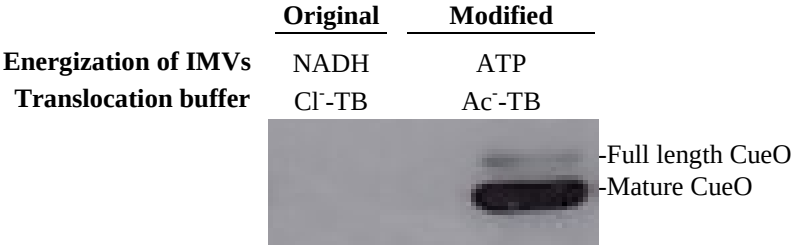


Figure 3- 33: Final comparison between original and modified Musser *in vitro* Tat transport assays

IMVs were all prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP or NADH and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection of the C-terminal His-tagged CueO was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution. Original Musser assay was done using NADH energized IMVs in Cl^--containing translocation buffer (Cl^--TB) (Table 3- 2) while the modified Musser assay was done using ATP energized IMVs in Ac^--containing translocation buffer (Ac^--TB). Note that the modified protocol was used to test IMVs in section 3.4.

The main differences between the protocols are:

- The exchange of Cl^- ions for Ac^- ions in the translocation buffer
- Energization of IMVs by ATP instead of NADH

The collective effect of these changes resulted in a higher transport efficiency.

3.6 Alternative *in vitro* Tat transport detection methods

Detection of transport in my *in vitro* Tat transport assay was based on a one-step immunoblot protocol. I explored two alternative detection methods: two-step immunoblotting and fluorescence detection.

3.6.1 Two-step immunoblotting using α-CueO

I had previously used a commercially available monoclonal α-His antibody conjugated to horse radish peroxidase (HRP) for detection of transport of CueO bearing a C-terminal His-tag. The sensitivity of detection can be improved by using a two-step immunoblotting protocol as the signal for detection can be amplified by the polyclonal secondary (2°) antibody. In addition, the α-CueO primary (1°) antibody could also be polyclonal and this would enable detection of multiple epitopes on CueO which allows binding of several 1° antibody per CueO (Figure 3- 34).

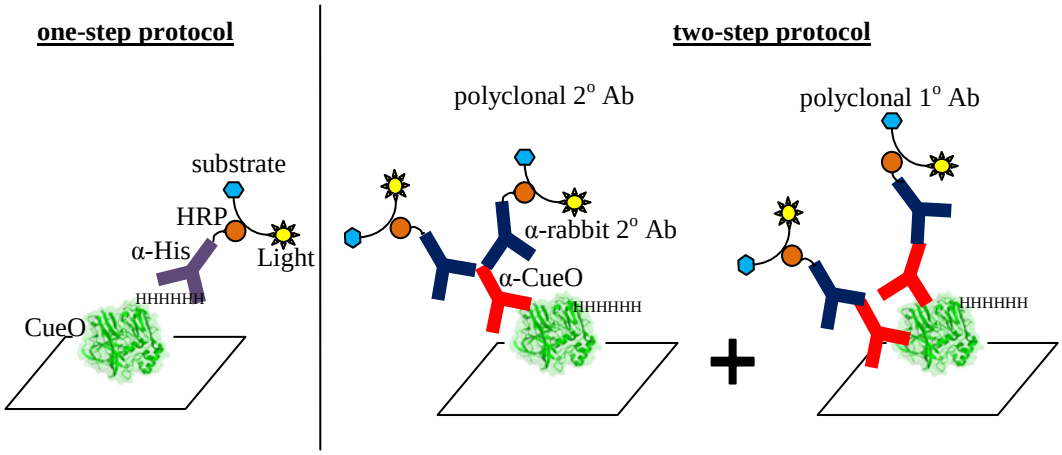


Figure 3- 34: One-step vs two-step immunoblot
In a one-step immunoblot protocol, $CueO_{His}$ is recognized by a monoclonal α-His 1° antibody (Ab) that is conjugated with horse radish peroxidase (HRP). HRP then reacts with a chromogenic substrate to produce a signal. In the two-step protocol, $CueO_{His}$ is recognized by the polyclonal α-CueO 1° Ab which was raised in rabbit. This is followed by the polyclonal α-rabbit 2° Ab that is conjugated to HRP. Since more than one α-rabbit 2° Ab can bind onto the α-CueO 1° Ab, the signal for detection is amplified. In addition, the α-CueO 1° Ab can also recognize several epitopes on CueO.

To make 1° antibodies, I purified the mature domain of $CueO_{His}$ from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 containing the plasmid pSC17. $CueO_{His}$ was purified by Ni^{2+} affinity chromatography followed by gel filtration chromatography. N-terminal sequencing (data not shown) verified that the mature domain of $CueO_{His}$ was purified. Purified mature $CueO_{His}$ was sent to Davids Biotechnology where the protein was injected into rabbits for α-CueO antibody production. The polyclonal α-CueO obtained was purified using a HiTrap NHS-activated HP column that had been coupled to the mature domain of $CueO_{His}$.

Using the purified α-CueO 1° antibody, I was able to detect transported CueO by using a two-step immunoblotting protocol. Figure 3- 35 showed that several bands that were absent when probed with α-His antibodies were detected by the α-CueO antibody. For example, the band running at ~40kDa (Figure 3- 35B, lane: -PK) in the absence of PK could be a degradation product of CueO. As mentioned in section 3.2.1, full length CueO is prone to degradation. Therefore, using α-CueO to probe for CueO would provide an internal control of the degree of purity of full length CueO that has been added to the assay.

The two extra bands (~36kDa and ~29kDa) observed by α-CueO after PK digestion (Figure 3- 35B, lane: +PK) could be the products of CueO digested by PK. These fragments were not detected by the α-His antibody, suggesting that they do not contain the His-tag. The molecular weights of these two bands add up to approximately the molecular weight of CueO. From these observations, it is likely that CueO was cleaved by PK into three fragments; a 36kDa fragment, a 29kDa fragment and a C-terminal fragment with the His-tag. The C-terminal fragment was not detected because it was very small fragment and would have ran off the gel.

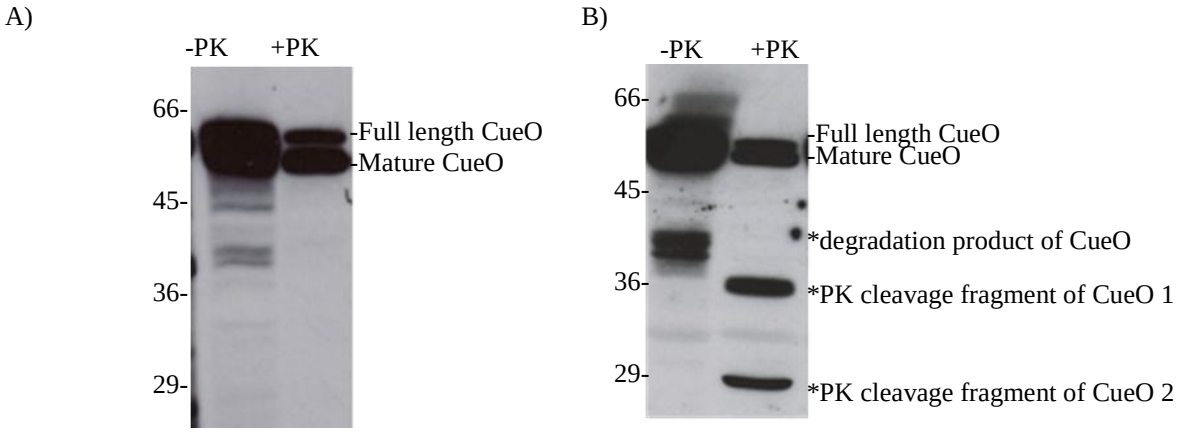


Figure 3- 35: Detection of transport by α-CueO

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. Detection was done using (A) one-step immunoblotting of the C-terminal His-tag (α-His antibody: 1/5,000 dilution) or (B) two-step immunoblot (α-CueO as primary antibody (1/2,000), followed by α-rabbit (1/5,000) as secondary antibody). * shows additional bands observed when the polyclonal α-CueO was used as the 1° antibody. Descriptions of the additional bands are provided on the right of the gel.

3.6.2 Detection of transport by fluorescence

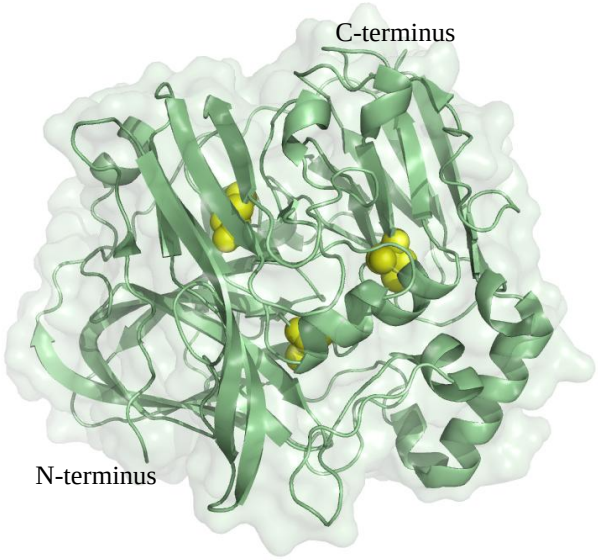
Western blot detection of substrates transported by the *in vitro* Tat transport assay only provides semi-quantitative analysis of the Tat transport. To quantitatively asses Tat transport, a fluorescence based method could be used. A fluorescence-based assay could also be used for real time experiments and single molecule fluorescence studies in the future. In order to use a fluorescence-based assay, I had to construct a fluorescently labelled Tat substrate that can be transported by the *in vitro* Tat system.

Proteins can be labelled with fluorescent probes via free amines, which are found on the N-terminal α-amino groups or the ε-amino groups of lysine residues, or via free sulfhydryls found on cysteines. I chose to label substrates using thiol-reactive fluorescent dyes, which target sulphydryls. Cys are rarely found on the surface of proteins. Therefore, the number of fluorophores bound per protein is low.

CueO was chosen as the candidate substrate because it was transported by the *in vitro* Tat system, as demonstrated in previous sections. However, CueO does not have any surface exposed sulphydryl groups that are accessible for labelling. All three native Cys are buried within the core of CueO (Figure 3- 36A). Hence, I had to engineer a modified CueO with an exposed sulphydryl group for labelling that is still transport competent in *in vitro* Tat transport assays.

I made several attempts to add an exposed sulphydryl group onto CueO. These included adding a Cys residue directly after the last residue on the C-terminus or adding a Cys that was separated from the C-terminus by linkers of various lengths. After several trials, I found a modified construct of CueO that was Tat transport amenable. This newly modified CueO had a 15 amino acid linker at the C-terminus followed by a cysteine residue. This linker consisted of three consecutive repeats of the peptide sequence GGGGS (GS3) and had been previously used to link a Tat signal peptide with the mature domain of a recombinant Tat substrate in the thylakoid system (Cline and Mccaffery 2007). I will term this modified substrate CueO_GS3 (Figure 3- 36B).

A)



B)

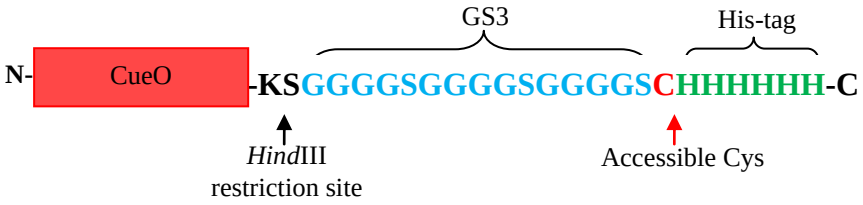


Figure 3- 36: CueO and construction of CueO_GS3
Illustrations showing (A) crystal structure of CueO with the three native cysteine residues in yellow (PDBID 1KV7). The N-terminus and C-terminus are labelled. (B) CueO_GS3 construct.

CueO_GS3 was shown to be Tat transport amenable (Figure 3- 37). CueO_GS3 can be labelled with the thiol-reactive dye, Alexa Fluor 633 C_5 maleimide (CueO_GS3^{633}). CueO_GS3^{633} was transported in *in vitro* Tat transport assays as detected by immunoblotting as well as by in-gel fluorescence scan of the gel on a fluorescence image scanner (Excitation λ: 633nm, Emission λ: 675nm).

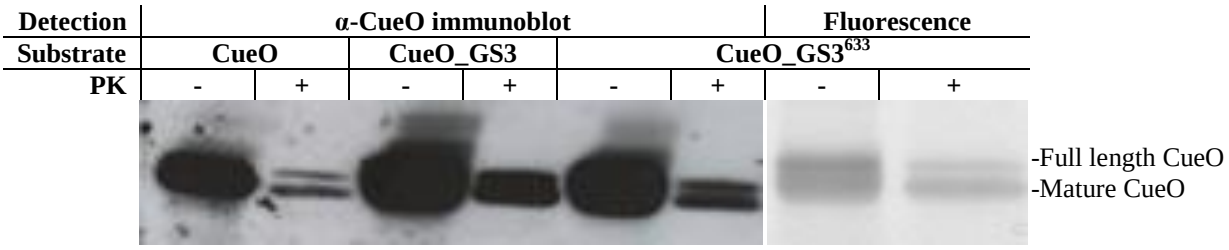


Figure 3- 37: Transport competency of CueO_GS3^{633}

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. Transport assay of CueO, CueO_GS3 and CueO_GS3^{633} were done under similar conditions. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM CueO. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. α-CueO immunoblots were probed with α-CueO 1° Ab (1/2,000) followed by α-rabbit 2° Ab (1/5,000). Fluorescence of CueO_GS3^{633} was detected by in gel fluorescence scanning (Excitation λ: 633nm, Emission λ: 675nm).

Now that a fluorescence-based *in vitro* Tat transport assay has been established, the next step was to ensure that the fluorescence detected had a linear relationship with the amount of fluorescently labelled substrates added. Quantitative analyses of in gel fluorescence scans were carried out using the ImageReader program, where each band was represented by a peak. The area under the peak was plotted against the known concentration of CueO_GS3^{633}. As can be seen from Figure 3- 38, the relationship between the area under the peak vs concentration of CueO_GS3^{633} was linear.

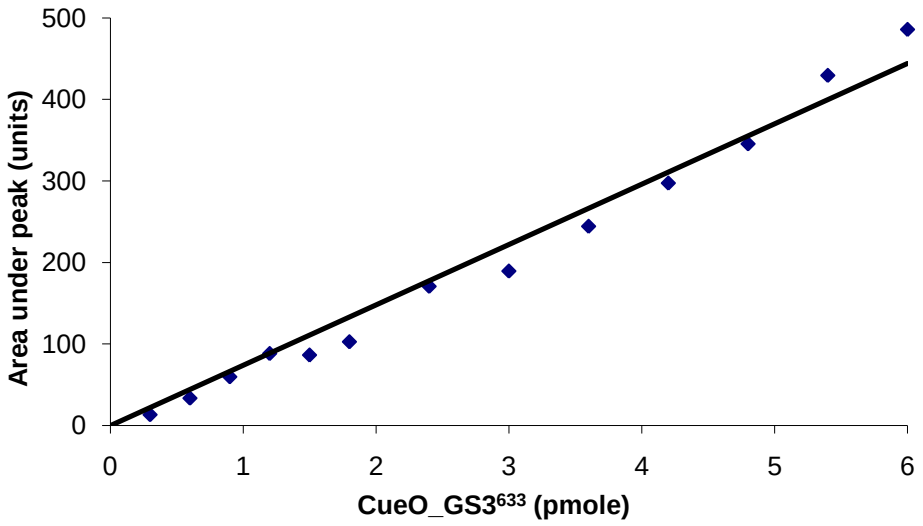


Figure 3- 38: Representative standard curve used for quantitative analyses of fluorescence-based *in vitro* Tat transport assays

A serial dilution of CueO_GS3^{633} was loaded onto a 12.5% SDS-PAGE gel. The gel was scanned on the Fuji FLA-3000 (Excitation λ: 633nm, Emission λ: 675nm) set at resolution of 50 and sensitivity of F1000. Quantitative analyses were carried out using the corresponding ImageReader program, where each band was represented by a peak. The area under the peak was plotted against the known concentration of CueO_GS3^{633}.

3.6.3 Dependency of Tat transport efficiency on substrate concentration

Transport efficiencies of *in vitro* Tat transport assays are calculated based on the percentage of substrate transported out of the total amount of substrate added. To test if the transport efficiency varies according to the amount of substrate added, I added different amounts of substrates to the *in vitro* Tat transport assay. Transport was detected even when only 0.2pmole of CueO_GS3^{633} was added to the assay (Figure 3- 39A, top). It was observed that when increasing amounts of CueO_GS3^{633} was added into the assay, the actual amount of CueO_GS3^{633} transported increased (Figure 3- 39B, **red**). However, the percentage of CueO_GS3^{633} transported (i.e. transport efficiency) decreased when more CueO_GS3^{633} was added (Figure 3- 39C, **red**). This result showed that the transport efficiency was dependent on the amount of substrate added into the *in vitro* transport assay. This means that transport

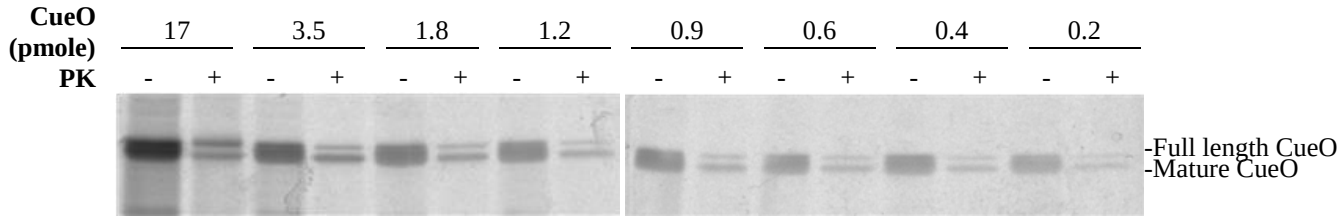
efficiencies can only be compared when equal amounts of substrate are used for *in vitro* Tat transport assays.

To determine if my modified *in vitro* transport assay (ATP energized IMVs in Ac^- TB) was better than the original Musser transport assay (NADH energized IMVs in Cl^--TB), I repeated the experiment using the original Musser translocation assay. As can be seen from Figure 3- 39, the Tat transport efficiency when using the original Musser transport assay (A bottom, B and C: blue) was lower when compared to the modified transport assay (A top, B and C: red). In the original Musser transport assay, transport could not be detected when the amount of CueO_GS3^{633} added was below 1.1pmole (Figure 3- 39, A bottom). This means that either CueO_GS3^{633} was not transported at such low concentration or the amount of CueO_GS3^{633} transported was below the limit of detection. On the other hand, transport was still detected when less than 1.1pmole of CueO was added for the modified protocol (Figure 3- 39, A top). The transport efficiency of the original Musser transport assay, as with the modified Musser assay, depends on the concentration of substrates added to the assay (Figure 3- 39, **blue**).

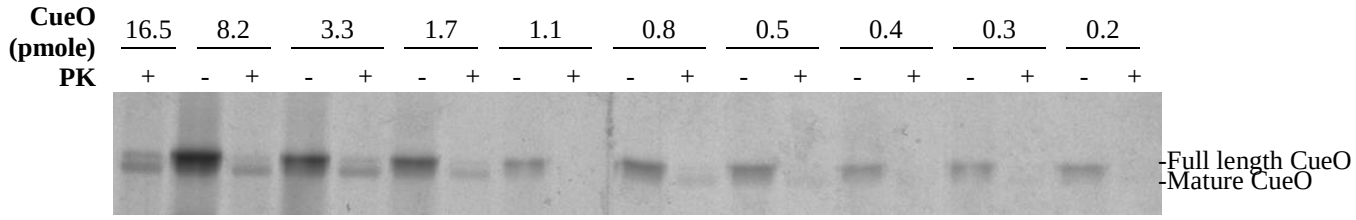
Another observation that should be brought to attention is that transport was observed for the original Musser transport assay when 3.3pmole of CueO_GS3^{633} was added to the reaction. Immunoblotting was unable to detect transport when this amount of CueO_GS3^{633} was used (Figure 3- 33). Therefore, fluorescence detection is more sensitive than immunoblot detection.

A)

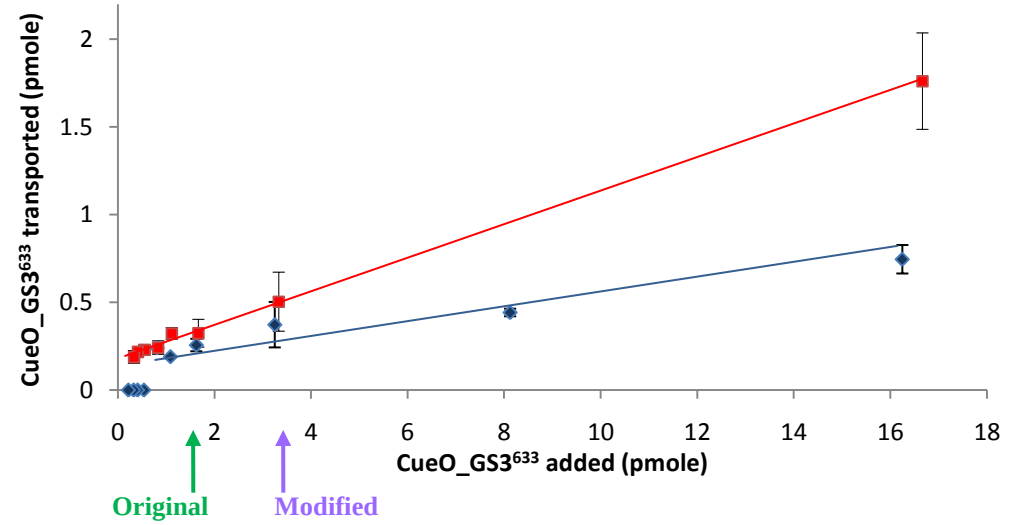
Modified Musser transport assay protocol



Original Musser transport assay protocol



B)



C)

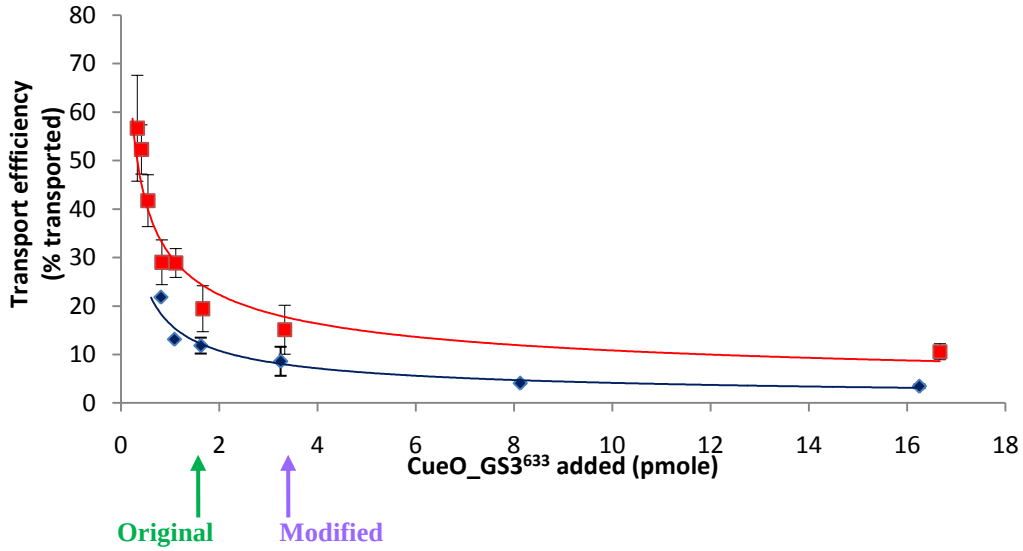


Figure 3- 39 (next page): Dependency of Tat transport efficiency on substrate concentration

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737 according to the modified Musser IMV isolation protocol. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and different amounts of CueO_GS3^{633}. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO that remained in the reaction mixture. (A) In-gel fluorescence scan of CueO_GS3^{633} after transport (Excitation λ: 633nm, Emission λ: 675nm). Top scan corresponds to the *in vitro* Tat transport assay following the modified Musser translocation protocol (ATP energized IMVs in acetate containing translocation buffer). Bottom scan corresponds to the *in vitro* Tat transport assay following the original Musser translocation protocol (NADH energized IMVs in Cl^- containing translocation buffer). Transport of CueO is graphically represented in (B) as pmole of CueO transported and (C) as transport efficiency (i.e. % CueO transported out of total CueO added to the assay). Experiments were done in triplicates and the average plotted. Error bars show standard error. **Red**: modified protocol; **blue**: Musser's original protocol. **Green** arrow indicates the amount of substrate that was used for the original Musser protocol. **Purple** arrow indicates the amount of substrate that was used for the modified Musser protocol. The plotted lines and curves are indicative rather than a fit to the data.

Quantitative analyses of *in vitro* Tat transport showed that the transport efficiency of my modified protocol was higher than that of the original Musser protocol. However, I have yet to achieve the 70% transport efficiency reported for the original Musser protocol. The purple arrows in Figure 3- 39B and C show the amount of substrate (3.3pmole) used in my modified assay while the green arrows show the amount of substrate (1.8pmole) used in the original Musser protocol. From the graph, it can be inferred that when 1.8pmole of substrate was added, less than 20% of the substrate was transported when my modified protocol was used but only ~12% was transported when the original Musser protocol was used. This is far below the 70% that was reported. This is likely due to the different substrate used in my experiments. SufI was used in the original Musser experiments while CueO was used in my experiments as I was unable to purify full length SufI. The dependency of *in vitro* Tat transport efficiency on the choice of substrate used had been observed before. Bageshwar *et. al.* (2007) have used several native or non-native substrates in the original Musser transport assays, but only the native substrate, SufI, gave a 70% transport efficiency while the other substrates were transported at much lower efficiency (Bageshwar and Musser 2007). Nonetheless, when my modified assay was carried out in parallel with the original Musser IMV protocol using CueO as the substrate, my modified assay gave a higher transport efficiency.

3.7 *In vitro* Tat transport of other Tat substrates

There are at least 30 native *E. coli* substrates (Palmer *et. al.* 2010) but I had only focused on one of the substrates, CueO. Although CueO can be transported *in vitro*, it is prone to proteolytic degradation upon storage (section 3.2.2). Therefore, I attempted to purify an alternative Tat substrate which is better-behaved than CueO.

My modified *in vitro* Tat transport is only applicable for Tat substrates that can be purified in biochemical amounts and remain folded. The purified Tat substrates must have an

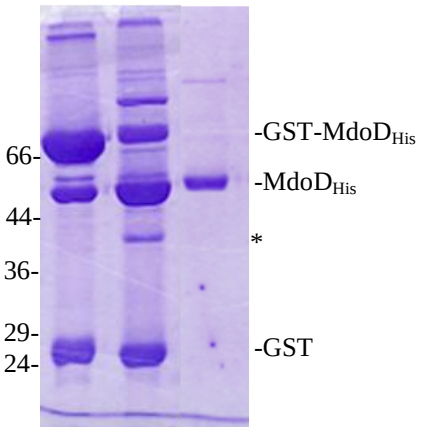
intact N-terminal Tat signal peptide for recognition by the Tat translocase. In addition, my assay can only assess transport if the protein is digestible by proteases that will be added to the translocation assay to distinguish between transported and untransported substrates.

3.7.1 Transport of MdoD

One of the Tat substrates tested as an alternative model substrate was MdoD. MdoD is involved in the synthesis of osmoregulated periplasmic glucans (OPGs) and is most likely a regulatory subunit of the OPG biosynthetic machinery (Lequette *et. al.* 2004).

MdoD was purified in a similar manner to CueO. It was expressed as a N-terminal GST-fusion protein with a C-terminal His-tag for immunoblot detection. The N-terminal GST domain is separated from MdoD by a 3C cleavage site. The N-terminal GST domain protects the Tat signal peptide from being cleaved during protein purification and ensures that full length MdoD is purified. The GST domain can then be cleaved off by addition of GST-3C protease (Figure 3- 40A). The GST cleaved full length MdoD was further purified by a second GSTrap affinity chromatography step. SDS-PAGE analysis showed a major band between 44kDa and 66kDa. Analysis of this protein by mass spectrometry showed that full length MdoD was purified (MW=64240kDa) (Figure 3- 40). The second highest peak (MW=64352kDa) in the MS spectrum corresponds to a trifluoroacetic acid (TFA) adduct of full length MdoD.

A)



B)

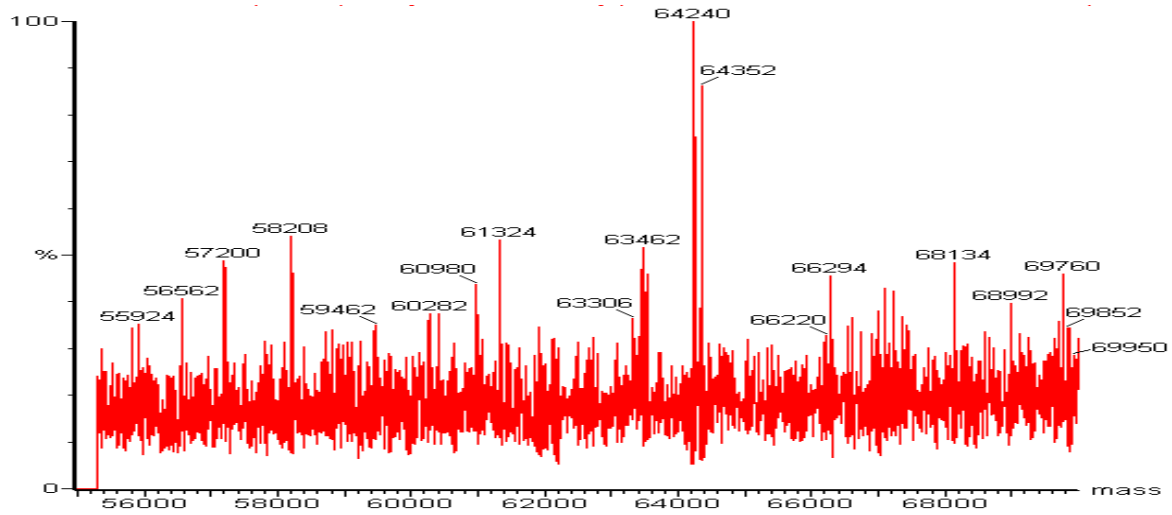


Figure 3- 40: Characterization of MdoD

(A) A cleared cell lysate was prepared from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 overproducing GST-MdoD$_{His}$ from the plasmid pGEX- MdoD$_{His}$. The lysate was applied to a GSTrap affinity column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5). Bound GST-MdoD$_{His}$ was eluted with 10mM glutathione. Glutathione was removed by ultrafiltration using a 30kDa molecular weight cut off (MWCO) membrane. GST was cleaved from GST-MdoD$_{His}$ by GST-3C protease at 4°C overnight. The cleaved sample was applied to a GSTrap affinity column equilibrated in resuspension buffer. MdoD$_{His}$ will elute in the flow through while free GST and GST-3C protease remained on the column. Fractions collected from GSTrap affinity columns were analysed by 12.5% SDS-PAGE and proteins visualised by Coomassie Brilliant Blue staining. Lane1: Eluate from first GSTrap affinity column purification; Lane2: After 3C protease cleavage but before the second GSTrap affinity column purification; Lane 3: Flow through of second GSTrap affinity column purification. The molecular masses (kDa) of standard proteins are given to the left of the gel. MdoD$_{His}$ degradation product is indicated by an asterisk (*). (B) MdoD in 0.1% TFA was analyzed by ESI-MS to determine its molecular weight.

MdoD is suitable for *in vitro* assays as it can be digested by PK (Figure 3- 41, lane 2). Similar to CueO, it was transported by the Tat system *in vitro* (Figure 3- 41, lane 4). A small amount of protease resistant MdoD was observed for non-energized IMVs (Figure 3- 41, lane 3) and de-energized IMVs (Figure 3- 41, lane 5). This most probably resulted from non-specific binding to the membranes since these bands disappeared when the IMVs were treated with Triton X-100 (Figure 3- 41, lane 6).

Lane	1	2	3	4	5	6
IMV	-	-	+	+	+	+
ATP	-	-	-	+	+	+
FCCP	-	-	-	-	+	-
Triton X-100	-	-	-	-	-	+
PK	-	+	+	+	+	+



Figure 3- 41: *In vitro* Tat transport assay using MdoD
IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM MdoD. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added following transport. PK was added to digest any untranslocated MdoD that remained in the reaction mixture. Detection of the C-terminal His-tagged MdoD was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

3.7.2 Transport of $MdoD_{C17I}$

To be able to perform quantitative analyses on the transport of MdoD using fluorescence-based *in vitro* Tat transport assays, MdoD had to be labelled with a fluorescent dye. Similar to the strategy employed for CueO, Alexa Fluor 633 C_5 maleimide was used to target exposed Cys residues. Unlike CueO, MdoD has a Cys in the signal peptide at position 17 (C17). This Cys had to be removed since the signal peptide would be cleaved off after Tat transport and labelling of the signal peptide with a relatively large and hydrophobic dye could hinder the recognition of MdoD by the Tat translocase. Therefore, the first step was to remove the Cys found in the signal peptide by mutating it to Ile ($MdoD_{C17I}$). I chose to mutate

the Cys to Ile as it has been shown that a Cys → Ile mutation in the signal peptide of SufI resulted in higher transport efficiency compared to wild type SufI (Bageshwar *et. al.* 2009). It can be seen that $MdoD_{C17I}$ was transported by the Tat system *in vitro* (Figure 3- 42). The band seen in lane 2 where IMVs were not energized is most likely due to non-specific binding of $MdoD_{C17I}$ on the surface of the IMVs as this band disappeared when the assay was treated with Triton X-100 (lane 4).

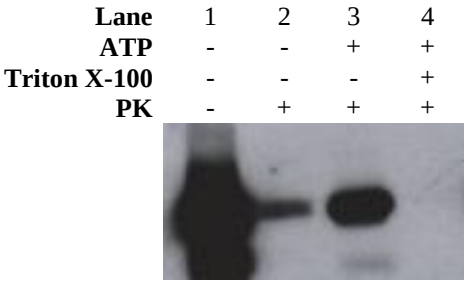


Figure 3- 42: *In vitro* Tat transport assay using $MdoD_{C17I}$
IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM $MdoD_{C17I}$. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added following transport as required. PK was added to digest any untranslocated $MdoD_{C17I}$ that remained in the reaction mixture. Detection of the C-terminal His-tagged $MdoD_{C17I}$ was performed using penta-His HRP conjugate antibody (Qiagen) at a 1:5,000 dilution.

Next, I had to determine if MdoD had any intrinsic exposed Cys residues that are accessible for labelling. Since the structure of MdoD is unavailable, I used Swiss-MODEL to build a homology model of MdoD based on its close structural homologue, MdoG (Bordoli *et. al.* 2008). Only residues 48-538 were modelled for MdoD (i.e. the Tat signal peptide was excluded). This model of MdoD showed that at least 4 (C250, C291, C422, C521) out of the 7 Cys residues in the mature domain are exposed on the surface and might be susceptible to labelling.

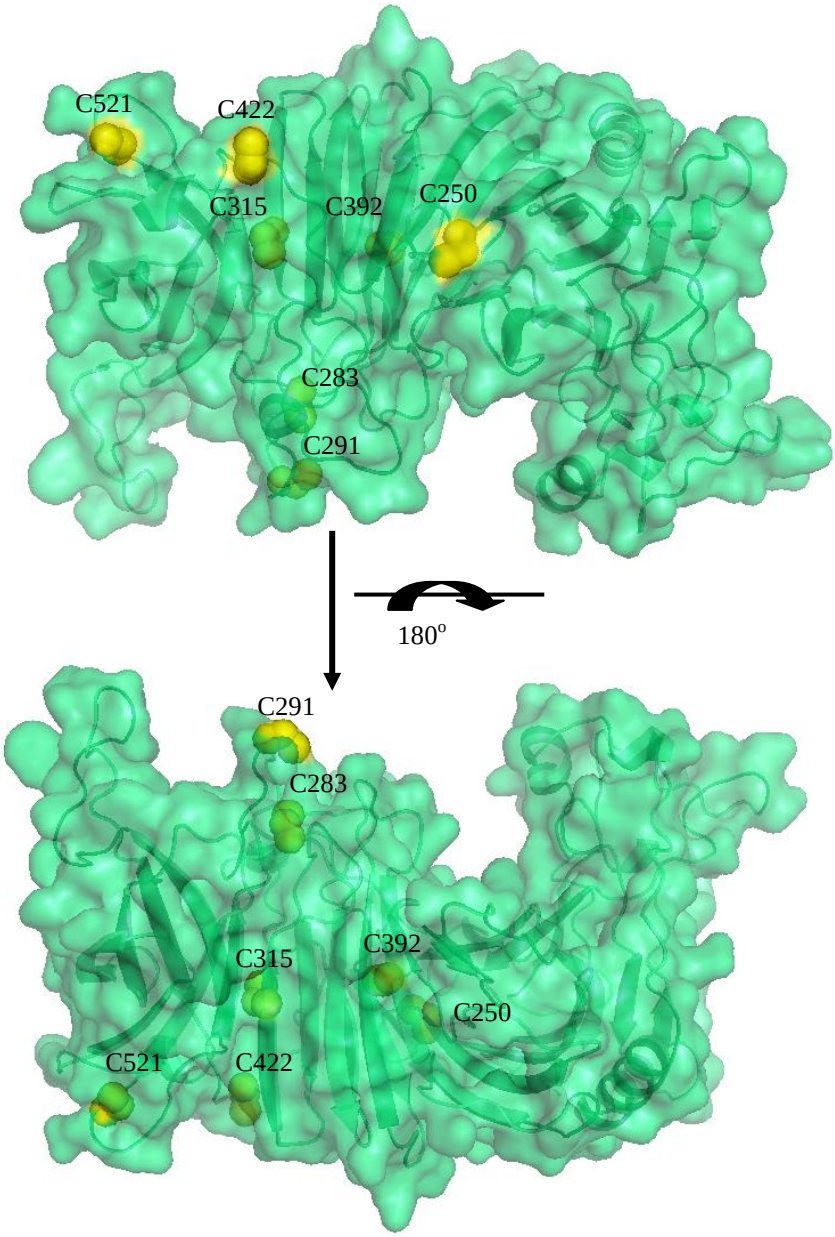


Figure 3- 43: Model of MdoD
Surface representation of MdoD which was modeled using MdoG as template by Swiss-Model (Bordoli *et. al.* 2008). Cys residues are coloured in yellow. According to this model, the four exposed Cys are C250, C291, C422 and C521.

$MdoD_{C17I}$ can be labelled by Alexa Fluor 633 C_5 maleimide (i.e. $MdoD_{C17I}^{633}$) (Figure 3- 44B). However, fluorophore labelling of $MdoD_{C17I}$ abolished Tat transport (Figure 3- 44C and D). This could be a result of multiple labelling of the dye on the four exposed Cys residues. The presence of a high number of hydrophobic dyes on the surface of MdoD might hinder transport as the tolerance of the Tat system against hydrophobic patches is reported to

be relatively low (Richter *et. al.* 2007). A possible solution to this problem of hydrophobic patches on the surface of MdoD as a result of labelling would be to substitute the surface exposed the Cys residues and introduce only one exposed Cys residue at the C-terminus in a construct that is similar to CueO_GS3.

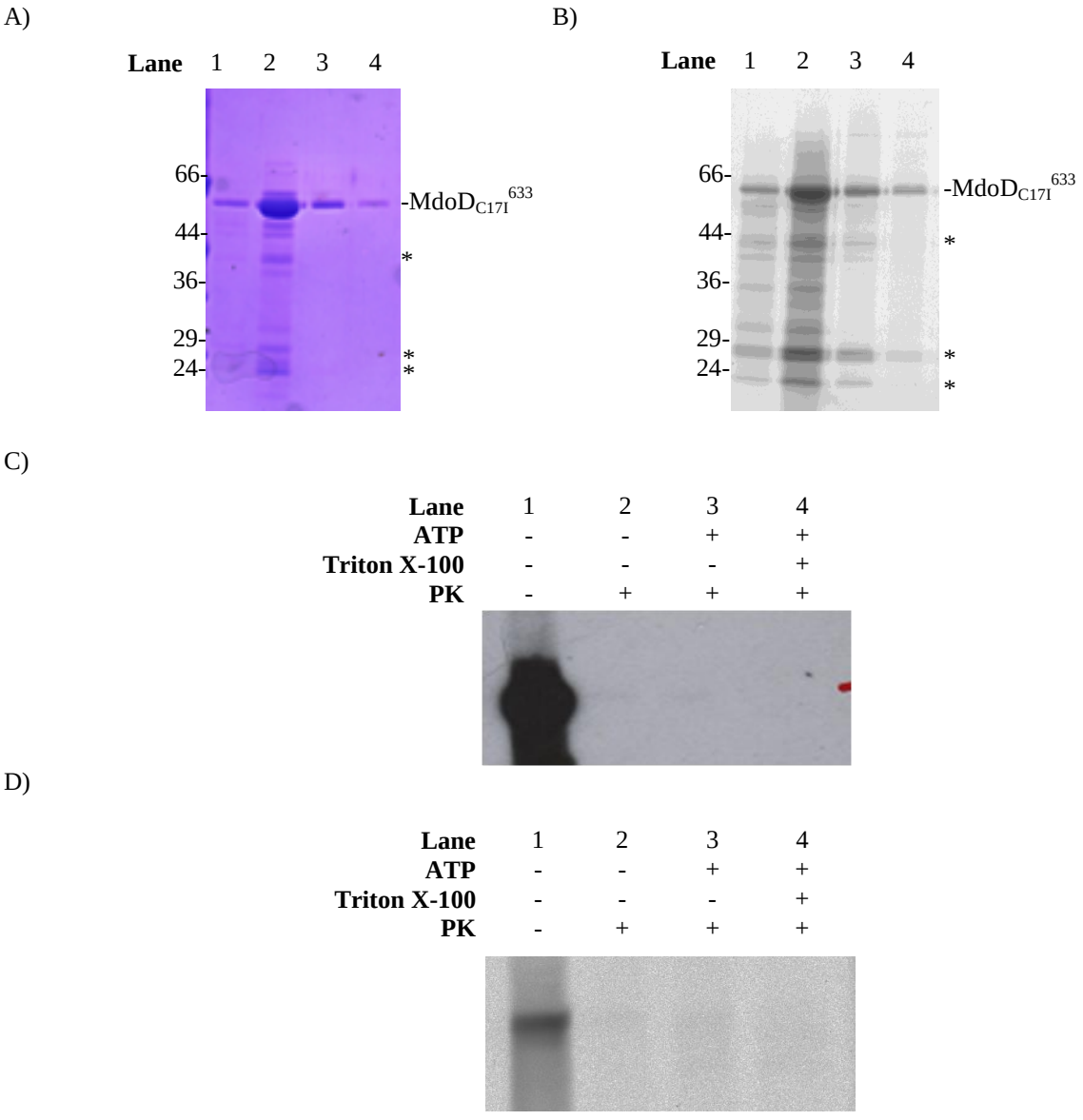


Figure 3- 44: Labelling and transport of MdoD$_{C17I}$

MdoD$_{C17I}$ was labelled with Alexa Fluor 633 C_5 maleimide (1:20 protein:dye ratio) overnight at 4°C and free dye removed by loading the sample on a PD-10 desalting column (GE Healthcare). Fractions were collected and loaded onto a 10% SDS-PAGE gel. (A) Coomassie blue stained gel and (B) in-gel fluorescence (Excitation λ: 633nm, Emission λ: 675nm) of Alexa Fluor 633 labelled MdoD$_{C17I}$ (MdoD$_{C17I}^{633}$) after purification on a PD-10 column. The molecular masses (kDa) of standard proteins are given to the left of the gel. Degradation fragments are market as *. (C-D) *In vitro* Tat transport assay of MdoD$_{C17I}^{633}$. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP and 100nM MdoD$_{C17I}^{633}$. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added following transport as required. PK was added to digest any untranslocated MdoD$_{C17I}^{633}$ that remained in the reaction mixture. Transport was analyzed by (C) immunoblot detection of C-terminal His-tagged MdoD (α-His: 1/5,000) and (D) in-gel fluorescence scan.

3.8 Discussion

In this chapter, I have compared the transport efficiencies of *in vitro* Tat transport assays following the Müller protocol (Alami *et. al.* 2002) with assays following the Musser protocol (Bageshwar and Musser 2007). The IMVs used in the two methods were purified in a different manner and the conditions in which the transport assays were done were different. I have shown that IMVs produced by the Musser protocol had a higher transport efficiency than the Müller protocol. EM images showed that this difference in transport efficiency could be due to the morphological difference between the IMVs. In particular, the Musser IMVs were larger than the Müller IMVs. Smaller IMVs would have less Tat components. Hence, the chances of forming active Tat translocases would be reduced, thereby affecting the Tat transport efficiency.

The Tat transport competency of IMVs was shown to be sensitive towards mechanical stress including French press pressure and ultracentrifugation steps which could cause damage to the IMVs. The bacterial strain from which the IMVs were isolated from also affected the quality of the IMVs. The transport efficiency of *in vitro* Tat transport assays depended on the source of energization where ATP was a better energization source than NADH. It is believed that this is because energization by NADH required a complete electron transport chain, which is made up of several components and sufficient O_2 as a final electron acceptor for generation of a PMF, whereas generation of PMF by ATP only required a functional and intact ATP synthase. Tat transport was shown to be inhibited by the presence of >10mM Cl^- ions. However, it is still unclear how Cl^- ions affect transport.

By understanding these parameters that affect the Tat transport *in vitro*, I have successfully increased the *in vitro* Tat transport efficiency by modifying the Musser protocol. This modified version of the *in vitro* Tat transport was shown to have a higher transport efficiency than the original Musser protocol. I have also developed a fluorescence based *in vitro* Tat transport assay for quantitative analysis of Tat transport. I have used this to

quantitatively verify that my modified *in vitro* Tat transport assay has a higher transport efficiency than the original Musser protocol. It was also observed that the transport efficiency is dependent on the amount of substrates added to the *in vitro* assay. This has to be taken into account when comparing transport efficiencies of different transport assays.

Chapter 4:

Characterization of *in vitro* Tat transport

4.1 Introduction

The transport efficiency of *in vitro* Tat transport assays has been a limiting factor for *in vitro* studies of the Tat system. In the previous chapter, I have reported how I improved the transport efficiencies of current *in vitro* Tat transport assays. I have also shown that quantitative analysis of Tat transport can be achieved by detecting transport of fluorescently labelled substrates. This chapter will focus on using the modified *in vitro* Tat transport assay to characterize the Tat system *in vitro*. This includes bioenergetic and kinetic studies of the Tat system *in vitro*.

4.2 Bioenergetic studies of the Tat machinery

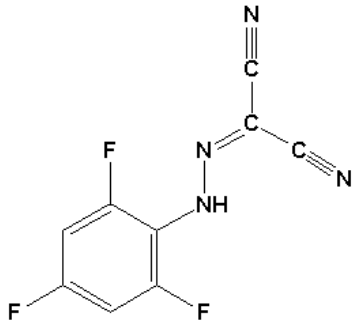
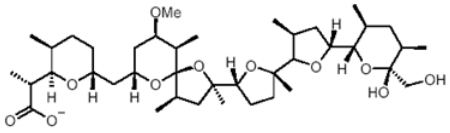
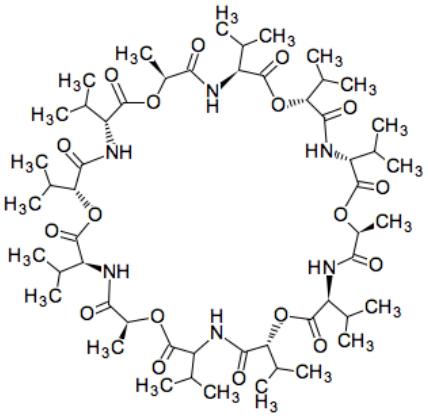
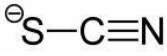
Tat transport in *E. coli* depends on the presence of a PMF (Yahr and Wickner 2001; Bageshwar and Musser 2007). Bageshwar and colleagues (2007) reported that the Tat transport of the native Tat substrate, SufI, was independent of the ΔpH component of the PMF. They hypothesized that a substantial but short duration Δψ is required for an early transport step. Following that, a small but long duration Δψ is needed to drive a later transport step.

4.2.1 Effect of protonophores and ionophores on Tat transport

I wanted to test if I could reproduce the observations from the bioenergetic study reported by Bageshwar and Musser (2007) using my modified *in vitro* Tat transport assay and a different Tat substrate, CueO. To investigate whether the ΔpH or the ΔΨ had a more significant effect on the Tat transport, I added ionophores into my *in vitro* Tat transport assays. Ionophores are charge carrying, small amphipathic molecules which can freely cross the phospholipid bilayers and increase the permeability of membranes to specific ions. Protonophores are ionophores that are proton specific. Therefore, protonophores and

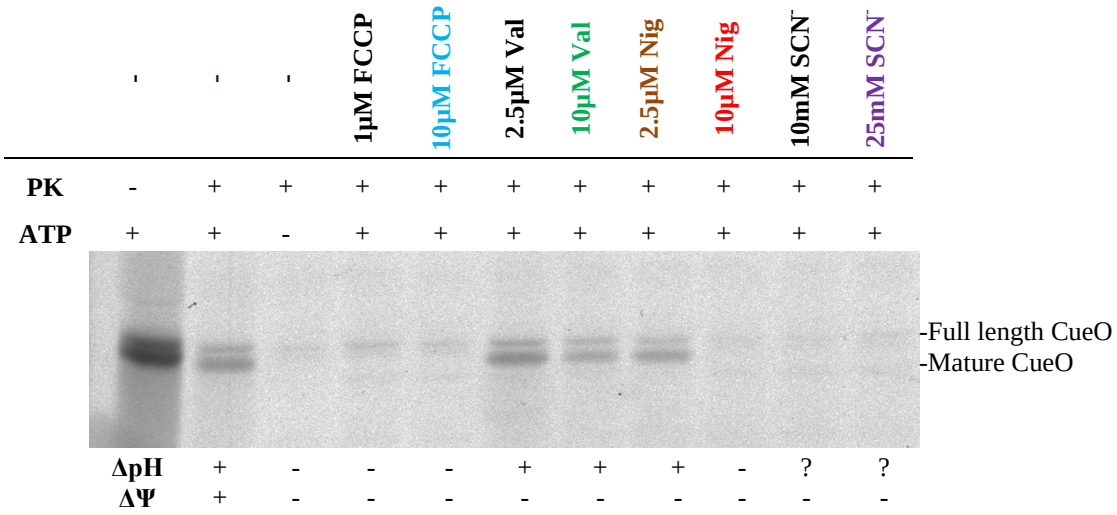
ionophores are capable of dissipating the ΔpH and/or the ΔΨ component of the PMF. Table 4- 1 provides a brief description of the ionophores and protonophores used.

Table 4- 1: Reagents used to dissipate ΔpH and/or ΔΨ in *in vitro* Tat transport assays

	Structure	Description	ΔpH / ΔΨ affected
FCCP (Carbonyl cyanide p-(trifluoromethoxy) phenylhydrazone)		Protonated and anionic form of FCCP diffuse across the membranes, discharging the PMF	ΔpH and ΔΨ
Nigericin		Electroneutral K^+/H^+ exchange across membranes	ΔpH
Valinomycin		Moves K^+ across membranes	ΔΨ
Thiocyanate		Permeable anion	ΔΨ

To ensure that the protonophores and ionophores were collapsing the ΔpH and/or ΔΨ, the ΔpH and ΔΨ across the IMVs were monitored by the ΔpH sensitive dye, ACMA, and the ΔΨ sensitive dye, oxonol V. These energization tests using ACMA and oxonol V are only semi-quantitative and do not reflect the exact magnitude of the ΔpH and ΔΨ generated. Since the effects of valinomycin and nigericin on the PMF depend on the movement of K^+ across membranes, I created a potassium gradient across the IMVs in my translocation assay where $[K^+]_{in}$: $[K^+]_{out}$ was 1mM: 50mM.

In vitro Tat transport assays done in the presence of FCCP showed no transport (Figure 4- 1). This is because FCCP is a protonophore and addition of FCCP under these experimental conditions dissipated both the ΔpH and ΔΨ components of the PMF (Figure 4- 2; **light blue**). This confirms that Tat transport requires the PMF.

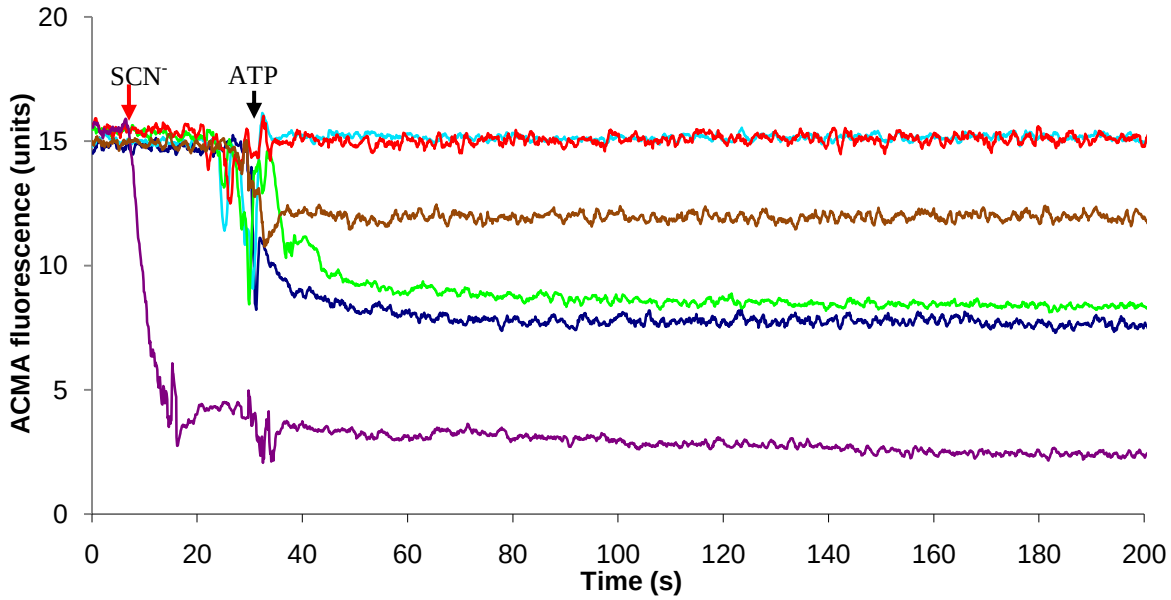


*? – uncertain of the effect of SCN^- on ACMA

Figure 4- 1: Transport assay in the presence of ionophores and protonophores

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to the translocation buffer containing 4mM ATP, ionophores/protonophores and 100nM CueO_GS3^{633}. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. Detection of transported CueO_GS3^{633} was performed by in-gel fluorescence scan (Excitation λ: 633nm, Emission λ: 675nm). The presence of ΔpH and ΔΨ was determined using the ΔpH sensitive dye, ACMA, and the ΔΨ sensitive dye, oxonol V, as described in Figure 4- 2. The labels on top of the gel are colour correlated to Figure 4- 2. Val: Valinomycin; Nig: Nigericin; SCN^-: Thiocyanate

A)



B)

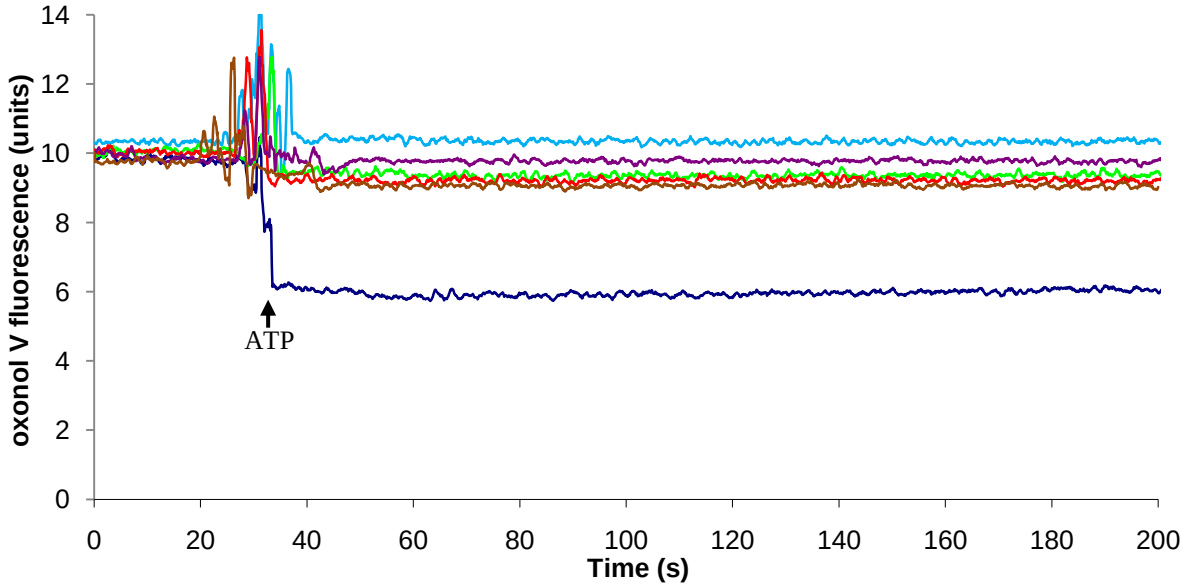


Figure 4- 2: Effect of ionophores on ΔpH and ΔΨ of IMVs

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were incubated at 37°C in 100µl translocation buffer in the presence of protonophores or ionophores before addition of fluorescent dyes and the absorbance was followed. When a stable fluorescent baseline was obtained, 4mM of ATP was added. (A) ACMA dye (0.1µM) was used to monitor the ΔpH across membranes (Excitation λ: 419nm, Emission λ: 483nm). In the case of thiocyanate (SCN^-), SCN^- was added as indicated by the red arrow, followed by ATP. (B) Oxonol V dye (0.2µM) was used to monitor the ΔΨ (Excitation λ: 625nm, Emission λ: 645nm). The experiment was done in triplicate and the average plotted. **Dark blue**: no uncouplers; **light blue**: 10µM FCCP; **green**: 10µM Valinomycin; **brown**: 2.5µM Nigericin; **red**: 10µM Nigericin; **purple** : 25mM SCN^-.

To test if Tat transport can occur in the absence of ΔΨ, valinomycin, which is a K^+ ionophore, was added. CueO was transported in the presence of up to 10μM valinomycin (Figure 4- 1). Energization tests showed that the addition of valinomycin resulted in dissipation of ΔΨ while maintaining the ΔpH (Figure 4- 2; **green**). This suggests that ΔpH can support Tat transport in the absence of ΔΨ. This contradicts with the data obtained using the original Musser *in vitro* transport assay which claimed that Tat transport *in vitro* required ΔΨ (Bageshwar and Musser 2007). The authors did not use valinomycin for their experiments because they claimed that valinomycin was not able to dissipate ΔΨ or ΔpH. Instead, the authors used thiocynate (SCN^-), which is a membrane permeable anion, to abolish ΔΨ. The authors reported that transport was abolished in the presence of SCN^-, which according to their data, SCN^- only selectively eliminated the ΔΨ and not the ΔpH. In my modified transport assay, transport was also not observed in the presence of SCN^- (Figure 4-1). According to the oxonol V energization tests, SCN^- was able to dissipate ΔΨ. However, it was observed that the fluorescence of ACMA was reduced in the presence of SCN^- even before addition of ATP (Figure 4- 2; **purple**). A reduction in ACMA fluorescence was also observed when SCN^- was added to the buffer in the absence of IMVs (data not shown). This phenomenon has been reported before and it was speculated that the interaction between ACMA and SCN^- resulted in the quenching of ACMA (Huang *et. al.* 1983). This quenching of ACMA fluorescence could be misinterpreted to be due to the presence of a ΔpH. Therefore, although there was a decrease in ACMA fluorescence in the presence of SCN^-, it cannot be concluded that ΔpH was present. Moreover, experiments done in the presence of valinomycin showed that transport can occur in the absence of ΔΨ. Since no transport was observed in the presence of SCN^-, this suggests that the ΔpH could be affected by SCN^-. A possible explanation for this could be that SCN^- is a known chaotropic agent which is capable of separating the soluble F_1 subunit from the membrane bound F_o subunit. Although the F_1

subunit is still capable of hydrolyzing ATP after separation, proton movement by the F_o subunit is no longer coupled to ATP hydrolysis. The F_o subunit now acts as a bidirectional and passive channel for protons (Patel *et. al.* 1975). This means that the membrane is now leaky to protons, resulting in a dissipation of the PMF. This could explain why Tat transport was not observed in the presence of SCN^-.

To test if the Tat transport can occur in the absence of ΔpH, nigericin, an electroneutral K^+/H^+ exchanger, was added to the translocation assay. When the translocation assay was done using the original Musser *in vitro* Tat translocation assay, transport was observed in the presence of 2.5µM nigericin (Bageshwar and Musser 2007). In my modified *in vitro* Tat transport assays, transport was also observed in the presence of 2.5µM nigericin (Figure 4- 1). However, energization tests by ACMA showed that this was because 2.5µM nigericin was not sufficient to fully dissipate the ΔpH (Figure 4- 2A; **brown**). However, it was observed that there was no reduction in the fluorescence of oxonol V (Figure 4- 2B; **brown**). This means that nigericin, which was supposed to specifically collapse ΔpH, was instead dissipating the ΔΨ. Since transport was still observed in the absence of ΔΨ, this means that Tat transport can occur without ΔΨ. However, when I increased the concentration of nigericin used four fold to 10µM, Tat transport was not observed (Figure 4- 1). Energization tests showed that at 10µM, the ΔpH was completely dissipated (Figure 4- 2A, **red**). However, the fluorescence of oxonol V did not decrease, suggesting that ΔΨ was also dissipated. Since both the ΔpH and ΔΨ were affected by nigericin, it is not possible to draw a valid conclusion based on these results to address the question of whether Tat transport can occur in the absence of ΔpH.

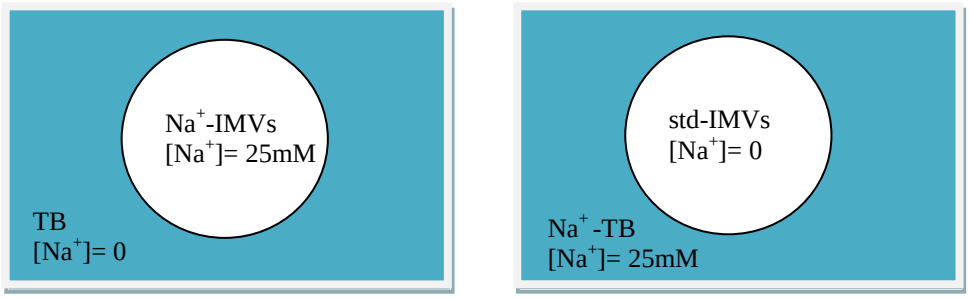
4.2.2 Sodium motive force (SMF)

Protons are not the only coupling ions used in energy transduction. Another commonly used ion is sodium. Recently, it has been reported that the transport of α-amylase (AmyH) by the Tat system from the halophilic archaea *H. hispanica* depends on the sodium motive force (SMF) (Kwan *et. al.* 2008). I wanted to test if *E. coli* can utilize the SMF as well as the PMF for Tat transport.

Standard IMVs were made without Na^+ in the lumen (std-IMVs). To address the issue of whether SMF affects the transport efficiency, IMVs were made with 25mM Na^+ inside the lumen (Na^+-IMVs). Transport assays using Na^+-IMVs were done in translocation buffer (no Na^+) where the ratio of Na^+ inside the IMVs to outside the IMVs (*i.e.* translocation buffer) is 25:0 (Figure 4- 3A). This ratio is similar to the reported ratio of Na^+ in the cytoplasm of *E. coli* and its environment (Castle *et. al.* 1986). A sodium gradient in the opposite direction was generated by carrying out the transport assays using std-IMVs in translocation buffer containing 25mM Na^+ (Na^+-TB) (*i.e.* ratio of Na^+ in:out is 0:25).

As can be seen from Figure 4- 3B, no transport was observed in the absence of ATP regardless of the direction of the sodium gradient (lanes 1 and 5). As a control experiment, I added ATP to the reactions to initiate Tat transport. It was observed that the addition of ATP resulted in successful transport and that the transport efficiency of Na^+-IMVs was similar to that of std-IMVs in translocation buffer (TB) (Figure 4- 3B, lanes 2 and 4). Transport was also observed when std-IMVs were energized in Na^+-TB (Figure 4- 3B, lane 6). This suggests that the lack of transport of non-energized Na^+-IMVs and std-IMVs in the presence of a Na^+ gradient was not due to loss of transport competency of the IMVs under those experimental conditions.

A)



Lanes 1 and 2 in (B) Lanes 5 and 6 in (B)

B)

Lanes	1	2	3	4	5	6
[Na+] in IMV lumen (mM)	25	25	0	0	0	0
[Na+] in TB (mM)	0	0	0	0	25	25
ATP	-	+	+	+	-	+
PK	+	+	-	+	+	+

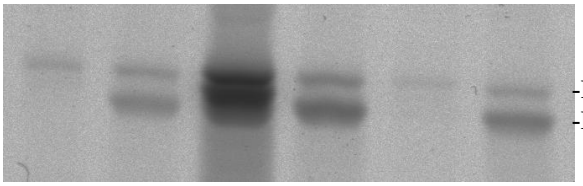


Figure 4- 3: Effect of sodium motive force on *in vitro* Tat transport

A difference in Na^+ concentration was used to test the effect of SMF on Tat transport in *E. coli* IMVs. (A) Schematic representation of the conditions of the *in vitro* transport assay and the difference in Na^+ concentration. To generate an outwardly directed sodium gradient (figure on the left), IMVs were French pressed in buffer containing 25mM Na^+. Therefore, these IMVs have 25mM Na^+ in the lumen (Na^+-IMVs). Na^+-IMVs were added to translocation buffer (TB) without Na^+. To generate an inwardly directed Na^+ gradient (figure on the right), IMVs were made without Na^+ in the lumen (std-IMVs) and transport assays were done in translocation buffer with 25mM Na^+ (Na^+-TB). (B) *In vitro* Tat transport assay was carried out using CueO_GS3^{633} as substrate. IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to their respective translocation buffers that contained CueO_GS3^{633} and 4mM ATP if required. The reaction was done at 37°C for 30min and quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. Detection of transported CueO_GS3^{633} was performed by in-gel fluorescence scan (Excitation λ: 633nm, Emission λ: 675nm).

This experiment shows that *E. coli* Tat transport is not dependent on SMF and still requires generation of a PMF by addition of ATP for transport even in the presence of a sodium gradient. This suggests that the SMF alone was not sufficient to initiate Tat transport in *E. coli*.

This is in contrast to the *H. hispanica* Tat system in which the Tat system is solely SMF dependent (Kwan *et. al.* 2008). The reason for this difference may be because the halophilic *H. hispanica* has to live in high saline conditions and could have adapted its Tat system to utilize the SMF generated under such extreme conditions.

4.3 Transport kinetics

This section will focus on the transport kinetics of the Tat system *in vitro*. An understanding of the kinetics of *in vitro* Tat transport would be useful for elucidating the mechanism of the Tat system under physiological conditions. These studies use the fluorescence based transport assay described in Chapter 3.

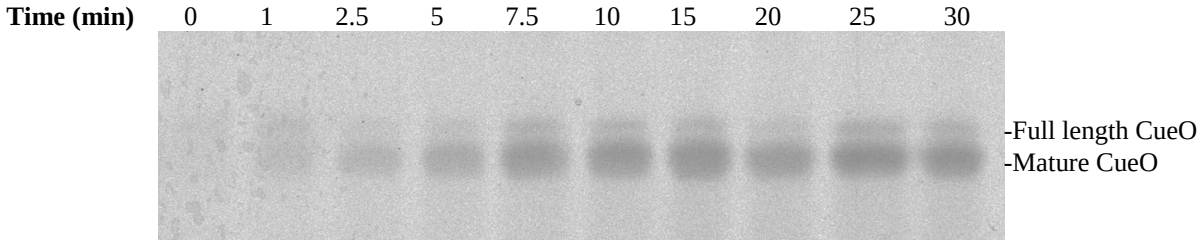
4.3.1 Time trial assay of *in vitro* Tat transport

I examined the transport kinetics of CueO_GS3^{633} import into TatABC-enriched IMVs. Transport reactions were initiated by addition of IMVs to translocation buffer containing ATP and CueO_GS3^{633}. Transport reactions were quenched at various time points by incubating sample tubes on ice for at least 2min. Transport was detected by in-gel fluorescence (Figure 4- 4A). Since no bands were observed on the gel after PK digestion at time= 0min (when the reaction was immediately stopped after addition of IMVs into the translocation buffer by incubating the reaction on ice), this showed that incubation on ice was sufficient to stop further transport. This observation also means that the full length CueO_GS3^{633} bands observed in other lanes represented transported full length CueO_GS3^{633} that was not processed by the signal peptidase, LepB. Hence, both the full length and mature

CueO_GS3^{633} detected by in-gel fluorescence after PK digestion were summed to give the total amount of CueO_GS3^{633} transported into the IMVs. It was observed that the transport of CueO_GS3^{633} occurred over a 10min period but the transport slowed down and eventually ceased after 10min (Figure 4- 4).

This result suggests that the *in vitro* Tat transport assay loses transport competency over time. By contrast, Tat transport in living cells is expected to occur continuously. To understand why transport competency is lost over time, I investigated the involvement of factors known to be required for Tat transport to occur.

A)



B)

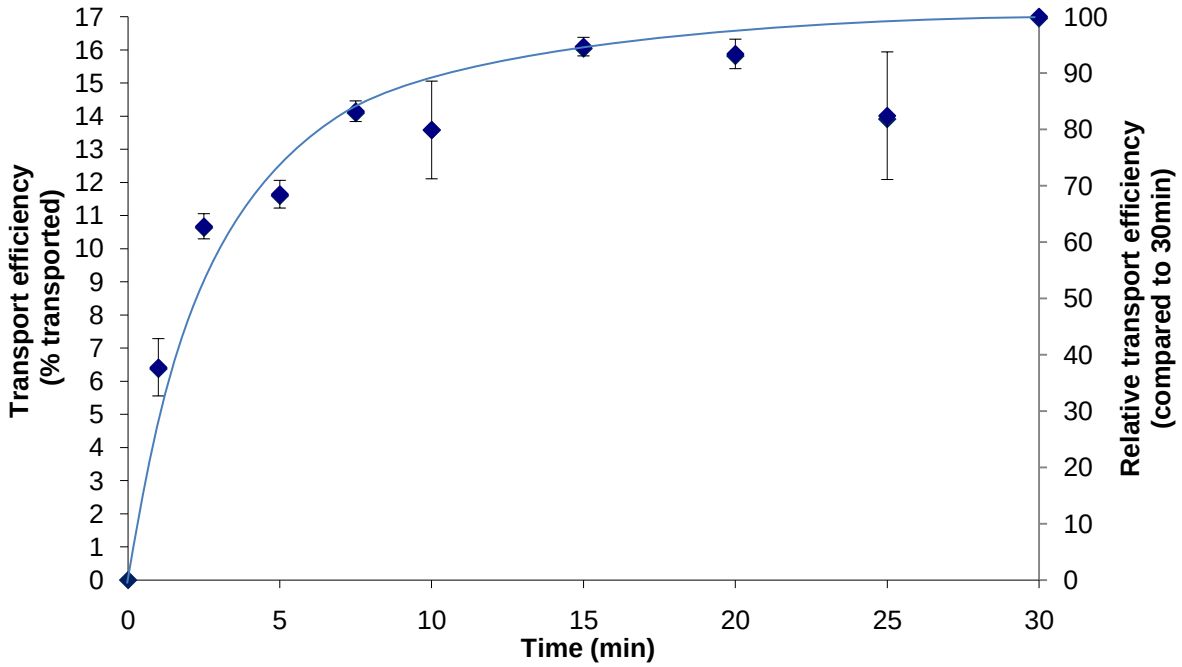


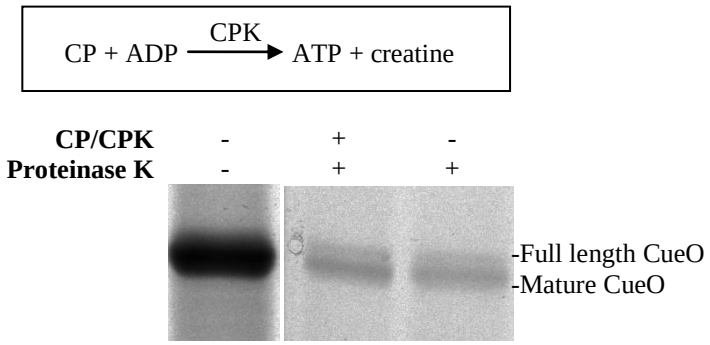
Figure 4- 4: *In vitro* transport kinetics of CueO_GS3^{633}

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer containing 4mM ATP and 100nM CueO_GS3^{633}.The reaction was done at 37°C for the specified time period and quenched by placing the reaction on ice for 2min. PK was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. Quantitation analysis was correlated to a standard curve of known concentrations of CueO_GS3^{633} on the same gel. (A) Detection of CueO_GS3^{633} transported was performed by in-gel fluorescence scan (Excitation λ: 633nm, Emission λ: 675nm). (B)Transport efficiency at various time points was plotted. Left vertical axis represents the actual transport efficiency (i.e. % transported out of total CueO_GS3^{633} added to the reaction), right vertical axis represents the transport efficiency relative to the actual transport efficiency at 30min (i.e transport efficiency at t=30min was normalized to 100%). Both full length and mature CueO_GS3^{633} were summed during quantitative analyses. The experiment was done in triplicate and the average plotted. Error bars show the standard error of actual transport efficiency. The plotted curve is indicative rather than a fit to the data.

4.3.2 Effect of ADP

One possibility explanation for the gradual loss of Tat transport efficiency would be the build up of inhibitory reaction by-products. The most obvious candidate would be the ATP hydrolysis product ADP. It has been reported that ADP inhibits ATP synthase by binding onto ATP synthase (Fischer *et. al.* 2000). To test if ADP inhibits the Tat transport *in vitro*, I tried to 'remove' ADP from the transport assay and compared the transport efficiency to assays done under standard conditions. 'Removal' of ADP can be achieved by adding an ATP regenerating system, which reforms ATP from ADP, into the reaction assay. Two ATP regenerating systems were used: the creatine phosphate/ creatine phospho kinase (CP/CPK) system and the phosphoenol pyruvate/ pyruvate kinase (PEP/PyK) system. In the CP/CPK system, CPK catalyses the conversion of creatine phosphate (CP) and ADP to creatine and ATP. In the PEP/PyK system, PyK catalyses the conversion of phosphoenol pyruvate (PEP) and ADP to pyruvate and ATP. Figure 4- 5 shows that the Tat transport efficiency was not increased by addition of either ATP regenerating system. For reasons unknown, full length CueO was not observed in the presence of the PEP/PyK.

A)



B)

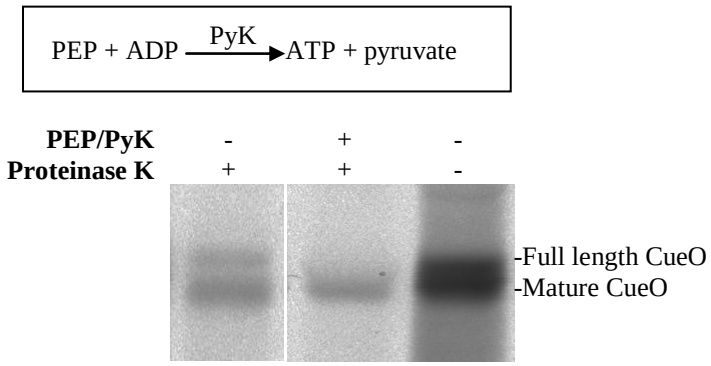


Figure 4- 5: Effect of ATP regenerating systems on Tat transport

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer containing 4mM ATP, 100nM CueO_GS3^{633} and the ATP regeneration systems CP/CPK or PEP/PyK. The reactions were done at 37°C for a total of 30min and quenched by placing the reaction on ice for 2min. Proteinase K was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. Detection of transported CueO_GS3^{633}was performed by in-gel fluorescence scan (Excitation λ: 633nm, Emission λ: 675nm). Transport assays when done in the presence of (A) CPK (8mM) and CP (0.04mg/ml) or (B) PEP (2.5mM) and PyK (0.1mg/ml).

To further test if ADP inhibited Tat transport, I added ADP into the reaction in the presence of ATP. The total concentration of AxP was maintained at 4mM. It was observed that the reaction assay was able to tolerate up to 3mM of ADP without significantly affecting the Tat transport efficiency (Figure 4- 6) . It was also observed that addition of 4mM ADP without ATP supported slight Tat transport. This could be due to contamination of ATP in commercially available ADP. This result demonstrated that the presence of ADP in the reaction assay does not affect Tat transport *in vitro*.

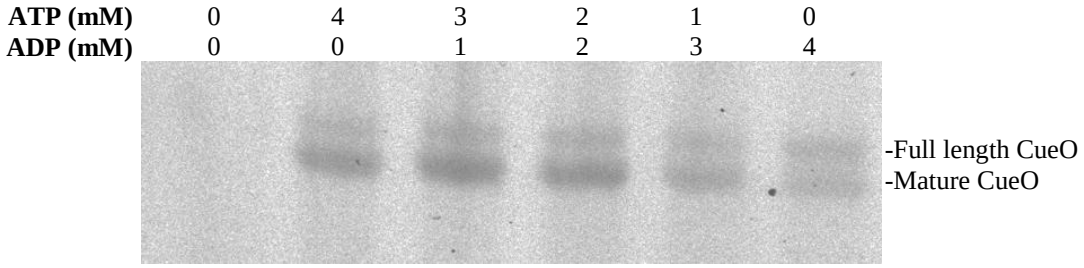


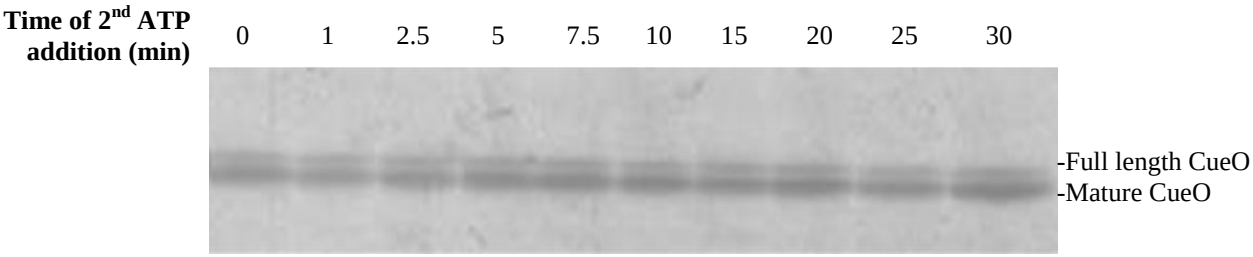
Figure 4- 6: Effect of ADP on Tat transport efficiency

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer containing a total 4mM AxP at different ratios and 100nM CueO_GS3^{633}. The reaction was left at 37°C for 30min and quenched by placing it on ice for 2min. PK was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. Detection of CueO_GS3^{633} after transport was performed by in-gel fluorescence scan (Excitation λ: 633nm, Emission λ: 675nm).

4.3.3 Role of ATP

In my *in vitro* Tat transport assays, ATP hydrolysis by ATP synthase was used to generate the PMF required for Tat transport. Therefore, it is possible that the concentration of ATP is the limiting factor in the translocation assay. To test this hypothesis, I added additional ATP to the assay at different time points. Initially, IMVs were added into translocation buffer containing 4mM ATP and CueO_GS3^{633} to initiate transport. I then added a second dose of 4mM ATP at different time points during the course of the assay. All reactions were terminated by incubation on ice 30min after the transport reaction was initiated. As Figure 4- 7 shows, there was no significant increase in Tat transport efficiency regardless of when additional ATP was added to the assay.

A)



B)

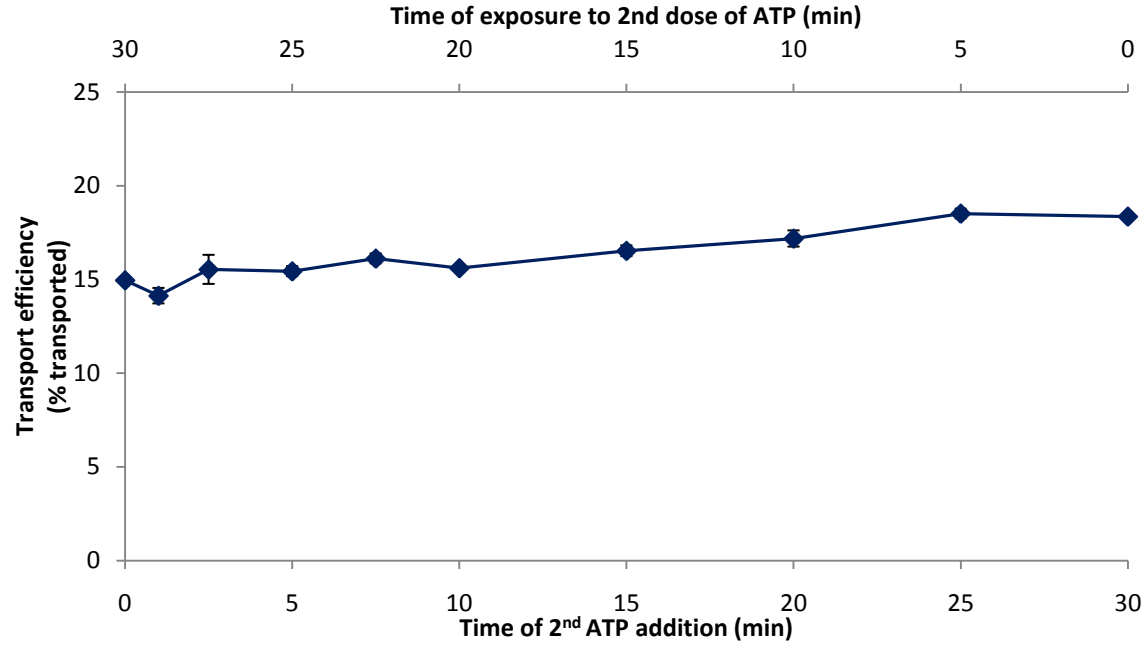
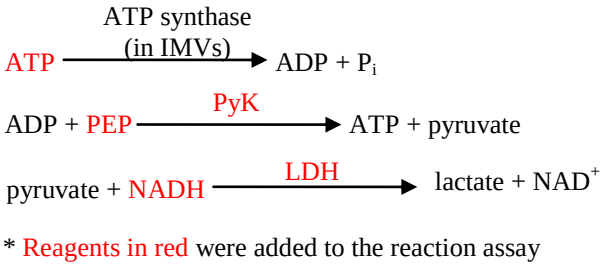


Figure 4- 7: Effect of addition of second dose of ATP on Tat transport efficiency

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer containing 4mM ATP and 100nM CueO_GS3^{633}. A second dose of 4mM ATP was added at the indicated time points. The reaction was done at 37°C for a total of 30min and quenched by placing the reaction on ice for 2min. PK was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. (A) Detection of transported CueO_GS3^{633} was performed by in-gel fluorescence scan (Excitation λ: 633nm, Emission λ: 675nm). (B) The transport efficiency of IMVS at the time points when the second dose of ATP was added was plotted. Top horizontal axis shows the time of exposure of the IMVs to the second dose of ATP. The experiment was done in triplicate and the average plotted. Error bars show standard error.

This result means that the *in vitro* transport assay was not limited by the availability of ATP. Alternatively, it is possible IMVs were unable to utilize that excess ATP for transport because the ATP synthase was inactivated during the assay. To test this possibility, I measured the activity of ATP synthase under transport conditions. The activity of ATP synthase was determined by measuring the rate of ATP hydrolysis. The method chosen to measure ATP hydrolysis was by monitoring NADH absorption at 340nm in the presence of a phosphoenolpyruvate (PEP)/ pyruvate kinase (PyK) ATP regenerating system (Scheme 1). In this assay, ATP would be hydrolyzed by active ATP synthase to produce ADP. ADP is converted back to ATP in the presence of PEP and PyK. The by-product of this reaction is pyruvate. Pyruvate would oxidize NADH to NAD^+ in a reaction catalyzed by lactate dehydrogenase (LDH) to form lactate. Since NAD^+ would not absorb at 340nm, while NADH does, A_{340nm} would decrease if ATP synthase was active.



Scheme 1: PEP/PyK ATP regenerating system

As seen in Figure 4- 8, the rate of ATP hydrolysis by IMVs was maintained for up to an hour under transport conditions. This implies that ATP synthase was able to hydrolyze ATP throughout the duration of the *in vitro* transport assay. However, ATP hydrolysis might not be coupled to generation of PMF if the F_1F_o subunits of the ATP synthase were separated during IMV preparation. The F_1 subunit would still be able to hydrolyze ATP but hydrolysis would not be coupled to proton movement by the F_o subunit.

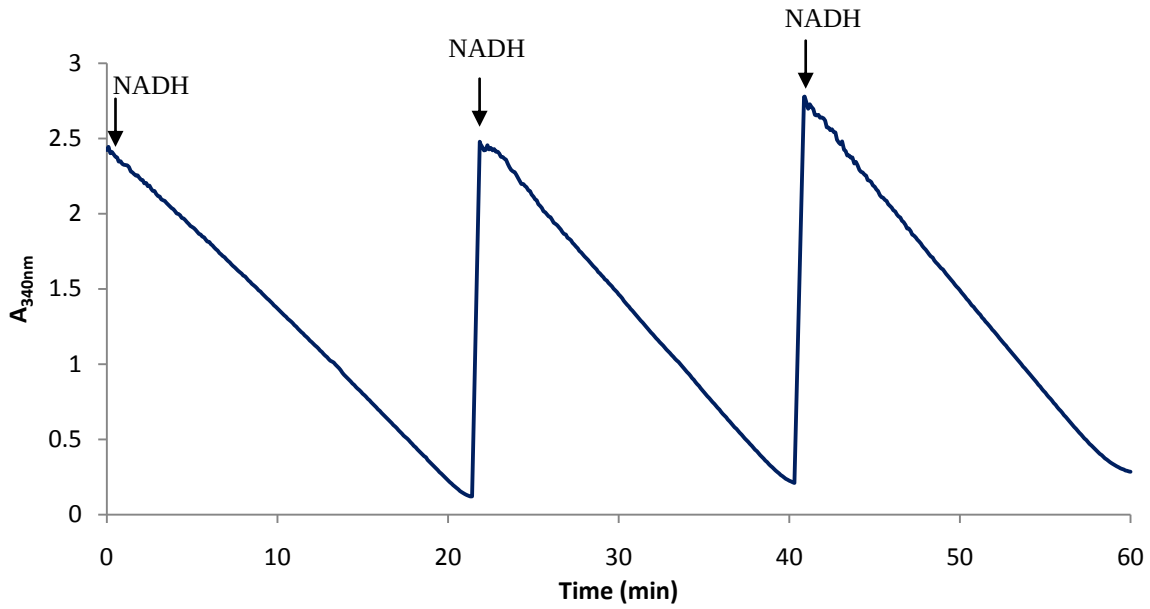


Figure 4- 8: Measurement of ATP synthase activity

ATP hydrolysis rate was measured by following the decrease in NADH absorbance at 340nm. IMVs were diluted to 0.02mg/ml in assay buffer (20mM HEPES, 5mM $MgCl_2$, 100mM KCl, 5mM KCN, 2.5mM PEP, 0.1mg/ml PyK, 0.1mg/ml LDH, pH7.5). A blank reading was obtained before 0.7mM NADH was added. The reaction was initiated by addition of 1mM ATP. When NADH was exhausted (A_{340nm} <0.1), another 0.7mM NADH was added to continue the measurement (as denoted by arrows). The experiment was done in triplicate.

Taken together, these results suggest that the IMVs should have sufficient ATP that can be hydrolysed by the ATP synthase throughout the whole duration of an *in vitro* Tat transport assay. However, these experiments do not provide information on the sustainability of the PMF during the assay.

4.3.4 Sustainability of PMF during *in vitro* transport

It is known that PMF is required in at least one step of the transport cycle although it is still unclear if the PMF is required for the entire cycle or specific steps in the cycle. Therefore, if the PMF was not maintained in the *in vitro* experiments, new cycles of Tat transport cannot occur and Tat transport would be stalled.

ACMA and oxonol V based energization tests carried out in chapter 3 showed that both the ΔpH and ΔΨ components of the PMF generated in IMVs by ATP hydrolysis were maintained throughout a time period corresponding to the whole duration of the transport assay (Figure 3- 25). However, these energization tests were carried out in the absence of Tat substrates. This time, I tested if active transport of Tat substrates will eventually lead to a collapse in PMF, resulting in loss of further transport. As can be seen from Figure 4- 9, addition of substrates after energization of the IMVs did not result in a detectable decrease in ΔpH or ΔΨ. This means that PMF was present throughout the assay duration. Therefore, the loss of PMF as a result of active transport of substrates would not explain the cessation of transport observed after 10min.

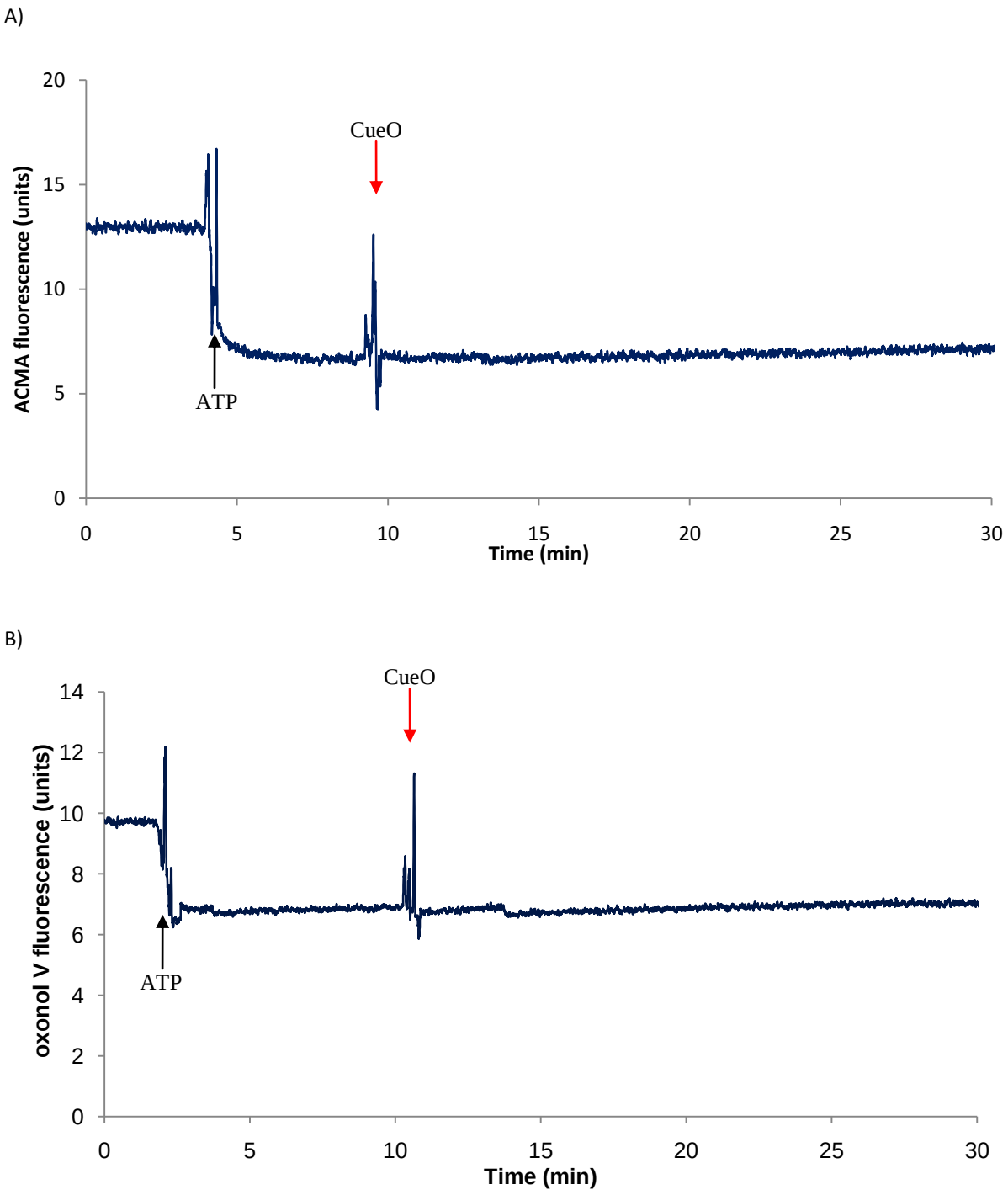


Figure 4- 9: Effect of active Tat transport of CueO on ΔpH and ΔΨ

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were incubated at 37°C in 100µl translocation buffer before addition of fluorescent dyes and the absorbance was followed. When a stable fluorescent baseline was obtained, 4mM of ATP was added followed by addition of 100nM CueO at the indicated time point. (A) ACMA dye (0.1µM) was used to monitor ΔpH across membranes (Excitation λ: 419nm, Emission λ: 483nm). (B) Oxonol V dye (0.2µM) was used to monitor ΔΨ (Excitation λ: 625nm, Emission λ: 645nm). The experiments were done in triplicate and the average plotted.

4.3.5 Transport competency of Tat substrates and IMVs

The Tat system is very specific for substrates that are properly folded (Sargent *et. al.* 1998). Substrates could lose their Tat transport compatibility when they become unfolded or are degraded. It is therefore possible that the gradual decrease in the rate of transport in *in vitro* Tat transport assays was because the substrate, CueO_GS3^{633}, had lost its Tat transport competency over time at 37°C. To test this, I incubated CueO_GS3^{633} with the IMVs for 30min at 37°C before initiating Tat transport by energizing the IMVs with ATP. The transport efficiency of pre-incubated CueO_GS3^{633} was similar to that of CueO_GS3^{633} without prior incubation (Figure 4- 10, lanes 1 and 3). This suggests that CueO_GS3^{633} remained transport compatible throughout the assay. Most IMVs also retained their Tat transport competency even after a 30min incubation at 37°C (Figure 4- 10, lane 2). These results suggest that the transport competency of the Tat substrates and IMVs were not compromised throughout the duration of the *in vitro* transport assay.

4.3.6 Order of substrate addition

Figure 4- 10 also showed that the Tat transport efficiency was affected by the order of substrate addition into the reaction assay relative to energization of IMVs. The transport efficiency was greatly reduced when substrates were added to pre-energized IMVs (Figure 4- 10, lane 4). Since Tat transport still occurs in the inner membrane which is constantly energized in a living cell, this observation does not fully reflect *in vivo* conditions. One theory to explain this observation is that the IMVs have de-energized during the 30min incubation time. However, there are two lines of evidence that disprove this theory. Firstly, this phenomenon could not be reversed by re-energizing the IMVs with ATP (Figure 4- 10, lane 5). Secondly, as shown earlier in Figure 4- 9, IMVs are still energized after 30min. Therefore, loss of PMF is insufficient to explain the reason behind the reduced transport efficiency of pre-energized IMVs.

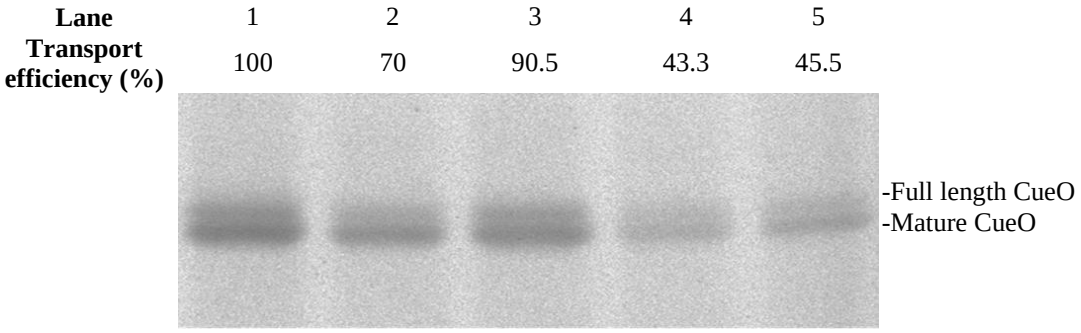


Figure 4- 10: Effect of order of IMV energization and substrate addition

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. The reactions were quenched by placing them on ice for 2min. PK was added to digest any untranslocated CueO_GS3^{633} that remained in the reaction mixture. Transported CueO_GS3^{633} was detected by in-gel fluorescence scan of CueO_GS3^{633} (Excitation λ: 633nm, Emission λ: 675nm). Transport efficiency was defined as % CueO transported relative to standard assay conditions (i.e. lane 1). Lane 1: CueO_GS3^{633} and ATP were added to translocation buffer and incubated at 37°C for 5min before IMVs were added. The reaction was stopped after 30min [i.e. standard assay condition]; Lane 2: IMVs in translocation buffer were incubated at 37°C for 30min before the addition of ATP and CueO_GS3^{633}. The reaction was left for a further 30min; Lane 3: IMVs and CueO_GS3^{633} were added to translocation buffer and incubated at 37°C for 30min before addition of ATP. The reaction was left for a further 30min; Lane 4: IMVs and ATP were added to translocation buffer and incubated at 37°C for 30min before addition of CueO_GS3^{633}. The reaction was left for a further 30min; Lane 5: IMVs and ATP were added to translocation buffer and incubated at 37°C for 30min before addition of CueO_GS3^{633} and a second dose of ATP. The reaction was left for a further 30min. (IMVs A_{280nm}=10, CueO=100nM; ATP=4mM)

4.4 Discussion of the bioenergetic and kinetic studies on Tat transport *in vitro*

I have carried out bioenergetic studies on the Tat system using my modified *in vitro* Tat transport assay. My results showed that the Tat transport can occur in the absence of $\Delta\Psi$. This contradicts the results reported by Bageshwar and Musser (2007) where it was shown that Tat transport *in vitro* required $\Delta\Psi$. A possible reason for this difference could be because the *in vitro* transport assays were done under different conditions. I have also carried out experiments to test if the *E. coli* Tat transport could use the sodium motive force, which was shown to drive Tat transport in the haloarcheon *H. hispanica*. My results showed that the *E. coli* Tat transport cannot be driven by the sodium motive force.

Kinetic studies on the Tat system using my modified *in vitro* assay showed that the transport of CueO ceased ~10min after the transport reaction was initiated. A similar observation had been reported when IMVs were energized with NADH (Bageshwar and Musser 2007). Further investigation into the time dependent loss of Tat transport competency revealed that this limitation was neither due to the inability of the IMVs to maintain a PMF during the assay nor was it restricted by the transport competency of IMVs and CueO_GS3^{633}. Bageshwar *et. al.* reported an increase in transport efficiency when oxygenated buffer was added to the assay after 10min. This is because the IMVs in Bageshwar's experiments were energized by NADH and generation of PMF by NADH is restricted by the limited amount of oxygen in solution. On the other hand, I had used ATP to energize my IMVs. However, I did not observe significant increases in transport when a second dose of ATP was added at different time points. This could be because there was sufficient ATP in my modified *in vitro* assay to maintain a PMF throughout the entire duration of the experiment. I also tested if ADP, a by-product of ATP hydrolysis that is required to generate the PMF for *in vitro* Tat transport, inhibited Tat transport. Although my results showed that ADP did not have a

significant inhibitory effect on Tat transport, there is still the possibility that there are inhibitory effects from other unidentified by-products of Tat transport. It is important to note that although ATP hydrolysis was used to generate a PMF in these *in vitro* experiments, ATP hydrolysis is not required for Tat transport *in vivo*. It was also observed that CueO has to be added before or during energization of IMVs for most efficient transport. This suggests that imposition of a PMF in the absence of substrates resulted in loss of Tat transport competency.

This series of experiments clearly demonstrates that the *in vitro* transport assay does not fully reflect physiological conditions because in living cells, the Tat translocase would continuously transport substrate across the inner membrane when required. Moreover, the inner membrane is constantly energized. Therefore, transport of Tat substrates does occur after membrane energization in cells. This raises the question as to why the *in vitro* system does not behave like the *in vivo* conditions. One reason could be that the Tat system *in vitro* could not revert back to its rest phase to reset the whole transport cycle after transport. It has been hypothesized that TatABC could undergo conformational changes upon energization to facilitate transport. For example, there is evidence that the topology of TatA changes during the transport cycle (Gouffi *et. al.* 2004; Chan *et. al.* 2007). However, not much detail is known about what triggers this conformation change and what is required to reset the system (Figure 4- 11). It is also possible that cytosolic proteins are required in this process. Since the cytoplasmic components have been removed during purification of IMVs, the Tat transport cycle would not be able to reset itself if cytoplasmic components are involved. Although it has been shown in Chapter 3 (Figure 3- 19) that reconstitution of IMVs with cellular fractions during the transport assay did not increase the transport efficiency, this could be because there was insufficient amounts of these proteins that can be used by the overproduced TatABC in the IMVs.

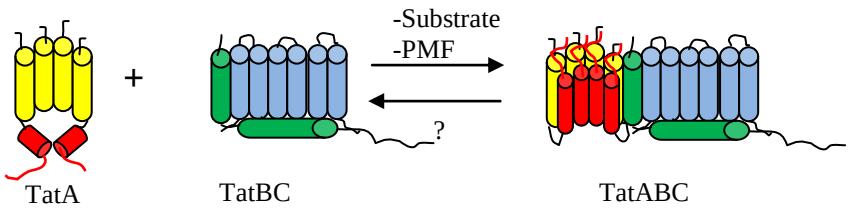


Figure 4- 11: Schematic representation of hypothesized conformational changes of TatABC upon substrate binding

In the resting state, the TatA complex is not attached to TatBC. Upon binding of substrate and in the presence of PMF, TatBC forms a complex with TatA. It is believed that the amphilathic helix (red) of TatA flips inwards to line to pore of the Tat translocase. The TatABC complex has to reset itself following a series of unknown events to support further translocation processes.

There have been reports of soluble proteins that are involved in the Tat transport. For example, it has been shown that PspA, a phage shock protein found in *E. coli*, was able to overcome saturation of Tat transport. It was also shown that the expression of *pspA* was increased in *tat* mutants. This suggests that PspA plays a role in Tat transport especially when the Tat system is under stress (Delisa *et. al.* 2004). There has also been suggestions that TatD could be involved in Tat transport as part of the quality control system (Matos *et. al.* 2009). However, this hypothesis was disproved by Lindenstrauß *et. al.* (2010). Nonetheless, these findings suggest that it is possible that there are other components that could play a supporting role in Tat transport and have yet to be identified.

4.5 Development of a fluorescence-based assay for fluorescence microscopy analysis of the Tat transport *in vitro*

So far, all the experiments carried out were based on bulk analysis of the Tat transport system. These experiments fail to resolve the heterogeneity between individual Tat substrates or Tat translocases. One method to overcome this issue is to observe the transport of Tat substrates at the single molecule level using single molecule fluorescence microscopy. This could be achieved using the droplet-on-hydrogel bilayers (DHB) method where IMVs can be incorporated into large, planar bilayers (Heron *et. al.* 2009). This method was developed by a collaborator, Dr. Mark I. Wallace, in the Department of Chemistry, University of Oxford. In this system, a thin layer of agarose containing the IMVs is formed on a glass coverslip. A poly(methyl methacrylate) microchannel device containing wells of ~500μm in diameter is then placed on top of the glass coverslip. The thin layer of agarose is rehydrated by filling the microchannel device with agarose dissolved in translocation buffer (TB) that is used for *in vitro* Tat transport assays. The device wells are then filled with a solution of lipid in oil. An aqueous droplet containing fluorescently labelled Tat substrate in TB (approximately 50nl in volume) that has been coated by lipids is placed on top of the hydrated agarose underneath the oil. A lipid bilayer (approximately 200μm in diameter) is formed at the interface between the two aqueous phases (Thompson *et. al.* 2007) (Figure 4- 12). The IMVs enriched with the Tat components will diffuse into the bilayer. Once the lipids form a bilayer, any fluorescent molecules that associate with the bilayers can be imaged using total internal reflection fluorescence (TIRF) microscopy. A PMF to initiate Tat transport can also be introduced by addition of ATP to drive ATP synthase activity in the IMVs or by placing electrodes in the aqueous compartments. The fluorescently labelled Tat substrate, CueO_GS3^{633}, that had been developed in chapter 3 (see 3.6.2) can be used for this system.

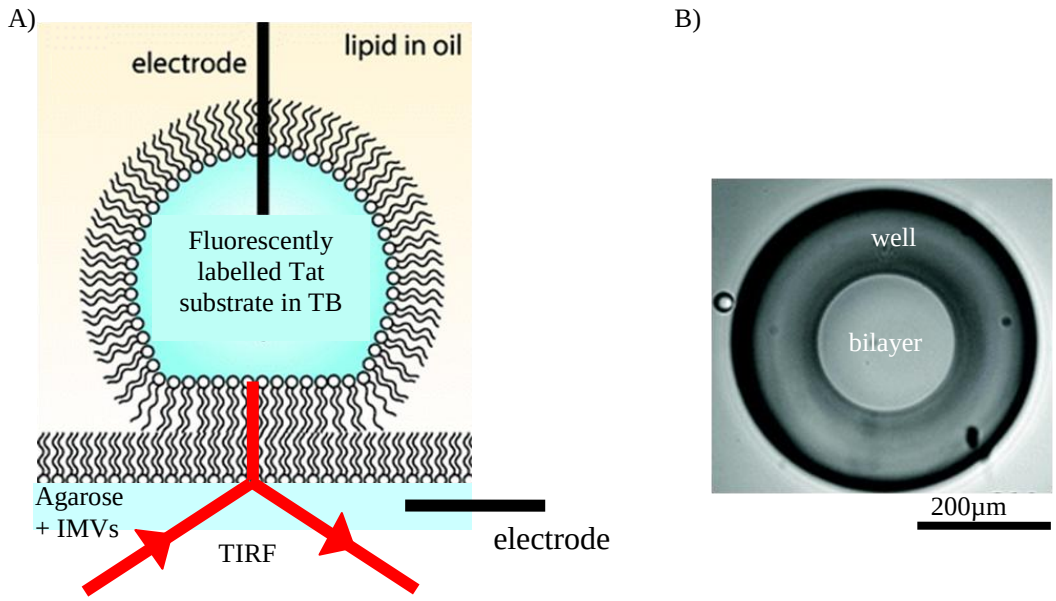


Figure 4- 12: Schematic representation of TIRF setup for single molecule fluorescence microscopy to investigate Tat transport

(A) IMVs enriched with TatABC are incorporated into the agarose layer and will diffuse into the lipid bilayer. Fluorescent substrates are inside the aqueous droplet coated with lipids. Once a bilayer is formed, any fluorescent substrate associated with the membrane can be detected by TIRF. A PMF can be generated by electrodes or addition of ATP. TB: translocation buffer. (B) A single well from the device with a bilayer formed in the middle. The outer ring is the edges of the well while the inner ring is the bilayer formed. Images adapted from (Thompson *et. al.* 2007).

Although fluorescently labelled substrates are useful for studying binding events between the Tat components and substrates, it may not be suitable for studying the transport of substrates across the membrane in this setup. This is because it will be difficult to determine which side the fluorescently labelled protein is on. Therefore, it is difficult to determine if Tat transport has occurred. One way to address this issue would be to find a Tat substrate that can act on a fluorogenic compound to produce a fluorescent product. This fluorogenic compound will be incorporated into the agarose layer. The rationale behind this approach is that a fluorescent signal would only be observed when the Tat substrate is transported across the membrane to the agarose layer where the compound is. This eliminates false positive results as fluorescent signals detected will only be due to true transport events.

In addition, this method provides signal amplification as a single transported Tat substrate can act on many fluorogenic compounds.

A possible candidate for such a reporter substrate would be a Tat transported phosphatase. A fluorogenic substrate for alkaline phosphatases (APase), AttoPhos®, is commercially available and produces a highly fluorescent product (Figure 4- 13). Tat transportable alkaline phosphatases include PhoA which is fused to the Tat signal peptide from TorA (TorA-PhoA) (section 4.5.1) and the native Tat substrate PhoX (section 4.5.2).

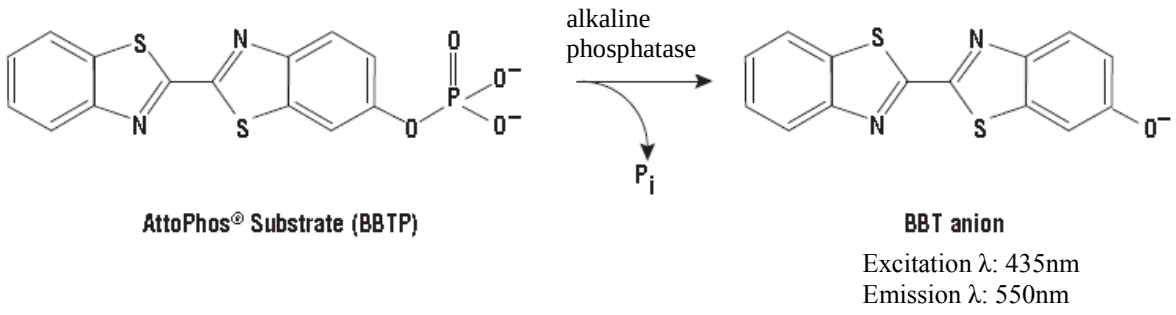


Figure 4- 13: The reaction of AttoPhos® substrate with alkaline phosphatase

4.5.1 TorA-PhoA

PhoA is the major APase in *E. coli*. In *E. coli*, PhoA is transported by the Sec system as it requires the formation of disulfides bonds which cannot be achieved in the reducing environment in the cytoplasm. However, TorA-PhoA, which is a fusion construct of PhoA with the Tat signal peptide from TorA, has been shown to be transported *in vitro* by IMVs. TorA-PhoA used in reported experiments was synthesized by *in vitro* translation/ transcription in the presence of the oxidizing agent, glutathione disulfide (GSSG), to allow formation of disulfide bonds in TorA-PhoA. Transport was only observed when TorA-PhoA was properly folded with its disulfide bonds formed (Panahandeh *et. al.* 2008). I tried to

purify TorA-PhoA in its oxidized form either by adding GSSG during the purification process or by expressing *torA-phoA* in an *E. coli* mutant strain with an oxidizing cytoplasm. However, oxidized TorA-PhoA could not be easily purified (data not shown) because expression levels were low and the TorA-PhoA produced was mostly found in inclusion bodies. Hence, TorA-PhoA could not be used as a substrate for *in vitro* assays.

4.5.2 PhoX

It has been reported that some organisms such as *Pseudomonas fluorescens*, *Vibrio cholerae* and *Pasteurella multocida* use a different APase known as PhoX (Roy *et. al.* 1982; Monds *et. al.* 2006; Wu *et. al.* 2007). Unlike PhoA, PhoX has been shown to be dependent on the Tat system for transport.

4.5.2.1 *P. fluorescens* PhoX

I tried to purify *P. fluorescens* PhoX using *E. coli* as the host cell. *P. fluorescens* is a γ-proteobacterium like *E. coli*. Therefore, Tat substrates of *P. fluorescens* are likely to be compatible with the *E. coli* Tat machinery. *P. fluorescens* PhoX was expressed with a N-terminally fused GST. The N-terminal GST was separated from PhoX by a 3C protease cleavage site and served to protect the Tat signal peptide to allow purification of full length *P. fluorescens* PhoX. A His-tag was placed at the C-terminus to allow for immunodetection. Figure 4- 14 (lane 1) showed that GST-*P. fluorescens* PhoX was purified but it seemed that cleavage of GST-*P. fluorescens* PhoX to form full length *P. fluorescens* PhoX had occurred prior to addition of 3C-protease. The remaining GST-*P. fluorescens* PhoX was cleaved in a further step of the purification by 3C-protease to produce full length *P. fluorescens* PhoX. The majority of the purified protein migrated as a band of ~66kDa by SDS-PAGE (Figure 4-14, lane 3). This was assumed to be full length *P. fluorescens* $PhoX_{His}$ (theoretical molecular weight: 69.5kDa).

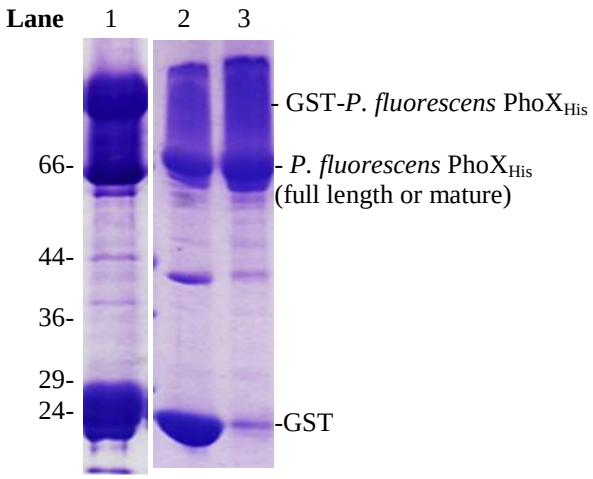


Figure 4- 14: Purification of *P. fluorescens* PhoX$_{His}$ by GSTrap affinity chromatography
A cleared cell lysate was prepared from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 overproducing GST-PhoX$_{His}$ from the plasmid pGEX-PF_PhoX$_{His}$. The lysate was applied to a GSTrap affinity column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5). Bound GST-PhoX$_{His}$ was eluted with 10mM glutathione. Glutathione was removed by ultrafiltration using a 30kDa molecular weight cut off (MWCO) membrane. GST was cleaved from GST-PhoX$_{His}$ by GST-3C protease at 4°C overnight. The cleaved sample was reapplied to a GSTrap affinity column equilibrated in resuspension buffer. PhoX$_{His}$ will elute in the flow through while free GST and GST-3C protease remained on the column. Fractions collected from both GSTrap affinity columns were analysed by 12.5% SDS-PAGE Proteins were visualised by Coomassie Brilliant Blue staining. The molecular masses (kDa) of standard proteins are given to the left of the gel. Lane1: Eluate from first GSTrap affinity column purification; Lane2: After 3C protease cleavage but before the second GSTrap affinity column purification; Lane 3: Flow through of second GSTrap affinity column purification.

The Tat transport competency of *P. fluorescens* PhoX was tested using the modified *in vitro* Tat transport assay (Figure 4- 15). Two bands were observed when *P. fluorescens* PhoX was added to non-energized IMVs (Figure 4- 15, lane 2). The approximate molecular weights of the two bands suggested that Band 1 (>66kDa) could be full length *P. fluorescens* *P. fluorescens* PhoX while Band 2 (~66kDa) could be mature *P. fluorescens* PhoX (theoretical molecular weight: 64.5kDa). This could mean that not all of the *P. fluorescens* PhoX added to the assay was Tat transport compatible. Addition of proteinase K (PK) resulted in digestion of Band 1 but not Band 2 (Figure 4- 15, lane1). As the molecular weight of Band 2 was similar to that expected for mature *P. fluorescens* PhoX, it was impossible to distinguish between transported and non-transported *P. fluorescens* PhoX. Therefore, the results obtained from this experiment were inconclusive.

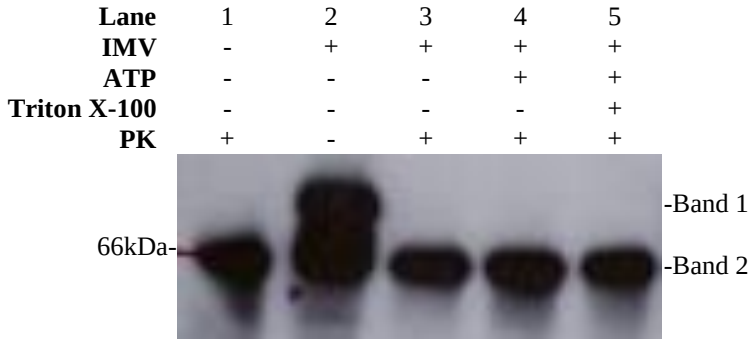


Figure 4- 15: *In vitro* Tat transport of *P. fluorescens* PhoX
IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer containing 4mM ATP and 100nM *P. fluorescens* PhoX. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added following transport if required. PK was added to digest any untranslocated *P. fluorescens* PhoX that remained in the reaction mixture. Detection of C-terminal His-tagged *P. fluorescens* PhoX was performed using Penta-His HRP conjugate (Qiagen) at a 1:5,000 dilution. The position of the 66kDa standard protein is indicated on the left.

4.5.2.2 *V. fischeri* PhoX

Since transport results obtained using *P. fluorescens* PhoX was inconclusive due to incomplete digestion by proteinase K, I had to search for an alternative substitute. I chose to test PhoX from *V. fischeri*, an organism which is reasonably closely related to enterobacteria. Similar to *P. fluorescens* PhoX, *V. fischeri* PhoX was also produced with a N-terminal GST domain and a C-terminal His-tag. N-terminal GST was cleaved at the 3C cleavage site between the GST domain and *V. fischeri* PhoX to give full length *V. fischeri* PhoX. The final purified sample was analysed by SDS-PAGE (Figure 4- 16). The purified sample migrated as a band of molecular weight ~66kDa and was assumed to be full length *V. fischeri* PhoX (theoretical molecular weight: 69.1kDa).

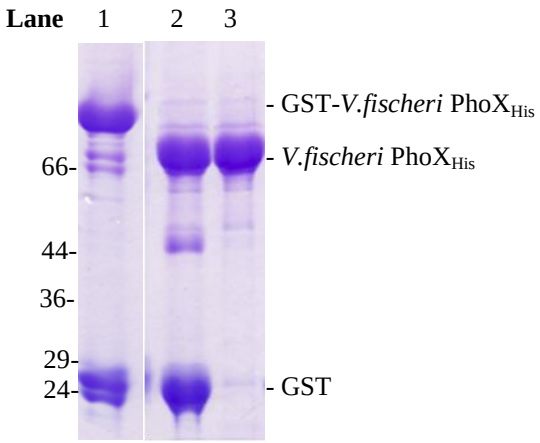


Figure 4- 16: Purification of *V. fischeri* PhoX$_{His}$ by GSTrap affinity chromatography

A cleared cell lysate was prepared from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 overproducing GST-PhoX$_{His}$ from the plasmid pGEX-VF_PhoX$_{His}$. The lysate was applied to a GSTrap affinity column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5). Bound GST-PhoX$_{His}$ was eluted with 10mM glutathione. Glutathione was removed by ultrafiltration using a 30kDa MWCO membrane. GST was cleaved from GST-PhoX$_{His}$ by GST-3C protease at 4°C overnight. The cleaved sample was reapplied to a GSTrap affinity column equilibrated in resuspension buffer. PhoX$_{His}$ will elute in the flow through while free GST and GST-3C protease remained on the column. Fractions collected from both GSTrap affinity columns were analysed by 12.5% SDS-PAGE. Proteins were visualised by Coomassie Brilliant Blue staining. Lane1: Eluate from first GSTrap affinity column purification; Lane2: After 3C protease cleavage but before the second GSTrap affinity column purification; Lane 3: Flow through of second GSTrap affinity column purification. The molecular masses (kDa) of standard proteins are given to the left of the gel.

V. fischeri PhoX was subsequently used for *in vitro* transport assays. Unlike *P. fluorescens* PhoX, *V. fischeri* PhoX was completely digested by proteinase K (Figure 4- 17, lane 2). Although *V. fischeri* PhoX is not found in *E. coli* and is therefore not a native *E. coli* Tat substrate, it was successfully transported when IMVs were energized by NADH (lanes 6 and 7). However, *V. fischeri* PhoX was not transported when IMVs were energized by ATP (lanes 3 and 4). This contradicted the observations in chapter 3 (Figure 3- 23) where it was shown that IMV energization by ATP resulted in higher transport efficiency. A simple explanation is that PhoX, being a monophosphoesterase, will dephosphorylate ATP but not NADH. This is confirmed in chapter 5.

Lane	1	2	3*	4*	5	6*	7*	8
ATP	-	-	+	+	+	-	-	-
NADH	-	-	-	-	-	+	+	+
Triton X-100	-	-	-	-	+	-	-	+
PK	-	+	+	+	+	+	+	+



- Full length *V. fischeri* PhoX
- Mature *V. fischeri* PhoX

Figure 4- 17: *In vitro* transport assay of *V. fischeri* PhoX

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer containing 4mM ATP and 100nM *V. fischeri* PhoX. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added following transport if required. PK was added to digest any untranslocated *V. fischeri* PhoX that remained in the reaction mixture. Detection of C-terminal His-tagged *V. fischeri* PhoX was performed using Penta-His HRP conjugate (Qiagen) at a 1:5,000 dilution. *: Lanes 3 and 4 are duplicates of translocation assays done with IMVs energized by ATP; Lanes 6 and 7 are duplicates of translocation assays done with IMVs energized by NADH.

Although *V. fischeri* PhoX was transported, the transport efficiency was very low. This could partially be due to using NADH to energize IMVs. As discussed in chapter 3 (see 3.5.1), the transport efficiency of IMVs energized with NADH was lower than IMVs energized with ATP. Another explanation would be that not all of the *V. fischeri* PhoX added to the transport assay was transport competent. A simple alkaline phosphatase test using pNPP as substrate showed that this batch of *V. fischeri* PhoX was not enzymatically active even when supplemented with Ca^{2+} (Table 4- 2), which, as shown in chapter 5, is one of the metal cofactors of PhoX. It is hypothesized that this is due to the glutathione used to elute the GST fused *V. fischeri* PhoX from the GSTrap column. As I will show in the chapter 5 (see 5.4.1), besides Ca^{2+} ions, PhoX also requires two Fe^{3+} ions for activity. Reductants, such as glutathione, will reduce Fe^{3+} to Fe^{2+}, which results in loss of activity. As the Tat system is specific for only folded, functional proteins, inactive *V. fischeri* PhoX might be transport incompetent.

To avoid this issue, I purified the GST fused *V. fischeri* PhoX using the C-terminal His-tag instead. Figure 4- 18 (lane 1) showed that two major products were purified from the Ni^{2+} affinity column. The first band with a molecular weight of >66kDa is likely to be GST-*V. fischeri* $PhoX_{His}$ and the second band with a molecular weight of ~66kDa is likely to be *V. fischeri* $PhoX_{His}$ though it is not possible to distinguish between full length and mature *V. fischeri* $PhoX_{His}$. The sample was treated with GST-3C protease and loaded onto a GSTrap affinity column. The GST cleaved full length *V. fischeri* $PhoX_{His}$ was purified from free GST and GST-3C protease by GSTrap affinity chromatography. The purified sample had a molecular weight of ~66kDa, which is assumed to be full length *V. fischeri* PhoX. (Figure 4-18, lane 3). This purification method avoided the use of glutathione in the presence of *V. fischeri* PhoX. This batch of *V. fischeri* PhoX was enzymatically active but only in the presence of the metal cofactor Ca^{2+} (Table 4- 2).

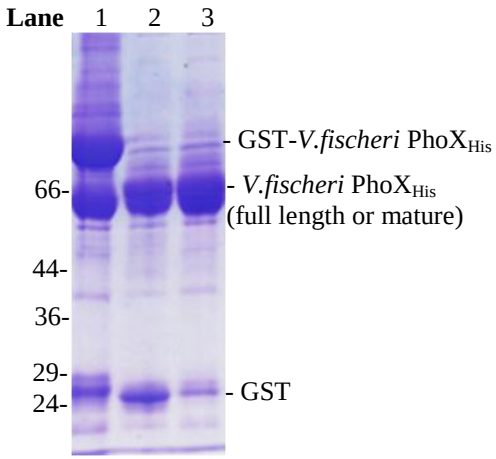


Figure 4- 18: Purification of *V. fischeri* PhoX by Ni^{2+} affinity chromatography

A cleared cell lysate was prepared from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 overproducing GST-$PhoX_{His}$ from the plasmid pGEX-VF_$PhoX_{His}$. The lysate was applied to a Ni^{2+} affinity column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5) with 30mM imidazole. Non-specifically bound proteins were removed by washing the column with 30mM imidazole and bound proteins were eluted in a linear imidazole gradient from 30mM to 500mM imidazole. Peak fractions were collected and imidazole was removed by ultrafiltration using a 30 kDa MWCO membrane. GST was cleaved from GST-$PhoX_{His}$ by GST-3C protease in resuspension buffer at 4°C overnight. The cleaved sample was applied to a GSTrap affinity column equilibrated in resuspension buffer. $PhoX_{His}$ will elute in the flow through while free GST and GST-3C protease remained on the column. Peak fractions from the Ni^{2+} affinity column and GSTrap affinity column were analysed by 12.5% SDS-PAGE. Proteins were visualised by Coomassie Brilliant Blue staining. Lane1: Eluate from first Ni^{2+} affinity column purification; Lane2: After 3C protease cleavage but before GSTrap affinity column purification; Lane 3: Flow through of GSTrap affinity column purification. The molecular masses (kDa) of standard proteins are given to the left of the gel.

Table 4- 2: Activity assay of *V. fischeri* PhoX purified via different protocols

Activity assay was done in 1M Tris-HCl buffer, pH8.0, using pNPP as substrate. Reactions were done in the absence or presence of 1mM Ca^{2+}. The reaction was quenched after 10min by addition of 1M NaOH and measurement obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX.

Purification method	Specific activity (µmol/min/mg PhoX)	
	$-Ca^{2+}$	$+Ca^{2+}$
Via N-terminal GST domain (in the presence of glutathione)	0	0
Via C-terminal His-tag	0	1.85

Figure 4- 19 shows that *V. fischeri* PhoX purified in the absence of glutathione was transported. A higher transport efficiency was achieved in the presence of Ca^{2+}, suggesting that *V. fischeri* PhoX has to be enzymatically active for more efficient transport. *V. fischeri* PhoX can also cleave the fluorogenic AttoPhos® substrate (data not shown). Hence, PhoX could be used in TIRF based single molecule fluorescence techniques.

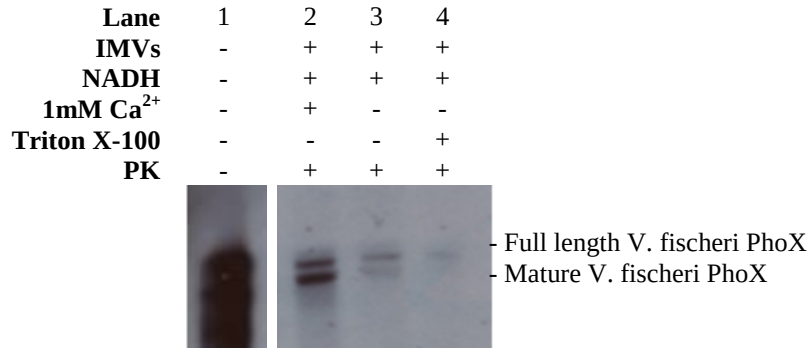


Figure 4- 19: *In vitro* transport assay of *V. fischeri* PhoX purified via C-terminal His-tag

IMVs were prepared from *E. coli* strain BL21 (DE3)-pLysS overproducing TatABC from plasmid p8737. IMVs (A_{280nm}=10) were added to translocation buffer (supplemented with 1mM Ca^{2+} when necessary) which contained 4mM ATP and 100nM *V. fischeri* PhoX purified via the C-terminal His-tag. The reaction was left at 37°C for 30min before the reaction was quenched by placing it on ice for 2min. Triton X-100 was added following transport if required. PK was added to digest any untranslocated *V. fischeri* PhoX that still remained in the reaction mixture. Detection of C-terminal His-tag of *V. fischeri* PhoX was performed using Penta-His HRP conjugate (Qiagen) at a 1:5,000 dilution.

One complication of purifying a Tat substrate via the C-terminal His tag alone is that not all of the purified protein will have an intact Tat signal sequence as the unstructured Tat signal sequence is often spontaneously cleaved. It was difficult to determine what proportion of *V. fischeri* PhoX that was purified via the C-terminal His tag was full length as the main band was not sufficiently resolved by SDS-PAGE analysis (Figure 4- 18). Cleavage of Tat signal peptide occurs even when Δ*tatABCD*Δ*tatE E. coli* mutant strains are used as host cells for protein expression. This problem was slightly alleviated by fusing the full length Tat substrate with a bulky N-terminal protein, such as GST. In addition, purification of the protein via the N-terminal GST will ensure that only full length protein is purified. Therefore, when *V. fischeri* PhoX was instead purified using its C-terminal His tag, the purified protein would be a mixture of full length and mature protein. Separation of the full length protein from the mature protein would be difficult as the difference in size was not significant enough to be resolved on a gel filtration column.

The transport efficiency of *V. fischeri* PhoX could be improved if it was purified using a different N-terminal cleavable tag, such as SUMO. This would ensure purification of full length *V. fischeri* PhoX without the complication of glutathione associated with elution of the GST fused *V. fischeri* PhoX. Another option would be to express *V. fischeri tatABC* in *E. coli* as the *V. fischeri* Tat components might recognize its own native substrate better than the *E. coli* Tat system.

4.5.3 Discussion

In vitro assays can be applied to mechanistic studies of Tat transport. However, the data obtained is an average of all molecules in the sample. This method does not take into account the heterogeneity of the sample and the contribution of each individual component. Therefore, single molecule fluorescence experiments are required to resolve these issues. This could be achieved by incorporating IMVs enriched with TatABC into artificial lipid bilayers that are in contact with an aqueous droplet containing the fluorescent substrate, CueO_GS3^{633}, which I have previously made (see 3.6.2). Detection of membrane associated or transported CueO_GS3^{633} using TIRF microscopy should then be possible.

An alternative method to detect Tat transport by fluorsecence microscopy is to use a Tat substrate that converts a precursor compound to a fluorescent product. This compound would be placed on the *trans* side of the membrane bilayers. Therefore, any fluorescence detected will be due to the transport of the Tat substrate across the membrane by the Tat system. To achieve this, I have successfully transported purified *V. fischeri* PhoX, an alkaline phosphatase which is a native Tat substrate into IMVs enriched with *E. coli* TatABC. *V. fischeri* PhoX was able to hydrolyze the fluorogenic AttoPhos® to produce a fluorescent product. Hence, I now have the biochemical tools to study the mechanism of the Tat transport cycle using single molecule fluorescence microscopy methods in the future.

Chapter 5:

Biochemical and structural studies of *Pseudomonas fluorescens* PhoX

5.1 Introduction

Phosphorus is one of the main mineral nutrients required by plants, animals and microorganisms. It is a major constituent of phospholipids in cell membranes, forms the phosphate-ester backbone of nucleic acids and is a part of energy storing molecules such as ATP. Unfortunately, phosphates are not abundant in soil or in the marine environment and are often the limiting factor for growth of plants and microorganisms. One of the methods employed by microorganisms to overcome this issue is to increase the synthesis of phosphatases to release inorganic phosphate (P_i), which is the preferred source of phosphate for bacteria, from organic phosphates in the environment.

Alkaline phosphatases (APases), as their name suggests, function optimally at alkaline pH. Synthesis of this class of enzymes is increased under phosphorous-limiting conditions. In Gram-negative bacteria, APases are usually found in the periplasm. Phosphorylated substrates diffuse into the periplasm via porins in the outer membrane. Phosphorylated substrates are then hydrolyzed in the periplasm and P_i is transported into the cytoplasm via the P_i channel in the inner membrane (Torriani 1990).

The most extensively studied APase is PhoA from *Escherichia coli*. PhoA is a homodimer and contains two Zn^{2+} ions and a Mg^{2+} ion in its active site. A covalent phosphoenzyme intermediate is produced during catalysis on a conserved Ser residue (Kim and Wyckoff 1991; Holtz and Kantrowitz 1999; Holtz *et. al.* 1999). Synthesis of PhoA is regulated by the PhoR/PhoB two-component signal transduction system. PhoB becomes phosphorylated under P_i-limiting conditions by the histidine kinase PhoR. Phosphorylated PhoB binds to the Pho box, which is located upstream of *phoA*, and promotes interaction with the σ70 subunit of RNA polymerase. This results in an increase in the transcription of *phoA* (Torriani 1990; Wanner 1993).

However, not all organisms have PhoA or homologues of PhoA. Most organisms that lack PhoA have an alternative APase, known as PhoX. A few organisms such as *Kineococcus radiotolerans* and *Citrobacter koseri* have both PhoX and a PhoA-related protein. PhoX and PhoX-related proteins can be found in all three super kingdoms, eubacteria, archaea and eukarya. Unlike its bacterial counterparts, the eukaryotic PhoX has two PhoX domains, which is believed to have originated from gene duplication and fusion. All characterized PhoX so far have shown that it is a calcium–dependent phosphatase (Hallmann 1999; Sebastian and Ammerman 2009; Zaheer *et. al.* 2009).

Although there is no sequence similarity between the PhoA and PhoX families, these alkaline phosphatases appear to be functionally equivalent. However, PhoA is believed to act as a homodimer while studies so far have shown that PhoX is most likely a monomer (Coleman 1992; Zaheer *et. al.* 2009). In addition, PhoX is transported to the periplasm via the Tat pathway (Monds *et. al.* 2006; Wu *et. al.* 2007; Van Mourik *et. al.* 2008). Results from section 4.5.2.2 in chapter 4 also showed that *V. fischeri* PhoX is transported *in vitro* by the *E. coli* Tat transport system. On the other hand, PhoA is transported to the periplasm by the Sec system in its unfolded state as PhoA contains two intramolecular disulfide bonds required for folding which cannot form in the reducing cytoplasm of the cell (Sone *et. al.* 1997).

The PhoX family represents a new class of alkaline phosphates that are distinct from the classic PhoA family. This work will mostly focus on PhoX from the Gram-negative gamma proteobacteria, *Pseudomonas fluorescens*. *P. fluorescens* is a common soil bacterium that can improve plant health through nutrient cycling, pathogen antagonism and induction of plant defences (Silby *et. al.* 2009). Since the concentration of soluble phosphate in soil is very low (less than 10µM) (Bieleski 1973), *P. fluorescens* has to adapt to this limiting P_i environment by up-regulating the synthesis of APases. Although *P. fluorescens* does not have a PhoA homologue, at least two other phosphatases were found to be up-regulated in *P.*

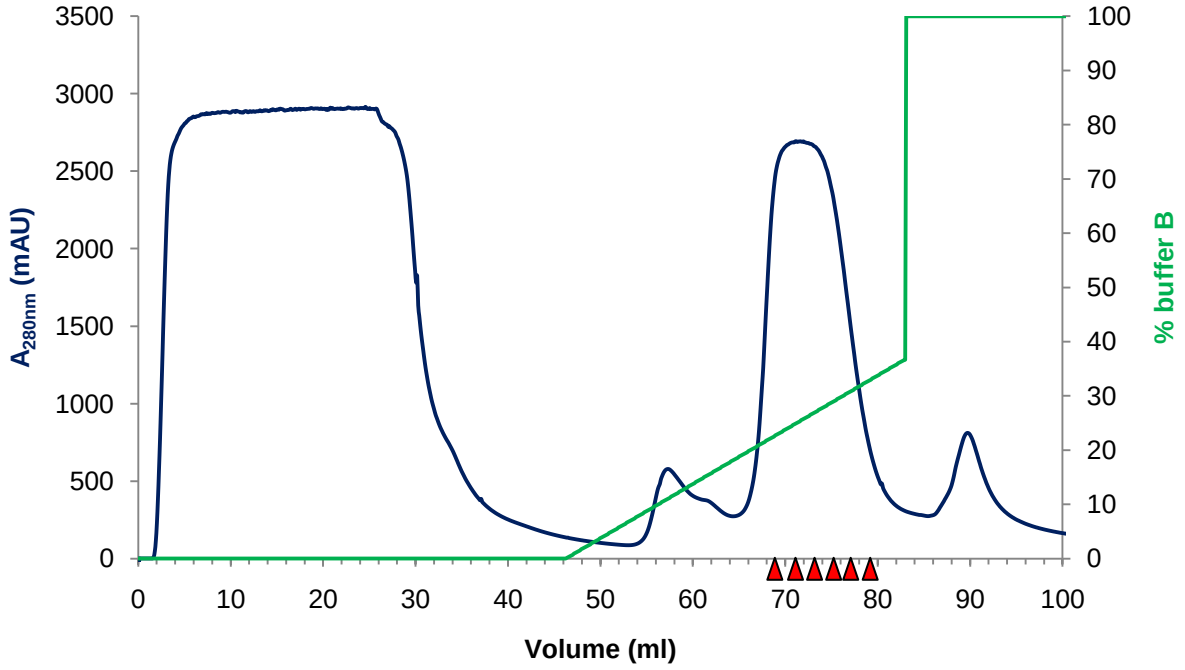
fluorescens under P_i-limiting environments. This up-regulation is controlled by a two-component regulatory system, similar to the PhoR/PhoB systems in PhoA. These phosphatases were identified as PhoD and PhoX, with PhoX being the main alkaline phosphatase (Monds *et. al.* 2006).

As shown in the previous chapter, PhoX can be used to study the Tat system *in vitro* by single molecule fluorescence microscopy due to its ability to dephosphorylate a fluorogenic compound. Knowledge of the characteristics of PhoX would aid in understanding the relationship between PhoX and the Tat system. However, not much is known about PhoX. The main aim of this chapter is to learn more about the structure and function of PhoX.

5.2 Purification of *P. fluorescens* PhoX

Full length *P. fluorescens* PhoX was expressed in the *ΔtatABC E. coli* strain C43 (DE3) (pRE4) and purified via the C-terminal His-tag by Ni^{2+} affinity chromatography (Figure 5- 1). This was followed by purification by size exclusion chromatography (Figure 5-2). The final yield of PhoX obtained was ~50mg per liter of cells. PhoX was purified as a purple coloured protein with an absorbance maximum at ~550nm (Figure 5- 3). This suggests that PhoX probably contains at least one transition metal as the colour of the protein is likely due to ligand-to-metal charge transfer.

A)



B)

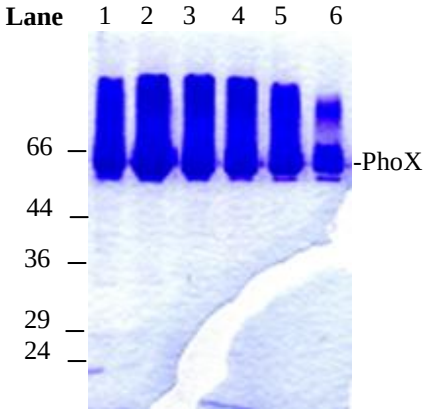
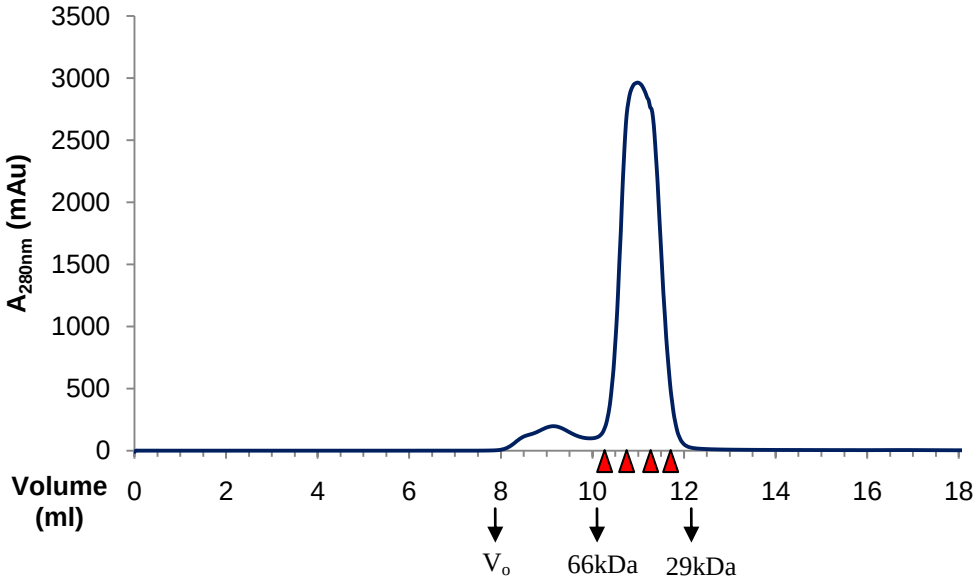


Figure 5- 1: Purification of PhoX by Ni^{2+} affinity chromatography

(A) A cleared cell lysate was prepared from the Δ*tatABC E. coli* strain C43 (DE3)-pREP4 overproducing $PhoX_{His}$ from the plasmid pQE60-PF_$PhoX_{His}$. The lysate was applied to a Ni^{2+} affinity column equilibrated in buffer A (20mM MOPS, 200mM NaCl, 30mM imidazole, pH7.5). Non-specifically bound proteins were removed by washing the column with buffer A and bound proteins were eluted at 1ml/min with buffer B (20mM MOPS, 200mM NaCl, 1M imidazole, pH7.5) in a linear imidazole gradient from 30mM to 500mM imidazole. 2ml fractions were collected. Chromatography was done at room temperature. Green line: % buffer B, blue line: A_{280nm} (B) Peak fractions from the imidazole gradient elution were analysed by 10% SDS-PAGE and visualised by Coomassie Brilliant Blue staining. Lanes 1-6 correspond to successive fractions from the imidazole elution gradient (Fractions were collected at volume 68ml-80ml as marked with red triangles in A). The molecular masses (kDa) of standard proteins are given to the left, and the position of $PhoX_{His}$ to the right, of the gel.

A)



B)

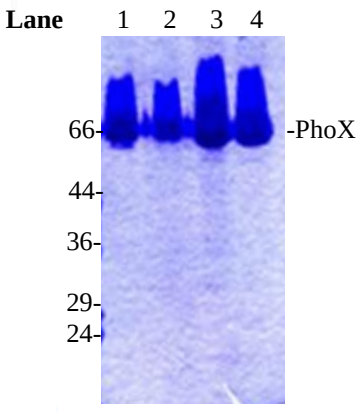


Figure 5- 2: Size exclusion chromatography of PhoX
(A) Lanes 1-6 from the Ni^{2+} affinity chromatography column shown in

Figure 5- 1were pooled. The sample was concentrated to 500µL by ultrafiltration using a 30 kDa MWCO membrane and loaded onto a Superdex-75 column equilibrated in resuspension buffer (20mM MOPS, 200mM NaCl, pH7.5). The column was run at a flow rate of 0.4ml/min at room temperature and 0.5ml fractions were collected. The elution positions of water soluble standard proteins are shown below the column profile and labelled with the protein mass (kDa). Standards used were (V_o) alcohol dehydrogenase, (66 kDa) bovine serum albumin and (29 kDa) carbonic anhydrase. (B) Fractions from the size exclusion chromatography column were analysed by SDS-PAGE and proteins visualised by Coomassie Brilliant Blue staining. Lanes 1-4 represent fractions collected from peak fractions (*i.e.* marked with red triangles). The molecular masses (kDa) of standard proteins are given to the left, and the position of $PhoX_{His}$ to the right, of the gel.

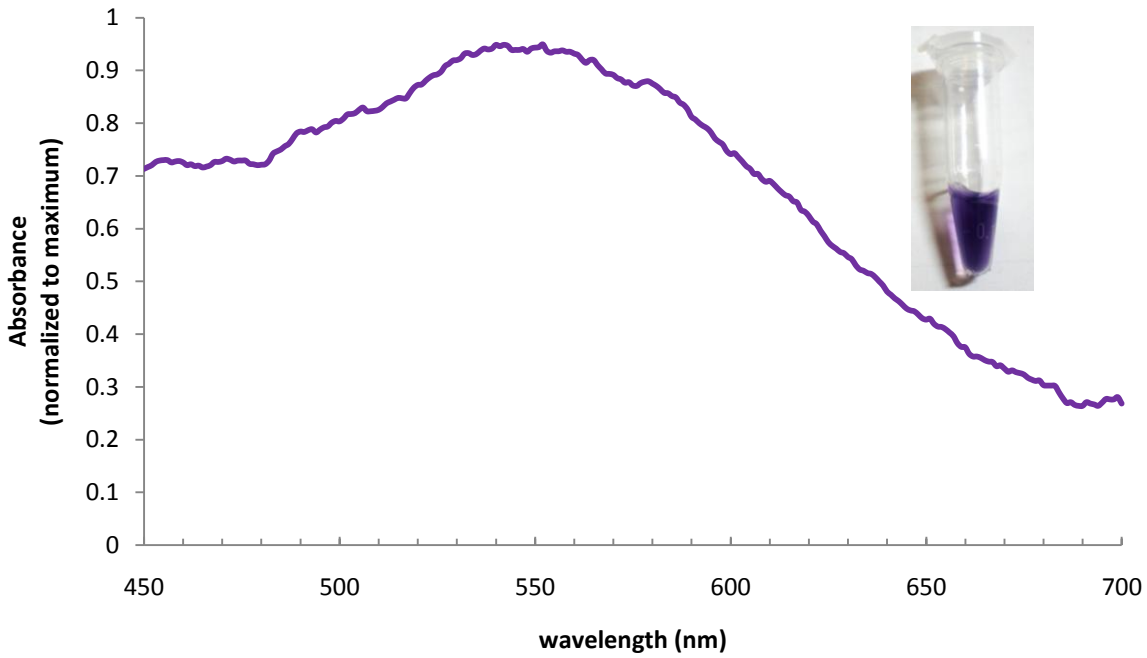
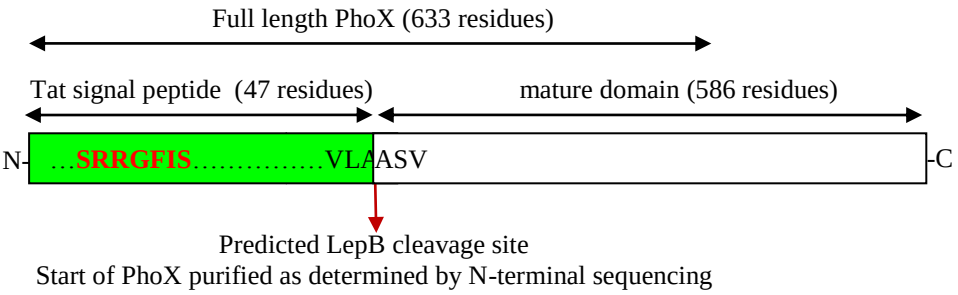


Figure 5- 3: PhoX is a coloured protein
Wavelength scan of PhoX (400nm-700nm). Insert shows the colour of PhoX after purification. Concentration of PhoX is 15mg/ml.

SDS-PAGE analysis showed that PhoX had an approximate molecular weight of 66kDa. I attempted to determine the exact molecular weight of PhoX by electrospray ionisation mass spectrometry (ESI-MS) but failed to obtain a clear peak. This could be because PhoX was not completely unfolded under ESI-MS conditions (1% trifluoroacetic acid, 80% acetonitrile) and could not be properly ionized. N-terminal sequencing data (Figure 5- 4A) showed that the purified PhoX lacked the Tat signal peptide although full length *phoX* was expressed in cells. N-terminal sequencing showed that the sequence of the purified PhoX started where the signal peptidase, LepB, was predicted to cleave after Tat substrates are transported to the periplasm (Lüke *et. al.* 2009). This prediction was made according to the TatP prediction server (Bendtsen *et. al.* 2005). The mature PhoX would have a molecular weight of 64.6kDa, which roughly correlated to that observed in gel filtration chromatography and SDS-PAGE analysis.

A)



B)

MSLLEENQSTDLEKMVGLSRRGFISAGALCGAAMFLGGNLLSRSVLAASVSAGNSRLLGFDSIPAATTDT
ISLPKGYKSSVLISWGQPLHKNGPAFDPSGNGTAAAQEVQFGDNNDGMSLFEFPGEKNRALMAINNEYTN
YRYLYPHGGMPQSAEDVRKALACEGVSVIEVQRKNGQWQFVQGSRYNRRIHGNSPLRISGPAAGHELMKT
SADKHGKKVLGTFQNCANGKTPWGTYLTCEENFTDCFGSSNAQQQFDPAQKRYGVSAASREINWHPFDPR
FDMAKNPNELNRHGWVVEIDPFDPQSTPVKRTALGRFKHENAALAETDDGRAVVYMGDDERGEFIYKFVS
RDKINHRNAKANRDILDHGTLYVARFDAGDGNPDHPKGQGQWIELTHGKNGIDASSGFADQAEVLIHARL
AASVVGATRMDRPEWIVVSPKDGQVYCTLTNNAKRGEDGQPVGGPNPREKNVYGQILRWRTDRDDHASKT
FAWDLFVVAGNPSVHAGTPKGGSSNITPQNMFNSPDGLGFDKAGRLWILTDGDSSNAGDFAGMGNNQMLC
ADPATGEIRRFMVGPIGCEVTGISFSPDQKTLFVGIQHPGENGGSTFPEHLPNGKPRSSVMAITREDGGI
VGA

Figure 5- 4: Sequence of *P. fluorescens* PhoX

Full length PhoX consists of the N-terminal signal peptide and the mature domain. (A) Although PhoX was expressed as full length PhoX, the purified PhoX had lost its N-terminal Tat signal peptide during the purification process. Only the mature domain remained, as determined by N-terminal sequencing. The cleavage site of the Tat signal peptide is determined by TatP prediction server (Bendtsen *et. al.* 2005). (B) Sequence of full length sequence of PhoX with the Tat signal sequence in green.

I tried to overcome the problem of cleavage of full length PhoX by LepB by mutating the LepB cleavage site of full length PhoX from LAAS to LSWS. According to the TatP prediction server, LSWS is not recognized by LepB. Full length PhoX was purified in the same way as before. However, N-terminal sequencing of this purified protein (data not shown) showed that only the mature domain remained after purification. This meant that the cleavage of the N-terminal Tat signal peptide was not solely due to LepB. This could be because the Tat signal peptide was unstructured and was therefore exposed to other non-specific proteases.

Cleavage of Tat signal peptide from full length Tat substrates during purification is a common phenomenon, especially if purified in the absence of a N-terminal fusion tag. Although PhoX was expressed in a *ΔtatABCDΔtatE* background strain and would not have been exported to the periplasm during growth, PhoX would be exposed to LepB and other periplasmic proteases during cell lysis prior to protein purification. Addition of a mixture of protease inhibitors failed to resolve this issue.

5.3 Characterization of *P. fluorescens* PhoX

5.3.1 Determination of optimum pH for PhoX activity

Phosphatases are classified as an alkaline phosphatase if they are active under alkaline conditions. The activity of *P. fluorescens* PhoX was determined using p-nitrophenyl phosphate (pNPP) as substrate. pNPP is a chromogenic substrate which is cleaved by phosphatases to form yellow p-nitrophenyl that absorbs at 420nm. PhoX from *P. fluorescens* was shown to be most active at pH8.0 (Figure 5- 5). This is similar to the pH optimum of PhoX from *Vibrio cholera* (Roy *et. al.* 1982) but is lower than the optimum pH of PhoX from *Campylobacter jejuni* and *Pasteurella multocida*, which have optimum activities at pH10.0 (Wu *et. al.* 2007; Van Mourik *et. al.* 2008).

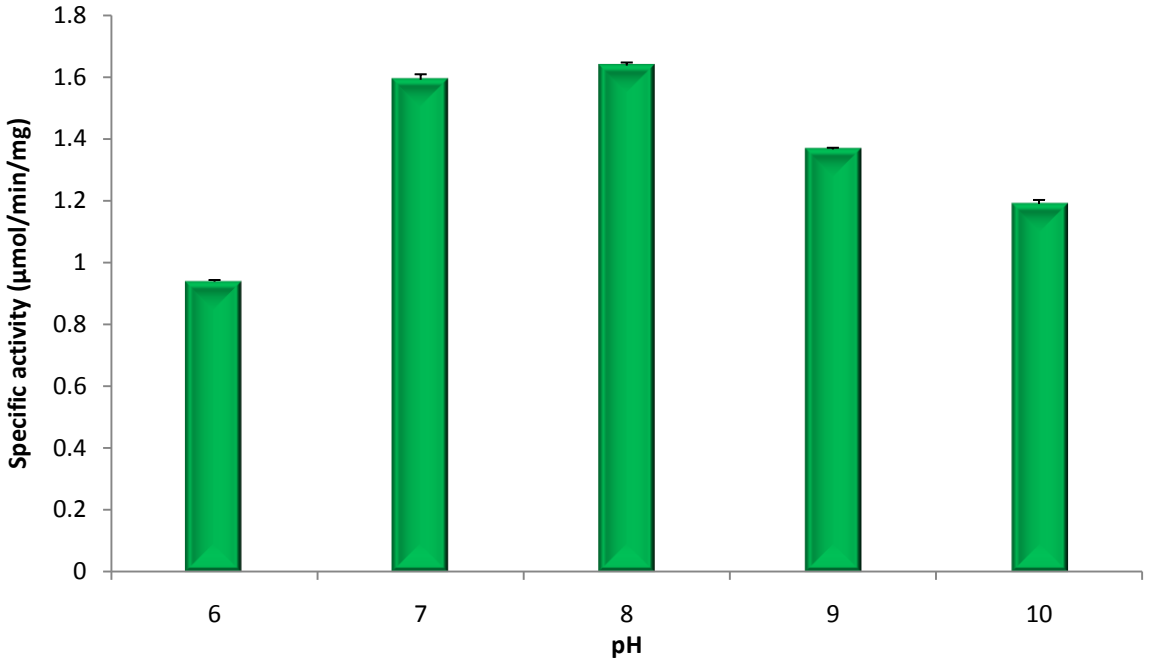


Figure 5- 5: pH optimum of PhoX activity

The pH optimum of PhoX activity was determined in the presence of 5mM Ca^{2+} in 1M Tris-HCl buffer of the indicated pH using pNPP (0.04% w/v) as substrate. The reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Formation of p-nitrophenyl was determined by measuring the absorbance at 420nm. Specific activity was calculated as µmol of P_i released per min per mg of PhoX. Error bar represents standard error for three independent experiments.

5.3.2 Metal dependency of PhoX

Since most alkaline phosphatases require metal ions to function, I wanted to test if this was also true for PhoX. EDTA, a chelating agent normally used to sequester metal ions, was added to the pNPP reaction assay to strip any metal ions bound to PhoX. Figure 5- 6 shows that addition of 10µM EDTA significantly reduced the activity of PhoX while 100µM EDTA completely abolished the activity of PhoX. This suggests that PhoX requires metal ions for activity.

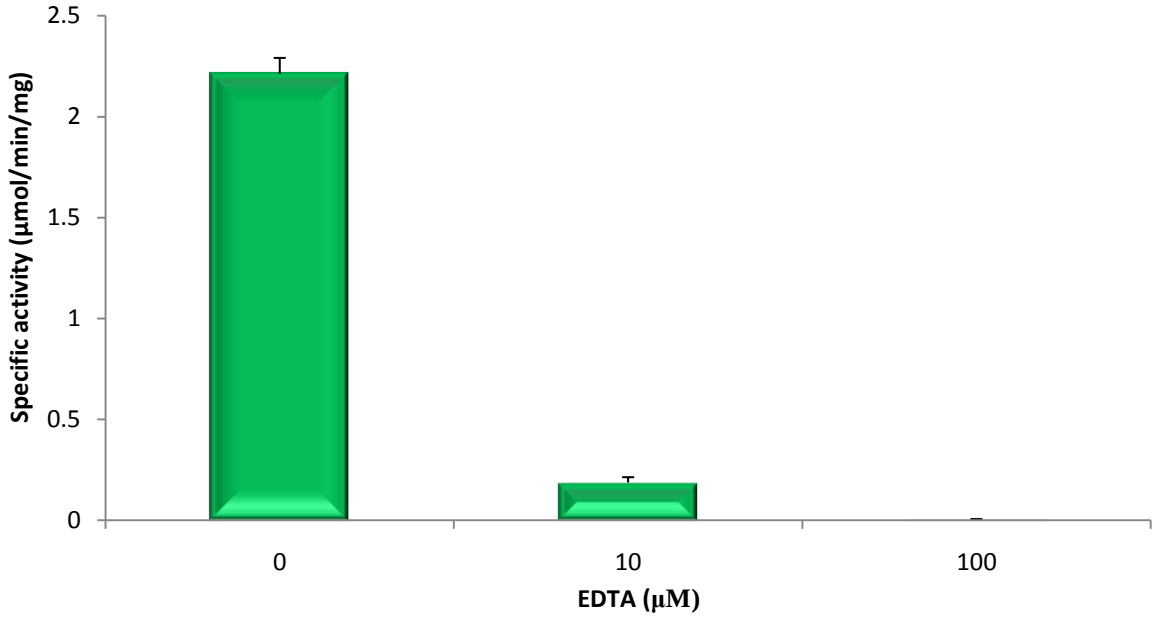


Figure 5- 6: PhoX activity in the presence of EDTA

The reaction assay contained 0.04% w/v pNPP and 5µg of PhoX in 1M Tris-HCl, pH8.0. EDTA was added to the reaction if required. The reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Measurement was obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX. Error bar represents standard error for three independent experiments.

It was also observed that the absorbance maximum of *P. fluorescens* PhoX was shifted in the presence of EDTA from ~550nm (purple) to ~570nm (blue) (Figure 5- 7). This shift in maximum wavelength absorbance suggests that some of the metal ions in PhoX were not removed by EDTA as native PhoX would be colourless due to the lack of metal to ligand charge transfer. However, the environment in which the remaining metal ions are in could have been perturbed by the removal of other metal ions. This could have an effect on the ligand-to-metal charge transfer, thereby resulting in a change in colour of PhoX.

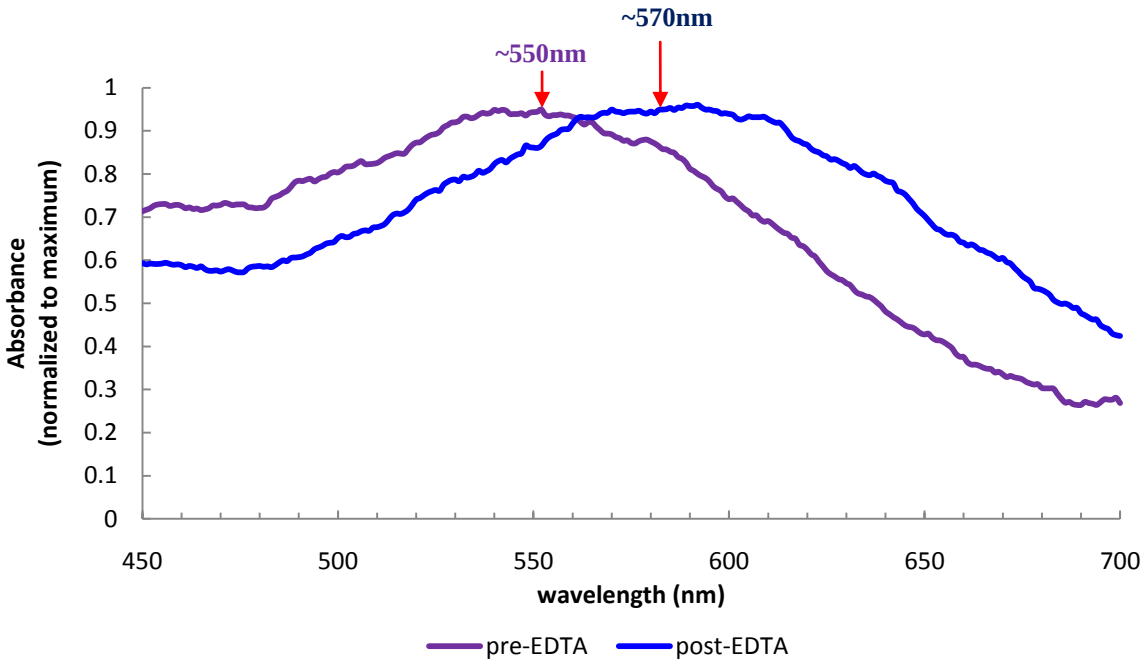


Figure 5- 7: Effect of EDTA on the visible absorption spectrum of PhoX
The absorbance maximum of PhoX (15mg/ml) before and after treatment with 100µM EDTA. PhoX was in 20mM MOPS, 200mM NaCl, pH7.5.

Although having EDTA present in the reaction assay abolished the activity of PhoX, once EDTA was removed from the solution, PhoX regained most of its enzymatic activity (Figure 5- 8). It was also observed that the colour of PhoX returned to purple. One explanation would be that EDTA was unable to completely remove all the metal ions from PhoX and instead formed a complex with the PhoX-metal complex which prevented substrates from entering the active site. Therefore, when EDTA was removed, access to the active site was restored. Another possibility is that the milli-Q water used for the reaction assay was contaminated with trace amounts of metal ions and PhoX was able to recruit these metal ions for activity.

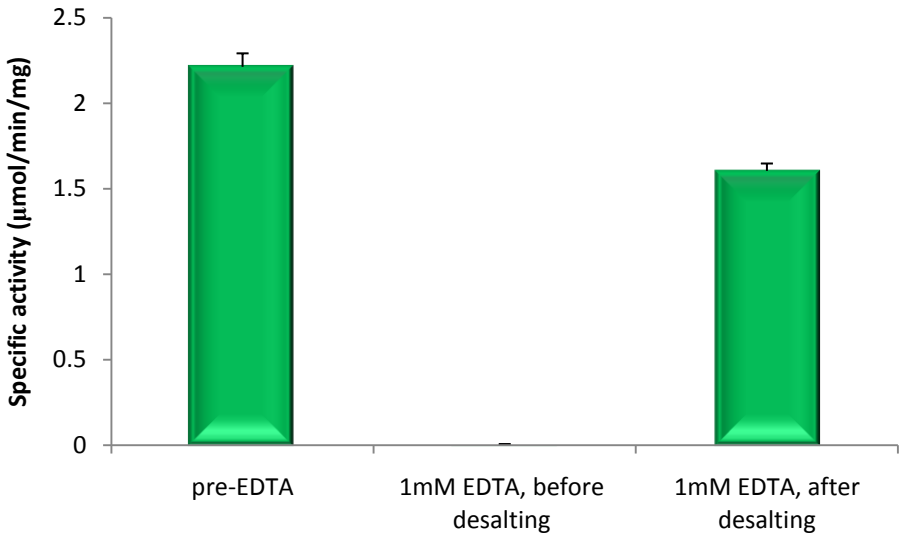


Figure 5- 8: Activity of EDTA-treated PhoX

PhoX was treated with 1mM of EDTA for 1h on ice before removing EDTA by desalting using a PD-10 column. PhoX (5µg) phosphatase activity was determined using the pNPP assay in 1M Tris-HCl, pH8.0. The reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Measurement was obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX. Error bars represent standard error for three independent experiments.

Next, I looked at cations that can re-activate PhoX after EDTA treatment. As mentioned in Figure 5- 6, addition of 100µM EDTA completely abolished PhoX activity. It was observed that the activity EDTA treated PhoX was not only restored but was also increased by the addition of divalent metal cations Ca^{2+} and Mn^{2+}. This increase in activity of PhoX compared to the activity of PhoX before EDTA treatment suggests that the PhoX purified was not fully loaded with Ca^{2+} or Mn^{2+} that is required for PhoX activity. Other cations such as Cu^{2+}, Co^{2+}, Ni^{2+} and Zn^{2+} slightly restored the activity of PhoX while some others such as Fe^{2+}, Mg^{2+} were unable to restore enzymatic activity (Figure 5- 9).

The results obtained agree with the published results on PhoX from different species in that Ca^{2+} was the best metal ion to reconstitute PhoX enzymatic activity while Mg^{2+} was the worst. However, the effect of the other divalent metal ions differed between different species. For example, Co^{2+} was able to reactivate PhoX from *Pastuerella multocida* and *Vibrio cholera* (Roy *et. al.* 1982; Wu *et. al.* 2007) but not from *Sinorhizobium meliloti*

(Zaheer *et. al.* 2009) while the activity of PhoX from *P. fluorescens* was only partially restored. Taken together, these data suggest that although PhoX can utilize a wide range of divalent metal ions, the optimal activity of *P. fluorescens* PhoX is in the presence of Ca^{2+} and Mn^{2+}.

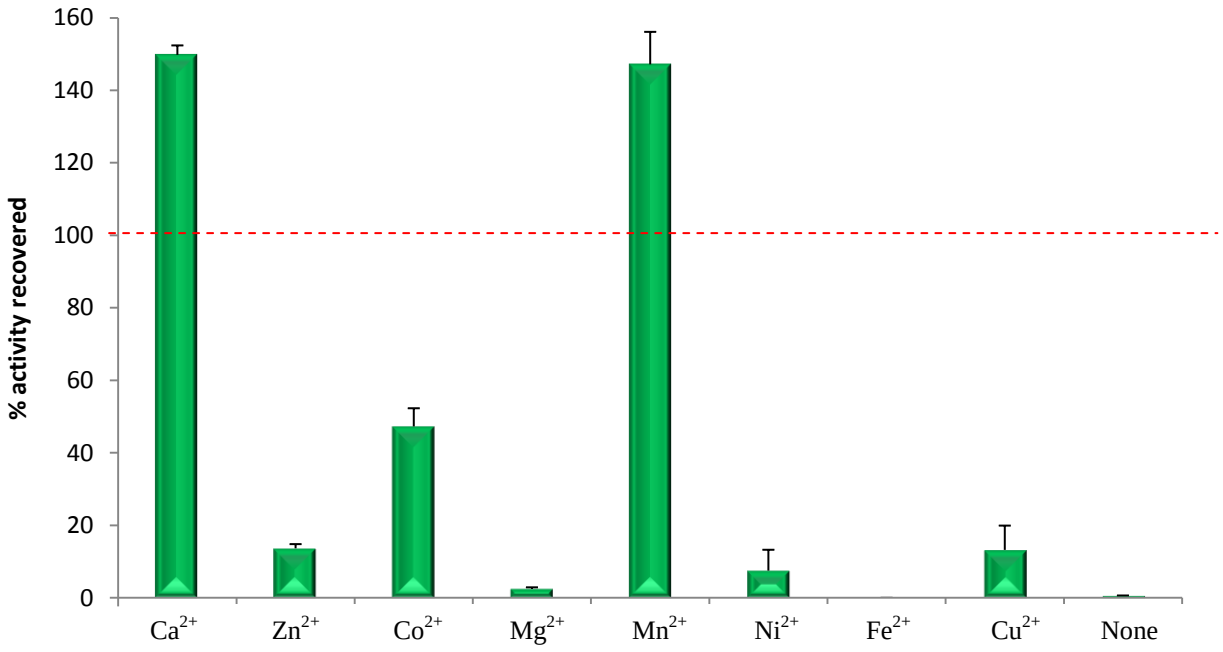


Figure 5- 9: Cation requirement for enzymatic activity of PhoX
PhoX requirement of divalent cation for activity was determined using the pNPP assay. 5µg of PhoX was added to 1M Tris-HCl, pH8.0. 100µM of EDTA was added to the pNPP reaction assay for 30s to abolish PhoX activity before addition of 1mM divalent metal ions with Cl^- as counterion. The reaction was followed at A_{420nm} at room temperature. PhoX activity in the presence of metal ion was plotted relative to the activity of PhoX prior to EDTA treatment. Red dotted line indicates activity of PhoX prior to EDTA treatment (*i.e* 100%). Error bars represent standard error for three independent experiments.

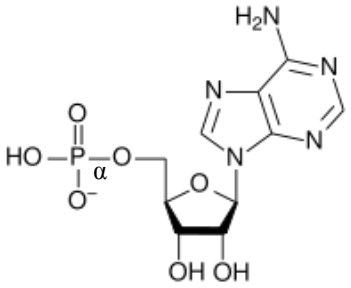
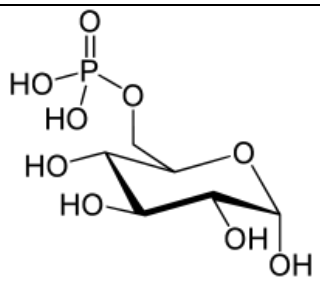
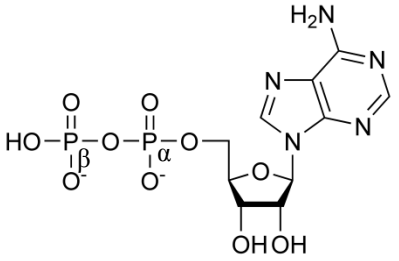
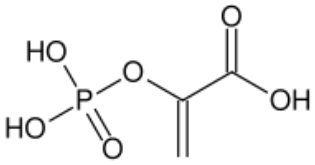
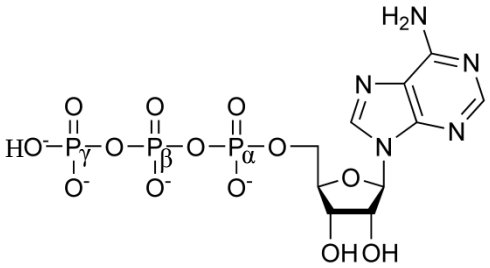
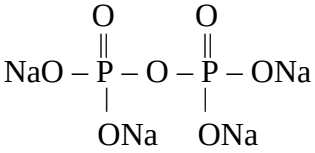
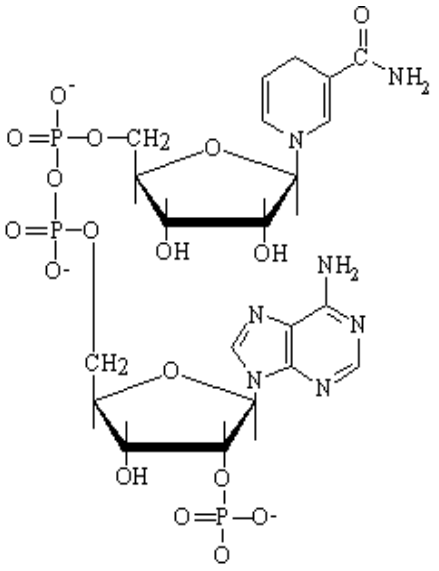
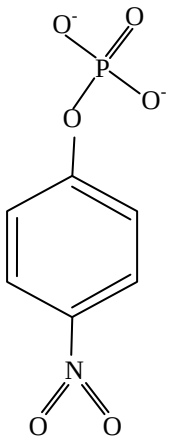
5.3.3 Substrate specificity of PhoX

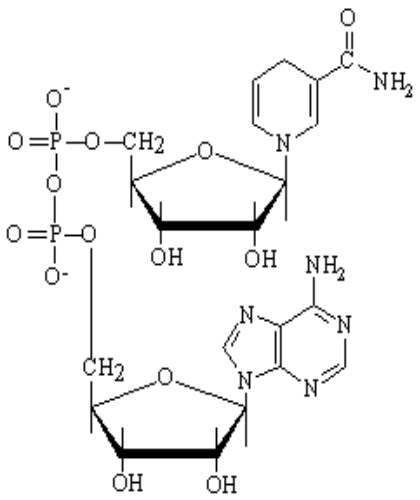
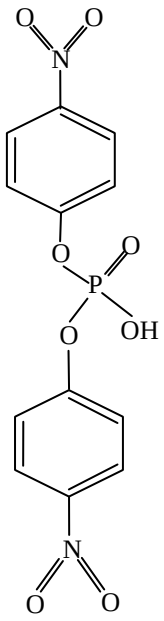
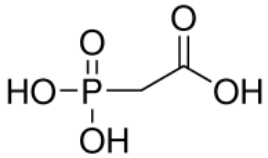
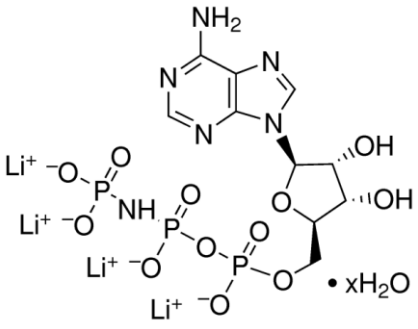
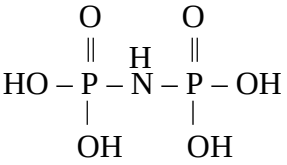
The main function of alkaline phosphatases is to release P_i for bacterial growth from various exogenous organophosphate compounds, especially under P_i limiting conditions. To assess which type of compounds can be utilized by PhoX, I tested various phosphate containing compounds. The compounds used are listed in Table 5- 1. Activity was measured as release of P_i from the compounds using the malachite green oxalate assay as explained in Figure 5- 10A. Figure 5- 10B shows that PhoX was able to hydrolyze the C-O-P phosphate

ester bond in adenosine monophosphate (AMP). Cleavage of C-O-P bonds was also observed in phosphoenol pyruvate, glucose-6-phosphate and pNPP. P_i was also released from nicotinamide adenine dinucleotide phosphate (NADPH). To determine which of the three phosphates in NADPH was cleaved by PhoX, NADH was tested. NADH is a phosphodiester and lacks the terminal phosphate in NADPH. NADH was not hydrolyzed by PhoX. This means that the phosphate that was cleaved from NADPH was the terminal phosphate. The inability of PhoX to dephosphorylate NADH also means that PhoX is a phosphomonoesterase. This is further confirmed by the inability of PhoX to use bis-pNPP, another example of a phosphodiester, as substrate. PhoX was also able to cleave phosphate groups from sodium pyrophosphate, which has a phosphoanhydride (P-O-P) bond. Adenosine diphosphate (ADP) and adenosine triphosphate (ATP) were also used as substrates by PhoX. Since PhoX can only act on terminal phosphate groups, it can be deduced that the terminal phosphate groups of ADP (*i.e.* P_β) and ATP (i.e. P_γ) has to be cleaved before the inner phosphates. This hypothesis when taken together with the observation that PhoX can use as substrate adenylyl-imidodiphosphate (AMP-PNP), which has its terminal phosphate group linked to a nitrogen atom, suggests that PhoX can cleave nitrogen-phosphate (N-P) bonds. This is consistent with the ability of PhoX to cleave imidodiphosphate. However, PhoX was unable to cleave phosphonoacetate which carries a stable carbon-phosphorous(C-P) bond.

From the results above, it can be concluded that PhoX is solely a phosphomonoesterase with broad substrate specificity against phosphate monoesters. Although most PhoX from various species have similar substrate specificity, there were some exceptions. For example, *P.multocida* PhoX was able to cleave the phosphodiester bis-pNPP (Wu *et. al.* 2007). Surprisingly, it was reported that *V.cholerae* PhoX could not use ATP and ADP as substrates (Roy *et. al.* 1982).

Table 5- 1: List of substrate for PhoX activity

O-P bonds			
AMP		glucose-6-phosphate	
ADP		phosphoenol pyruvate	
ATP		sodium pyrophosphate	
NADPH		pNPP	

phosphodiesters			
NADH		bis-pNPP	
C-P bond			
phosphonoacetate			
N-P bond			
AMP-PNP		imidodiphosphate	

A)

Step 1: P_i + ammonium molybdate ➔phosphomolybdate

Step 2: Phosphomolybdate + malachite green oxalate ➔Green complex (A_{620nm})

B)

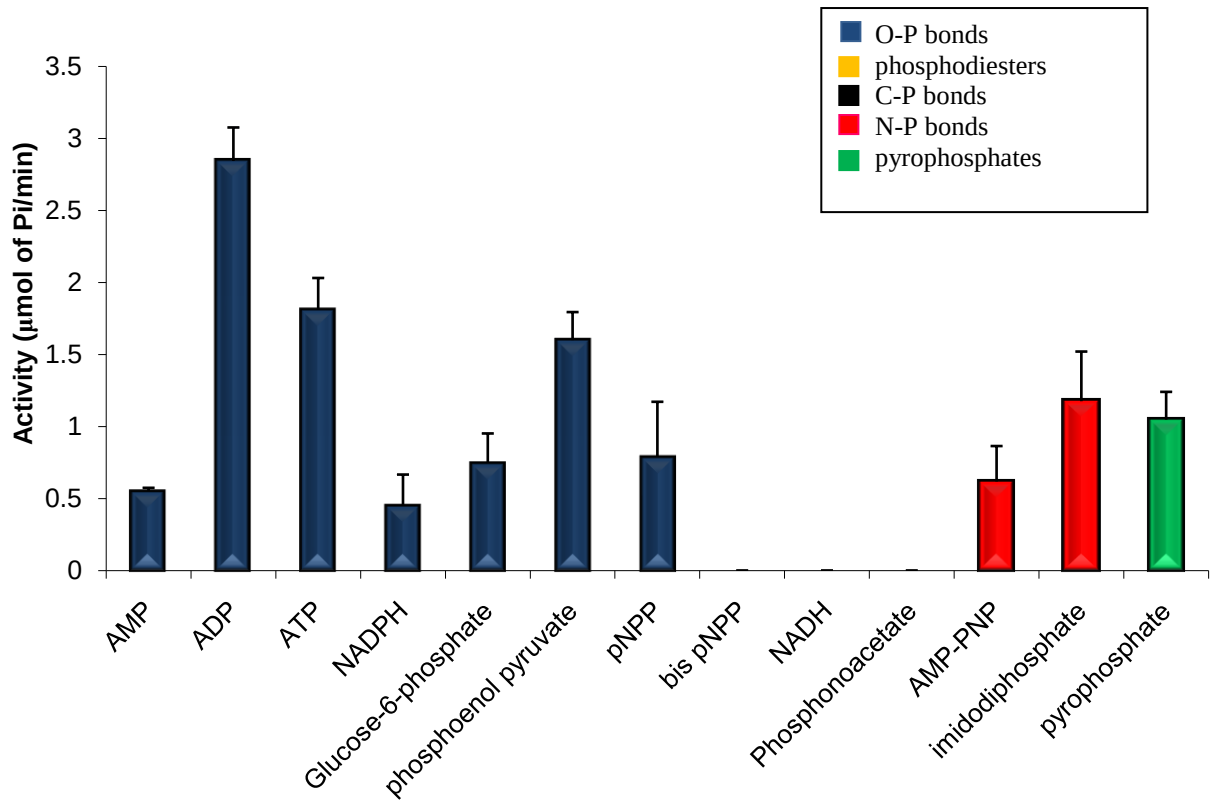


Figure 5- 10: Substrate specificity of PhoX

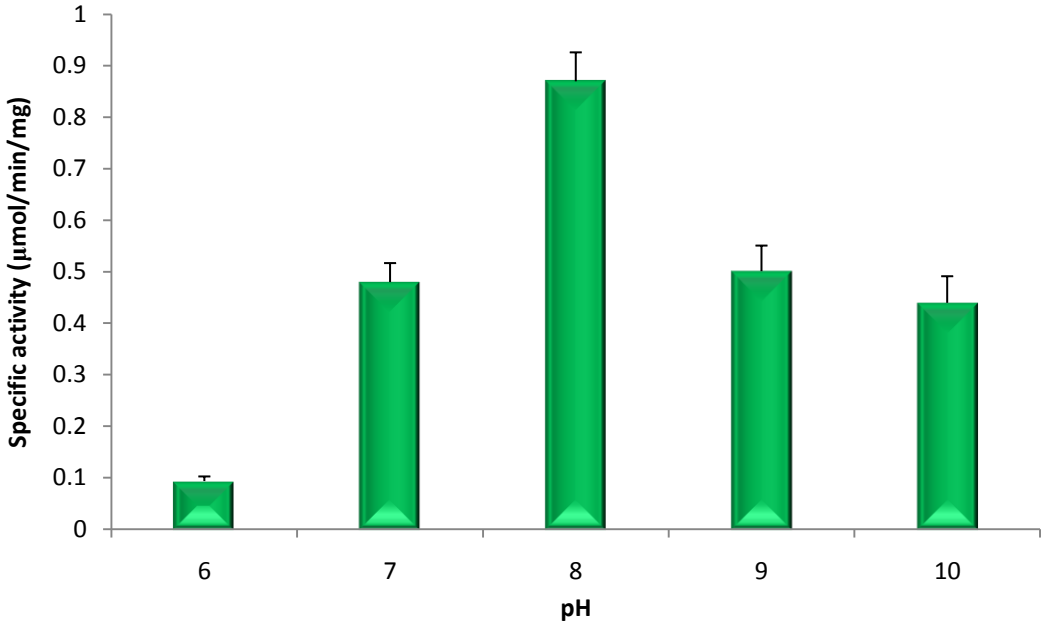
(A) Reaction of inorganic phosphate with malachite green reagent to produce a coloured product with an absorbance maximum of 620nm. (B) 13 phosphate containing compounds (as listed in Table 5- 1) were used to assess substrate specificity of PhoX. 0.02mM of substrate was added into the reaction containing 1µg of PhoX in 50mM Tris-HCl, pH8.0 supplemented with 0.5mM $CaCl_2$. Assays were done in a 96-well plate and left to incubate at 37°C for 10min. The reaction was quenched by the addition of the malachite green assay reagent. The samples were incubated at room temperature for 30min to allow formation of the green complex before measuring its absorbance at 620nm. The amount of P_i released was normalized against a standard phosphate curve. Error bar represents standard error for three independent experiments.

5.3.4 Characterization of *Vibrio fischeri* PhoX

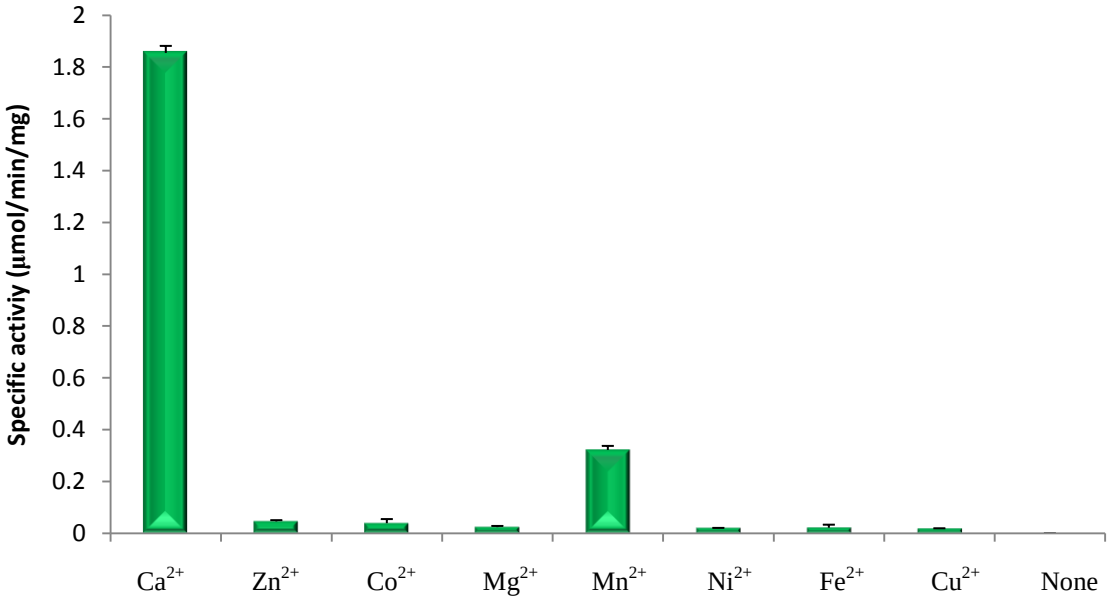
Since PhoX from different species appear to have slightly different properties, it would be interesting to characterize PhoX from another species under the same reaction conditions as *P. fluorescens* PhoX to determine if there were any differences. The species that I have chosen to work on was *Vibrio fischeri*. As reported in Chapter 4 (see 4.5.2.2), *V. fischeri* PhoX could be purified and was successfully used for *in vitro* Tat transport assays. Therefore, characterization of *V. fischeri* PhoX would be useful for understanding the relationship between PhoX and the Tat system.

Since *V. fischeri* PhoX was transported by the Tat system *in vitro*, this showed that the majority of purified *V. fischeri* PhoX must have contained the Tat signal peptide at the N-terminus. By contrast, *P. fluorescens* PhoX lacked the N-terminal Tat signal peptide. Nonetheless, from the enzymatic point of view, *V. fischeri* PhoX behaved in a similar manner to *P. fluorescens* PhoX. Both exhibited maximum activity at pH8.0 (Figure 5- 11A) and required Ca^{2+} for activity (Figure 5- 11B). However, unlike *P. fluorescens* PhoX which was active even in the absence of Ca^{2+}, *V. fischeri* PhoX was only active in the presence of Ca^{2+}. Therefore, EDTA treatment was not necessary prior to reconstitution of divalent cations to determine which cation was required for PhoX activity. *V. fischeri* PhoX was slightly active in the presence of Mn^{2+} though the activity was ninefold less than that of Ca^{2+}. Similar to *P. fluorescens* PhoX, *V. fischeri* PhoX was also shown to be a phosphomonoesterase that cleaved N-P bonds, P-O-P bonds and pyrophosphates (Figure 5- 11C).

A)


Specific activity (µmol/min/mg)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
6
7
8
9
10
pH

B)


Specific activiy (µmol/min/mg)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Ca2+
Zn2+
Co2+
Mg2+
Mn2+
Ni2+
Fe2+
Cu2+
None

C)

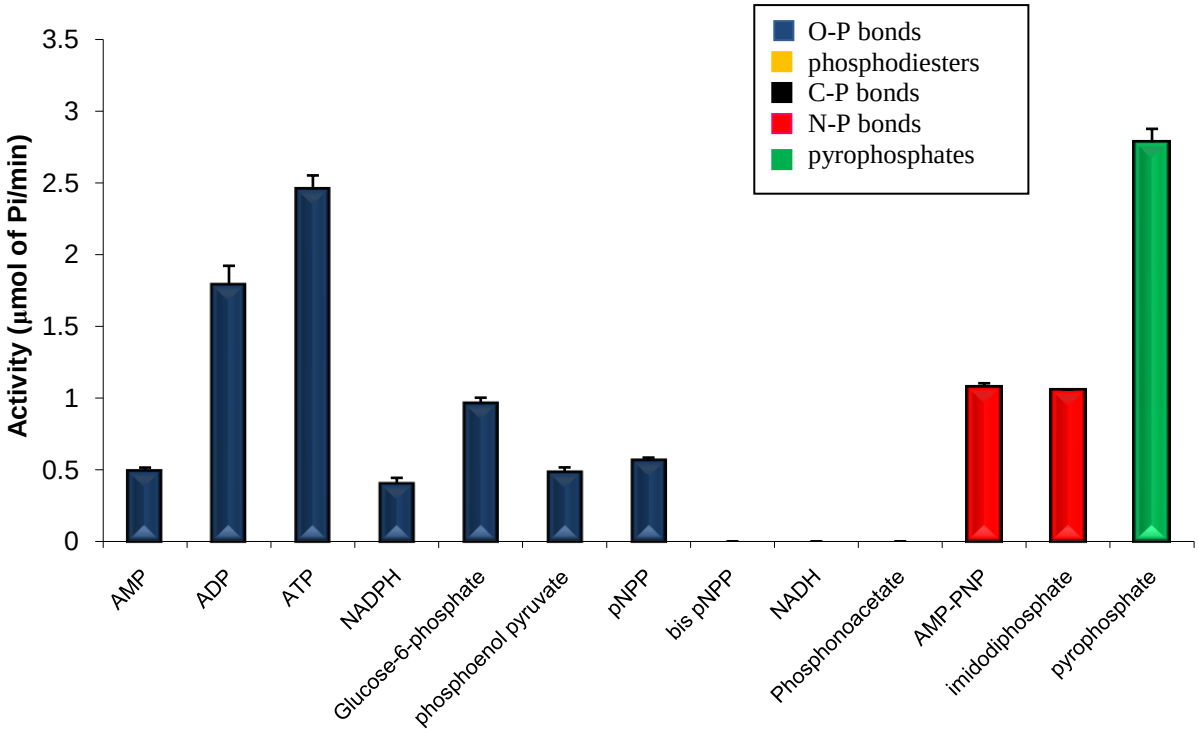


Figure 5- 11: Characterization of *V. fischeri* PhoX

(A) (Previous page)The optimum pH for *V. fischeri* PhoX (5µg) activity was determined in the presence of 5mM Ca^{2+} in 1M Tris-HCl buffer of the indicated pH using pNPP (0.04% w/v) as substrate. (B) (Previous page) Since *V. fischeri* PhoX was not active in the absence of metal ions, 1mM of divalent metal cations were added into the pNPP reaction assay in 1M Tris-HCl, pH 8.0 without EDTA treatment. In (A) and (B), the reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Measurement was obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX. (C) Substrate specificity assay was done based on the malachite green phosphatase assay. 0.02mM of substrate was added into the reaction containing 1µg of PhoX in 50mM Tris-HCl, pH8.0 supplemented with 0.5mM $CaCl_2$. Assays were done in a 96-well plate and left to incubate at 37°C for 10min. The reaction was quenched by the addition of the malachite green assay reagent and left at room temperature for 30min to allow the formation of green complex before measuring its absorbance at 620nm. The amount of P_i released was normalized against a standard phosphate curve. Error bar represents standard error for three independent experiments.

5.3.5 Comparison of biochemical properties between PhoX and PhoA

PhoA and PhoX are both alkaline phosphomonoesterases. The optimum pH for *V. fischeri* and *P. fluorescens* PhoX (pH8.0) is similar to that of *E. coli* PhoA (pH8.5) (Nesmeyanova *et. al.* 1981) though PhoX from other species such as *Campylobacter jejuni* and *Pasteurella multocida*, are reported to have optimum activity at pH10.0 (Wu *et. al.* 2007; Van Mourik *et. al.* 2008). I also carried out kinetic studies on *V. fischeri* PhoX and *P. fluorescens* PhoX to compare the enzymatic kinetics between PhoX and PhoA. As seen in Table 5- 2, PhoA is >100 more efficient than PhoX when pNPP was used as substrate. Therefore, PhoA is a more efficient enzyme compared to PhoX. This is true for *V. fischeri* PhoX and *P. fluorescens* PhoX and the published data for *S. meliloti* PhoX (Zaheer *et. al.* 2009).

Table 5- 2: Kinetic properties of PhoX and PhoA

Activity of *P. fluorescens* and *V. fischeri* PhoX was determined by pNPP reaction assay in 5mM $CaCl_2$, 1M Tris-HCl, pH 8.0. The reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Measurement was obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX.

	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
E. coli* PhoA	32.7	167.5	$5.1x10^6$
***P. fluorescens* PhoX**	240	3.07	$12.8x10^3$
***V. fischeri* PhoX**	150	1.94	$12.9x10^3$
***S. meliloti* PhoX**‡	85.3	2.3	$27.6x10^3$

Values of *P. fluorescens* and *V. fischeri* PhoX were obtained from a Lineweaver-Burke plot.
* Data obtained from (Wang *et. al.* 2005)
‡Data obtained from (Zaheer *et. al.* 2009)

5.4 Structural studies of *P. fluorescens* PhoX

5.4.1 Crystallization of *P. fluorescens* PhoX

P. fluorescens PhoX was initially purified to be used for *in vitro* Tat transport assays. However, it could not be used as a Tat substrate as it lacked the Tat signal peptide. Since the structure of PhoX was unknown when I started the project and I had high yields of pure PhoX, I attempted to determine its crystal structure.

Crystallisation conditions for PhoX were screened with the sparse-matrix approach. Sitting drops of PhoX at 7mg/ml in 20mM MOPS, 200mM NaCl, pH7.5 and solutions from JCSG and PACT Premiere screens (Molecular Dimensions) were set up in 96 well crystallization vapour diffusion trays to identify conditions for crystal growth. Each crystallization drop contained 60% protein solution and 40% precipitant solution. The total volume of the drop was 250nl. Crystals of PhoX were obtained at 21°C in 20% v/v PEG 6000, 0.1M lithium chloride, 0.1M Tris, pH8.0. This coincides with the optimum pH for PhoX activity (Figure 5- 5). The native form of PhoX was grown in the presence of 180µM Ca^{2+} while the PhoX-vanadate complex (PhoX:VO_4) was grown in the presence of 100µM Na_3VO_4 and 180µM Ca^{2+}. To obtain phasing information, a subset of proteins was co-crystallized with 0.1mM lutetium chloride ($LuCl_3$) and 180µM Ca^{2+}. The crystals for the native PhoX, with and without Lu^{3+}, and PhoX:VO_4 had dimensions of approximately 100µm×50µm×25µm. The crystals were cryoprotected in mother liquor containing 20% w/v ethylene glycol and flash cooled in liquid nitrogen.

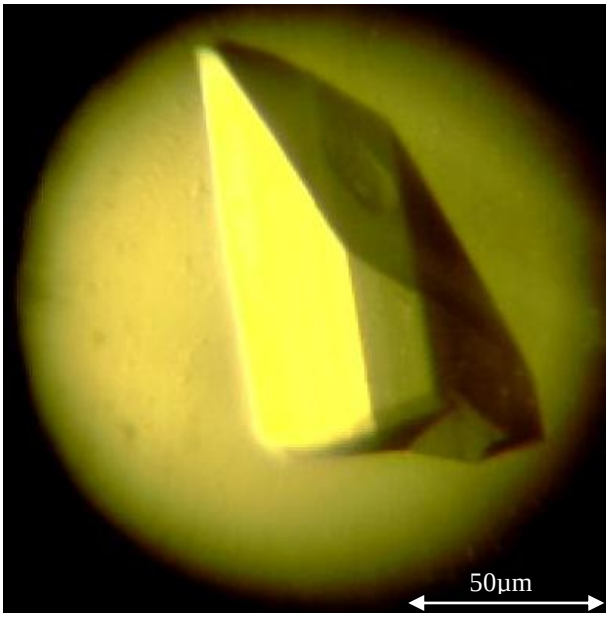


Figure 5- 12: Crystal of PhoX
PhoX crystal was grown in 20% v/v PEG 6000, 0.1M lithium chloride, 0.1M Tris, pH8.0 at 21°C.

5.4.2 Crystallographic data

Datasets were collected at 100K on beamlines BM14 and ID14-1 at the ESRF. This work was done in collaboration with Dr. Pietro Roversi and Prof. Susan Lea at Dunn School of Pathology, University of Oxford. The native form of the PhoX crystal diffracted to a resolution of 2.3Å while the PhoX:VO_4 crystal diffracted to a resolution of 1.4Å. The $LuCl_3$ cocrystals diffracted to 1.7 Å. All crystals belonged to the space group $P2_12_12_1$ of dimensions a=69.14Å, b=70.64Å, c=229.47Å for the native PhoX crystal and a=68.79Å, b=71.47Å, c=227.83Å for the PhoX:VO_4 crystal. The Single-wavelength Anomalous Diffraction (SAD) method was used to obtain the initial phasing information, by using native PhoX co-crystallized with Lu^{3+}. The structure was determined by initially locating 16 partially occupied Lu^{3+} sites based on the anomalous differences of a SAD dataset collected at the Lu L-III edge (resolution=1.7Å, lambda=1.3411Å) with the program SHELXD, and then phasing with the computer program SHARP. The average anomalous phasing power was 4.0

across the resolution range 38.0-1.7Å; the average figure of merit for acentrics (centrics) was 0.6 (0.1).

Each asymmetric unit contains two PhoX monomers. Solvent flattening in SOLOMON with 41% solvent content produced an interpretable map in which the autobuilding software Arp-wArp traced main chain and correctly docked side chains for all but 10 residues (*i.e.* autobuilt 98.3% of the structure except for residues 8-10, 126-132 and 587-588). These residues were then easily built manually in the same solvent-flattened map. Initially, an autoBUSTER (Bricogne *et. al.* 2008) refinement was performed on PhoX. Further model building was done with Coot (Emsley and Cowtan 2004) and additional refinement with Refmac5 (Murshudov *et. al.* 1997). The R_{work}/ R_{free} factors of the final model were 0.19/0.21 for PhoX and 0.16/0.18 for PhoX:VO_4. Details of data collection and crystallographic statistics are described in Table 5- 3.

Table 5- 3: PhoX X-ray data collection and crystallographic statistics

	PhoX	PhoX:VO_4
Data collection		
Beamline, λ(Å)	ESRF BM14, 1.602	ESRF ID14-4, 0.9395
Space group (Z)	$P2_12_12_1$(8)	
Molecules/asu	2	
Cell dimensions (*a*, *b*, *c* (Å))	69.14, 70.64, 229.47	68.79, 71.47, 227.83
Resolution (Å)	45.0-2.3 (2.35-2.30)	38.4-1.4 (1.43-1.40)
R_{sym} or R_{merge}	0.08 (0.46)	0.06 (0.64)
I / σI	12.4(2.2)	14.4(1.7)
Completeness (%)	95.5(94.0)	99.7(100.0)
Redundancy	3.1(3.1)	5.8(3.6)
Refinement		
Resolution (Å)	45.0-2.3	38.4-1.4
No. reflections	49287	223750
R_{work} / R_{free}	0.19 / 0.21	0.16/0.18
No. atoms	9851	10703
Protein	8974	8974
Ligand/ion	6 Ca^{2+} ions 4 Fe^{3+} ions	6 Ca^{2+} ions 4 Fe^{3+} ions
Water	857	1705
***B*-factors (Å^2)**		
Protein	24.0	16.4
Ligand/ion	27.6	25.0
R.m.s. deviations		
Bond lengths (Å)	0.007	0.007
Bond angles (Å)	1.01	0.98

*Values in parentheses are for highest-resolution shell.

5.4.3 Overall architecture of PhoX

PhoX crystallized as a monomer and forms a β-propeller approximately 50Å in diameter and 55Å in height. As the purified PhoX lacked the N-terminal Tat signal sequence, residues were numbered from the N-terminal residues of the start of the mature domain. The propeller contains six blades, labelled from β1 to β6, arranged around a funnel-shaped central cavity (Figure 5- 13). Each blade has 4 anti-parallel β strands, except for blade 2 (six strands) and blade 3 (five strands). The strands in blade x are labeled βxa to βxf; in each blade x, βxa is the innermost β-strand, bordering the central cavity. Blades 3 and 4 are fused into a single 9-stranded antiparallel β-sheet. Ring closure occurs at the β6 blade, with strands β6a, β6b and β6c belonging to the C-terminal part of the polypeptide, and strand β6d belonging to the N-terminal part. The loop between βxb and βxc (βxb-βxc) and the loop connecting βxd to βxa of the following blade (βxd-β(x+1)a) are at the rim of the cavity while the βxa-βxb and βxc-βxd loops are at the opposite face of the structure. The N- and C-termini are adjacent to one another, on the face of the protein distal to the cavity. The central cavity is plugged at the bottom by Trp388 (Figure 5- 14) from blade 4. Above it, the active site contains five metal ions, which were assigned as three calcium (Ca^{2+}) ions and two ferric (Fe^{3+}) ions, near the rim of the cavity (spheres in Figure 5- 13). These metal ions lie in the central pocket whose electrostatic potential (computed in absence of the ions) is largely negative. None of the seven cysteines of PhoX forms disulfide bonds.

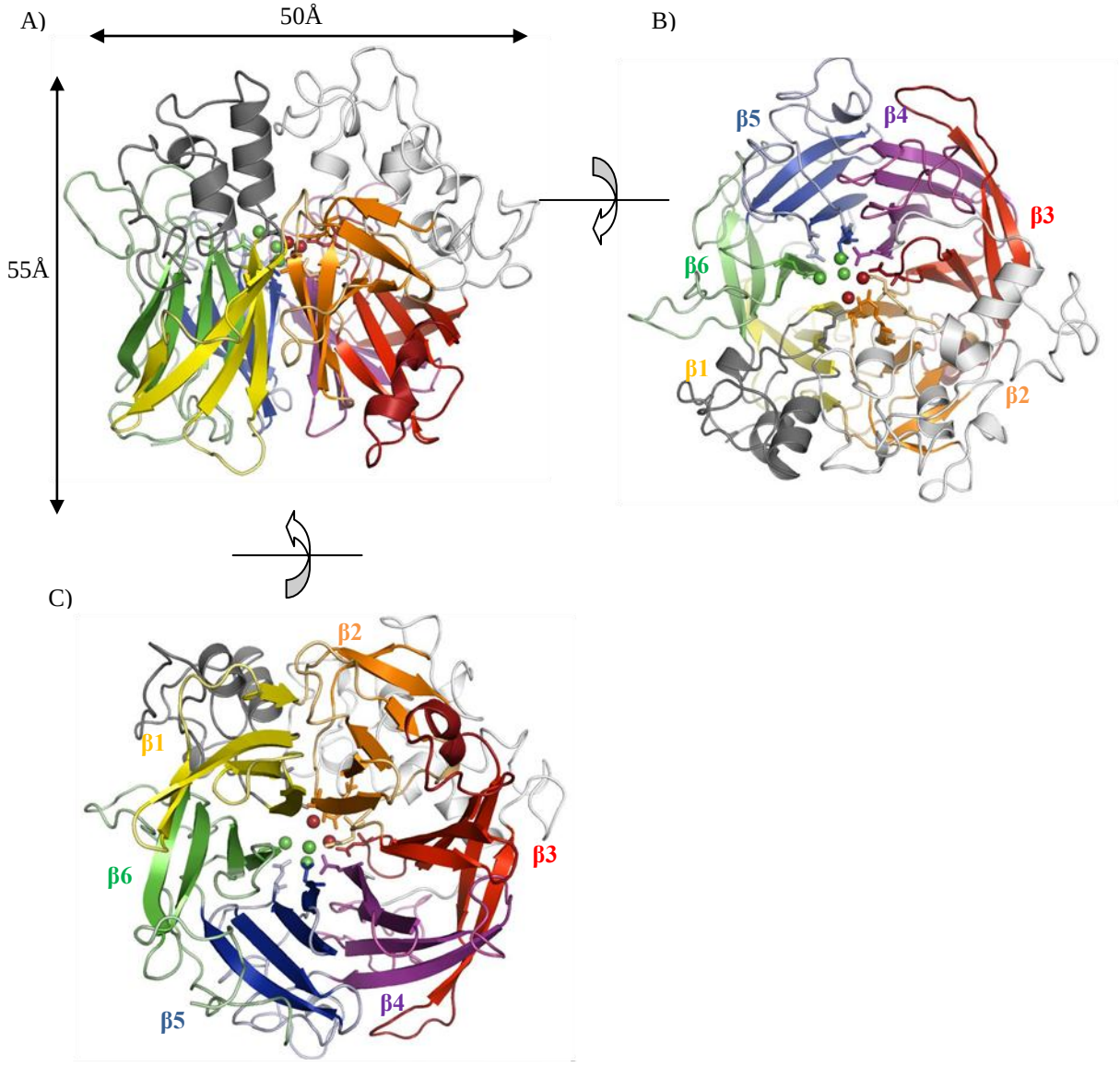


Figure 5- 13: Overall structure of PhoX

Cartoon representation of PhoX (based on molecule A in the crystal). (A) side view. (B) top view, related to (A) by a 90˚ rotation around the horizontal axis; (C) bottom view, related to (A) by a -90˚ rotation around the same axis. Blade 1, yellow: residues 69-143; blade 2, orange: residues 144-152, 170-193 and 249-275; blade 3, red: residues 276-386; blade 4, purple: residues 387-451; blade 5, blue: residues 452-526; blade 6, green: residues 527-588 and 8-68. The alpha-subdomains are in grey (38-68 and 90-118) and white (153-169,194-248 and 348-386). The three Ca^{2+} ions are green spheres and the two Fe^{3+} ions red spheres. The residues coordinating the ions in the active site are in stick representation. Figures were produced using Pymol (DeLano Scientific LLC).

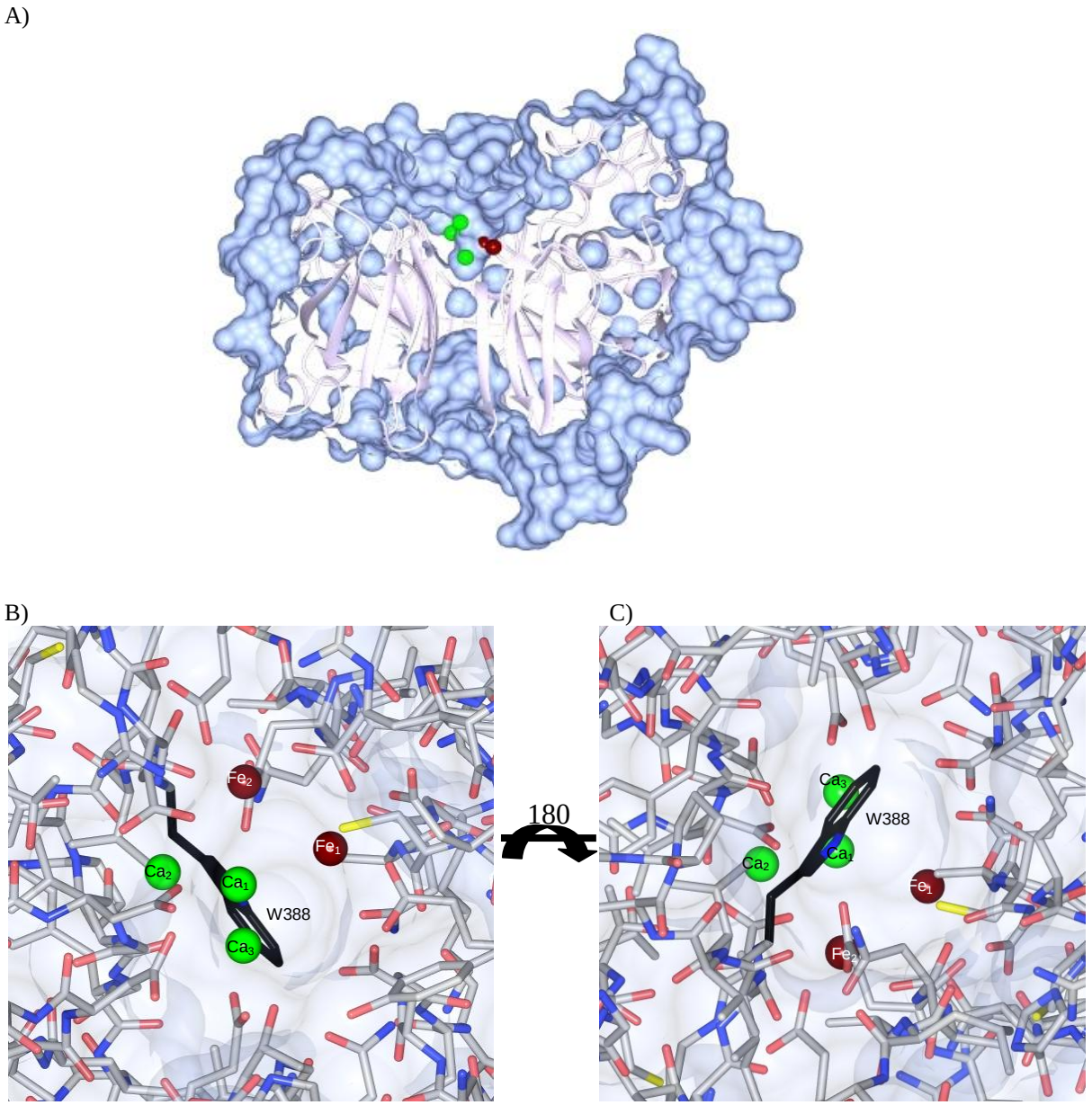


Figure 5- 14: The active site cavity of PhoX

(A) Cross section view of the space filling model of PhoX with same orientation as Figure 5- 13A shows the active site near the rim of the cavity. A ribbon representation of PhoX is overlaid on top to show the β-propeller domain. The cavity appears to be plugged from the bottom. Zoomed in view of the cavity with (B) metals on top and (C) metals below showing Trp388 in black. Other residues surrounding the cavity are represented as sticks. Green sphere: Ca^{2+}; brown sphere: Fe^{3+}. Figures were produced using CCP3mg (Potterton *et. al.* 2004).

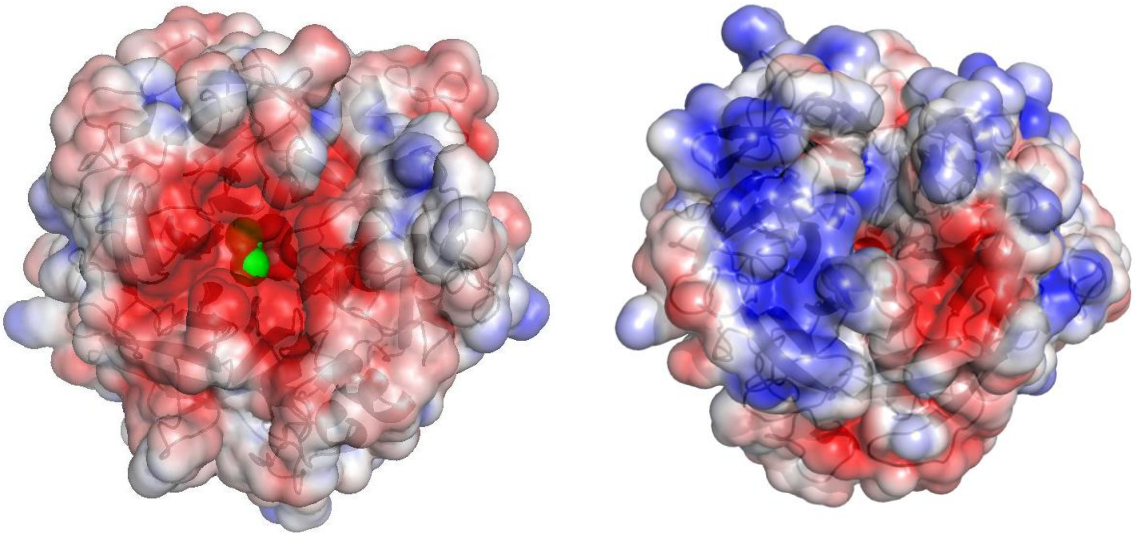


Figure 5- 15: Electrostatic surface potential of PhoX

Surface representation of PhoX (cartoon in transparency). (A) top view (same orientation as Figure 5- 13B) looking down towards the active site. (B) bottom view (same orientation as Figure 5- 13C), related to (A) by a 180˚ rotation around the horizontal direction. Contour levels: red, -3 kT/e; blue, +3 kT/e. The three Ca^{2+} ions are green spheres and the two Fe^{3+} ions are red spheres. Figures were produced using Pymol (DeLano Scientific LLC).

5.4.3.1 Blade alignment

β-propellers are believed to have evolved from duplication of a single blade, followed by functional differentiation. Thus, there is a tendency for β-propellers to share sequence similarity between the blades such as the WD40 motif, RCC1 repeat, KELCH motif, YWTD repeat, NHL motif and YVTN repeat (Chaudhuri *et. al.* 2008). However, alignment of the six β blades of PhoX showed a lack of sequence similarity between blades (Figure 5- 16). Hence, *P. fluorescens* PhoX could have evolved differently from other β-propellers or diverged extensively from the ancestral β-propeller.

```
                 βxa          βxb          βxc          βxd
Blade1       NNDGMSLFEFPGEKNRALMAINNEYTN*EGVSVIEVQRKNG----QWQFVQGSR (67-93, 117-138)
Blade2       FQNCANGKT-PW---GTYLTCEEN*NRHGWVVEIDPFDPQSTP—VKRTALGRFK (176-195, 244-271)
Blade3       ENAALAET-----DDGRAVVYMGDDERGEFIYKFVSRDK*DILDHGTLYVARFD (273-306, 371-332)
Blade4       DRPEWIVVS-P--KDGQVYCTLTNN*VYGQILRWRTDRDDHASKTFAWDLFVVA (384-405, 425-452)
Blade5       NSPDGLGFD----KAGRLWILTDGD*GNNQMLCADPAT------GEIRRFMVGP (476-496, 507-528)
Blade6(a-c)  IGCEVTGISFSPDQKTLFVGIQHPG*PRSSVMAITREDGGIVGA---------- (529-553, 569-586)
Blade6 (d)   ----------------------------------------KGYKSSVLIS (28-37)
```

Figure 5- 16: Amino acid sequence alignment of the six β-propeller blades of PhoX from *P. fluorescens*.
Sequences contributing to each blade are in red. * indicates where a segment of the sequence is removed due to its length.

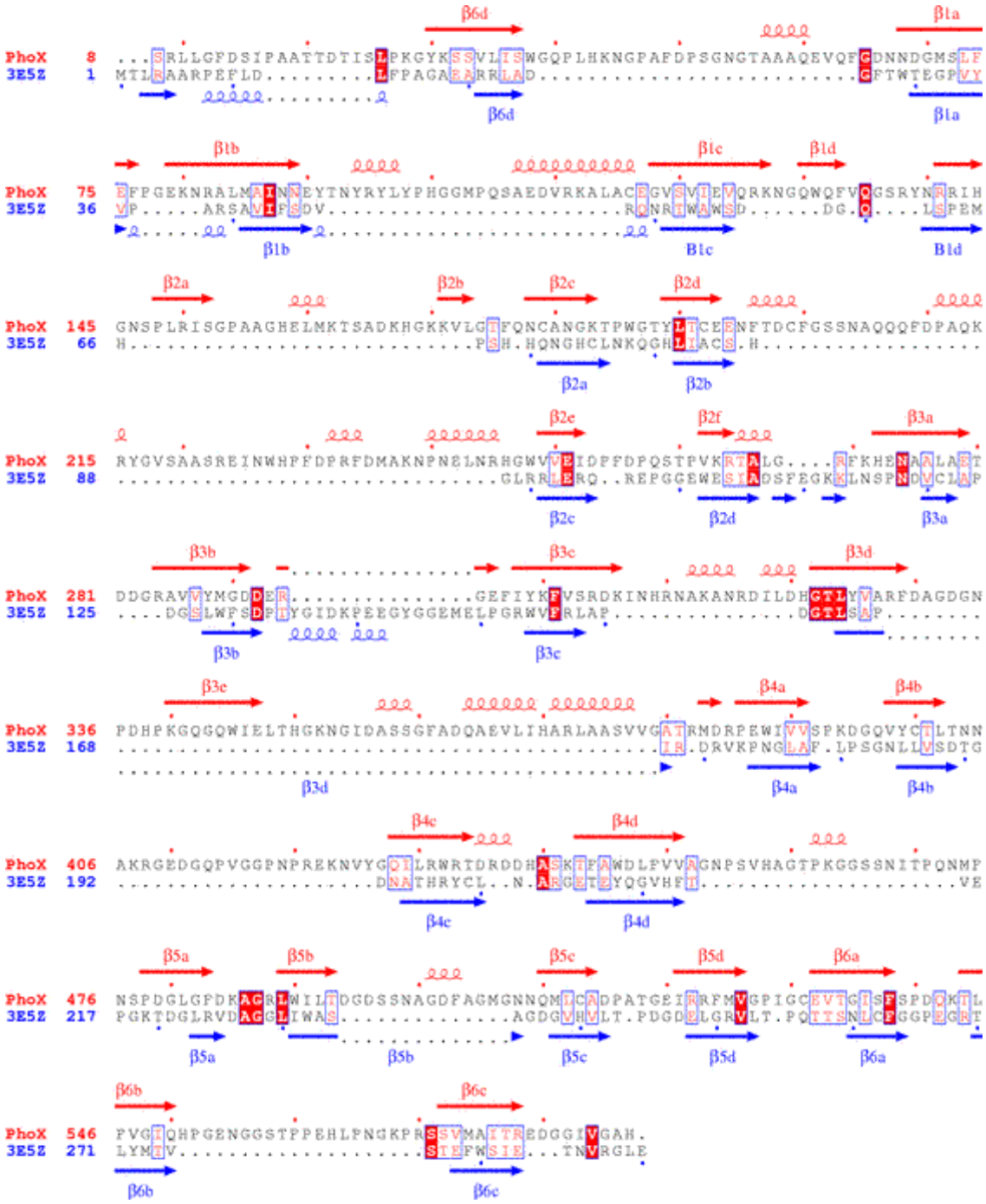
5.4.4 Sequence and structural comparison with close structural homologues of PhoX

The closest structural homologues of PhoX is a putative gluconolactonase from *Deinococcus radiodurans*, (PDB ID 3e5z, rmsd= 2.3Å over 202 aligned Cαs). Gluconatonases hydrolyze D-glucono-lactone to D-gluconate. There is one Mg^{2+} ion assigned to this structure. This Mg^{2+} ion is located at the rim of the β-propeller, which is most likely the active site. However, not much is known about this enzyme.

In addition, there is a phosphatase which is a close structural homologue of PhoX. This phosphatase is known as diisopropyl fluorophosphatase (DFPase) and is from the European squid, *Loligo vulgaris* (PDB ID 2gvv, rmsd= 2.6Å over 207 aligned Cαs). DFPase was shown to be a calcium-dependent phosphotriesterase which hydrolyzes a wide range of organophosphorus nerve agents such as DFP. This leads to the formation of a phosphate or phosphonate and a fluoride ion, resulting in the detoxification of the organophosphorous agent (Blum and Chen 2010). Two calcium ions were identified in DFPase – one of which is located within the central water tunnel and the other at the base of the active site pocket.

These structural homologues of PhoX were identified through similarity between their β-propeller domains. However, neither gluconolactonase nor DFPase have the extra α-helical subdomains. Figure 5- 17A reports the annotated sequence alignment between PhoX and the gluconolactonase homologue. The sequence of PhoX is almost twice as long as that of the gluconolactonase. The extra residues are interspersed between the blades and contribute to two subdomains made of loops and α-helices. The subdomains are situated on the same face of the propeller and overhang the central cavity of the protein (Figure 5- 17B). Neither of these α-helical subdomains is structurally similar to any other domain of known structures. The first α-helical subdomain is formed by two insertions in blades 1 and 6 (residues 38-68 and 90-118, painted grey in Figure 5- 13); the second subdomain is inserted in blades 2, 3 and 4 (residues 153-169, 194-248 and 348-386, painted white in Figure 5- 13).

A)


PhoX 8 ...SRLLGFDSIPAATTDTISLPKGYKSSVLTSWGQPLHKNGPAFDPSGNGTAAAQEVQFGDNNDGMSLF
3E5Z 1 MTLRAARPEFLD.....LFPAGAEARRLAD.....GFTWTEGPVY
PhoX 75 EFFGEKNRALMAINNEYTNYRYLYPHGGMPQSAEDVRKALACEGVSVIEVQRKNGQWQFVQGSRYNRRIH
3E5Z 36 VP.....ARSAVIFSDV.....RQNRTWAWSD.....DG.Q.....LSPEM
PhoX 145 GNSPLRISGPAAGHELMKTSADKHGKKVLGTFQNCANGKTPWGTYLICEENFTDCFGSSNAQQQFDPAQK
3E5Z 66 H.....PSH.HQNGHCLNKQGHLIACS.H.....
PhoX 215 RYGVSAASREINWHPFDPRFDMAKNPNELNRHGWVVEIDPFDPQSTPVKRTALG....RFKHENAALAET
3E5Z 88 .....GLRRLERQ..REPGGEWESIADSFEGKKLNSPNDVCLAP
PhoX 281 DDGRAVVYMGDDER.....GEFIYKFVSRDKINHRNAKANRDILDHGTLYVARFDAGDGN
3E5Z 125 ....DGSLWFSDPIYGIDKPEEGYGGEMELPGRWVFRLAP.....DGTLSAP.....
PhoX 336 PDHPKGQGQWIELTHGKNGIDASSGFADQAEVLIHARLAASVVGATRMDRPEWIVVSPKDGQVYCTLTNN
3E5Z 168 .....IH.DRVKPNGLAF.LPSGNLLVSDTG
PhoX 406 AKRGEDGQPVGGPNPREKNVYGQILRWRTDRDDHASKTFAWDLPVVAGNPSVHAGTPKGGSSNITPQNMP
3E5Z 192 .....DNATHRYCL..N.ARGETEYQGVHFT.....VE
PhoX 476 NSPDGLGFDKAGRLWILTDGDSSNAGDFAGMGNNQMLCADPATGEIRRFMVGPIGCEVTGISFSPDQKTL
3E5Z 217 PGKTDGLRVDAGGLIWAS.....AGDGVHVLT.PDGDELGRVLT.PQTTSNLCFGGPEGRT
PhoX 546 PVGIQHPGENGGSTPPEHLPNGKPRSSVMAITREDGGIVGAH.
3E5Z 271 LYMIV.....STEFWSIE...TNVRGLE

B)

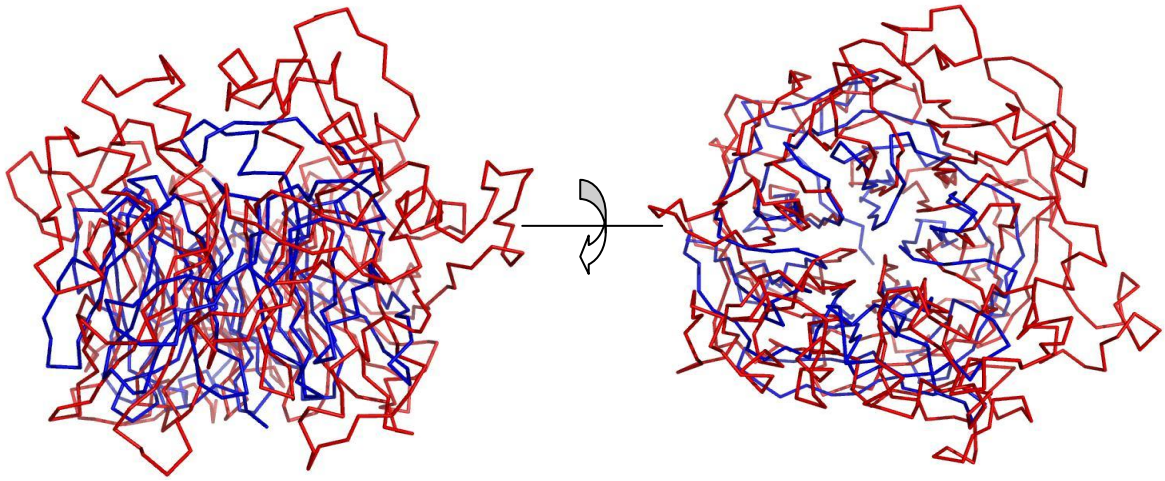


Figure 5- 17: Sequence and structural alignment of PhoX with gluconolactonase

(A) Structure-based sequence alignment of PhoX and gluconolactonase (PDBID 3e5z) using STRAP (Gille and Frommel 2001). Corresponding secondary structure is annotated. (B) Structural superposition of PhoX and gluconolactonase from the side view and top view. It can be seen that PhoX has extra domains that are lacking in gluconolactonase. PhoX is in red while gluconolactonase is in blue. Figures were generated using Pymol (DeLano Scientific LLC).

5.4.5 Multiple sequence alignment of PhoX homologues

Multiple sequence alignment of PhoX from 50 different species including Gram-negative α, β, γ, ε proteobacteria, cyanobacteria and Gram-positive bacteria shows that most of the residues that make up the six blades of the β-propeller are conserved (Appendix 1). In addition, they all have similar lengths of residues interspersed between the blades which make up the α-helical subdomains. However, these residues are not conserved between different species. This observation is reaffirmed by structural alignment of PhoX from the same 50 species based on CONSURF (Landau *et. al.* 2005) (Figure 5- 18).

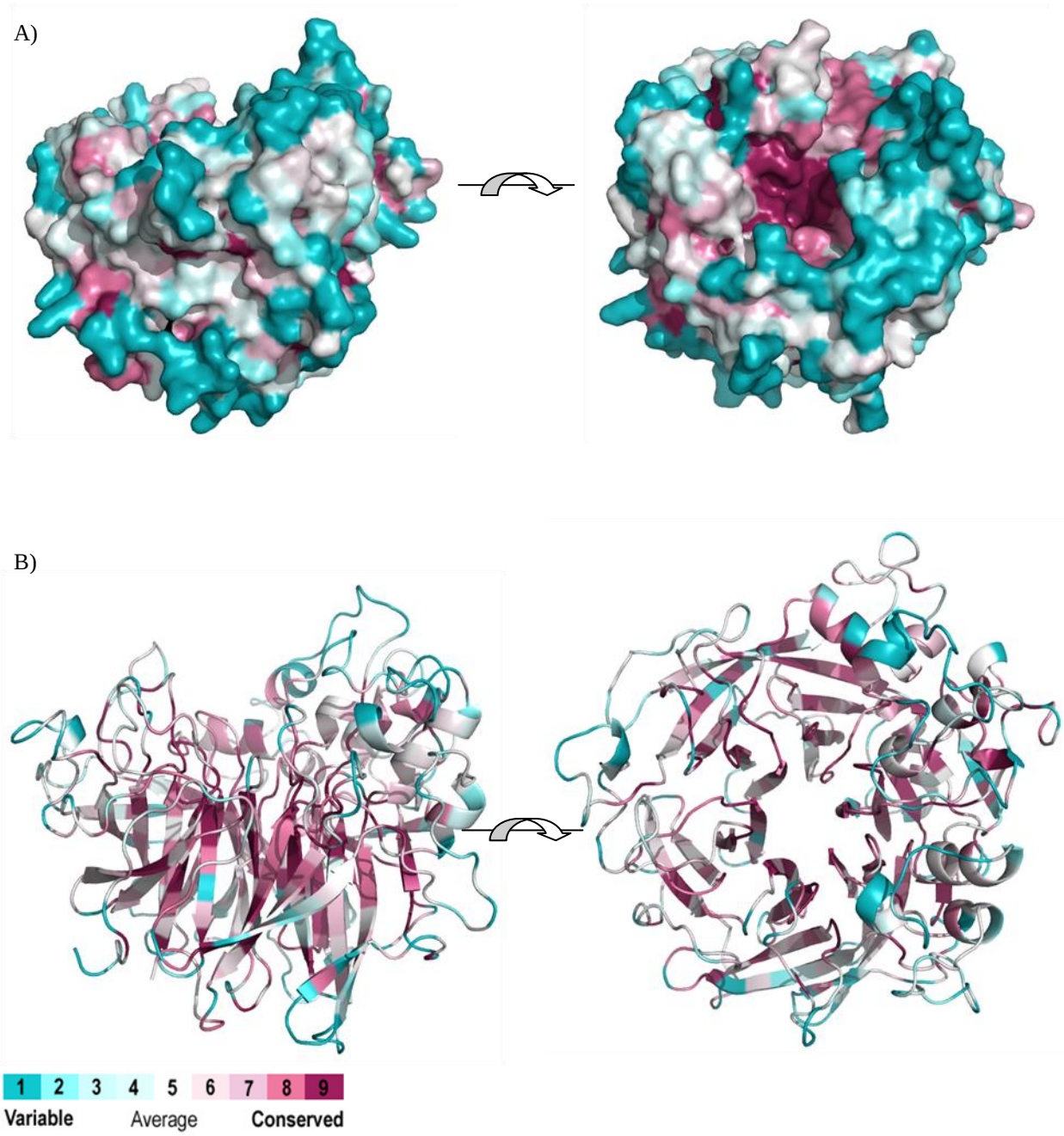


Figure 5- 18: PhoX amino acid conservation displayed on the *P. fluorescens* PhoX structure.
Colour-coded (A) surface representation and (B) cartoon representation of PhoX as calculated by CONSURF (Landau *et. al.* 2005) showing the invariant residues in PhoX from 50 different species. PhoX is oriented so that the α-helical subdomains are on the right side of the structure. Figures on the right represent the side view (same orientation as Figure 5- 13A) while figures on the left represent the top view (same orientation as Figure 5- 13B). The colour key is as noted below the figures.

5.4.6 Metal binding sites

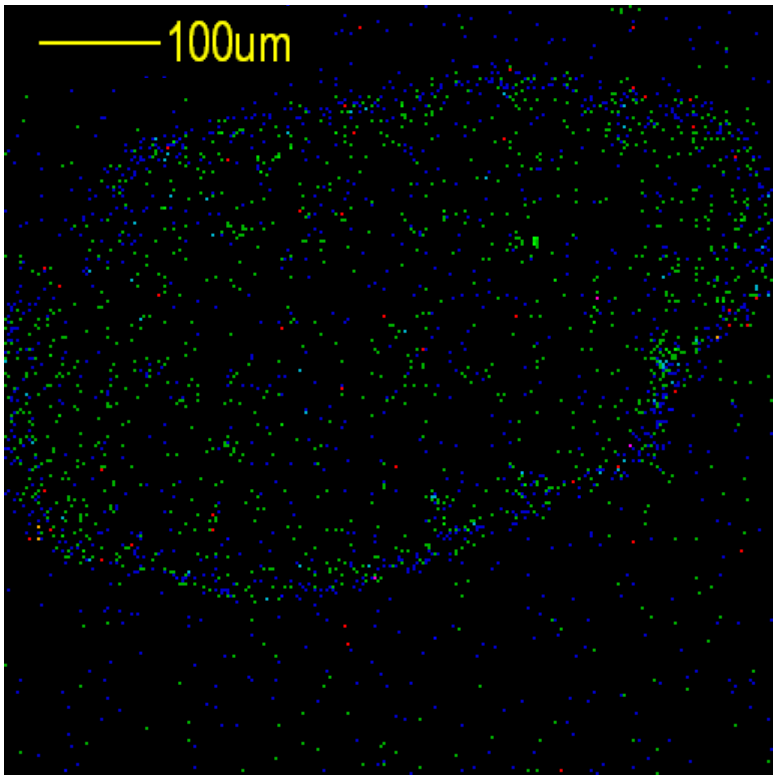
PhoX contains metal ion cofactors. There are five metal ions in the crystal structure of *P. fluorescens* PhoX. These metal ions have been assigned as two Fe^{3+} ions and three Ca^{2+} ions by X-ray anomalous scattering and analysis of their coordination geometry (Müller *et. al.* 2003). Each of the metals in PhoX has an octahedral geometry, which is relatively common for Ca^{2+} and Fe^{3+}. Occupancy refinement confirms full occupancy for all ions except for one Fe^{3+} ion, Fe_1, which has an occupancy of 0.9 and one Ca^{2+}, Ca_3, which is 0.7-0.8 occupied. The total content as estimated by crystal occupancies is therefore 2.7 calcium ions and 1.8 iron ions per molecule. No other significant anomalous peaks (but for sulfurs on Cys and Met residues) are found in the maps.

This assignment of metals is further confirmed by proton induced X-ray emission (microPIXE). This work was done in collaboration with Mr. Oliver Zeldin and Prof. Elspeth Garmen. Fresh samples of PhoX were prepared at 20mg/ml using 20mM Tris, 200mM KBr, pH7.5 instead of 20mM MOPS and 200mM NaCl, pH7.5 (the chlorine peak from NaCl can overlap with the sulphur signal, increasing the error on the measurement, and so was replaced by KBr). A 0.2µl drop of the sample was pipetted onto a 4µm thick polypropylene backing film held over a hole in an aluminium holder and then air dried. Figure 5- 19A shows the overlaid elemental microPIXE maps. The sulphur signal (blue) clearly defines the position of the PhoX protein droplet due to the presence of the 12 methionines and 7 cysteines in each PhoX molecule. Quantitative information was obtained by collecting spectra at selected points on the drop. These spectra were analysed to extract the area density of each element of interest in the sample (Figure 5- 19B). The number of atoms of each element per protein molecule was computed from the ratio of the known sulphur content per molecule of PhoX. Three points on the protein drop were measured for iron (red) and calcium (green) content, giving an average of 3.4 (individual measurements of 3.05, 3.57, 3.56) atoms of calcium and

1.6 (individual measurements of 1.54, 1.76 and 1.53) atoms of iron per PhoX molecule. These numbers were slightly different from the number of ions deduced from occupancy refinement of the PhoX crystal. This could be because the PhoX used for microPIXE was purified in a different buffer and these measurements were carried out on PhoX in solution instead of using the PhoX crystal.

The metal assignments are supported by the biochemical data reported in section 5.3.2 which showed that PhoX activity was reconstituted by the addition of Ca^{2+}. In addition, previous studies have demonstrated that iron must be present during the growth of *P. fluorescens* for phosphatase activity (Monds *et. al.* 2006). They have shown that the addition of iron back to crude extracts failed to recover the activity of PhoX, suggesting that iron needs to be incorporated during folding. The metal ions do not lie in the middle of the central cavity but are instead found at the rim of the cavity (Figure 5- 13). The residues responsible for coordinating the five metal ions are the side chain oxygens of Glu90, Glu194, Glu273, Asp292, Glu387, Asp479, Asp494, Glu532 and the sulphydryl group of Cys179 (Table 5- 4). It can be seen from Figure 5- 20 and Figure 5- 21 that many of the ligands coordinate more than one metal ion. Glu194, Glu273, Asp479, Asp494 each coordinates two different metals while Glu387 coordinates three metals. Seven water molecules in the active site are also involved in the coordination of the metal ions (Figure 5- 20, Figure 5- 21). One of the water molecules is most likely in the activated hydroxyl form (OH^-). This will be explained in more details in section 5.4.10. All the protein residues are conserved in PhoX homologues in eubacteria (Figure 5- 22). All ligands are within 3Å of the metal ions (Table 5- 4).

A)



B)

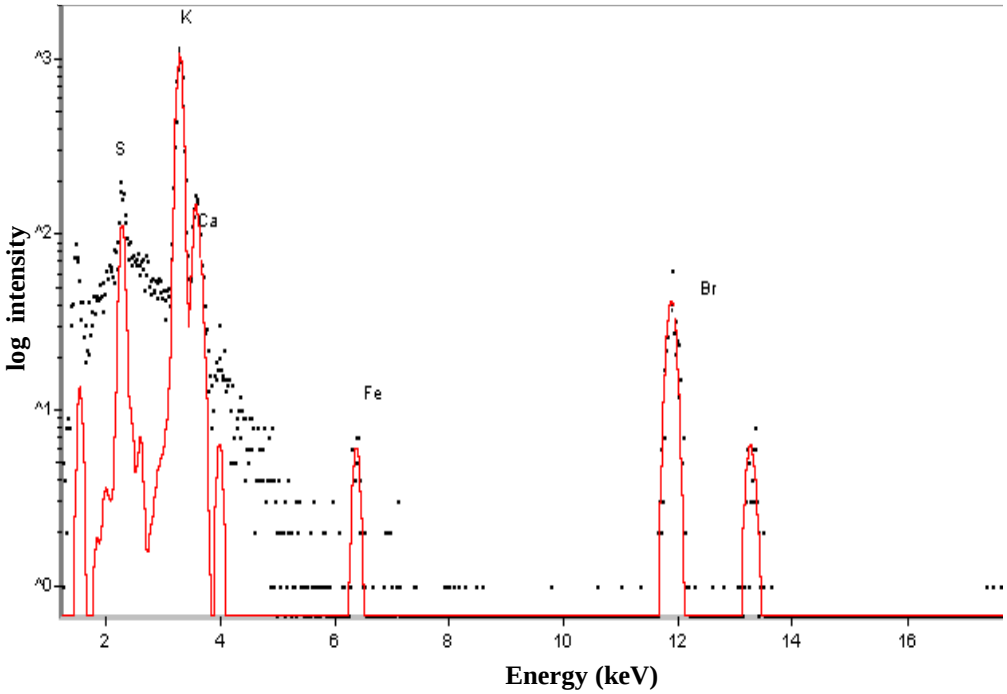


Figure 5- 19: MicroPIXE analysis of *P. fluorescens* PhoX

Elemental analysis by microPIXE of PhoX in 20mM Tris, 200mM KBr, pH7.5. (A) Overlay of two-dimensional surface maps of the sulphur (blue), calcium (green) and iron (red) distribution in the PhoX protein drop. The area of the map is 750microns x 750microns. (B) The fitted PIXE spectrum from a point on the protein droplet, showing sulphur, calcium and iron. The potassium and bromide were from the buffer. Black dots are the raw data and the red lines are the simulated peaks after background signal subtraction. Note the log scale of the y axis.

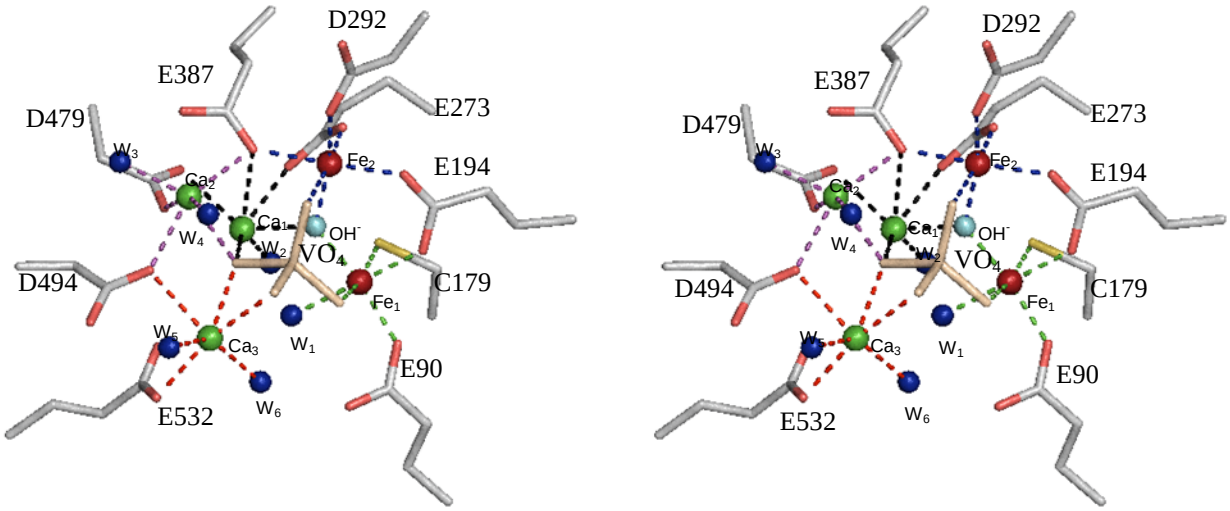


Figure 5- 20: The metal ions and their coordinating ligands in PhoX

Stereo view of the active site observed from the top of the active site of PhoX:VO_4 showing all 9 coordinating residue ligands (E90, E194, C179, E273, D292, E387, D479, D494 and E532), coordinating water molecules and vanadate. All five metal ions have octahedral coordination. Figures were generated using Pymol (DeLano Scientific LLC).

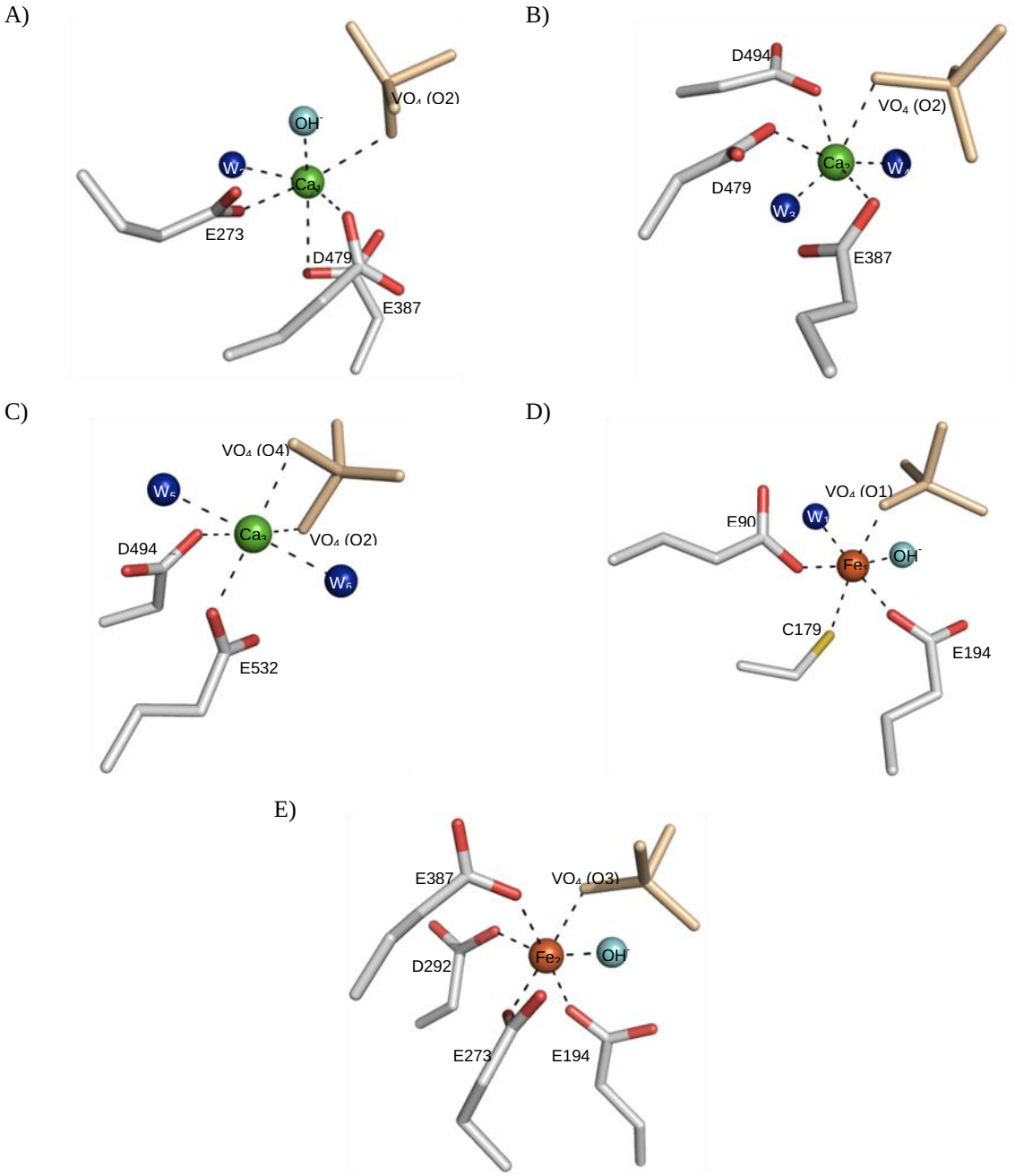


Figure 5- 21: Metal binding sites in PhoX

PhoX:VO_4 structure showing individual metal ion with 6 coordinating ligands contributed by a combination of residues on PhoX, water molecules and oxygen atoms of vanadate. Note that the images were differently oriented so as to give a better view of the octahedral geometry. (A) Ca_1, (B) Ca_2, (C) Ca_3, (D) Fe_1, (E) Fe_2. Figures were generated using Pymol (DeLano Scientific LLC).

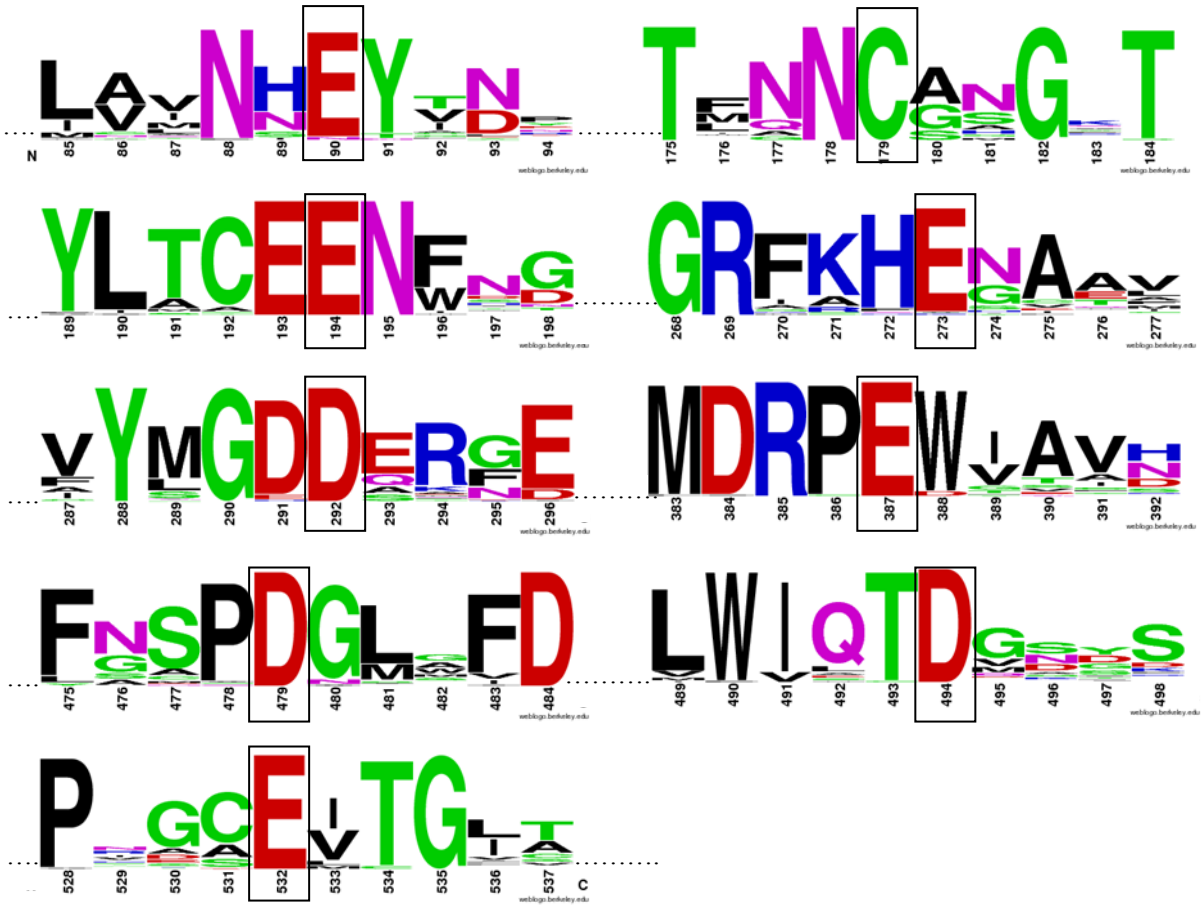


Figure 5- 22: Sequence alignment of metal coordinating ligands in bacterial PhoX homologues
Multiple sequence alignment of metal coordinating ligands as generated by Weblogo (Crooks *et. al.* 2004). This dataset is obtained from (Appendix 1). Ligands in *P. fluorescence* PhoX are boxed.

Table 5- 4: Bond distances between metal ions and coordinating ligands

Metal	Coordinating ligand	Bond distance (Å)			
		Vanadate-PhoX		native PhoX*	
		chain A	chain B	chain A	chain B
Fe_1	E90Oε2	2.00	1.99	2.02	2.01
	C179Sγ	2.53	2.50	2.62	2.56
	E194Oε1	2.06	2.04	2.12	2.11
	OH^-	1.98	1.94	NA	NA
	W_1	2.19	2.18	NA	NA
	$VO_4(O_1)$	2.07	2.12	-	-
	$PO_4(O_1)$	-	-	2.16	2.25
Ca_1	E273Oε2	2.35	2.35	2.34	2.34
	E387Oε1	2.86	2.85	2.83	2.82
	D479Oδ1	2.50	2.50	2.37	2.38
	OH^-	2.36	2.39	NA	NA
	W_2	2.56	2.47	NA	NA
	$VO_4(O_2)$	2.62	2.61	-	-
	$PO_4(O_2)$	-	-	4.38	4.33
Fe_2	E194Oε2	2.07	2.08	2.04	2.05
	E273Oε1	2.04	2.04	2.05	2.08
	D292Oδ2	2.00	1.99	1.92	1.95
	E387Oε1	2.14	2.13	2.13	2.13
	OH^-	1.91	1.92	NA	NA
	$VO_4(O_3)$	2.27	2.27	-	-
	$PO_4(O_3)$	-	-	2.30	2.34
Ca_2	E387Oε1	2.57	2.57	2.57	2.57
	D479Oδ2	2.30	2.29	2.32	2.32
	D494Oδ2	2.37	2.37	2.39	2.38
	W_3	2.54	2.53	NA	NA
	W_4	2.50	2.44	NA	NA
	$VO_4(O_2)$	2.36	2.36	-	-
	$PO_4(O_2)$	-	-	3.05	3.03
Ca_3	D494Oδ2	2.48	2.49	2.44	2.44
	E532Oε2	2.27	2.26	2.27	2.28
	W_5	2.40	2.47	NA	NA
	W_6	2.44	2.47	NA	NA
	$VO_4(O_2)$	2.39	2.39	-	-
	$VO_4(O_4)$	2.52	2.51	-	-
	$PO_4(O_2)$	-	-	4.07	3.97
	$PO_4(O_4)$	-	-	2.06	1.95

* the distances between the metal ions and water molecules (W) could not be accurately determined due to the lower resolution of the native PhoX structure.

Before the structure of PhoX was solved, PhoX was predicted to be a Ca^{2+} binding protein based on significant homology to the Ca^{2+} binding site of diisopropylfluotophosphate (DFPase) which shows a conserved Asp residue (D479 in *P. fluorescens*) (Wu *et. al.* 2007). Like PhoX, DFPase is part of the 6 blade β propeller family. It has two Ca^{2+} ions in its active

site but only one Ca^{2+} is involved in catalysis while the second Ca^{2+} stabilizes the enzyme (Scharff *et. al.* 2001). According to this sequence homology, it has been suggested that there are two groups of PhoX, PhoX-I and PhoX-II. This is based on the whether the PhoX sequence has a Glu (PhoX-I) or Asn (PhoX-II) following D479 (Wu *et. al.* 2007; Zaheer *et. al.* 2009).

However, the structure of PhoX clearly shows that the active site of PhoX is more complex than previously hypothesized. As predicted, D479 coordinates Ca^{2+} but it coordinates two Ca^{2+}, Ca_1 and Ca_2, instead of just one Ca^{2+}. Moreover, D479 is only one of the nine residues involved in coordinating the five metal ions in the active site.

Studies on the eukaryotic PhoX from *Volvox carteri* also showed that it depends on Ca^{2+} for activity. The authors predicted that there is at least one EF-hand motif for calcium binding per PhoX domain (Hallmann 1999). The EF-hand domain is a helix-loop-helix structural domain in which the Ca^{2+} ions are coordinated by ligands within the loop. However, the EF-hand motif that was predicted was not in the active site but was in the unique α-helical subdomain of *P. fluorescens* PhoX and not involved in Ca^{2+} binding.

5.4.7 Site directed mutagenesis studies of active site residues

In an attempt to determine the role of individual metal ions in the phosphatase activity of PhoX, variants lacking single ligand-binding residues were created. Since most metal ions are coordinated by overlapping ligand residues, I was not able to create only variants which specifically target one metal ion. Table 5- 5 shows that all the variants were inactive. Circular dichroism (CD) analysis shows that all of the variants were folded (Figure 5- 23). However, there was a slight difference in the CD spectra at far-UV range for variants affecting the Fe^{3+} coordination (C179S, E194A, D292A) where these variants had a stronger negative ellipticity between 200-203nm. However, since I was not able to crystallize the PhoX variants, it cannot

be determined if these variants affected the fold of PhoX in a subtle way which cannot be detected by CD. In addition, it was observed that all the PhoX variants were less stable than wild type PhoX (Table 5- 5). This is based on the midpoint transition temperatures for unfolding (i.e. T_m) of each mutant. To eliminate the possibility of a direct effect of the mutated residue on the stability of PhoX, I repeated the unfolding experiment using EDTA treated wild type PhoX. Addition of EDTA also reduced the stability of wild type PhoX, suggesting that the removal of metals was most likely affecting the stability of PhoX and not the residues themselves.

Table 5- 5: Activity and stability of PhoX variants

The metals coordinating of each ligand are listed below. Specific activities were based on pNPP assays. T_m of variants were measured by difference scanning fluorimetry (DSF) using SYPRO Orange dye. Standard errors for specific activity were <0.02µmol/min/mg.

	Coordinated metal(s)	Specific activity (µmol/min/mg)	T_m (°C)
Wild type		2.65	49 °C
C179S	Fe_1	0.00	40 °C
E194A	Fe_1, Fe_2	0.00	44.5 °C
D292A	Fe_2	0.00	44.5 °C
D494A	Ca_2, Ca_3	0.04	41.5 °C
E532A	Ca_3	0.13	41.5 °C
200µM EDTA		0.00	43 °C

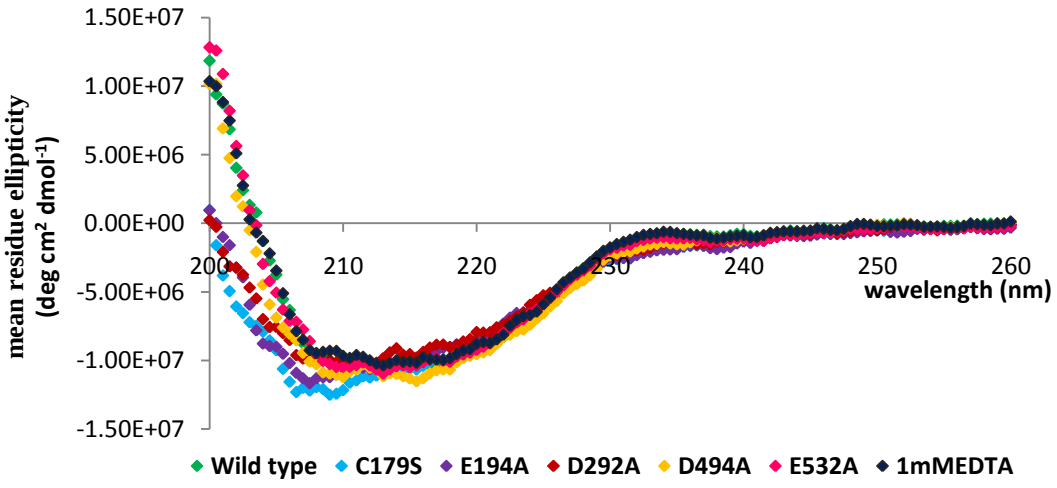


Figure 5- 23: CD spectra of PhoX variants

CD spectra of 0.5mg/ml PhoX variants in the far-UV range (200-260nm) in 1.3mM MOPS, 13mM NaCl, pH7.5.

5.4.8 Oxidation state of the iron cofactors

Analysis of coordination geometry (Müller *et. al.* 2003) hinted that the irons in PhoX were in the +3 oxidation state. In an attempt to confirm the oxidation state of PhoX, electron paramagnetic resonance (EPR) was used. This work was done in collaboration with Mr. James Lillington and Prof. Susan Lea, Dunn School of Pathology, University of Oxford. Initially, native PhoX showed no EPR signal. This can be expected from either a $Fe^{2+}Fe^{2+}$ antiferromagnetically coupled system or a $Fe^{3+}Fe^{3+}$ antiferromagnetically coupled system via the bridging carboxylate E194. Addition of EDTA, which would strip metal ions from PhoX, resulted in a Fe^{3+} characteristic peak at g=4.3 (Figure 5- 24). However, it is possible that irons, once complexed with EDTA, would be oxidized to Fe^{3+} (Seibig and Van Eldik 1997). If this was true, the oxidation state of the irons in PhoX would be +2 instead. To exclude this possibility, this experiment was repeated in the absence of oxygen. A Fe^{3+} peak was still observed, confirming that PhoX has Fe^{3+}. This peak could have originated from two sources. First, it could be from the EDTA-Fe^{3+} complex after Fe^{3+} was stripped from PhoX. The second source of the signal could be from the single Fe^{3+} that remained bound in PhoX after EDTA treatment. It was shown in Figure 5- 7 that treatment of EDTA resulted in a shift in absorbance maximum of PhoX from 550nm (purple) to 570nm (blue). Since PhoX was still coloured after EDTA treatment, this suggested that there was still ligand-to-metal charge transfer. Therefore, it can be hypothesized that the Cys coordinated Fe^{3+} remained bound in PhoX while the non-Cys coordinated Fe^{3+} was removed by EDTA. Since this Cys-Fe^{3+} remaining in PhoX is no longer antiferromagnetically coupled, a Fe^{3+} peak will be detected by EPR.

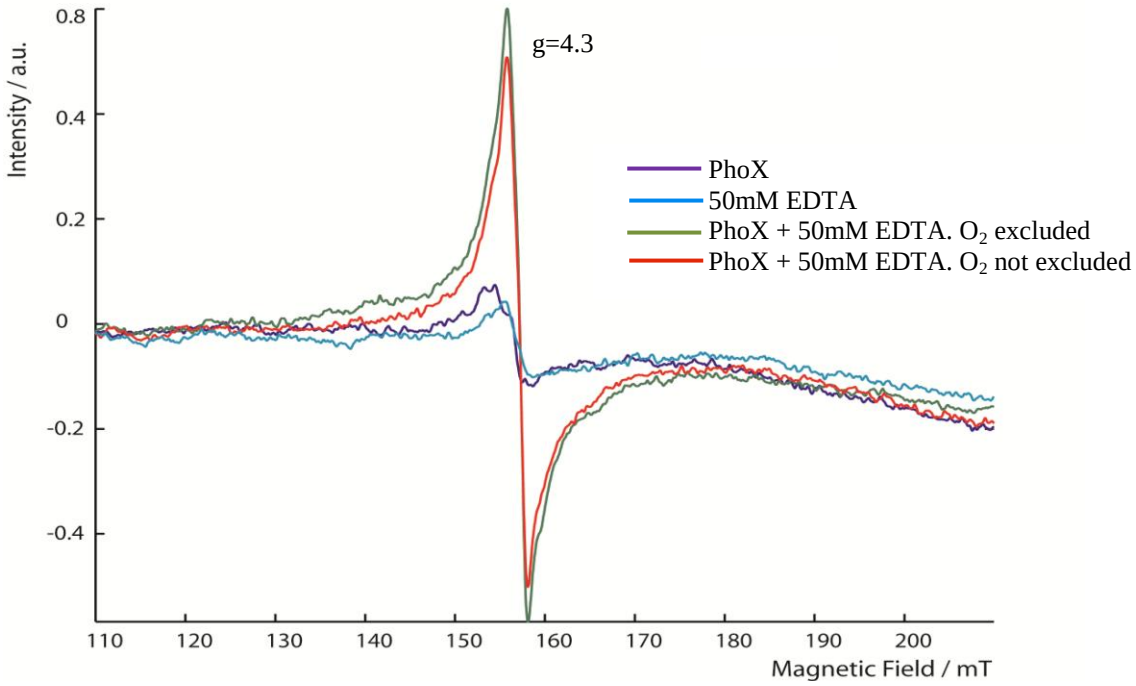
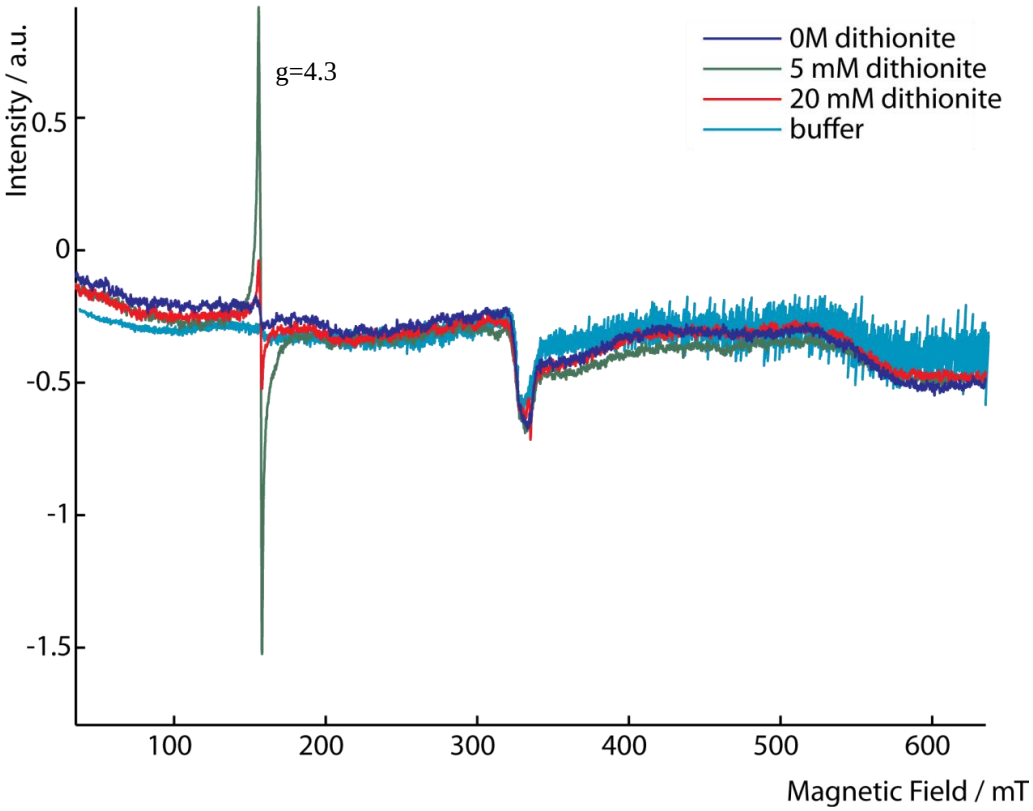


Figure 5- 24: EPR spectrum of EDTA treated PhoX

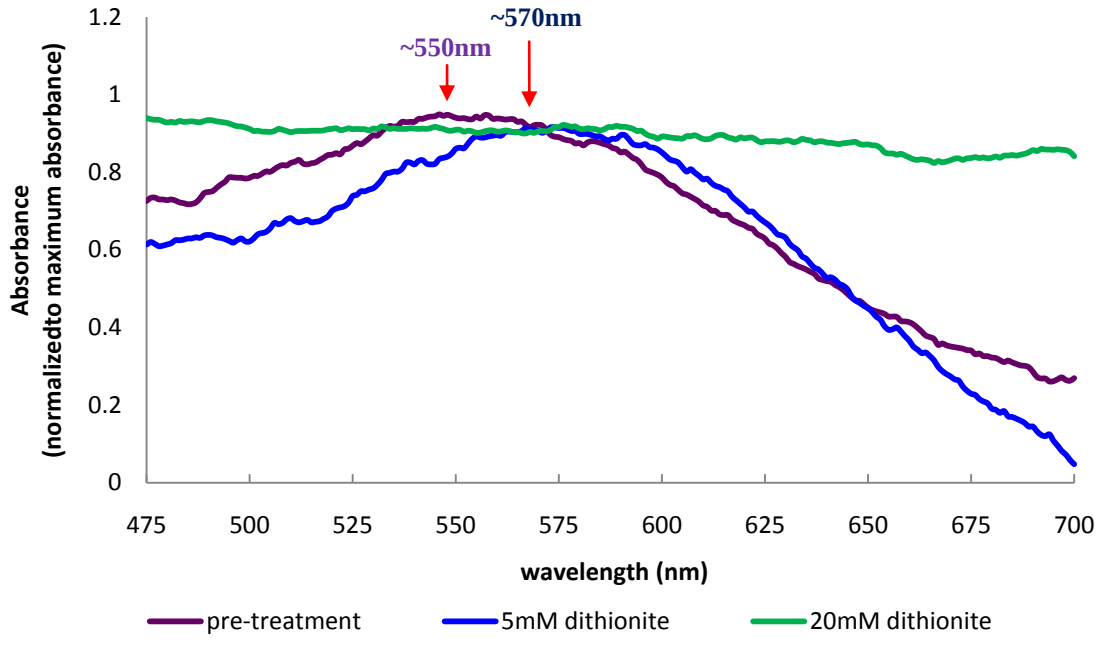
PhoX (10mg/ml) in 20mMOPS pH7.5, 200mM NaCl was treated with EDTA in the presence or absence of oxygen. 30% v/v glycerol was subsequently added. EPR operating conditions: microwave power, 2mW; microwave frequency, 9.425GHz; modulation, 5.0G at 100 kHz; temperature 10 K.

EDTA data seem to show that two Fe^{3+} are found in PhoX. In an attempt to investigate this further, I treated PhoX with the reducing agent, sodium dithionite. Like EDTA, addition of sodium dithionite resulted in the emergence of a Fe^{3+} characteristic peak (g=4.3) in the EPR spectrum (Figure 5- 25A). However, the characteristic spectrum of a coupled $Fe^{3+}Fe^{2+}$ system, which would have g = 1.86, 1.74, and 1.58 signals (Merkx and Averill 1998) , was not observed. This suggests that one of the Fe^{3+} was reduced to Fe^{2+} and the Fe^{2+} was subsequently released from PhoX. This could mean that Fe^{2+} is not able to bind stably to the active site. At this stage, a colour change of PhoX from purple (absorbance maximum ~550nm) to blue (absorbance maximum ~570nm) was observed (Figure 5- 25B), possibly due to ligand-to-metal charge transfer between Fe^{3+} and Cys. This suggests that the non-Cys coordinated Fe^{3+} was reduced first. A similar shift in wavelength of absorbance maximum was observed for EDTA. These observations in the presence of dithionite coincided with the loss of PhoX activity (Figure 5- 25C) suggesting that the non-Cys coordinated Fe^{3+} ion is required for the activity of PhoX. Figure 5- 25 also showed that titration of more sodium dithionite into the sample ultimately resulted in loss of EPR signal as the remaining Cys coordinated Fe^{3+} was reduced to Fe^{2+}. At this stage, PhoX became colourless due to the loss of ligand-to-metal charge transfer between Fe^{3+} and Cys .

A)


Intensity / a.u.
g=4.3
0M dithionite
5 mM dithionite
20 mM dithionite
buffer
0.5
0
-0.5
-1
-1.5
100
200
300
400
500
600
Magnetic Field / mT

B)


~570nm
~550nm
Absorbance
(normalizedto maximum absorbance)
1.2
1
0.8
0.6
0.4
0.2
0
475
500
525
550
575
600
625
650
675
700
wavelength (nm)
pre-treatment
5mM dithionite
20mM dithionite

C)

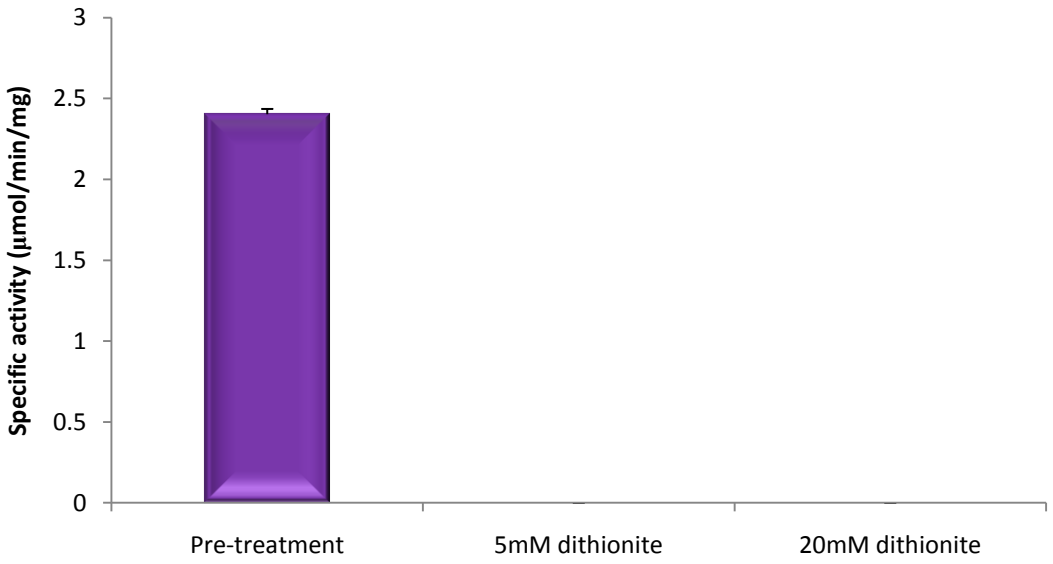


Figure 5- 25: Effect of dithionite on PhoX

(A) Absorbance spectra and (B) EPR spectra of PhoX (10mg/ml) treated with dithionite. PhoX was in 20mM MOPS pH 7.5, 200mM NaCl, 30% v/v glycerol. EPR operating conditions: microwave power, 2mW; microwave frequency, 9.425GHz; modulation, 5.0G at 100kHz; temperature 10K. (C) Activity assay of PhoX in the presence of dithionite. Activity of PhoX was determined by pNPP reaction assay in 5mM $CaCl_2$, 1M Tris-HCl, pH 8.0 in the absence or presence of dithionite. The reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Measurement was obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX. Error bar represents standard error for three independent experiments.

However, reduction of Fe^{3+} by dithionite did not affect the stability of PhoX as measured by DSF (Table 5- 6). In contrast, treatment of PhoX with EDTA reduced the stability of PhoX (Table 5- 5). This is because EDTA would remove the Ca^{2+} as well as the Fe^{3+} in PhoX while dithionite would not affect Ca^{2+}. This suggests that Ca^{2+} might be needed for stabilizing PhoX.

Table 5- 6: Effect of dithionite on stability of PhoX

Dithionite was added to 0.5mg/ml PhoX in 20mM MOPS, 200mM NaCl, pH7.5. T_m were measured by DSF using SYPRO Orange dye.

concentration of dithionite (mM)	Tm
0	49 ^oC
0.01	49 ^oC
1	49 ^oC
5	49 ^oC
10	49 ^oC

5.4.9 Cys-Fe^{3+} coordination

Most Cys-Fe bonds are found in haem groups or iron-sulfur clusters and are involved in redox reactions (Beinert 2000). It is hypothesized that the Fe-Cys bond found in PhoX is one of the rare few occasions where the Fe^{3+} acts solely as a Lewis acid and is not involved in any electron transfer reactions. This Cys residue (C179) is highly conserved across all organism including archaea and eukarya (Table 5- 7). In section 5.4.7, C179 was mutated to Ser (C179S), resulting in loss of PhoX activity. Since Ser is not a Fe^{3+} coordinating ligand, the result obtained using the C179S variant only implies that Fe_1 is required for activity of PhoX but is unable to provide information on whether C179 is required to coordinate Fe_1 or if other Fe^{3+} coordinating ligands can replace C179 as the coordinating ligand. To test if C179 is required for coordinating Fe_1, C179 was mutated to Asp, which is a common Fe^{3+} ligand and also coordinates Fe_2 in PhoX. Table 5- 8 shows that PhoX C179D variant is inactive. This suggests that C179 is likely required to coordinate Fe_1 and cannot be replaced by other Fe^{3+} coordinating ligand.

Table 5- 7: Sequence alignment of PhoX around C179

Sequence alignment of PhoX in different organisms across the phylogenetic tree. The conserved Cys is highlighted in red.

Phylogenetic position	Example species	PhoX sequence
α proteobacteria	*Labrenzia alexandrii*	…GTWNNCGNGVTPWGTYLTCEEN…
β proteobacteria	*Aromatoleum aromaticum*	…GTLNNCANGFTPWGTYLACEEN…
γ proteobacteria	*Pseudomonas fluorescens*	…GTFQNCANGKTPWGTYLTCEEN…
ε proteobacteria	*Campylobacter jejunI*	…GTFANCANGQTPWRTYITCEEN…
Gram positive bacteria	*Thermomonospora curvata*	…GTLNNCSNGQTPWGTYLACEEN…
Cyanobacteria	*Anabaena variabilis*	…GTLNNCANGYTPWGTYLTCEEN…
Archaea	*Halogeometricum borinquense*	…GTRINCYGDRSPWGTPLSAEEE…
Eukarya	*Volvox carteri* (Domain I)	…GVWTPCAGSMSPWGTHLGSEEY…
	Volvox carteri (Domain II)	…GLLTPCAGSITPWGTRLGSEEY…

Table 5- 8: Specific activity of PhoX_C179D

Activity of PhoX was determined by pNPP reaction assay in 5mM $CaCl_2$, 1M Tris-HCl, pH 8.0. The reaction was done at room temperature and quenched after 10min by addition of 1M NaOH. Measurement was obtained at A_{420nm}. Specific activity was calculated as µmol of P_i released per min per mg of PhoX.

	Specific activity (µmol/min/mg)
WT	2.651±0.021
C179D	0.004±0.001

5.4.10 PhoX active site

The native form of PhoX was crystallized in the absence of any additives. However, the electron density map showed density in the active site (Figure 5- 26A). This density appears to have a tetrahedral shape. It is hypothesized that this represents a tightly bound phosphate. However, since the buffer did not contain any phosphates, this phosphate is most likely incorporated into PhoX while still in the cell.

The crystal structure shows that Fe_1, Ca_1 and Fe_2 are arranged in an equilateral triangle (2×Fe-Ca triangular base) in the active site. A water molecule lies inside the triangle. This water molecule is less than 2Å from Fe_1 and Fe_2 and around 2.3Å from Ca_1 (Figure 5-26). As this water molecule is in close proximity with three metal ions that can cause

polarization of the H-O bonds in the water molecule, this water molecule is most likely in the activated OH^- state. Ca_2 and Ca_3 are located above the 2xFe-Ca triangular base.

In an attempt to further probe the mechanism of PhoX, PhoX was co-crystallized with vanadate. Vanadate is a close structural and chemical mimic of phosphate and has been extensively used to crystallize transition state mimics of phosphoryl transfer reactions (Davies and Hol 2004). Vanadate is a competitive inhibitor of PhoX with K_i of 98μM (Figure 5- 27A). Note that although PhoX was most active at pH8.0, the inhibition assay was done at pH10.0 as vanadate tends to polymerize at lower pH. In comparison, the K_i of phosphate at pH8.0 was 193μM (Figure 5- 27B). This means that vanadate binds more tightly to PhoX compared to phosphate.

The crystal structure of the PhoX:VO_4 complex was determined at a resolution of 1.4Å (Figure 5- 26B). The structure likely shows the intermediate state of PhoX. All four oxygen atoms of vanadate are coordinated by the metal ions. Three of the terminal oxygen atoms, O_1, O_2 and O_3, are coordinated to Fe_1, Ca_1 and Fe_2 respectively. O_2 of vanadate is also coordinated to Ca_2 and Ca_3. At the same time, Ca_3 coordinates O_4 of vanadate. The geometry of vanadate is trigonal bipyramidal instead of its usual tetrahedral. OH^- is directly underneath the vanadium and is in a position in line with the V-O_4 bond, where it mimics the attack to the phosphorus atom of phosphorylated substrates.

In contrast to vanadate, the phosphate in the native crystal maintained its tetrahedral geometry and is further away from the 2xFe-Ca triangular base than vanadate in the PhoX:VO_4 crystal. In addition, Ca_1 is no longer coordinated to O_2 of phosphate and Ca_3 appears to have lost contact with O_4 of phosphate (Figure 5- 26, Table 5- 4).

In both phosphate- and vanadate-bound PhoX, the substrate did not interact directly with the protein. Instead, the active site appears to be structured for positioning the metal ions, which in turn hydrolyze phosphates from substrates.

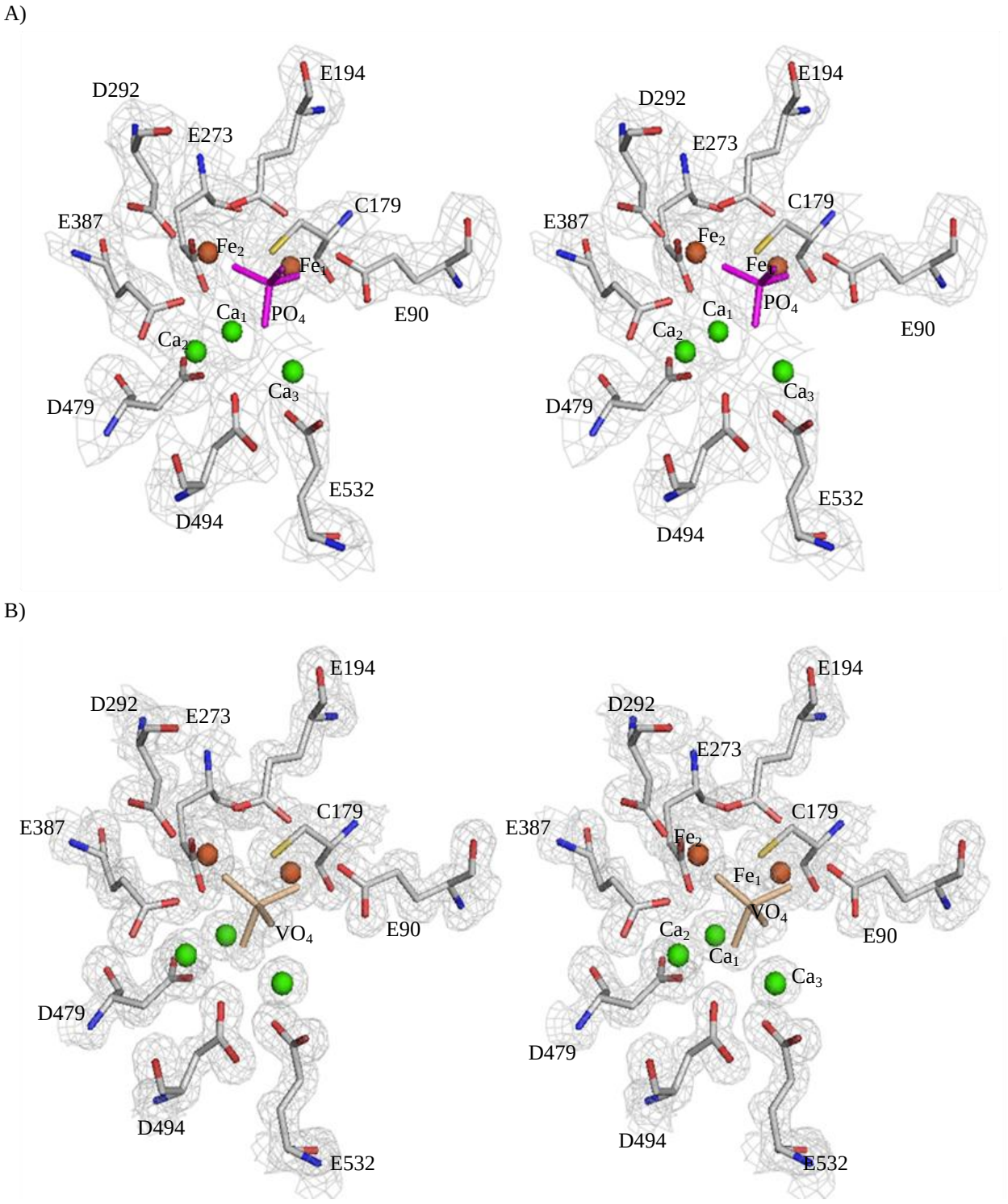
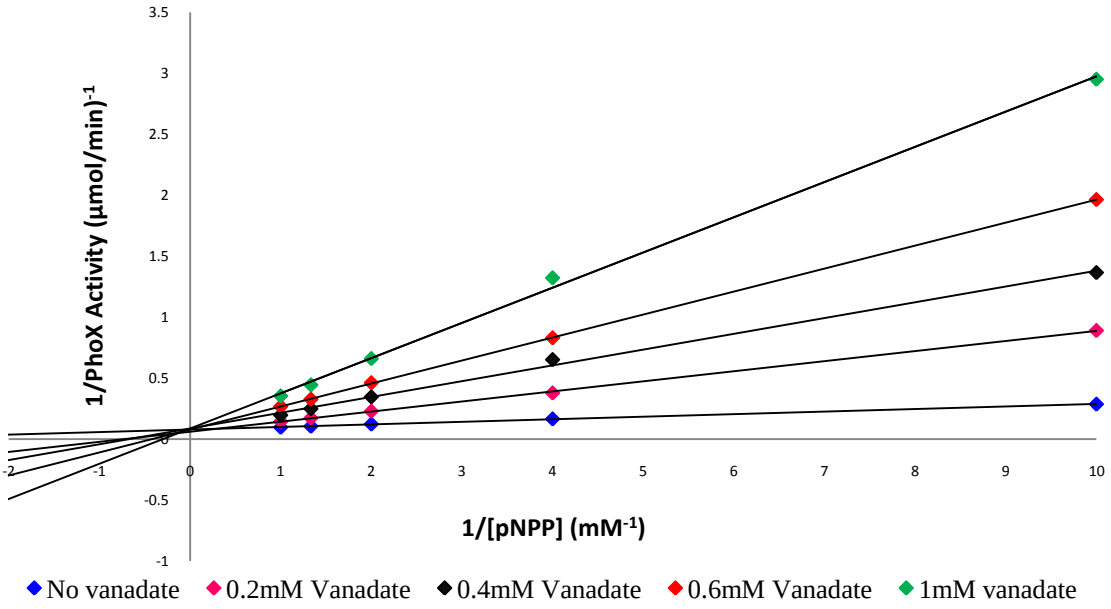


Figure 5- 26: Active site of PhoX in the presence of phosphate or vanadate

Stereo views of active site showing (A) phosphate or (B) vanadate bound. Overlaid in grey is the electron density, contoured at 1.0σ. Each oxygen atom of vanadate or phosphate is coordinated by at least one metal ion. Fe_1, Ca_1 and Fe_2 form a triangular base. Vanadate docks like a tripod on top of the 2xFe-Ca triangular base where its tetrahedral geometry is distorted to a trigonal bipyramidal geometry. On the other hand, phosphate maintains its tetrahedral geometry.

A)



B)

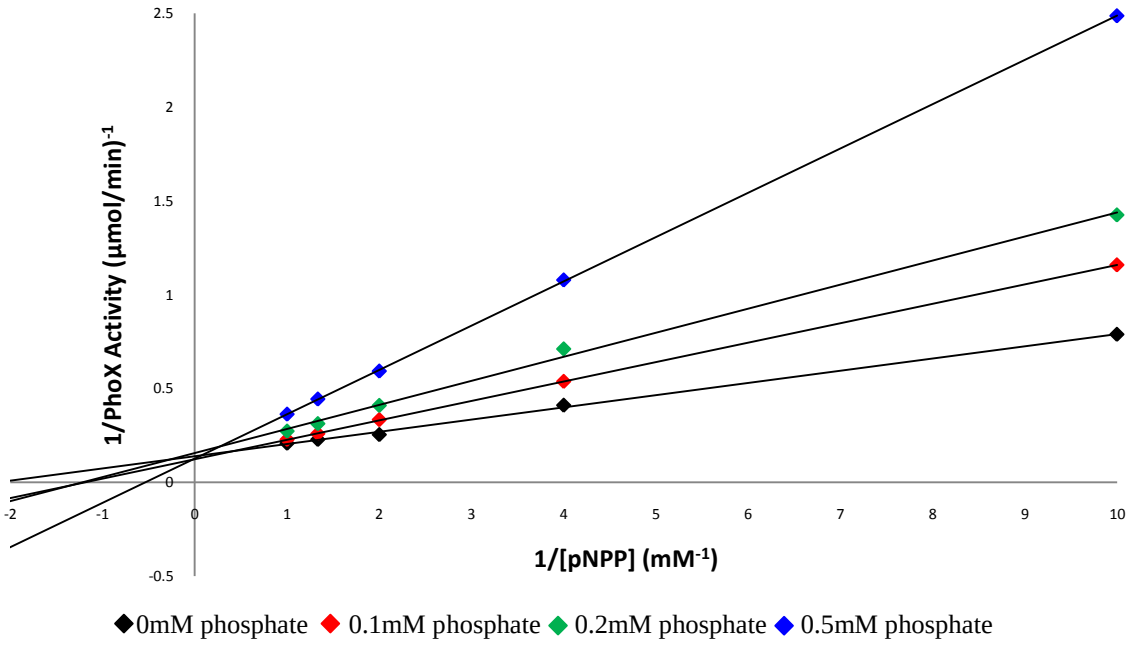


Figure 5- 27: Vanadate and phosphate inhibition assays

Double reciprocal plot of the activity of PhoX in the presence of increasing concentration of vanadate or phosphate based on pNPP phosphatase assay. (A) A 50mM sodium orthovanadate (Na_3VO_4) stock was prepared and adjusted to pH10.0. Vanadate inhibition assays were performed in 1M Tris-HCl, pH10.0. (B) Similar to vanadate inhibition assay except that the assay was done at 1M Tris-HCl, pH8.0. Measurements were taken at A_{420nm}. Activity was calculated as µmol of P_i released per min. Experiments were carried out in triplicates. The standard errors for PhoX activities are <0.03µmol/min/mg.

5.5 Discussion

In this chapter, *P. fluorescens* PhoX was shown to be an alkaline phosphomonoesterase with optimum activity at pH8.0. PhoX has a broad range of substrate, including substrates with C-O-P phosphate esters bonds, P-O-P bonds and N-P bonds. The X-ray structure of PhoX shows that it has a 6-blade β-propeller domain and two α-helical subdomains.

PhoX has a novel active site which is located at the rim of the β-propeller. It has two Fe^{3+} and three Ca^{2+} in its active site. The presence of a pentanuclear metal centre in the active site of a catalytic enzyme is very rare. As previously described, all five metal ions are required not only for enzymatic activity but also for the stability of PhoX. This could be explained by electrostatic interactions. Out of the nine residues coordinating the metals in the active site, there are eight carboxylates. This makes the active site highly negatively charged. The presence of metal ions would reduce the ionic repulsion between the acidic residues, thereby stabilizing PhoX. In addition, the presence of metal ions could be required to attract the substrates. The 8 carboxylates give a net 8- charge. The negative charge is offset by the three Ca^{2+} and two Fe^{3+} to give a net 4+ charge. This net positive charge would attract the negatively charged phosphate-containing substrate ($[R\text{-}PO_4]^{2-}$).

It is hypothesized that both Fe^{3+} in PhoX are not used for their redox activity but rather for their Lewis acid properties. Fe_2 is coordinated by the common Fe^{3+} coordinating ligands, Asp and Glu, while Fe_1 is coordinated by Glu and more interestingly, a Cys. Most Cys-Fe^{3+} bonds are found in haem groups or iron-sulfur proteins with iron-sulfur clusters that are involved in redox reactions (Beinert 2000). Examples of iron-sulfur proteins include ferredoxins, rubredoxins, hydrogenases, and nitrogenases. Nonetheless, not all iron-sulfur proteins are involved in electron transfer functions. Irons, mostly Fe^{3+}, can also serve as a Lewis acid in the active sites of enzymes. An example of this is nitrile hydratase. In this case,

two out of the three Cys thiol groups coordinating Fe were post-translationally modified to Cys-sulfenic and –sulfinic acid (Endo *et. al.* 2001). Another example of a Cys-Fe^{3+} coordination is found in YcfI, a putative structural protein from *Rhodopseudomonas palustris* (PDBID 2GYQ). The role of Fe^{3+} in this protein is not yet established. PhoX contains the first example of a Cys coordinated Fe^{3+} that is not involved in any redox activity and is not modified as a result of the enzymatic reaction. It is proposed that Fe^{3+} activates the water molecule directly beneath the phosphorous atom to OH^- to catalyze dephosphorylation of substrates.

The only other phosphatases that are known to use iron as a cofactor are purple acid phosphatases (PAPs). The best characterized PAPs are the mammalian enzymes uteroferrin and bovine spleen PAP, as well as the plant enzyme from red kidney beans. Generally, PAPs contain a dinuclear $Fe^{3+}M^{2+}$ centre in the active site, with M = Fe^{2+}, Mn^{2+} or Zn^{2+}. Mammalian PAPs, in particular, have two irons in the active site. It has been shown that PAP is only active when it is in the $Fe^{3+}Fe^{2+}$ form. The enzyme loses its activity when both irons are oxidized to give $Fe^{3+}Fe^{3+}$. When both irons are completely reduced to $Fe^{2+}Fe^{2+}$, the active site is disrupted (Halleen *et. al.* 1998). In contrast, PhoX is active when both irons are in the +3 oxidation state. This suggests that although PhoX has two irons similar to PAPs, they act via different mechanisms.

5.5.1 Proposed mechanism of PhoX

A model of the mechanism of PhoX can be proposed based on the PhoX:VO_4 complex. The crystal structure of PhoX:VO_4 shows a lack of covalent intermediate between vanadate and residues of PhoX. Instead, the reaction relies on the interaction between the substrate, the five metal ions and the water molecule within the 2xFe-Ca triangular base. The H-O bonds in this water molecule are likely to be polarized to form OH^-, which is a better

nucleophile for attacking the electrophilic phosphate, due its proximity to Fe_1, Fe_2 and Ca_1. The electrophilicity of the phosphorous atom is also improved by polarization of P-O bonds by the coordinating metal ions. This leads to the lowering of the activation energy required for the phosphatase reaction. It is also observed that the vanadate, which is a structural mimic of phosphate, in the crystallized PhoX:VO_4 complex had a trigonal bipyramidal geometry instead of the usual tetrahedral geometry. This could be because the metals in the 2xFe-Ca triangular base force the phosphate group in substrates to dock in the active site like a tripod with all of its non-bridging oxygen atoms coordinated. As a result, the P-O bonds in the phosphate group are strained. The strained P-O bonds could increase the reactivity of the phosphate containing substrate towards dephosphorylation by PhoX.

Since the central cavity is blocked by the bulky side chain of Trp388, phosphate-containing substrates can only enter the active site from above. The metal ions orientate OH^- and the substrate in the correct position for catalysis to occur. Fe_1, Ca_1 and Fe_2 coordinate the three terminal oxygen atoms of the substrate. It is proposed that hydrolysis of phosphate occurs via a S_N2 reaction mediated by OH^- as the nucleophile (Figure 5- 28). OH^- attacks the phosphorous atom at the opposite end of the leaving group. An intermediate state is formed and is stabilized by the metal ions. This backside attack also results in an inversion of configuration of the phosphate molecule. The leaving group then leaves the active site. This also weakens the bonds between O_1- Fe_1, O_2- Ca_1 and O_3-Fe_2. Inorganic phosphate is subsequently released from the active site to be utilized.

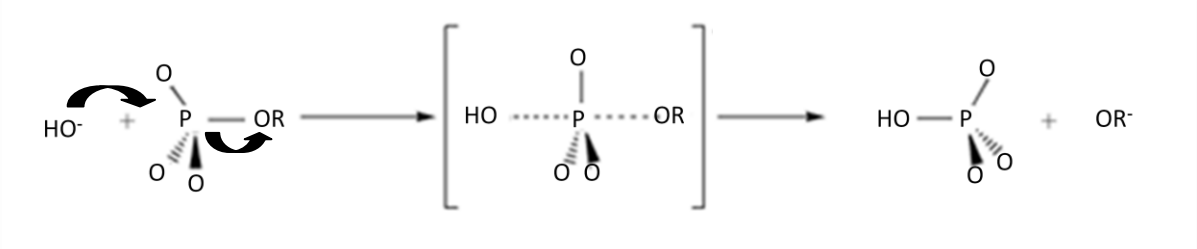


Figure 5- 28: Schematic representation of proposed mechanism of PhoX

Since all three terminal oxygens of phosphorous interact with metal ions in the active site, PhoX cannot hydrolyze phosphodiesters as phosphodiesters only have two terminal oxygens. However, it has also been published that PhoX from *P. multocida* is able to hydrolyze phosphodiesters (Wu *et. al.* 2007). This is the only example so far of PhoX having phosphodiesterase activity.

5.5.2 PhoX and the Tat system

Studies have shown that PhoX is transported by the Tat system in *P. fluorescens* (Monds *et. al.* 2006), *C. jejuni* (Van Mourik *et. al.* 2008) and *P. multocida* (Wu *et. al.* 2007). I have also shown in chapter 4 that *V. fischeri* PhoX can be transported by the *E. coli* Tat system *in vitro*. This raises the question as to why PhoX is transported by the Tat system. The Tat system only transports proteins that are fully folded. This means that PhoX has to be folded in the cytoplasm prior to transport. One possible explanation would be that PhoX requires cofactor insertion for folding. This is supported by previous studies which show that Fe^{3+} has to be incorporated during folding in the cytoplasm to produce active PhoX. PhoX that are purified from Fe^{3+} limiting conditions were inactive and its activity was not restored by addition of Fe^{3+} after purification (Monds *et. al.* 2006). This observation when taken together with the data from metal reconstitution experiments suggest that one of the Fe^{3+}, most likely the Cys liganded Fe^{3+} (i.e Fe_1), has to be incorporated during folding while the non-Cys liganded Fe^{3+} (i.e Fe_2) can be incorporated after folding.

5.5.3 Regulation of PhoX in the cytoplasm

Since PhoX is transported by the Tat system, PhoX has to be a properly folded, functional protein in the cytoplasm prior to transport. There are many phosphate-containing compounds in the cytoplasm such as ATP, GTP that can be dephosphorylated by PhoX. This could be undesirable for the organism. Hence, the activity of PhoX in the cytoplasm must

somehow be regulated. Regulation can be achieved by regulatory or chaperone proteins that could interact with the protein itself or the signal peptide. However, both *P. fluorescens* PhoX and *V. fischeri* PhoX can be overproduced in *E. coli* without severe defects in *E. coli* growth. This suggests that if PhoX is regulated by chaperone proteins, these chaperone proteins can also be found in *E. coli*. Another possible explanation is that the signal peptide itself could also have an intrinsic regulatory role in preventing PhoX from being active in the cytoplasm.

5.5.4 Comparison between PhoX and PhoA

PhoX and PhoA are both alkaline phosphatases that are induced under phosphate limited conditions to break down phosphate containing substrates to inorganic phosphate to be used by cells. However, PhoX and PhoA share neither sequence nor structural homology. PhoA is composed of mainly α-helices (Figure 5- 29) while PhoX has a 6 blade β-propeller domain with 2 α-helical subdomains (Figure 5- 13). In addition, the cofactors requirement is different between PhoA and PhoX. The cofactors for PhoA are two Zn^{2+} and one Mg^{2+}, while PhoX requires three Ca^{2+} and two Fe^{3+}. The active site of PhoA is less complex then PhoX. Unlike PhoX where all five metal ions are required for function and stability, in PhoA, only the two Zn^{2+} are directly involved in the catalysis while Mg^{2+} stabilizes its active conformation (Holtz and Kantrowitz 1999).

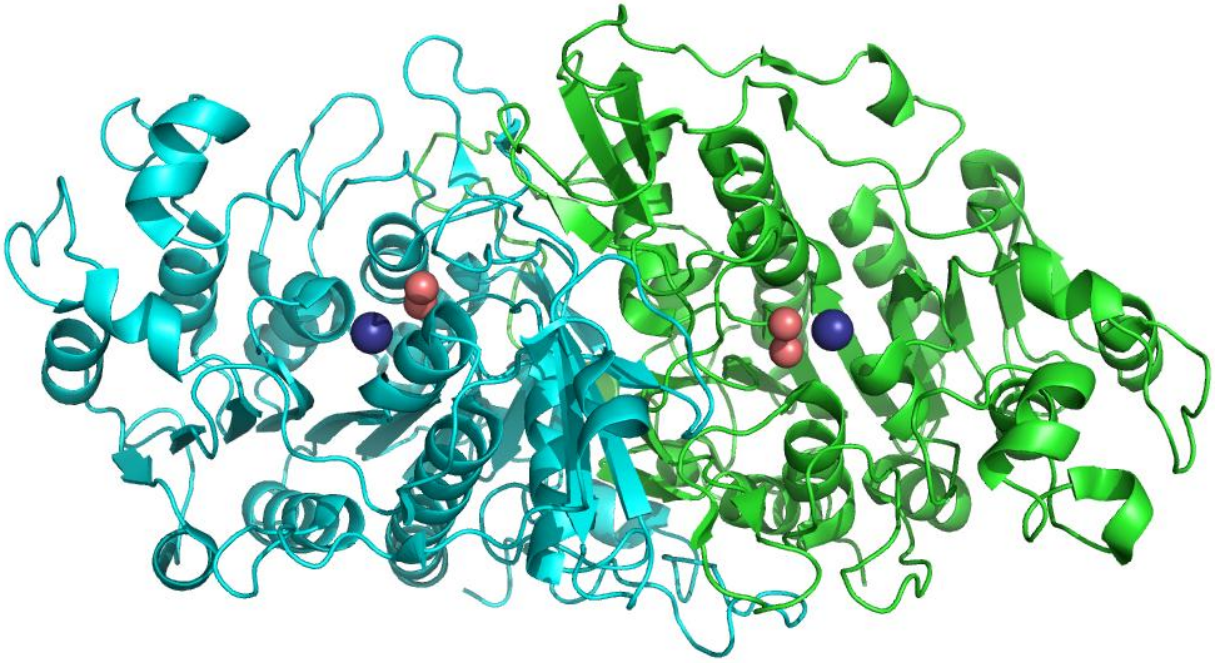


Figure 5- 29: X-ray crystal structure of PhoA
Crystal structure of PhoX (PDB ID: 1EW8) at 2.2Å resolution showing both chains A and B. Pink sphere: Zn^{2+}; blue sphere: Mg^{2+}.

Another distinction between PhoX and PhoA is the mechanism for phosphate hydrolysis. Phosphate hydrolysis in PhoA requires formation of a covalent intermediate with a conserved Ser residue (Holtz *et. al.* 1999) while the crystal structure of the PhoX:VO_4 complex did not show any indication of formation of a covalent intermediate. It is hypothesized that PhoX hydrolyses phosphates via a S_N2 reaction instead.

This study clearly shows that although PhoX is functionally equivalent to the *E. coli* alkaline phosphatase, PhoA, there is not similarity between PhoX and PhoA in protein sequence, structure, active site, metal cofactors or enzymatic mechanism. Hence PhoX represents a novel class of alkaline phosphatase with a unique active site.

Chapter 6:

Conclusions

and

future directions

6.1 Insights from *in vitro* Tat transport assays

The initial aim of this project was to optimize existing *in vitro* Tat transport assays for studying the bacterial Tat transport mechanism. *In vitro* assays are required for further biochemical investigations of the transport mechanism, such as bioenergetic requirements, kinetics and substrate recognition. The minimum components that are required for an *in vitro* Tat transport assay are IMVs enriched with TatABC, full length Tat substrates and a source of energization for the IMVs. The first *in vitro* Tat transport assay was developed by Yahr and Wickner almost a decade ago. However, the Tat transport efficiency was <1%. Since then, many improvements to the *in vitro* assay have been reported and transport efficiencies of up to 70% have been achieved (Yahr and Wickner 2001; Alami *et. al*. 2002; Bageshwar and Musser 2007). However, the 70% transport efficiency was only achieved with one substrate, SufI, while other substrates were transported at much lower efficiencies. Therefore, a more versatile *in vitro* assay is required to achieve high transport efficiencies when using various substrates.

In chapters 3 and 4, I have identified the parameters that contribute to the difference in transport efficiencies between different *in vitro* transport assay protocols. I found that the *in vitro* Tat transport assay is a relatively delicate system where transport efficiencies can be affected at all steps. This includes the process of IMV isolation, the choice of substrate used and assay conditions. The method used to isolate IMVs will dictate the amount of active Tat translocases in the IMVs and how well the IMVs can be energized. This was demonstrated when IMVs treated under harsh French press conditions produced smaller IMVs that are less transport competent.

The substrates used for *in vitro* Tat transport assays need to be Tat transport competent. This means that selected substrates have to be purified in its folded state, remain in solution and have an intact Tat signal peptide. Some proteins, such as SufI, are not easily

purified while other proteins, such as CueO, are degraded over time. In addition, the Tat signal peptide of Tat substrates is susceptible to cleavage during purification, especially when the N-terminus of the Tat substrate is not protected. This transport assay was set up to differentiate transported substrates from non-transported substrates by addition of proteases where non-transported substrates are digested by proteases while transported substrates are protected from protease digestion. Therefore, selected Tat substrates also need to be digested by proteases when proteases are added to the transport assay. A further complication is that susceptibility of Tat substrates towards protease digestion is shown to be affected by the presence of glycerol, which is required to store IMVs to maintain their Tat transport competency.

Transport of Tat substrates is shown to be inhibited by the presence of Cl^-. However, the inhibitory mechanism of Cl^- remains unknown. ATP was shown to be a better energization source for Tat transport *in vitro* compared to NADH. In addition, transport only occurs when substrates are added prior to or during energization of IMVs. Although this issue can be easily fixed by adding substrates in the right order, it would be interesting to understand why pre-energized IMVs, which still have a PMF when substrates are added minutes after energization, are not transport competent. Figure 6- 1 summarizes the issues faced in *in vitro* Tat transport assays. Substrates that were shown to be transported by this optimized *in vitro* Tat transport assay include the native *E. coli* Tat substrates, CueO and MdoD, and the native *V. fischeri* Tat substrate, PhoX.

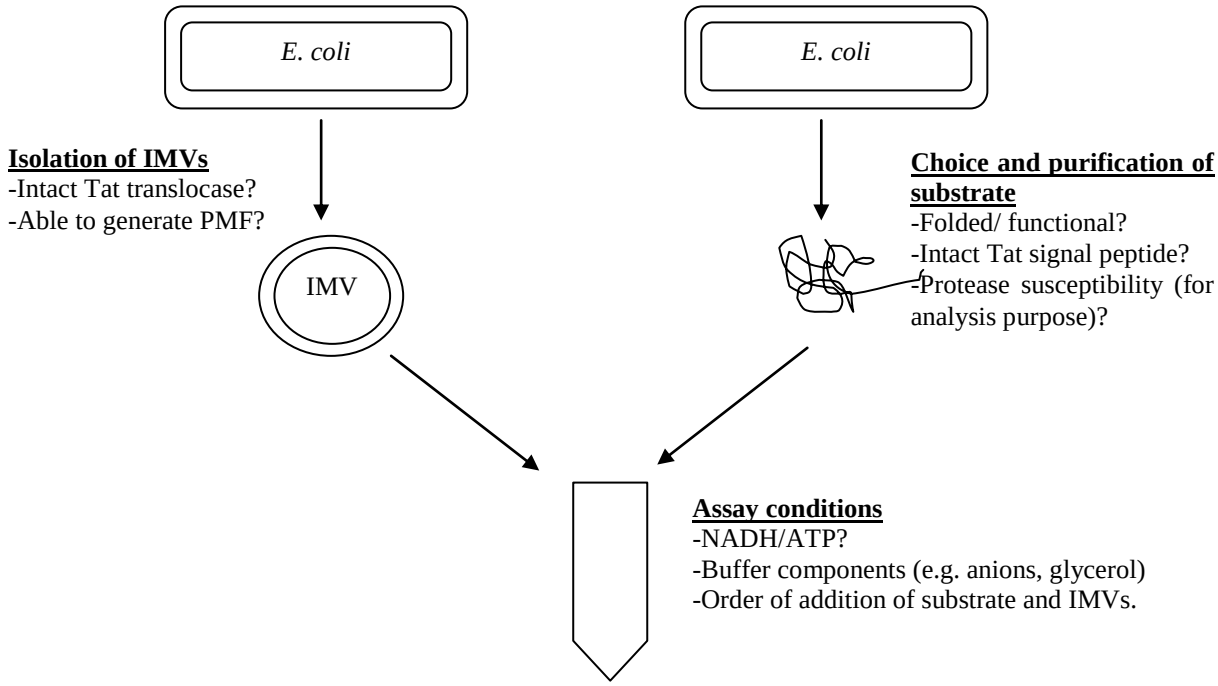


Figure 6- 1: Schematic representation of issues with the *in vitro* Tat transport system
A successful Tat transport depends on three main factors. First is how the IMVs are isolated. This would determine how many intact Tat translocases are in the IMVs and how efficient the IMVs are at generating a PMF. Second is the Tat substrate used. The Tat substrates chosen need to contain an intact Tat signal peptide, be folded correctly and be susceptible to protease degradation for analysis purposes. Finally, the assay conditions have to be taken into account. These include how the IMVs are energized, the buffer components and the order of addition of substrate.

I have also set up a fluorescence-based *in vitro* Tat transport assay which could be used for quantitative analyses. I did a time-course assay using this system and found that transport slows down and eventually ceases in a time-dependent manner *in vitro*. This phenomenon has also been observed in a previous study by Bageshwar and Musser (2007). This has two implications. First, the transport efficiency will not be improved unless this limitation is lifted; second, the *in vitro* transport system does not fully reflect *in vivo* conditions where transport is continuous. By understanding the reasons behind the limitations

of *in vitro* Tat transport assays, the transport efficiency may be significantly improved. More importantly, it could help shed some light on the mechanism of the Tat transport cycle.

I also applied the optimized *in vitro* assay to study the mechanism behind the Tat transport cycle. I showed that ΔpH can support Tat transport in the absence of ΔΨ. These results contradict previous findings where Tat transport was shown to be independent of ΔpH but requires at least two ΔΨ steps for transport (Bageshwar and Musser 2007). The issue of whether ΔpH or ΔΨ or both is required for Tat transport remains a controversial topic among researchers working on the bacterial or thylakoidal Tat system (Alder and Theg 2003; Di Cola *et. al.* 2005).

6.2 PhoX

In chapter 5, the structure and biochemical properties of PhoX were characterized. The 1.4Å resolution structure of PhoX was presented in this work. PhoX consists of a 6 blade β-propeller domain and two additional α-helical sub-domains. More work has to be done to determine the function of these α-helical sub-domains. PhoX has an active site which is located at the rim of the β-propeller domain. It contains a 2Fe-3Ca penta-nuclear metal binding site, which has never been reported before. These metal ions are coordinated by 8 acidic residues and 1 Cys. All five metal ions are required for both the stability and function of PhoX. Both Fe in PhoX have been identified as being in the +3 oxidation states. PhoX is active in the presence of a $Fe^{3+}Fe^{3+}$ system. The presence of a Cys liganded Fe^{3+} is quite interesting as the usual ligands for Fe^{3+} in proteins are Tyr, His, Glu and Asp. Cys-Fe^{3+} coordinations do exist but they are typically involved in redox reactions. PhoX is the first example of a protein with a Cys-Fe^{3+} coordination that is not involved in redox reactions. Hence, PhoX has a novel active site relative to other known enzymes.

PhoX was co-crystallized with vanadate, a chemical mimic of phosphate, to provide some mechanistic insights into the mode of action of PhoX. It is hypothesized that once in the active site, the phosphate group of the substrate interacts with all five of the metal ions. There is no direct interaction between the phosphate group and the protein residues. It is proposed that dephosphorylation occurs via a S_N2 reaction triggered by an activated OH^- which is in-line with the P-O-R bond in the substrate. The departure of the leaving group (R) occurs simultaneously with the backside attack by the OH^- on the phosphorous molecule. This leads to an inversion of the configuration of the phosphate group. As a result, free phosphate is released. In this mode of action, there are no covalent intermediates formed between the phosphate substrate and the protein

PhoX was shown to be a phosphomonoesterase with optimum activity at pH8.0. It is transported across the cytoplasmic membrane by the Tat system. This means that PhoX has to be folded prior to transport. This raises the question of why PhoX has to be folded in the cytoplasm. One possible reason is that it needs to incorporate at least one Fe^{3+} during folding. This has been demonstrated when reconstitution of folded but inactive apo PhoX with Fe^{3+} failed to restore its activity (Monds *et. al.* 2006).

PhoX may be functionally equivalent to the *E. coli* alkaline phosphatase, PhoA but I have shown that they have different structural properties, biochemical properties, metal cofactors and enzymatic mechanism. Hence, PhoX represents a novel class of alkaline phosphatase.

PhoX is also very interesting in the evolutionary sense. Although it has a β-propeller domain, the six blades share no sequence similarity. This is unlike other known β-propellers where blades tend to have a common motif due to blade amplification. It is possible that the β-propeller domain of PhoX had diverged extensively from ancestral β-propellers or this may be an example of convergence evolution.

There is an increasing interest in PhoX among environmental scientists since its discovery in the majority of marine bacteria. This discovery means that PhoX can be used as an additional P_i marker for P_i availability in the ocean to achieve a better understanding of the phosphate control in global biogeochemical cycles (Sebastian and Ammerman 2009).

6.3 Future directions

The *in vitro* Tat transport assay described in this thesis is based on analyzing the overall activity of the bulk of the Tat translocases in the IMVs while disregarding the heterogeneity between individual components. Hence, a better method to study the kinetics and mechanism of the individual components in Tat transport cycle could be to carry out single molecule fluorescence microscopy studies. To achieve this, two classes of Tat compatible substrates were developed in this work to be used in conjunction with a planar bilayer setup that has been developed by our collaborators which allows for incorporation of IMVs into large, planar bilayers. The first class of substrate is fluorescently labelled substrates. These can be used to visualize binding events of fluorescently labelled substrates to the Tat components. They can also be used for tracking the transport process by fluorescence. For this, I have successfully developed a fluorescently labelled substrate, CueO_GS3^{633}. The second class of substrates is substrates that can act on a fluorogenic compound. In this setup, only substrates that are transported will act on the fluorogenic compound on the *trans* side of the membrane and produce a fluorescent signal. Hence, any signal detected will be due to true transport events. The added advantage of this class of substrates over fluorescently labelled substrates is that the signal detected is amplified as one PhoX can act on many molecules of the fluorogenic compound. I have shown that the alkaline phosphatase PhoX, a native Tat substrate in *V. fischeri*, fits this profile. PhoX is transported by the *E. coli* Tat system and can act on commercially available fluorogenic

phosphatase substrates such as AttoPhos®. With the development of this system, it is now possible to use single molecule fluorescence techniques to answer questions regarding the Tat transport mechanism.

Chapter 7 :

Bibliography

Accardi, A. and Miller, C. (2004). "Secondary active transport mediated by a prokaryotic homologue of ClC Cl- channels." Nature 427(6977): 803-807.

Ahmed, I. and Krishnamoorthy, G. (1994). "Anomalous response of oxonol-V to membrane potential in mitochondrial proton pumps." Biochimica et Biophysica Acta (BBA) - Bioenergetics 1188(1-2): 131-138.

Alami, M., Lüke, I., Deitermann, S., Eisner, G., Koch, H.-G., Brunner, J. and Müller, M. (2003). "Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*." Molecular Cell 12(4): 937-946.

Alami, M., Trescher, D., Wu, L.F. and Müller, M. (2002). "Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli*." Journal of Biological Chemistry 277(23): 20499-503.

Alder, N.N. and Theg, S.M. (2003). "Energetics of protein transport across biological membranes: A study of the thylakoid ΔpH-dependent/cpTat pathway." Cell 112(2): 231-242.

Bachmann, J., Bauer, B., Zwicker, K., Ludwig, B. and Anderka, O. (2006). "The Rieske protein from *Paracoccus denitrificans* is inserted into the cytoplasmic membrane by the twin-arginine translocase." FEBS Journal 273(21): 4817-4830.

Bageshwar, U.K. and Musser, S.M. (2007). "Two electrical potential-dependent steps are required for transport by the *Escherichia coli* Tat machinery." The Journal of cell biology 179(1): 87-99.

Bageshwar, U.K., Whitaker, N., Liang, F. and Musser, S.M. (2009). "Interconvertibility of lipid- and translocon-bound forms of the bacterial Tat precursor pre-SufI." Molecular microbiology 74(1): 209-226.

Barnett, J.P., Eijlander, R.T., Kuipers, O.P. and Robinson, C. (2008). "A Minimal Tat System from a Gram-positive Organism: A Bifunctional TatA subunit participates in discrete TatAC and TatA complexes." Journal of Biological Chemistry 283(5): 2534-2542.

Barnett, J.P., van der Ploeg, R., Eijlander, R.T., Nenninger, A., Mendel, S., Rozeboom, R., Kuipers, O.P., van Dijl, J.M. and Robinson, C. (2009). "The twin-arginine translocation (Tat) systems from *Bacillus subtilis* display a conserved mode of complex organization and similar substrate recognition requirements." FEBS Journal 276(1): 232-243.

Barrett, C.M.L., Mathers, J.E. and Robinson, C. (2003). "Identification of key regions within the *Escherichia coli* TatAB subunits." FEBS Letters 537(1-3): 42-46.

Bauer, B.W. and Rapoport, T.A. (2009). "Mapping polypeptide interactions of the SecA ATPase during translocation." Proceedings of the National Academy of Sciences 106(49): 20800-20805.

Beinert, H. (2000). "Iron-sulfur proteins: ancient structures, still full of surprises." Journal of Biological Inorganic Chemistry 5(1): 2-15.

Bendtsen, J., Nielsen, H., Widdick, D., Palmer, T. and Brunak, S. (2005). "Prediction of twin-arginine signal peptides." BMC Bioinformatics 6(1): 167.

Berg, B.L., Li, J., Heider, J. and Stewart, V. (1991). "Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. I. Nucleotide sequence of the fdnGHI operon and evidence that opal (UGA) encodes selenocysteine." Journal of Biological Chemistry 266(33): 22380-22385.

Berks, B.C. (1996). "A common export pathway for proteins binding complex redox cofactors?" Molecular microbiology 22(3): 393-404.

Berks, B.C., Sargent, F. and Palmer, T. (2000). "The Tat protein export pathway." Molecular microbiology 35(2): 260-274.

Bernstein, H.D., Poritz, M.A., Strub, K., Hoben, P.J., Brenner, S. and Walter, P. (1989). "Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle." Nature 340(6233): 482-486.

Berthelmann, F., Mehner, D., Richter, S., Lindenstrauβ, U., Lunsdorf, H., Hause, G. and Brüser, T. (2008). "Recombinant expression of tatABC and tatAC results in the formation of interacting cytoplasmic TatA tubes in Escherichia coli." Journal of Biological Chemistry 283(37):25281-25289.

Bieleski, R.L. (1973). "Phosphate pools, phosphate transport, and phosphate availability." Annual Review of Plant Physiology 24(1): 225-252.

Blattner, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. (1997). "The Complete Genome Sequence of Escherichia coli K-12." Science 277(5331): 1453-1462.

Blaudeck, N., Kreutzenbeck, P., Müller, M., Sprenger, G.A. and Freudl, R. (2005). "Isolation and characterization of bifunctional *Escherichia coli* TatA mutant proteins that allow efficient Tat-dependent protein translocation in the absence of TatB." Journal of Biological Chemistry 280(5): 3426-3432.

Blum, M. M. and Chen, J.C.H. (2010). "Structural characterization of the catalytic calcium-binding site in diisopropyl fluorophosphatase (DFPase)--Comparison with related [beta]-propeller enzymes." Chemico-Biological Interactions 187(1-3): 373-379.

Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C. and Palmer, T. (1998). "An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria." Journal of Biological Chemistry 273(29): 18003-18006.

Bolhuis, A. (2002). "Protein transport in the halophilic archaeon Halobacterium sp. NRC-1: a major role for the twin-arginine translocation pathway?" Microbiology 148(11): 3335-3346.

Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C.M.L. and Robinson, C. (2001). "TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*." Journal of Biological Chemistry 276(23): 20213-20219.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. and Schwede, T. (2008). "Protein structure homology modeling using SWISS-MODEL workspace." Nature Protocols 4(1): 1-13.

Breyton, C., Haase, W., Rapoport, T.A., Kuhlbrandt, W. and Collinson, I. (2002). "Three-dimensional structure of the bacterial protein-translocation complex SecYEG." Nature 418(6898): 662-665.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Smart, O.S., Vonrhein, C. and Womack, T. (2008). BUSTER-TNT 2.5.1 and autoBUSTER 1.3.1. Cambridge., Global Phasing Ltd

Brüser, T. (2007). "The twin-arginine translocation system and its capability for protein secretion in biotechnological protein production." Applied microbiology and biotechnology 76(1): 35-45.

Buchanan, G., Sargent, F., Berks, B. and Palmer, T. (2001). "A genetic screen for suppressors of *Escherichia coli* Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif." Archives of Microbiology 177(1): 107-112.

Carter, C.W. Jr., Kraut, J., Freer, S.T., Xuong, N.H., Alden, R.A. and Bartsch, R.G. (1974). "Two-Angstrom crystal structure of oxidized chromatium high potential iron protein." Journal of Biological Chemistry 249(13): 4212-4225.

Casadaban, M.J. (1976). "Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu." Journal of Molecular Biology 104(3): 541-555.

Castle, A.M., Macnab, R.M. and Shulman, R.G. (1986). "Measurement of intracellular sodium concentration and sodium transport in *Escherichia coli* by 23Na nuclear magnetic resonance." Journal of Biological Chemistry 261(7): 3288-3294.

Chan, C.S., Zlomislic, M.R., Tieleman, D.P. and Turner, R.J. (2007). "The TatA subunit of *Escherichia coli* twin-arginine translocase has an N-in topology." Biochemistry 46(25): 7396-7404.

Chanal, A., Santini, C.L. and Wu, L.F. (2003). "Specific inhibition of the translocation of a subset of *Escherichia coli* TAT substrates by the TorA signal peptide." Journal of Molecular Biology 327(3): 563-570.

Chaudhuri, I., Söding, J. and Lupas, A.N. (2008). "Evolution of the beta-propeller fold." Proteins: Structure, Function, and Bioinformatics 71(2): 795-803.

Cline, K., Ettinger, W.F. and Theg, S.M. (1992). "Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP." Journal of Biological Chemistry 267(4): 2688-2696.

Cline, K. and McCaffery, M. (2007). "Evidence for a dynamic and transient pathway through the TAT protein transport machinery." The EMBO journal 26(13): 3039-3049

Cline, K. and Mori, H. (2001). "Thylakoid deltapH-dependent precursor proteins bind to a cpTat-Hcf106 complex before Tha4-dependent transport." The Journal of cell biology 154(4): 719-730.

Coleman, J.E. (1992). "Structure and mechanism of alkaline phosphatase." Annual Review of Biophysics and Biomolecular Structure 21(1): 441-483.

Cooper, C.E., Bruce, D. and Nicholls, P. (1990). "Use of Oxonol V as a probe of membrane potential in proteoliposomes containing cytochrome oxidase in the submitochondrial orientation." Biochemistry 29(16): 3859-3865.

Cristóbal, S., de Gier, J.W., Nielsen, H. and von Heijne, G. (1999). "Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*." The EMBO journal 18(11): 2982-90.

Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004). "WebLogo: A sequence logo generator." Genome Research 14(6): 1188-1190.

Dabney-Smith, C. and Cline, K. (2009). "Clustering of C-terminal stromal domains of Tha4 homo-oligomers during translocation by the Tat protein transport system." Molecular Biology of the Cell 20(7): 2060-2069.

Dabney-Smith, C., Mori, H. and Cline, K. (2006). "Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport." Journal of Biological Chemistry 281(9): 5476-5483.

Dalbey, R.E. and Robinson, C. (1999). "Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane." Trends in Biochemical Sciences 24(1): 17-22.

Davies, D.R. and Hol, W.G.J. (2004). "The power of vanadate in crystallographic investigations of phosphoryl transfer enzymes." FEBS Letters 577(3): 315-321.

De Buck, E., Lammertyn, E. and Anné, J. (2008). "The importance of the twin-arginine translocation pathway for bacterial virulence." Trends in Microbiology 16(9): 442-453.

De Buck, E., Vranckx, L., Meyen, E., Maes, L., Vandersmissen, L., Anné, J. and Lammertyn, E. (2007). "The twin-arginine translocation pathway is necessary for correct membrane insertion of the Rieske Fe/S protein in Legionella pneumophila." FEBS Letters 581(2): 259-264.

de Keyzer, J., van der Does, C. and Driessen, A.J.M. (2003). "The bacterial translocase: a dynamic protein channel complex." Cellular and Molecular Life Sciences 60(10): 2034-2052.

de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S.B., Müller, M., Sargent, F., Palmer, T. and Berks, B.C. (2002). "Oligomeric Properties and Signal Peptide Binding by *Escherichia coli* Tat Protein Transport Complexes." Journal of Molecular Biology 322(5): 1135-1146.

DeLisa, M.P., Lee, P., Palmer, T. and Georgiou, G. (2004). "Phage shock protein PspA of *Escherichia coli* relieves saturation of protein export via the Tat pathway." Journal of Bacteriology 186(2): 366-373.

DeLisa, M.P., Tullman, D. and Georgiou, G. (2003). "Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway." Proceedings of the National Academy of Sciences 100(10): 6115-6120.

Di Cola, A., Bailey, S. and Robinson, C. (2005). "The Thylakoid pH/psi Are Not Required for the Initial Stages of Tat-dependent Protein Transport in Tobacco Protoplasts." Journal of Biological Chemistry 280(50): 41165-41170.

Dilks, K., Gimenez, M.I. and Pohlschröder, M. (2005). "Genetic and Biochemical Analysis of the Twin-Arginine Translocation Pathway in Halophilic Archaea." Journal of Bacteriology 187(23): 8104-8113.

Dilks, K., Rose, R.W., Hartmann, E. and Pohlschröder, M. (2003). "Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey." Journal of Bacteriology 185(4): 1478-83.

Driessen, A.J.M. and Nouwen, N. (2008). "Protein translocation across the bacterial cytoplasmic membrane." Annual Review of Biochemistry 77(1).

Dubini, A. and Sargent, F. (2003). "Assembly of Tat-dependent [NiFe] hydrogenases: identification of precursor-binding accessory proteins." FEBS Letters 549(1-3): 141-146.

Emsley, P. and Cowtan, K. (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D 60(12 Part 1): 2126-2132.

Endo, I., Nojiri, M., Tsujimura, M., Nakasako, M., Nagashima, S., Yohda, M. and Odaka, M. (2001). "Fe-type nitrile hydratase." Journal of Inorganic Biochemistry 83(4): 247-253.

Feilmeier, B.J., Iseminger, G., Schroeder, D., Webber, H. and Phillips, G.J. (2000). "Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*." Journal of Bacteriology 182(14): 4068-4076.

Finazzi, G., Chasen, C., Wollman, F.A. and de Vitry, C. (2003). "Thylakoid targeting of Tat passenger proteins shows no [Delta]pH dependence in vivo." The EMBO journal 22(4): 807-815.

Fincher, V., Dabney-Smith, C. and Cline, K. (2003). "Functional assembly of thylakoid deltapH-dependent/Tat protein transport pathway components *in vitro*." European Journal of Biochemistry 270(24): 4930-4941.

Fischer, S., Graber, P. and Turina, P. (2000). "The activity of the ATP synthase from *Escherichia coli* is regulated by the transmembrane proton motive force." Journal of Biological Chemistry 275(39): 30157-30162.

Frielingsdorf, S. and Klösgen, R.B. (2007). "Prerequisites for terminal processing of thylakoidal Tat substrates." Journal of Biological Chemistry 282(33): 24455-24462.

Gerard, F. and Cline, K. (2006). "Efficient Twin Arginine Translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site." Journal of Biological Chemistry 281(10): 6130-6135.

Gille, C. and Frommel, C. (2001). "STRAP: editor for STRuctural Alignments of Proteins." Bioinformatics 17(4): 377-378.

Giménez, M., I., Dilks, K. and Pohlschröder, M. (2007). "*Haloferax volcanii* twin-arginine translocation substates include secreted soluble, C-terminally anchored and lipoproteins." Molecular microbiology 66(6): 1597-1606.

Gohlke, U., Pullan, L., McDevitt, C.A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H.R. and Berks, B.C. (2005). "The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter." Proceedings of the National Academy of Sciences 102(30): 10482-10486.

Gouffi, K., Gerard, F., Santini, C.L. and Wu, L.F. (2004). "Dual topology of the *Escherichia coli* TatA protein." Journal of Biological Chemistry 279(12): 11608-11615.

Grass, G. and Rensing, C. (2001). "CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*." Biochemical and Biophysical Research Communications 286(5): 902-908.

Graubner, W., Schierhorn, A. and Brüser, T. (2007). "DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone." Journal of Biological Chemistry 282(10): 7116-7124.

Groll, M., Bochtler, M., Brandstetter, H., Clausen, T. and Huber, R. (2005). "Molecular machines for protein degradation." ChemBioChem 6(2): 222-256.

Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R. and Sprenger, G.A. (1999). "The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding." European Journal of Biochemistry 263(2): 543-551.

Halleen, J.M., Kaija, H., Stepan, J.J., Vihko, P. and Väänänen, H.K. (1998). "Studies on the protein tyrosine phosphatase activity of tartrate-resistant acid phosphatase." Archives of Biochemistry and Biophysics 352(1): 97-102.

Hallmann, A. (1999). "Enzymes in the extracellular matrix of Volvox: an inducible, calcium-dependent phosphatase with a modular composition." Journal of Biological Chemistry 274 1691-1697.

Hartl, F.U., Lecker, S., Schiebel, E., Hendrick, J.P. and Wickner, W. (1990). "The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane." Cell 63(2): 269-279.

Hatzixanthis, K., Palmer, T. and Sargent, F. (2003). "A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase." Molecular microbiology 49(5): 1377-1390.

Heron, A.J., Thompson, J.R., Cronin, B., Bayley, H. and Wallace, M.I. (2009). "Simultaneous measurement of ionic current and fluorescence from single protein pores." Journal of the American Chemical Society 131(5): 1652-1653.

Hicks, M.G., de Leeuw, E., Porcelli, I., Buchanan, G., Berks, B.C. and Palmer, T. (2003). "The *Escherichia coli* twin-arginine translocase: conserved residues of TatA and TatB family components involved in protein transport." FEBS Letters 539(1-3): 61-67.

Hinsley, A.P., Stanley, N.R., Palmer, T. and Berks, B.C. (2001). "A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif." FEBS Letters 497(1): 45-49.

Holland, I.B. (2004). "Translocation of bacterial proteins--an overview." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1694(1-3): 5-16.

Holtz, K.M. and Kantrowitz, E.R. (1999). "The mechanism of the alkaline phosphatase reaction: insights from NMR, crystallography and site-specific mutagenesis." FEBS Letters 462(1-2): 7-11.

Holtz, K.M., Stec, B. and Kantrowitz, E.R. (1999). "A model of the transition state in the alkaline phosphatase reaction." Journal of Biological Chemistry 274(13): 8351-4.

Holzapfel, E., Eisner, G., Alami, M., Barrett, C.M.L., Buchanan, G., Lüke, I., Betton, J.M., Robinson, C., Palmer, T., Moser, M. and Müller, M. (2007). "The entire N-Terminal half of TatC is involved in Twin-Arginine precursor binding." Biochemistry 46(10): 2892-2898.

Hou, B., Frielingsdorf, S. and Klösgen, R.B. (2006). "Unassisted Membrane Insertion as the Initial Step in [Delta]pH/Tat-dependent Protein Transport." Journal of Molecular Biology 355(5): 957-967.

Hu, Y., Zhao, E., Li, H., Xia, B. and Jin, C. (2010). "Solution NMR structure of the TatA component of the twin-arginine protein transport system from Gram-positive bacterium *Bacillus subtilis*." Journal of the American Chemical Society 132(45): 15942-15944.

Huang, C.-S., Kopacz, S.J. and Lee, C.-P. (1983). "Mechanistic differences in the energy-linked fluorescence decreases of 9-aminoacridine dyes associated with bovine heart submitochondrial membranes." Biochimica et Biophysica Acta (BBA) - Bioenergetics 722(1): 107-115.

Hunt, J.F., Weinkauf, S., Henry, L., Fak, J.J., McNicholas, P., Oliver, D.B. and Deisenhofer, J. (2002). "Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA." Science 297(5589): 2018-2026.

Hutcheon, G.W. and Bolhuis, A. (2003). "The archaeal twin-arginine translocation pathway." Biochemical Society Transactions 31(Pt 3): 686-689.

Ignatova, Z., Hörnle, C., Nurk, A. and Kasche, V. (2002). "Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery." Biochemical and Biophysical Research Communications 291(1): 146-149.

Ito, K. and Mori, H. (2009). The Sec protein translocation pathway. Bacterial secreted proteins. Wooldridge, K. Norfolk, UK, Caister Academic Press**:** 3-22.

Iyer, R., Iverson, T.M., Accardi, A. and Miller, C. (2002). "A biological role for prokaryotic ClC chloride channels." Nature 419(6908): 715-718.

Ize, B., Porcelli, I., Lucchini, S., Hinton, J.C., Berks, B.C. and Palmer, T. (2004). "Novel phenotypes of *Escherichia coli* tat mutants revealed by global gene expression and phenotypic analysis." Journal of Biological Chemistry 279(46): 47543-47554.

Ize, B., Stanley, N.R., Buchanan, G. and Palmer, T. (2003). "Role of the *Escherichia col*i Tat pathway in outer membrane integrity." Molecular microbiology 48(5): 1183-1193.

Jack, R.L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T. and Sargent, F. (2004). "Coordinating assembly and export of complex bacterial proteins." The EMBO journal 23(20): 3962–3972.

Jack, R.L., Sargent, F., Berks, B.C., Sawers, G. and Palmer, T. (2001). "Constitutive expression of *Escherichia coli* tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth." Journal of Bacteriology 183(5): 1801-1804.

Jongbloed, J.D.H., Antelmann, H., Hecker, M., Nijland, R., Bron, S., Airaksinen, U., Pries, F., Quax, W.J., van Dijl, J.M. and Braun, P.G. (2002). "Selective contribution of the twin-arginine translocation pathway to protein secretion in *Bacillus subtilis*." Journal of Biological Chemistry 277(46): 44068-44078.

Jongbloed, J.D.H., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S. and van Dijl, J.M. (2004). "Two minimal Tat translocases in *Bacillus*." Molecular microbiology 54(5): 1319-1325.

Kim, E.E. and Wyckoff, H.W. (1991). "Reaction mechanism of alkaline phosphatase based on crystal structures : Two-metal ion catalysis." Journal of Molecular Biology 218(2): 449-464.

Kipping, M., Lilie, H., Lindenstrauß, U., Andreesen, J.R., Griesinger, C., Carlomagno, T. and Brüser, T. (2003). "Structural studies on a twin-arginine signal sequence." FEBS Letters 550(1-3): 18-22.

Kreutzenbeck, P., Kroger, C., Lausberg, F., Blaudeck, N., Sprenger, G.A. and Freudl, R. (2007). "*Escherichia coli* twin arginine (Tat) mutant translocases possessing relaxed signal peptide recognition specificities." Journal of Biological Chemistry 282(11): 7903-7911.

Kumamoto, C.A. (1989). "*Escherichia coli* SecB protein associates with exported protein precursors *in vivo*." Proceedings of the National Academy of Sciences 86(14): 5320-5324.

Kwan, D. and Bolhuis, A. (2010). "Analysis of the twin-arginine motif of a haloarchaeal Tat substrate." FEMS Microbiology Letters 308(2):138-143.

Kwan, D.C., Thomas, J.R. and Bolhuis, A. (2008). "Bioenergetic requirements of a Tat-dependent substrate in the halophilic archaeon *Haloarcula hispanica*." FEBS Journal 275(24): 6159-6167.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. and Ben-Tal, N. (2005). "ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures." Nucleic Acids Research 33(Web Server Issue): W299-302.

Leake, M.C., Greene, N.P., Godun, R.M., Granjon, T., Buchanan, G., Chen, S., Berry, R.M., Palmer, T. and Berks, B.C. (2008). "Variable stoichiometry of the TatA component of the twin-arginine protein transport system observed by in vivo single-molecule imaging." Proceedings of the National Academy of Sciences 105(40): 15376-15381.

Lee, P.A., Buchanan, G., Stanley, N.R., Berks, B.C. and Palmer, T. (2002). "Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation." Journal of Bacteriology 184(21): 5871-5879.

Lee, P.A., Orriss, G.L., Buchanan, G., Greene, N.P., Bond, P.J., Punginelli, C., Jack, R.L., Sansom, M.S.P., Berks, B.C. and Palmer, T. (2006). "Cysteine-scanning mutagenesis and disulfide mapping studies of the conserved domain of the twin-arginine translocase TatB component." Journal of Biological Chemistry 281(45): 34072-34085.

Lequette, Y., Odberg-Ferragut, C., Bohin, J.P. and Lacroix, J.-M. (2004). "Identification of mdoD, an mdoG paralog which encodes a twin-arginine-dependent periplasmic protein that controls osmoregulated periplasmic glucan backbone structures." Journal of Bacteriology 186(12): 3695-3702.

Lindenstrauß, U., Matos, C.F.R.O., Graubner, W., Robinson, C. and Brüser, T. (2010). "Malfolded recombinant Tat substrates are Tat-independently degraded in Escherichia coli." FEBS Letters 584(16): 3644-3648.

Liu, G., Topping, T.B. and Randall, L.L. (1989). "Physiological role during export for the retardation of folding by the leader peptide of maltose-binding protein." Proceedings of the National Academy of Sciences 86(23): 9213-9217.

Luirink, J. and Sinning, I. (2004). "SRP-mediated protein targeting: structure and function revisited." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1694(1-3): 17-35.

Lüke, I., Handford, J.I., Palmer, T. and Sargent, F. (2009). "Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB " Archives of Microbiology 191: 919–925.

Matos, C.F.R.O., Di Cola, A. and Robinson, C. (2009). "TatD is a central component of a Tat translocon-initiated quality control system for exported FeS proteins in *Escherichia coli* " EMBO reports 10(5): 474–479.

Maurer, C., Panahandeh, S., Jungkamp, A.C., Moser, M. and Müller, M. (2010). "TatB functions as an oligomeric binding site for folded Tat precursor proteins." Molecular Biology of the Cell 21(23): 4151-4161.

McDevitt, C.A., Buchanan, G., Sargent, F., Palmer, T. and Berks, B.C. (2006). "Subunit composition and in vivo substrate-binding characteristics of *Escherichia coli* Tat protein complexes expressed at native levels." FEBS Journal 273(24): 5656-5668.

Merkx, M. and Averill, B.A. (1998). "Ga3+ as a functional substitute for Fe3+: preparation and characterization of the Ga3+Fe2+ and Ga3+Zn2+ forms of bovine spleen purple acid phosphatase." Biochemistry 37(23): 8490-8497.

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(3): 289-298.

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., Ban, N. and Frank, J. (2005). "Structure of the *E. coli* protein-conducting channel bound to a translating ribosome." Nature 438(7066): 318-324.

Model, P., Jovanovic, G. and Dworkin, J. (1997). "The Escherichia coli phage-shock-protein (psp) operon." Molecular microbiology 24(2): 255-261.

Molik, S., Karnauchov, I., Weidlich, C., Herrmann, R.G. and Klösgen, R.B. (2001). "The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts. Missing link in the evolution of protein transport pathways in chloroplasts?" Journal of Biological Chemistry 276(46): 42761-42766.

Monds, R.D., Newell, P.D., Schwartzman, J.A. and O'Toole, G.A. (2006). "Conservation of the Pho regulon in Pseudomonas fluorescens Pf0-1." Applied and environmental microbiology 72(3): 1910-24.

Mori, H. and Cline, K. (1998). "A signal peptide that directs non-Sec transport in bacteria also directs efficient and exclusive transport on the thylakoid Delta pH pathway." Journal of Biological Chemistry 273(19): 11405-11408.

Mori, H. and Cline, K. (2002). "A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid {Delta}pH/Tat translocase." Journal of Cell Biology 157(2): 205-210.

Müller, P., Kopke, S. and Sheldrick, G.M. (2003). "Is the bond-valence method able to identify metal atoms in protein structures?" Acta Crystallographica Section D 59(1): 32-37.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallographica Section D 53(3): 240-255.

Nesmeyanova, M.A., Motlokh, O.B., Kolot, M.N. and Kulaev, I.S. (1981). "Multiple forms of alkaline phosphatase from *Escherichia coli* cells with repressed and derepressed biosynthesis of the enzyme." Journal of Bacteriology 146(2): 453-459.

Oates, J., Barrett, C.M.L., Barnett, J.P., Byrne, K.G., Bolhuis, A. and Robinson, C. (2005). "The *Escherichia coli* twin-arginine translocation apparatus incorporates a distinct form of TatABC complex, spectrum of modular TatA complexes and minor TatAB complex." Journal of Molecular Biology 346(1): 295-305.

Oates, J., Mathers, J., Mangels, D., Kühlbrandt, W., Robinson, C. and Model, K. (2003). "Consensus structural features of purified bacterial TatABC complexes." Journal of Molecular Biology 330(2): 277-286.

Oresnik, I.J., Ladner, C.L. and Turner, R.J. (2001). "Identification of a twin-arginine leader-binding protein." Molecular microbiology 40(2): 323-331.

Osborne, A.R., Clemons, W.M. and Rapoport, T.A. (2004). "A large conformational change of the translocation ATPase SecA." Proceedings of the National Academy of Sciences 101(30): 10937-10942.

Osborne, A.R. and Rapoport, T.A. (2007). "Protein translocation is mediated by oligomers of the SecY complex with one SecY copy forming the channel." Cell 129(1): 97-110.

Palmer, T., Sargent, F. and Berks, B.C. (2010). The Tat protein export pathway. EcoSal - Escherichia coli and Salmonella: cellular and molecular biology. http://www.ecosal.org. A. Böck, R.C., J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery Washington, DC, ASM Press.

Panahandeh, S., Maurer, C., Moser, M., DeLisa, M.P. and Müller, M. (2008). "Following the path of a twin-arginine precursor along the TatABC translocase of *Escherichia coli*." J. Biol. Chem. 283(48): 33267-33275.

Patel, L., Schuldiner, S. and Kaback, H.R. (1975). "Reversible effects of chaotropic agents on the proton permeability of *Escherichia coli* membrane vesicles." Proceedings of the National Academy of Sciences 72(9): 3387-3391.

Pérez-Rodríguez, R., Fisher, A.C., Perlmutter, J.D., Hicks, M.G., Chanal, A., Santini, C.L., Wu, L.F., Palmer, T. and DeLisa, M.P. (2007). "An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial Twin-arginine translocation pathway." Journal of Molecular Biology 367(3): 715-730.

Pommier, J., Mejean, V., Giordano, G. and Iobbi-Nivol, C. (1998). "TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in *Escherichia coli*." Journal of Biological Chemistry 273(26): 16615-16620.

Pop, O.I., Westermann, M., Volkmer-Engert, R., Schulz, D., Lemke, C., Schreiber, S., Gerlach, R., Wetzker, R. and Müller, J.P. (2003). "Sequence-specific binding of prePhoD to soluble TatAd Indicates protein-mediated targeting of the Tat export in *Bacillus subtilis*." Journal of Biological Chemistry 278(40): 38428-38436.

Porcelli, I., de Leeuw, E., Wallis, R., van den Brink-van der Laan, E., de Kruijff, B., Wallace, B.A., Palmer, T. and Berks, B.C. (2002). "Characterization and membrane assembly of the TatA component of the *Escherichia coli* twin-arginine protein transport system." Biochemistry 41(46): 13690-13697.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A. and Noble, M. (2004). "Developments in the CCP4 molecular-graphics project." Acta Crystallographica Section D 60(12 Part 1): 2288-2294.

Pradel, N., Ye, C., Livrelli, V., Xu, J., Joly, B. and Wu, L.-F. (2003). "Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7." Infection and Immunity 71(9): 4908-4916.

Pugsley, A.P. (1993). "The complete general secretory pathway in gram-negative bacteria." Microbiology and Molecular Biology Reviews 57(1): 50-108.

Punginelli, C., Maldonado, B., Grahl, S., Jack, R., Alami, M., Schroder, J., Berks, B.C. and Palmer, T. (2007). "Cysteine scanning mutagenesis and topological mapping of the *Escherichia coli* twin-arginine translocase TatC component." Journal of Bacteriology 189(15): 5482-5494.

Richter, S. and Brüser, T. (2005). "Targeting of unfolded PhoA to the Tat translocon of *Escherichia coli*." Journal of Biological Chemistry 280(52): 42723-42730.

Richter, S., Lindenstrauβ U., Lucke, C., Bayliss, R. and Brüser, T. (2007). "Functional Tat transport of unstructured, small, hydrophilic proteins." Journal of Biological Chemistry 282(46): 33257-33264.

Robinson, C. and Bolhuis, A. (2004). "Tat-dependent protein targeting in prokaryotes and chloroplasts." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1694(1-3): 135-147.

Robson, A. and Collinson, I. (2006). "The structure of the Sec complex and the problem of protein translocation." EMBO reports 7(11): 1099-1103.

Rodrigue, A., Chanal, A., Beck, K., Müller, M. and Wu, L.F. (1999). "Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial Tat pathway." Journal of Biological Chemistry 274(19): 13223-13228.

Rose, R.W., Brüser, T., Kissinger, J.C. and Pohlschröder, M. (2002). "Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway." Molecular microbiology 45(4): 943-950.

Rottenberg, H. (1989). Proton electrochemical potential gradient in vesicles, organelles, and prokaryotic cells. Methods in Enzymology, Academic Press. Volume 172: 63-84.

Roy, N.K., Ghosh, R.K. and Das, J. (1982). "Monomeric alkaline phosphatase of Vibrio cholerae." Journal of Bacteriology 150(3): 1033-1039.

San Miguel, M., Marrington, R., Rodger, P.M., Rodger, A. and Robinson, C. (2003). "An *Escherichia coli* twin-arginine signal peptide switches between helical and unstructured conformations depending on the hydrophobicity of the environment." European Journal of Biochemistry 270(16): 3345-3352.

Santini, C.L., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C. and Wu, L.-F. (2001). "Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock." Journal of Biological Chemistry 276(11): 8159-8164.

Santini, C.L., Ize, B., Chanal, A., Müller, M., Giordano, G. and Wu, L.F. (1998). "A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*." The EMBO journal 17: 101-112.

Sargent, F. (2007). "The twin-arginine transport system: moving folded proteins across membranes." Biochemical Society Transactions 035(5): 835-847.

Sargent, F., Berks, B.C. and Palmer, T. (2002). " Assembly of membrane-bound respiratory complexes by the Tat protein-transport system." Archives of Microbiology 178(2): 77-84.

Sargent, F., Berks, B.C. and Palmer, T. (2006). "Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins." FEMS Microbiology Letters 254(2): 198-207.

Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C. and Palmer, T. (1998). "Overlapping functions of components of a bacterial Sec-independent protein export pathway." The EMBO journal 17(13): 3640-3650.

Sargent, F., Gohlke, U., de Leeuw, E., Stanley, N.R., Palmer, T., Saibil, H.R. and Berks, B.C. (2001). "Purified components of the *Escherichia coli* Tat protein transport system form a double-layered ring structure." European Journal of Biochemistry 268(12): 3361-3367.

Sargent, F., Stanley, N.R., Berks, B.C. and Palmer, T. (1999). "Sec-independent protein translocation in *Escherichia coli*: a distinct and pivotal role for the TatB protein." Journal of Biological Chemistry 274(51): 36073-36082.

Schaerlaekens, K., Schierova, M., Lammertyn, E., Geukens, N., Anne, J. and Van Mellaert, L. (2001). "Twin-arginine translocation pathway in *Streptomyces lividans*." Journal of Bacteriology 183(23): 6727-6732.

Scharff, E.I., Koepke, J., Fritzsch, G., Lücke, C. and Rüterjans, H. (2001). "Crystal structure of diisopropylfluorophosphatase from *Loligo vulgaris*." Structure 9(6): 493-502.

Schlesier, R. and Klösgen, R.B. (2010). "Twin arginine translocation (Tat)-dependent protein transport: the passenger protein participates in the initial membrane binding step." Biological Chemistry 391(12): 1411-1417.

Scotti, P.A., Urbanus, M.L., Brunner, J., de Gier, J.W.L., von Heijne, G., van der Does, C., Driessen, A.J.M., Oudega, B. and Luirink, J. (2000). "YidC, the Escherichia coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase." The EMBO journal 19(4): 542-549.

Sebastian, M. and Ammerman, J.W. (2009). "The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA." ISME J 3(5): 563-572.

Seibig, S. and van Eldik, R. (1997). "Kinetics of [FeII(edta)] oxidation by molecular oxygen revisited. New evidence for a multistep mechanism." Inorganic Chemistry 36(18): 4115-4120.

Settles, A.M., Yonetani, A., Baron, A., Bush, D.R., Cline, K. and Martienssen, R. (1997). "Sec-independent protein translocation by the maize Hcf106 protein." Science 278(5342): 1467-1470.

Shanmugham, A., Wong, H.W.F.S., Bollen, Y.J.M. and Lill, H. (2006). "Membrane binding of twin arginine preproteins as an early step in translocation." Biochemistry 45(7): 2243-2249.

Silby, M.W., Cerdeno-Tarraga, A.M., Vernikos, G.S., Giddens, S.R., Jackson, R.W., Preston, G.M., Zhang, X.X., Moon, C.D., Gehrig, S.M., Godfrey, S.A., Knight, C.G., Malone, J.G., Robinson, Z., Spiers, A.J., Harris, S., Challis, G.L., Yaxley, A.M., Harris, D., Seeger, K., Murphy, L., Rutter, S., Squares, R., Quail, M.A., Saunders, E., Mavromatis, K., Brettin, T.S., Bentley, S.D., Hothersall, J., Stephens, E., Thomas, C.M., Parkhill, J., Levy, S.B., Rainey, P.B. and Thomson, N.R. (2009). "Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens." Genome biology 10(5): R51.

Sone, M., Kishigami, S., Yoshihisa, T. and Ito, K. (1997). "Roles of disulfide bonds in bacterial alkaline phosphatase." Journal of Biological Chemistry 272(10): 6174-6178.

Stanley, N.R., Palmer, T. and Berks, B.C. (2000). "The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*." Journal of Biological Chemistry 275(16): 11591-11596.

Strauch, E.M. and Georgiou, G. (2007). "*Escherichia coli* tatC mutations that suppress defective twin-arginine transporter signal peptides." Journal of Molecular Biology 374(2): 283-291.

Studier, F.W. (1991). "Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system." Journal of Molecular Biology 219(1): 37-44.

Studier, F.W. and Moffatt, B.A. (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." Journal of Molecular Biology 189(1): 113-130.

Sturm, A., Schierhorn, A., Lindenstrauβ, U., Lilie, H. and Brüser, T. (2006). "YcdB from *Escherichia coli* reveals a novel class of Tat-dependently translocated hemoproteins." Journal of Biological Chemistry 281(20): 13972-13978.

Tarry, M., Arends, S.J.R., Roversi, P., Piette, E., Sargent, F., Berks, B.C., Weiss, D.S. and Lea, S.M. (2009a). "The *Escherichia coli* cell division protein and model Tat substrate SufI (FtsP) localizes to the septal ring and has a multicopper oxidase-like structure." Journal of Molecular Biology 386(2): 504-519.

Tarry, M.J., Schafer, E., Chen, S., Buchanan, G., Greene, N.P., Lea, S.M., Palmer, T., Saibil, H.R. and Berks, B.C. (2009b). "Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system." Proceedings of the National Academy of Sciences 106(32): 13284-13289.

Tartoff, K.D. and Hobbs, C.A. (1987). " Improved media for growing plasmid and cosmid clones." Bethseda Research Laboratories Focus 9(12).

Thomas, J.R. and Bolhuis, A. (2006). "The tatC gene cluster is essential for viability in halophilic archaea." FEMS Microbiology Letters 256(1): 44-49.

Thompson, J.R., Heron, A.J., Santoso, Y. and Wallace, M.I. (2007). "Enhanced stability and fluidity in droplet on hydrogel bilayers for measuring membrane protein diffusion." Nano Letters 7(12): 3875-3878.

Torriani, A. (1990). "From cell membrane to nucleotides: The phosphate regulon in *Escherichia coli*." BioEssays 12(8): 371-376.

Tullman-Ercek, D., DeLisa, M.P., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T. and Georgiou, G. (2007). "Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides." Journal of Biological Chemistry 282(11): 8309-8316.

Valent, Q.A., De Gier, J.W.L., von Heijne, G., Kendall, D.A., Ten Hagen-Jongman, C.M., Oudega, B. and Luirink, J. (1997). "Nascent membrane and presecretory proteins synthesized in *Escherichia coli* associate with signal recognition particle and trigger factor." Molecular microbiology 25(1): 53-64.

van den Berg, B., Clemons, W.M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C. and Rapoport, T.A. (2004). "X-ray structure of a protein-conducting channel." Nature 427(6969): 36-44.

van Mourik, A., Bleumink-Pluym, N.M.C., van Dijk, L., van Putten, J.P.M. and Wosten, M.M.S.M. (2008). "Functional analysis of a Campylobacter jejuni alkaline phosphatase secreted via the Tat export machinery." Microbiology 154(2): 584-592.

Von Heijne, G. (1983). "Patterns of amino acids near signal-sequence cleavage sites." European Journal of Biochemistry 133(1): 17-21.

von Heijne, G. (1985). "Signal sequences : The limits of variation." Journal of Molecular Biology 184(1): 99-105.

von Heijne, G. (1990). "The signal peptide." Journal of Membrane Biology 115(3): 195-201.

Walther, T.H., Grage, S.L., Roth, N. and Ulrich, A.S. (2010). "Membrane alignment of the pore-forming component TatAd of the twin-arginine translocase from *Bacillus subtilis* resolved by solid-state NMR spectroscopy." Journal of the American Chemical Society 132(45): 15945-15956.

Wang, J., Stieglitz, K.A. and Kantrowitz, E.R. (2005). "Metal specificity is correlated with two crucial active site residues in *Escherichia coli* alkaline phosphatase PhoA." Biochemistry 44(23): 8378-8386.

Wanner, B.L. (1993). "Gene regulation by phosphate in enteric bacteria." Journal of cellular biochemistry 51(1): 47-54.

Warren, G., Oates, J., Robinson, C. and Dixon, A.M. (2009). "Contributions of the transmembrane domain and a key acidic motif to assembly and function of the TatA complex." Journal of Molecular Biology 388(1):122-132.

Wexler, M., Bogsch, E.G., Klösgen, R.B., Palmer, T., Robinson, C. and Berks, B.C. (1998). "Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the [Delta]pH pathway." FEBS Letters 431(3): 339-342.

Wexler, M., Sargent, F., Jack, R.L., Stanley, N.R., Bogsch, E.G., Robinson, C., Berks, B.C. and Palmer, T. (2000). "TatD Is a Cytoplasmic Protein with DNase Activity. No requirement for TatD family proteins in Sec-independent protein export." Journal of Biological Chemistry 275(22): 16717-16722.

White, G.F., Schermann, S.M., Bradley, J., Roberts, A., Greene, N.P., Berks, B.C. and Thomson, A.J. (2010). "Subunit organization in the TatA complex of the twin arginine protein translocase." Journal of Biological Chemistry 285(4): 2294-2301.

Widdick, D.A., Eijlander, R.T., van Dijl, J.M., Kuipers, O.P. and Palmer, T. (2008). "A facile reporter system for the experimental identification of twin-arginine translocation (Tat) signal peptides from all kingdoms of life." Journal of Molecular Biology 375(3): 595-603.

Woodbury, R.L., Hardy, S.J.S. and Randall, L.L. (2002). "Complex behavior in solution of homodimeric SecA." Protein Science 11(4): 875-882.

Wu, J.R., Shien, J.H., Shieh, H.K., Hu, C.C., Gong, S.R., Chen, L.Y. and Chang, P.C. (2007). "Cloning of the gene and characterization of the enzymatic properties of the monomeric alkaline phosphatase (PhoX) from Pasteurella multocida strain X-73." FEMS Microbiology Letters 267(1): 113-120.

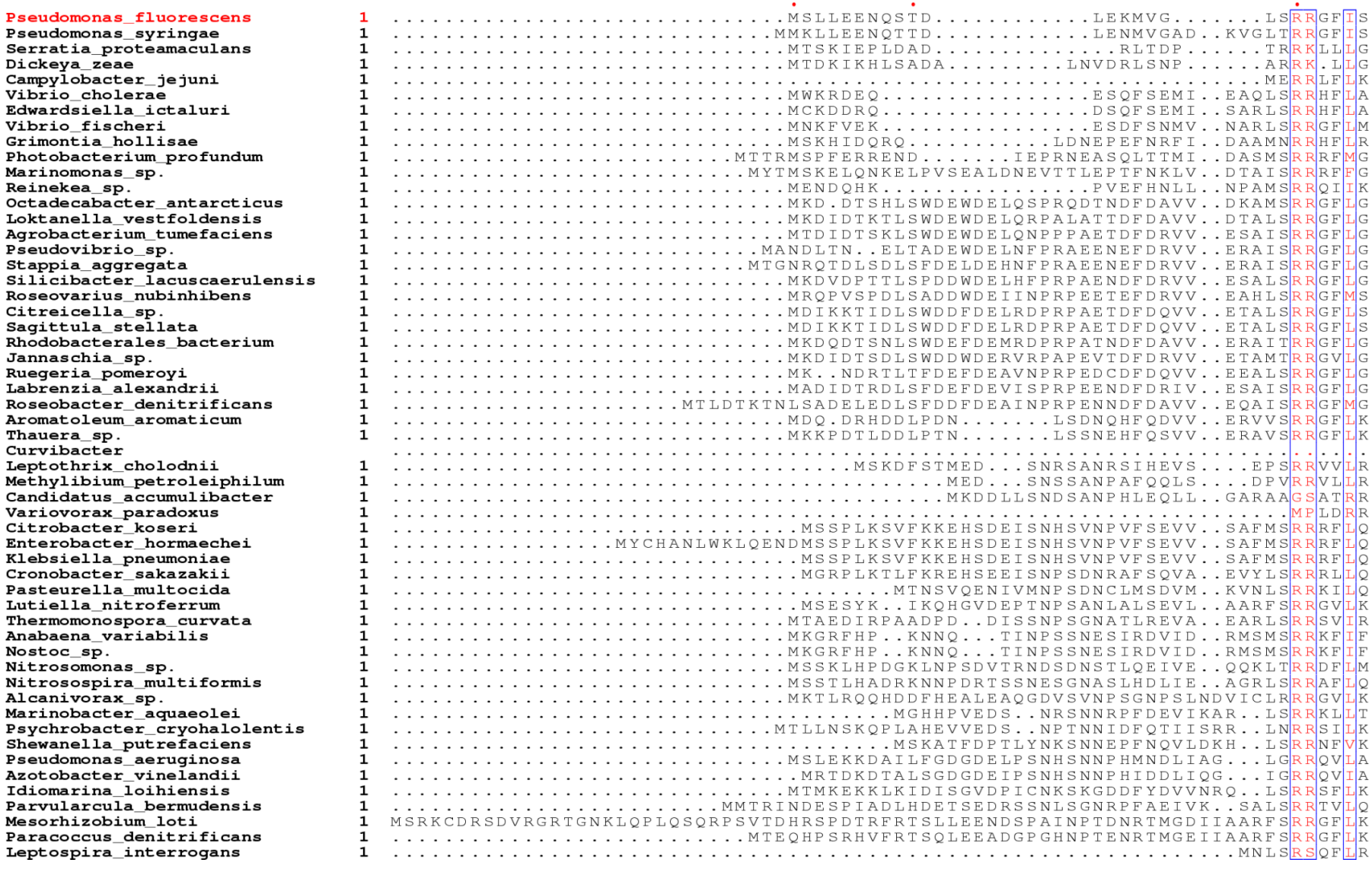
Yahr, T.L. and Wickner, W.T. (2001). "Functional reconstitution of bacterial Tat translocation *in vitro*." The EMBO journal 20(10): 2472-2479.

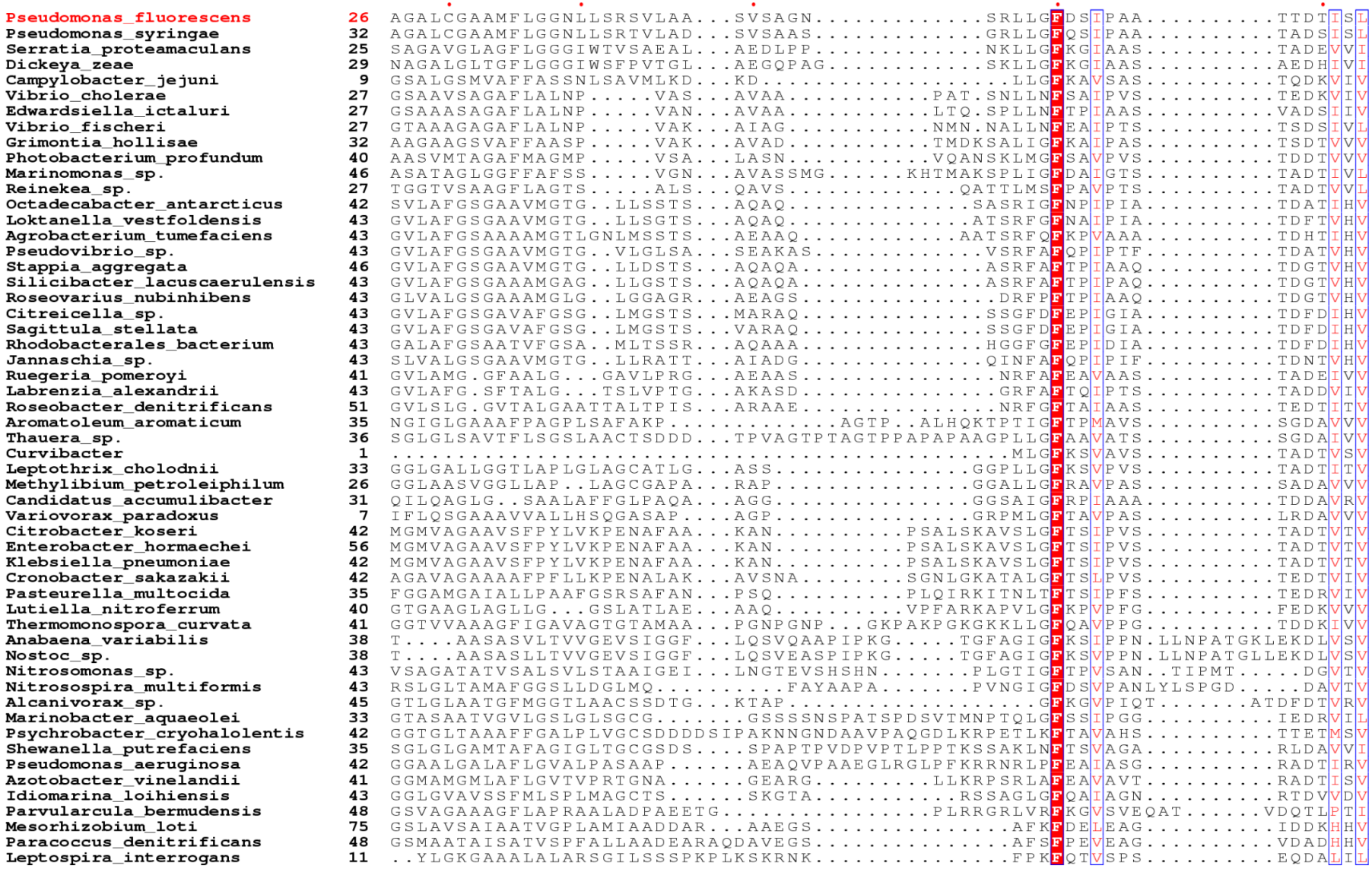
Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors." Gene 33(1): 103-119.

Zaheer, R., Morton, R., Proudfoot, M., Yakunin, A. and Finan, T.M. (2009). "Genetic and biochemical properties of an alkaline phosphatase PhoX family protein found in many bacteria." Environmental Microbiology 11(6): 1572-1587.

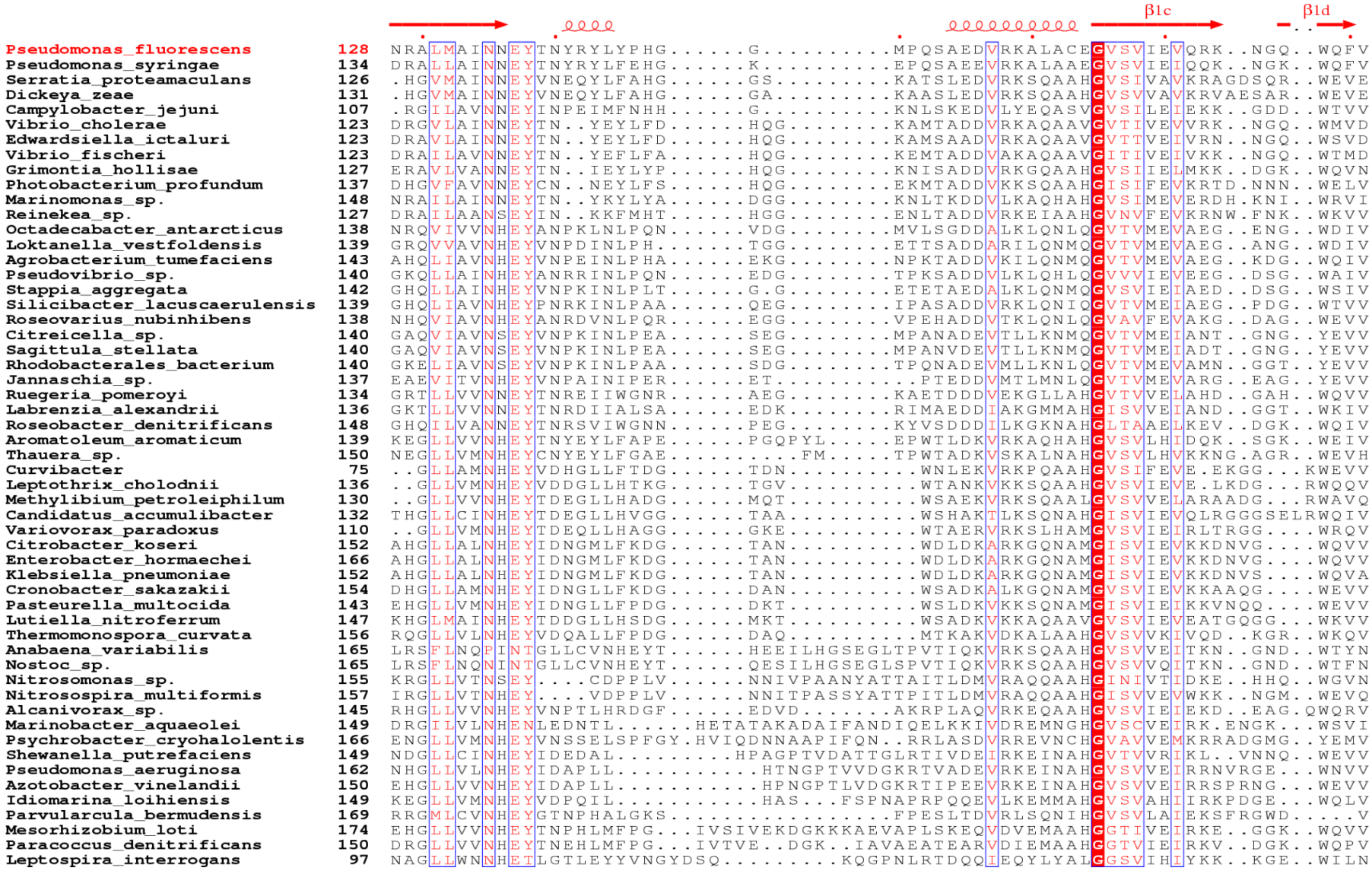
Chapter 8 :

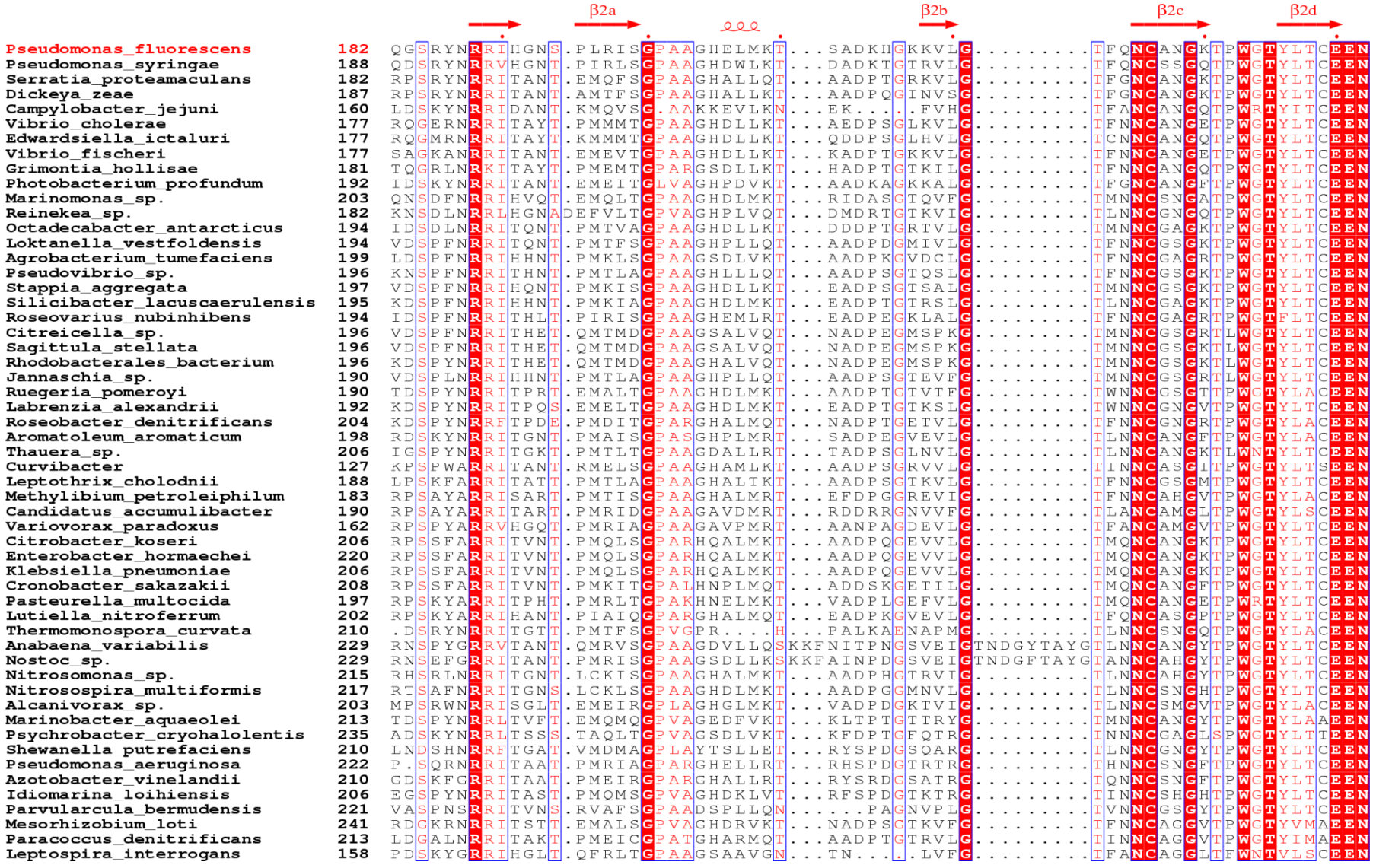
Appendix

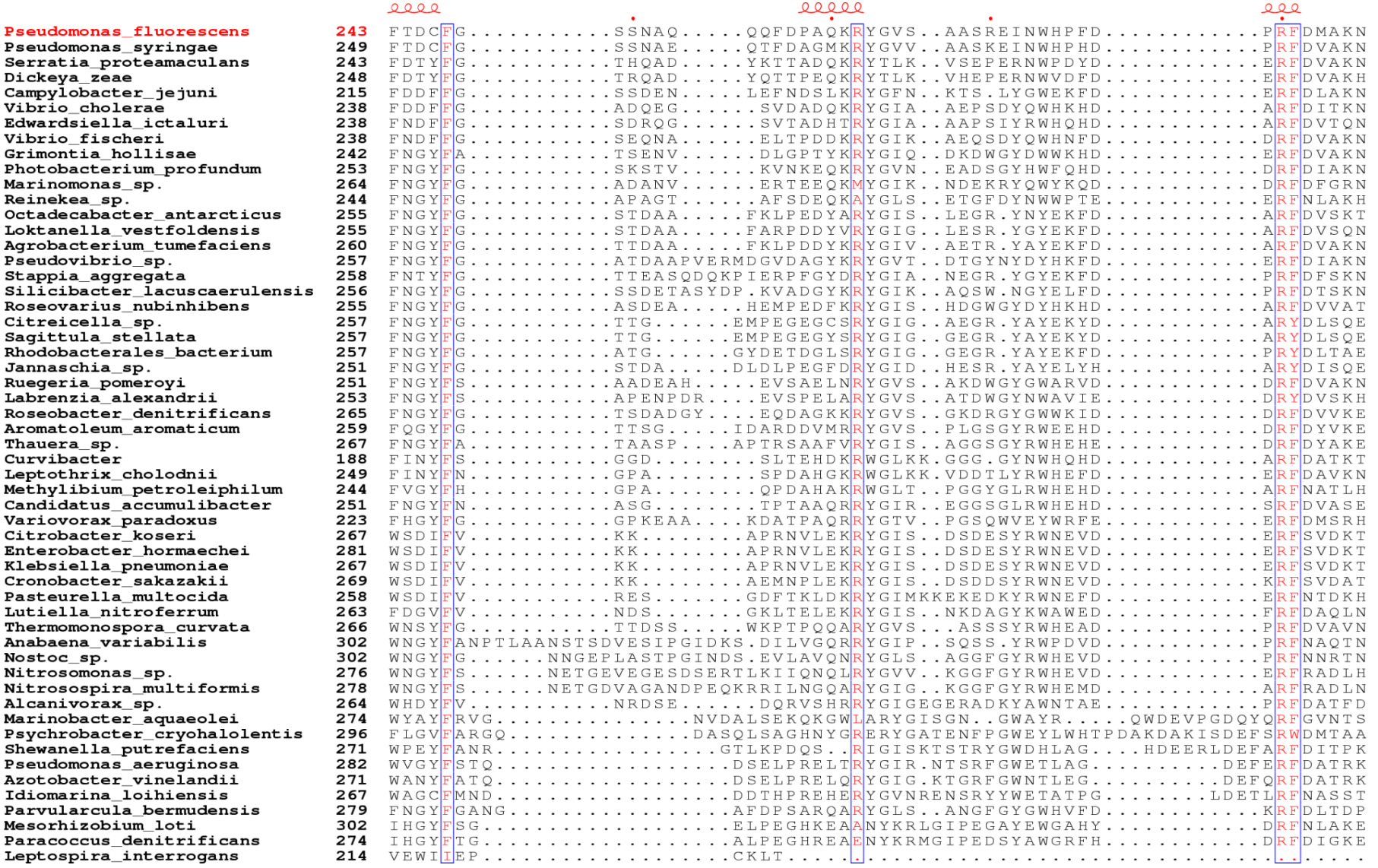

Pseudomonas_fluorescens 1
Pseudomonas_syringae 1
Serratia_proteamaculans 1
Dickeya_zeae 1
Campylobacter_jejuni 1
Vibrio_cholerae 1
Edwardsiella_ictaluri 1
Vibrio_fischeri 1
Grimontia_hollisae 1
Photobacterium_profundum 1
Marinomonas_sp. 1
Reinekea_sp. 1
Octadecabacter_antarcticus 1
Loktanella_vestfoldensis 1
Agrobacterium_tumefaciens 1
Pseudovibrio_sp. 1
Stappia_aggregata 1
Silicibacter_lacuscaerulensis 1
Roseovarius_nubinhibens 1
Citreicella_sp. 1
Sagittula_stellata 1
Rhodobacterales_bacterium 1
Jannaschia_sp. 1
Ruegeria_pomeroyi 1
Labrenzia_alexandrii 1
Roseobacter_denitrificans 1
Aromatoleum_aromaticum 1
Thauera_sp. 1
Curvibacter
Leptothrix_cholodnii 1
Methylibium_petroleiphilum 1
Candidatus_accumulibacter 1
Variovorax_paradoxus 1
Citrobacter_koseri 1
Enterobacter_hormaechei 1
Klebsiella_pneumoniae 1
Cronobacter_sakazakii 1
Pasteurella_multocida 1
Lutiella_nitroferrum 1
Thermomonospora_curvata 1
Anabaena_variabilis 1
Nostoc_sp. 1
Nitrosomonas_sp. 1
Nitrosospira_multiformis 1
Alcanivorax_sp. 1
Marinobacter_aquaeolei 1
Psychrobacter_cryohalolentis 1
Shewanella_putrefaciens 1
Pseudomonas_aeruginosa 1
Azotobacter_vinelandii 1
Idiomarina_loihiensis 1
Parvularcula_bermudensis 1
Mesorhizobium_loti 1
Paracoccus_denitrificans 1
Leptospira_interrogans 1

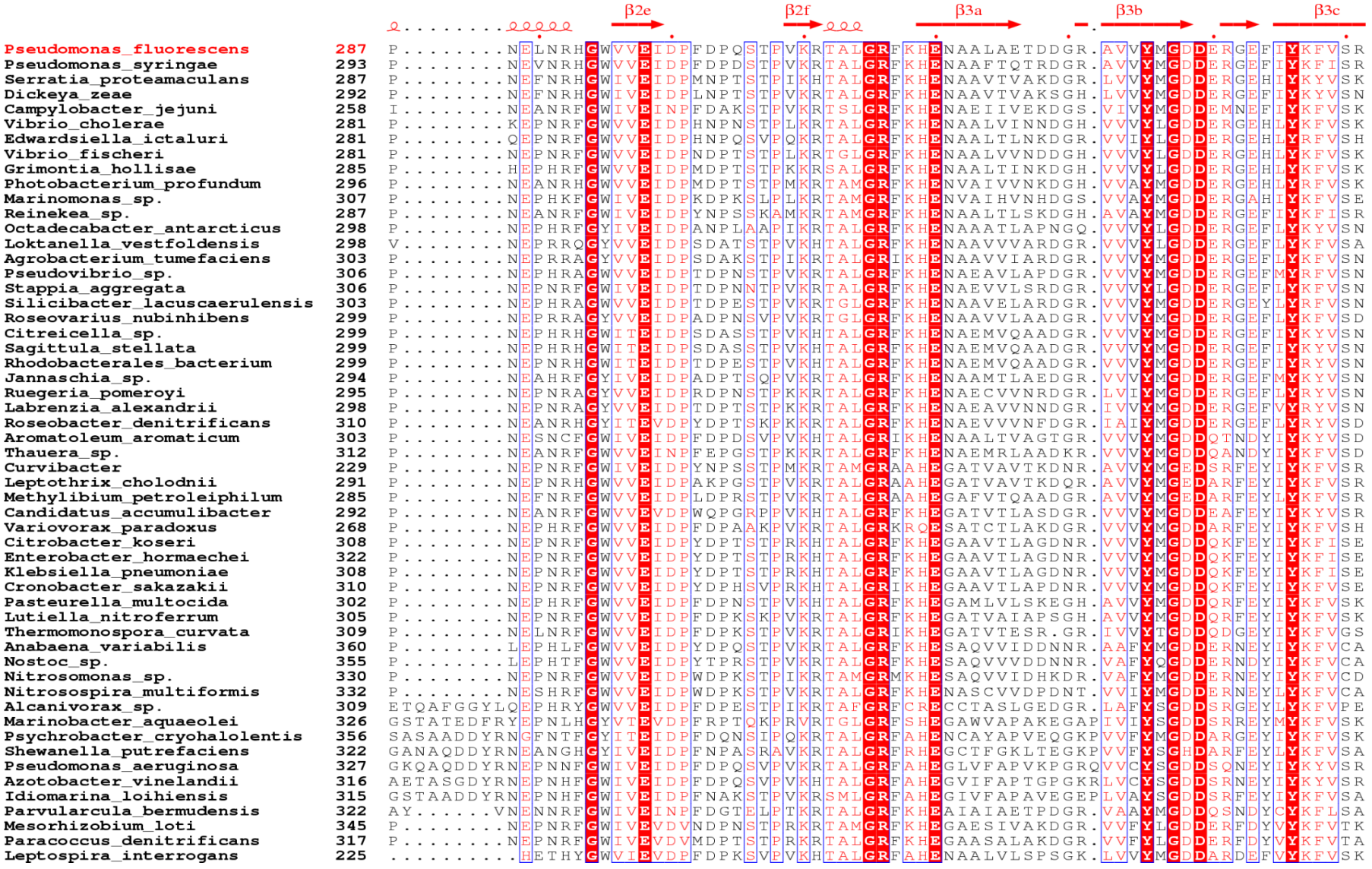

Pseudomonas_fluorescens 26 AGALCGAAMFLGGNLLSRSVLAA...SVSAGN.....SRLLGFDSIPAA.....TTDTISL
Pseudomonas_syringae 32 AGALCGAAMFLGGNLLSRTVLAD...SVSAAS.....GRLLGFQSIPAA.....TADSISL
Serratia_proteamaculans 25 SAGAVGLAGFLGGGIWTVSAEAL...AEDLPP.....NKLLGFKGIAAS.....TADEVVI
Dickeya_zeae 29 NAGALGLTGFLGGGIWSFPVTGL...AEGQPAG.....SKLLGFKGIAAS.....AEDHIVI
Campylobacter_jejuni 9 GSALGSMVAFFASSNLSAVMLKD...KD.....LLGFKAVSAS.....TQDKVIV
Vibrio_cholerae 27 GSAAVSAGAFLALNP.....VAS...AVAA.....PAT.SNLLNFSAIPVS.....TEDKVIV
Edwardsiella_ictaluri 27 GSAAASAGAFLALNP.....VAN...AVAA.....LTQ.SPLLNFTPIAAS.....VADSIIV
Vibrio_fischeri 27 GTAAAGAGAFLALNP.....VAK...AIAG.....NMN.NALLNFEAIPTS.....TSDSIVL
Grimontia_hollisae 32 AAGAAGSVAFFAASP.....VAK...AVAD.....TMDKSALIGFKAIPAS.....TSDTVVV
Photobacterium_profundum 40 AASVMTAGAFMAGMP.....VSA...LASN.....VQANSKLMGFSAVPVS.....TDDTVVV
Marinomonas_sp. 46 ASATAGLGGFFAFSS.....VGN...AVASSMG.....KHTMAKSPLIGFDAIGTS.....TADTIVL
Reinekea_sp. 27 TGGTVSAAGFLAGTS.....ALS...QAVS.....QATTLMSFPAVPTS.....TADTVVL
Octadecabacter_antarcticus 42 SVLAFGSGAAVMGTG..LLSSTS...AQAQ.....SASRIGFNPIPIA.....TDATIHV
Loktanella_vestfoldensis 43 GVLAFGSGAAVMGTG..LLSGTS...AQAQ.....ATSRFGFNAIPIA.....TDFTVHV
Agrobacterium_tumefaciens 43 GVLAFGSAAAAMGTLGNLMSSTS...AEAAQ.....AATSRFQFKPVAAA.....TDHTIHV
Pseudovibrio_sp. 43 GVLAFGSGAAVMGTG..VLGLSA...SEAKAS.....VSRFAFQPIPTF.....TDATVHV
Stappia_aggregata 46 GVLAFGSGAAVMGTG..LLDSTS...AQAQA.....ASRFAFTPIAAQ.....TDGTVHV
Silicibacter_lacuscaerulensis 43 GVLAFGSGAAAMGAG..LLGSTS...AQAQA.....ASRFAFTPIPAQ.....TDGTVHV
Roseovarius_nubinhibens 43 GLVALGSGAAAMGLG..LGGAGR...AEAGS.....DRFPFTPIAAQ.....TDGTVHV
Citreicella_sp. 43 GVLAFGSGAVAFGSG..LMGSTS...MARAQ.....SSGFDFEPIGIA.....TDFDIHV
Sagittula_stellata 43 GVLAFGSGAVAFGSG..LMGSTS...VARAQ.....SSGFDFEPIGIA.....TDFDIHV
Rhodobacterales_bacterium 43 GALAFGSAATVFGSA..MLTSSR...AQAAA.....HGGFGFEPIDIA.....TDFDIHV
Jannaschia_sp. 43 SLVALGSGAAVMGTG..LLRATT...AIADG.....QINFAFQPIPIF.....TDNTVHV
Ruegeria_pomeroyi 41 GVLAMG.GFAALG...GAVLPRG...AEAAS.....NRFAFEAVAAS.....TADEIVV
Labrenzia_alexandrii 43 GVLAFG.SFTALG...TSLVPTG...AKASD.....GRFAFTQIPTS.....TADDVIV
Roseobacter_denitrificans 51 GVLSLG.GVTALGAATTALTPIS...ARAAE.....NRFGFTAIAAS.....TEDTITV
Aromatoleum_aromaticum 35 NGIGLGAAAFPAGPLSAFAKP.....AGTP..ALHQKTPTIGFTPMAVS.....SGDAVVV
Thauera_sp. 36 SGLGLSAVTFLSGSLAACTSDDD...TPVAGTPTAGTPPAPAPAAGPLLGFAAVATS.....SGDAIVV
Curvibacter 1 .....MLGFKSVAVS.....TADTVSV
Leptothrix_cholodnii 33 GGLGALLGGTLAPLGLAGCATLG...ASS.....GGPLLGFKSVPVS.....TADTITV
Methylibium_petroleiphilum 26 GGLAASVGGLLAP..LAGCGAPA...RAP.....GGALLGFRAVPAS.....SADAVVV
Candidatus_accumulibacter 31 QILQAGLG..SAALAFFGLPAQA...AGG.....GGSAIGFRPIAAA.....TDDAVRV
Variovorax_paradoxus 7 IFLQSGAAAVVALLHSQGASAP....AGP.....GRPMLGFTAVPAS.....LRDAVVV
Citrobacter_koseri 42 MGMVAGAAVSFPYLVKPENAFAA...KAN.....PSALSKAVSLGFTSIPVS.....TADTVTV
Enterobacter_hormaechei 56 MGMVAGAAVSFPYLVKPENAFAA...KAN.....PSALSKAVSLGFTSIPVS.....TADTVTV
Klebsiella_pneumoniae 42 MGMVAGAAVSFPYLVKPENAFAA...KAN.....PSALSKAVSLGFTSIPVS.....TADTVTV
Cronobacter_sakazakii 42 AGAVAGAAAAFPFLLKPENALAK...AVSNA.....SGNLGKATALGFTSLPVS.....TEDTVIV
Pasteurella_multocida 35 FGGAMGAIALLPAAFGSRSAFAN...PSQ.....PLQIRKITNLTFTSIPFS.....TEDRVIV
Lutiella_nitroferrum 40 GTGAAGLAGLLG...GSLATLAE...AAQ.....VPFARKAPVLGFKPVPFG.....FEDKVVV
Thermomonospora_curvata 41 GGTVVAAAGFIGAVAGTGTAMAA...PGNPGNP...GKPAKPGKGKKLLGFQAVPPG.....TDDKIVV
Anabaena_variabilis 38 T....AASASVLTVVGEVSIGGF...LQSVQAAPIPKG.....TGFAGIGFKSIPPN.LLNPATGKLEKDLVSV
Nostoc_sp. 38 T....AASASLLTVVGEVSIGGF...LQSVEASPIPKG.....TGFAGIGFKSVPPN.LLNPATGLLEKDLVSV
Nitrosomonas_sp. 43 VSAGATATVSALSVLSTAAIGEI...LNGTEVSHSHN.....PLGTIGFTPVSAN..TIPMT.....DGVTV
Nitrosospira_multiformis 43 RSLGLTAMAFGGSLLDGLMQ.....FAYAAPA.....PVNGIGFDSVPANLYLSPGD.....DAVTV
Alcanivorax_sp. 45 GTLGLAATGFMGGTLAACSSDTG...KTAP.....GFKGVPIQT.....ATDFDTVRV
Marinobacter_aquaeolei 33 GTASAATVGVLGSLGLSGCG.....GSSSSNSPATSPDSVTMNPTQLGFSSIPGG.....IEDRVIL
Psychrobacter_cryohalolentis 42 GGTGLTAAAFFGALPLVGCSDDDDSIPAKNNGNDAAVPAQGDLKRPETLKFTAVAHS.....TTETMSV
Shewanella_putrefaciens 35 SGLGLGAMTAFAGIGLTGCGSDS....SPAPTPVDPVPTLPPTKSSAKLNFTSVAGA.....RLDAVVI
Pseudomonas_aeruginosa 42 GGAALGALAFLGVALPASAAP.....AEAQVPAAEGLRGLPFKRRNRLPFEAIASG.....RADTIRV
Azotobacter_vinelandii 41 GGMAMGMLAFLGVTVPRTGNA.....GEARG.....LLKRPSRLAFEAVAVT.....RADTISV
Idiomarina_loihiensis 43 GGLGVAVSSFMLSPLMAGCTS.....SKGTA.....RSSAGLGFQAIAGN.....RTDVVDV
Parvularcula_bermudensis 48 GSVAGAAAGFLAPRAALADPAEETG.....PLRRGRLVRFKGVSVEQAT.....VDQTLPTI
Mesorhizobium_loti 75 GSLAVSAIAATVGPLAMIAADDAR...AAEGS.....AFKFDELEAG.....IDDKHHV
Paracoccus_denitrificans 48 GSMAATAISATVSPFALLAADEARAQDAVEGS.....AFSFPEVEAG.....VDADHHV
Leptospira_interrogans 11 ..YLGKGAAALALARSGILSSSPKPLKSKRNK.....FPKFQTVSPS.....EQDALIL

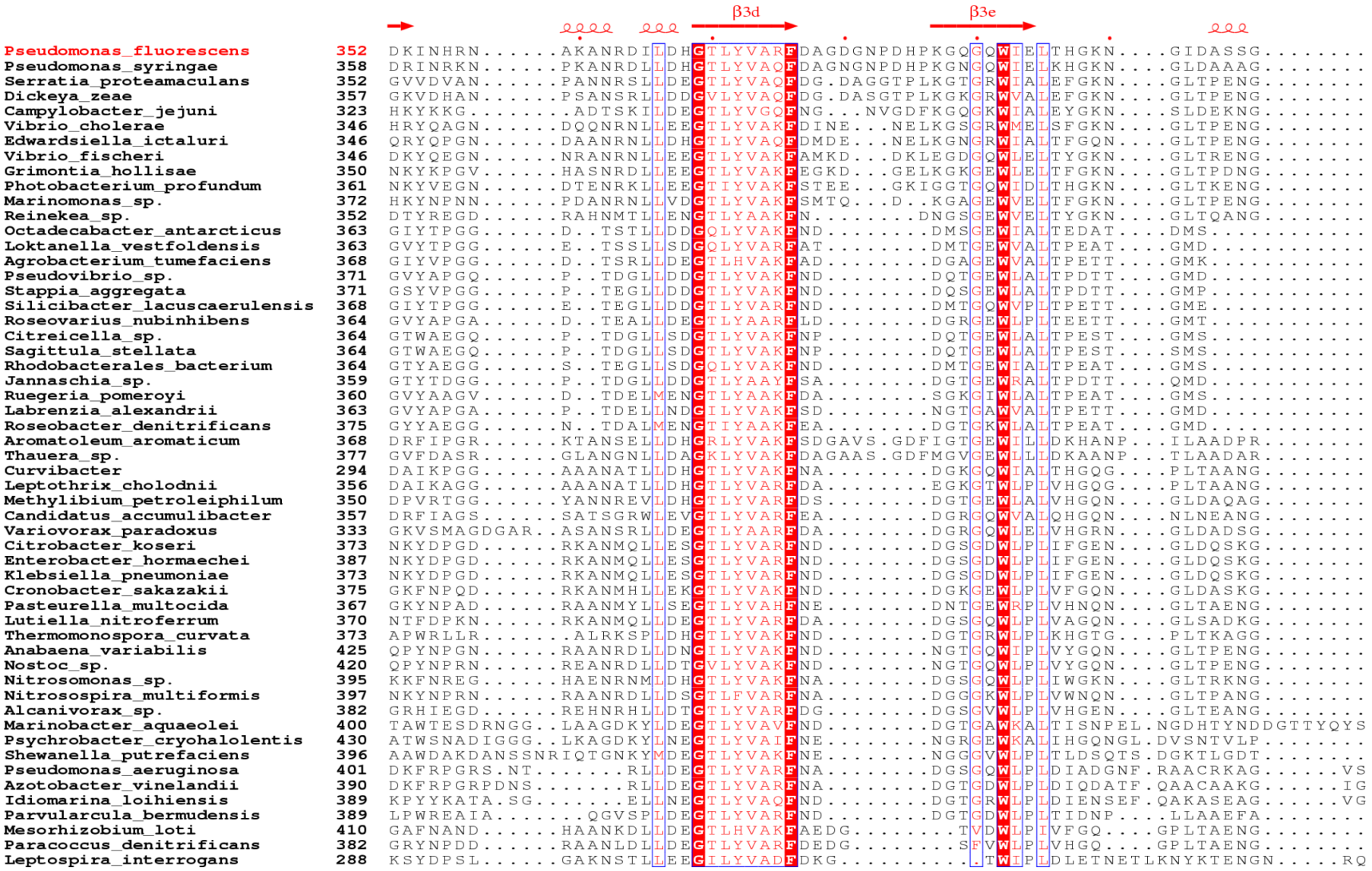
β6d
β1a
β1b
Pseudomonas_fluorescens 74 PKGYKSSVLISWGQPLHK....NGP.AFDPS...GNGTAAAQEVQFGDNNDGMSLFEFPGEK
Pseudomonas_syringae 80 PAGYTSSVLISWGQPLRA....GGP.AFDPS...GKGTAQAQEVQFGDNNDGMCLFPMPGRK
Serratia_proteamaculans 73 APGYRAEVLISWGEPLVE....GAA.AFDPQ...GNNSAADQEKQFGDNNDGMSFFPIDDN
Dickeya_zeae 78 TPGYRAEVLISWGEPLVD....GAP.AFHAE...GKNSAADQEKQFGDNNDGMSFFPIDDR
Campylobacter_jejuni 51 PQGYEAKVLISWGDPLFS....KAK.PYDENKIIDMNAVKNAHLVFGDNDDGMSFFPLSKN
Vibrio_cholerae 71 PKGYKATPLMSWGDPIFA....NAP.EFDQSG...KQDSKAQEKQFGDNTDGMSFFPISE
Edwardsiella_ictaluri 71 PPGYRATPLISWGDPIFA....DAP.EFTRPE...AQDAAAQERQFGDNTDGMSFFPISD
Vibrio_fischeri 71 PKGYSWNTLMSWGDPIFV....NAP.KFNPNG...KQDSKAQALQFGDNTDGMSVFPLSN
Grimontia_hollisae 77 PEGYVAEKLISWGDALFP....DSP.KFSQS.....NDSYAQEKQYGDNTDGMTFFPLSD
Photobacterium_profundum 85 PDDYNAKVLIRWGDGLFP....TAP.TFDPTG...NAPSSHQEVQFGDNTDGMTFFPLSD
Marinomonas_sp. 96 PEGYQWQSLVSWGDPLFP....NAP.KFDPSG...NANAASQAMQFGDNTDGMSLFPLSE
Reinekea_sp. 70 PEGYEYDVLVSWGDPILK....GAS.KFDVEN.....SAEDQLGQYGDNTDGMSLFHLTDKDGGVD.....E
Octadecabacter_antarcticus 87 PEGYSWHPVAKWGQPLFS....DAP.DLNHQT...GGSLASSDRVFGENTDGMELFA.IG
Loktanella_vestfoldensis 88 PEGYTWQPVAKWGQPLFS....DAA.PLSQET...GVDLASSDRVFGENTDGMELFS.IG
Agrobacterium_tumefaciens 92 PEGYSWKPVAKWGQPLFS....NVP.DLDPAK...GVSVENSDKVFGENTDGMELFM.VG
Pseudovibrio_sp. 89 PEGYNWKILARWGDALVD....GAE.EFDHET...GGSAAGADKVFGENTDGMESFQ.YQ
Stappia_aggregata 91 PEGYDWKLLARWGDPYFS....GVD.AFDPET...GVSLEASDKVFGENTDGMESFV.YN
Silicibacter_lacuscaerulensis 88 PEGYDWKVLIRWGDPLFS....DAA.GYDVTD...GGPVEKSDRVFGENTDGMEAFS.YR
Roseovarius_nubinhibens 87 PEGYSWEVLARWGDPLFS....DAP.AFDNAT...GGAAAGSDRIFGENTDGMELFS.IG
Citreicella_sp. 88 PAGYQWKPLVKWGQPLFS....EAEGAYSPET...GVSVAMSDKVFGENTDGMELFE.VD
Sagittula_stellata 88 PAGYQWKPLVKWGQPLFS....EAEGAFSPET...GVSVAMSDKVFGENTDGMELFE.VD
Rhodobacterales_bacterium 88 PEGYEWKTLVRWGDALWS....EADGAYDAET...GVSPAMSDKVFGENTDGMETFF.VD
Jannaschia_sp. 88 PDGYTWSTVARWGDPLFS....GAP...TFDN...GHGVDGSDRIFGENTDGMELFR.VG
Ruegeria_pomeroyi 83 PKGFKADVLVRWADPLWS....DGV.AFDQVT...RGTAASQARAFGDNTDGMEVFH.HE
Labrenzia_alexandrii 85 PEGYKSEMLVRWGDPLWS....DVP.EFDDST...RGTADSQARAFGDNTDGMDVFS.HD
Roseobacter_denitrificans 96 PDGYNWHVATRWGDPLFS....DVP.EFDHAT...RGTAASQARAFGDNNDGMAIFTDKD
Aromatoleum_aromaticum 84 PSGYRAQVLARWGEALFA....DSP.EWLPDG...SNTGEDQARQIGDNHDGMHFFPLNGQS.....S
Thauera_sp. 97 PAGYSAQIFTPWGSPLFS....DSP.AWRADG...TNTGEEQARQVGDNHDGMSYFPIDG.....S
Curvibacter 18 PEGYTVQVIAPWGDPVG...MSGENTAFKDDA...SNTAAHQETQFGMNHDGIHYFAQEGSKS
Leptothrix_cholodnii 79 PPGYVAQVIAAWGEPVG...IEGRMPAFKFDG...SNTFAEQEAQLGMHHDGIHFYAQDGAKR
Methylibium_petroleiphilum 70 PPGYVAEVIYAWGDAIGAPGLAAGQPAFRGDG...SHSAEDQQLQAGMHHDGMHFFPAGRGDR
Candidatus_accumulibacter 75 PEGYTLQLLYAWGDPIS.....LGPAFRGDA...SNSAPEQAEQAGMHHDGMQFFPFVNDGRPSS
Variovorax_paradoxus 52 PPEYEWQLLYPWGTPTG..IAGQPMPAFAPDG...SNSAADQALQAGMHHDGMHFFTLASPDR
Citrobacter_koseri 93 PEGYIARPFYRWGDAVG...IKGNLPEFKFDA...SNTTDEQAAQAGMHHDGMAWFSLPQGKENP
Enterobacter_hormaechei 107 PEGYIARPFYRWGDAVG...IKGNLPEFKFDA...SNTTDEQAAQAGMHHDGMAWFSLPQGKENP
Klebsiella_pneumoniae 93 PEGYIARPFYRWGDAVG...IKGNLPEFKFDA...SNTTDEQAAQAGMHHDGMAWFSLPQGKENP
Cronobacter_sakazakii 95 PEGYIAKPFYKWGDATG...LAGNLPVFKTDG...SNTTEEQAAQAGMHHDGMAWFSLPQGGNSA
Pasteurella_multocida 86 PKGYSAKAFYAWGDPVG...LENNNPQFKVDA...SNSAEEQAAQAGMNHDGMAYFPFAEHGN
Lutiella_nitroferrum 88 PEGYVADVLYAWGDPVG...IKGAMPAFKPDA...SNSAAEQELQAGMHHDGMAFFPLPLGSTSS
Thermomonospora_curvata 99 PPGYTAQVLIPWGTPLSS....RGP.KWRKDA...SNSAADQAKQIGMHHDGMAFFPLGAGPLGS
Anabaena_variabilis 99 PEGYTAKVLVAWGDPIAPG.....GPTWLADA...SQDAAAQEKQFGMHADGMHYFPISYGNPVGRTVSSQARS
Nostoc_sp. 99 PEGYTAKVLIAWGDPIMPT.....APNWLPDA...SQGAAAQEMQYGMHVDGMHYFPLSPGNAVGRTVGSQTRS
Nitrosomonas_sp. 100 PTGYSARILVSWGDSLTP.....ESHWDTVS...ALDETTQSHCYGAHTDGMHYFPFPG..PFG.....N
Nitrosospira_multiformis 95 PAGYTARVLVAWGDSLTQ.....APHWNPGS...AMTETVQLHAFGAHVDGMHLFPFP...PMGSSGSPGVAN
Alcanivorax_sp. 90 PPGYTARPFFSWGDAVLA....QAPAWRTDA...GNDWREQLMQAGQNNDGMHFFPFPDNPN
Marinobacter_aquaeolei 90 PPGYKADVLYAKGDPLFS.....GLAPFKN.....DGTDVDYDSRSGDEHDGMHFFGMNASG.....QYDPEAS
Psychrobacter_cryohalolentis 106 AAGYKAEMILPLGTPLIS.....GIDDWKDNR...EQSAESFEWRMGDNHDGMWFFGKKGS.....AYDPKAS
Shewanella_putrefaciens 95 PEGYTAQVLAPWGTPLNS....KAASWKSDG...SNTADDQANSVGMHHDGMHFFPLNGA.....G
Pseudomonas_aeruginosa 100 PAGYKATPFIPWGTPISG.....SYPGFLDDA...SNSAQDQAEQIGMHHDGMHFFPIDAQFG....FGGGHIS
Azotobacter_vinelandii 90 PEGYQAKTFIPWGTPICG.....SYPAYREDG...SNTAQEQAEQTGMHHDGMHYFPIDARN.....GKGHNS
Idiomarina_loihiensis 90 PVGYRAQPFLPWGTPLVP.....GAGEFKTDA...SNSALDQARQVGSHHDGMHFFPIDVKE.....GGASS
Parvularcula_bermudensis 100 SDDYTYQAFIPWGTPIEP....GATRPYNGEPD.RRPTPEEASKQVGLGHDGMWFFPEAPEFLDGPQRYALGSN
Mesorhizobium_loti 121 APGYDADVLLRWGDPLFAD.....SPEFDPTKQSAEAQARQFGYNNDYVGYIAIDGSA
Paracoccus_denitrificans 97 AQGYDADILLRWGDKVFAD.....APDFDPKNQSEAAQERQFGYNNDFVGFIPLDGAD
Leptospira_interrogans 58 PSGYQYNTIALYGDKINP.....QGDTFGFNSDFNCFFPLEGKK

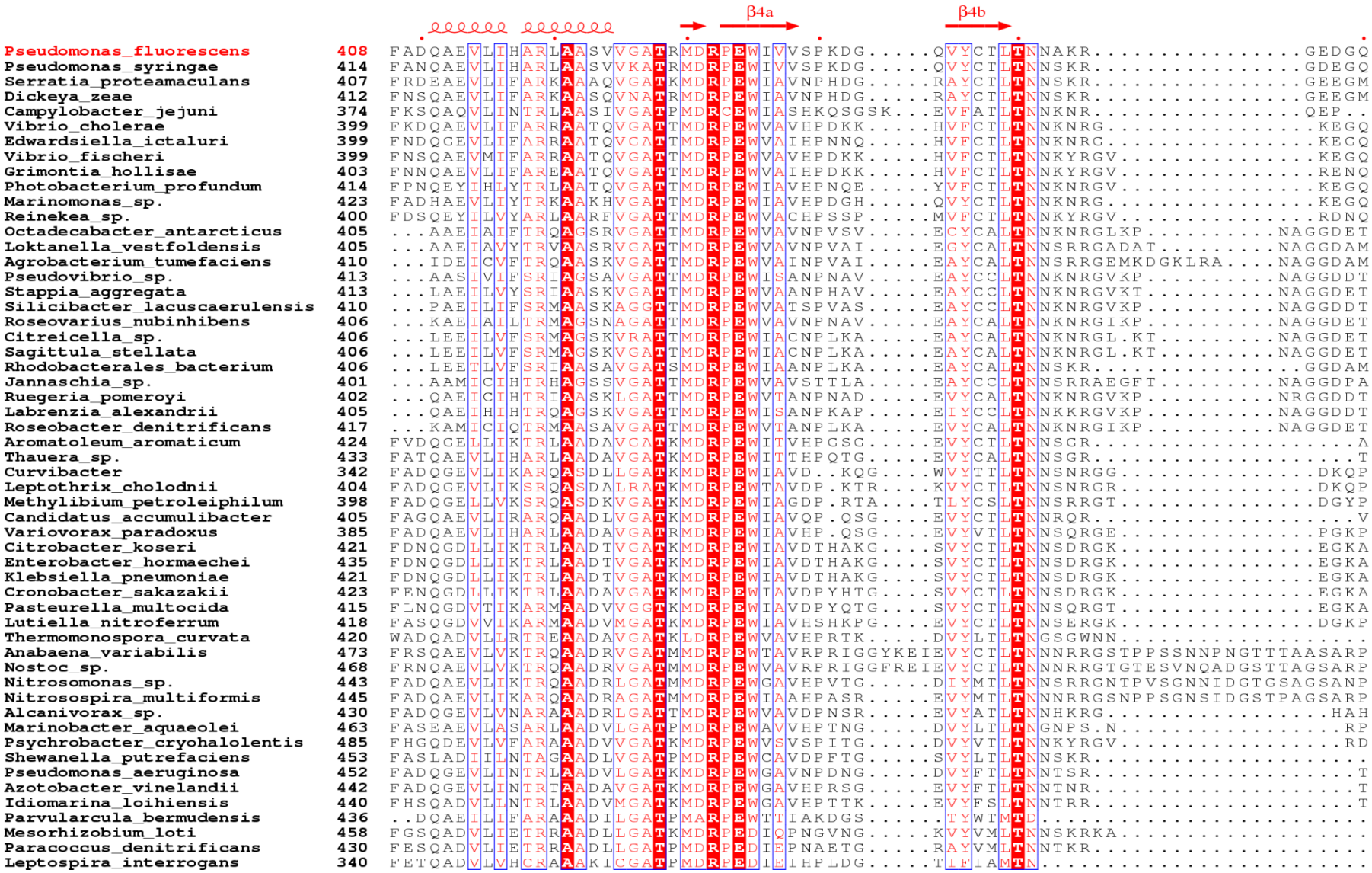

β1c
β1d
Pseudomonas_fluorescens 128 NRALMAINNEYTNYRYLYPHG.....G.....MPQSAEDVRKALACEGVSVIEVQRK..NGQ..WQFV
Pseudomonas_syringae 134 DRALLAINNEYTNYRYLFEHG.....K.....EPQSAEEVRKALAAEGVSVIEIQQK..NGK..WQFV
Serratia_proteamaculans 126 .HGVMAINNEYVNEQYLFAHG.....GS.....KATSLEDVRKSQAAHGVSIVAVKRAGDSQR..WEVE
Dickeya_zeae 131 .HGVMAINNEYVNEQYLFAHG.....GA.....KAASLEDVRKSQAAHGVSVVAVKRVAESAR..WEVE
Campylobacter_jejuni 107 .RGILAVNNEYINPEIMFNHH.....G.....KNLSKEDVLYEQASVGVSILEIEKK..GDD..WTVV
Vibrio_cholerae 123 DRGVLAINNEYTN..YEYLFD.....HQG.....KAMTADDVRKAQAAVGVTIVEVVRK..NGQ..WMVD
Edwardsiella_ictaluri 123 DRAVLAINNEYTN..YEYLFD.....HQG.....KAMSAADVRKAQAAVGVTTVEIVRN..NGQ..WSVD
Vibrio_fischeri 123 DRAILAVNNEYTN..YEFLFA.....HQG.....KEMTADDVAKAQAAVGITIVEIVKK..NGQ..WTMD
Grimontia_hollisae 127 ERAVLVANNEYTN..IEYLYP.....HQG.....KNISADDVRKGQAAHGVSIIELMKK..DGK..WQVN
Photobacterium_profundum 137 DHGVFAVNNEYCN..NEYLFS.....HQG.....EKMTADDVKKSQAAHGISIFEVKRTD.NNN..WELV
Marinomonas_sp. 148 NRAILAINNEYTN..YKYLYA.....DGG.....KNLTKDDVLKAQHAHGVSIMEVERDH.KNI..WRVI
Reinekea_sp. 127 DRAILAANSEYIN..KKFMHT.....HGG.....ENLTADDVRKEIAAHGVNVFEVKRNW.FNK..WKVV
Octadecabacter_antarcticus 138 NRQVIVVNHEYANPKLNLPQN.....VDG.....MVLSGDDALKLQNLQGVTVMEVAEG..ENG..WDIV
Loktanella_vestfoldensis 139 GRQVVAVNHEYVNPDINLPH.....TGG.....ETTSADDARILQNMQGVTVMEVAEG..ANG..WDIV
Agrobacterium_tumefaciens 143 AHQLIAVNHEYVNPEINLPHA.....EKG.....NPKTADDVKILQNMQGVTVMEVAEG..KDG..WQIV
Pseudovibrio_sp. 140 GKQLLAINHEYANRRINLPQN.....EDG.....TPKSADDVLKLQHLQGVVVIEVEEG..DSG..WAIV
Stappia_aggregata 142 GHQLLAINHEYVNPKINLPLT.....G.G.....ETETAEDALKLQNMQGVSVIEIAED..DSG..WSIV
Silicibacter_lacuscaerulensis 139 GHQLIAVNHEYPNRKINLPAA.....QEG.....IPASADDVRKLQNIQGVTVMEIAEG..PDG..WTVV
Roseovarius_nubinhibens 138 NHQVIAVNHEYANRDVNLPQR.....EGG.....VPEHADDVTKLQNLQGVAVFEVAKG..DAG..WEVV
Citreicella_sp. 140 GAQVIAVNSEYVNPKINLPEA.....SEG.....MPANADEVTLLKNMQGVTVMEIANT..GNG..YEVV
Sagittula_stellata 140 GAQVIAVNSEYVNPKINLPEA.....SEG.....MPANVDEVTLLKNMQGVTVMEIADT..GNG..YEVV
Rhodobacterales_bacterium 140 GKELIAVNSEYVNPKINLPAA.....SDG.....TPQNADEVMLLKNLQGVTVMEVAMN..GGT..YEVV
Jannaschia_sp. 137 EAEVITVNHEYVNPAINIPER.....ET.....PTEDDVMTLMNLQGVTVMEVARG..EAG..YEVV
Ruegeria_pomeroyi 134 GRTLLVVNNEYTNREIIWGNR.....AEG.....KAETDDDVEKGLLAHGVTVVELAHD..GAH..WQVV
Labrenzia_alexandrii 136 GKTLLVVNNEYTNRDIIALSA.....EDK.....RIMAEDDIAKGMMAHGISVVEIAND..GGT..WKIV
Roseobacter_denitrificans 148 GHQILVANNEYTNRSVIWGNN.....PEG.....KYVSDDDILKGKNAHGLTAAELKEV..DGK..WQIV
Aromatoleum_aromaticum 139 KEGLLVVNHEYTNYEYLFAPE.....PGQPYL.....EPWTLDKVRKAQHAHGVSVLHIDQK..SGK..WEIV
Thauera_sp. 150 NEGLLVMNHEYCNYEYLFGAE.....FM.....TPWTADKVSKALNAHGVSVLHVKKNG.AGR..WEVH
Curvibacter 75 ..GLLAMNHEYVDHGLLFTDG.....TDN.....WNLEKVRKPQAAHGVSIFEVE.EKGG...KWEVV
Leptothrix_cholodnii 136 ..GLLVMNHEYVDDGLLHTKG.....TGV.....WTANKVKKSQAAHGVSVIEVE.LKDG...RWQQV
Methylibium_petroleiphilum 130 ..GLLVVNHEYTDEGLLHADG.....MQT.....WSAEKVRKSQAALGVSVVELARAADG...RWAVQ
Candidatus_accumulibacter 132 THGLLCINHEYTDEGLLHVGG.....TAA.....WSHAKTLKSQNAHGISVIEVQLRGGGSELRWQIV
Variovorax_paradoxus 110 ..GLLVMNHEYTDEQLLHAGG.....GKE.....WTAERVRKSLHAMGVSVIEIRLTRGG....WRQV
Citrobacter_koseri 152 AHGLLALNHEYIDNGMLFKDG.....TAN.....WDLDKARKGQNAMGISVIEVKKDNVG....WQVV
Enterobacter_hormaechei 166 AHGLLALNHEYIDNGMLFKDG.....TAN.....WDLDKARKGQNAMGISVIEVKKDNVG....WQVV
Klebsiella_pneumoniae 152 AHGLLALNHEYIDNGMLFKDG.....TAN.....WDLDKARKGQNAMGISVIEVKKDNVS....WQVA
Cronobacter_sakazakii 154 DHGLLAMNHEYIDNGLLFKDG.....DAN.....WSADKALKGQNAMGVSVIEVKKAAQG....WEVV
Pasteurella_multocida 143 EHGLLVMNHEYIDNGLLFPDG.....DKT.....WSLDKVKKSQNAMGISVIEIKKVNQQ....WEVV
Lutiella_nitroferrum 147 KHGLMAINHEYTDDGLLHSDG.....MKT.....WSADKVKKAQAAVGVSVIEVEATGQGG....WKVV
Thermomonospora_curvata 156 RQGLLVLNHEYVDQALLFPDG.....DAQ.....MTKAKVDKALAAHGVSVVKIVQD..KGR..WKQV
Anabaena_variabilis 165 LRSFLNQPINTGLLCVNHEYT.....HEEILHGSEGLTPVTIQKVRKSQAAHGVSVEITKN..GND..WTYN
Nostoc_sp. 165 LRSFLNQNINTGLLCVNHEYT.....QESILHGSEGLSPVTIQKVRKSQAAHGVSVVQITKN..GND..WTFN
Nitrosomonas_sp. 155 KRGLLVTNSEY....CDPPLV.....NNIVPAANYATTAITLDMVRAQQAAHGINIVTIDKE..HHQ..WGVN
Nitrosospira_multiformis 157 IRGLLVTNHEY....VDPPLV.....NNITPASSYATTPITLDMVRAQQAAHGISVVEVWKK..NGM..WEVQ
Alcanivorax_sp. 145 RHGLLVVNHEYVNPTLHRDGF.....EDVD.....AKRPLAQVRKEQAAHGVSVIEIEKD..EAG.QWQRV
Marinobacter_aquaeolei 149 DRGILVLNHENLEDNTL.....HETATAKADAIFANDIQELKKIVDREMNGHGVSCVEIRK.ENGK...WSVI
Psychrobacter_cryohalolentis 166 ENGLLVMNHEYVNSSELSPFGY.HVIQDNNAAPIFQN..RRLASDVRREVNCHGVAVVEMKRRADGMG..YEMV
Shewanella_putrefaciens 149 NDGLLCINHEYIDEDAL.....HPAGPTVDATTGLRTIVDEIRKEINAHGVTVVRIKL..VNNQ..WEVV
Pseudomonas_aeruginosa 162 NHGLLVLNHEYIDAPLL.....HTNGPTVVDGKRTVADEVRKEINAHGVSVVEIRRNVRGE...WNVV
Azotobacter_vinelandii 150 EHGLLVVNHEYIDAPLL.....HPNGPTLVDGKRTIPEEVRKEINAHGVSVVEIRRSPRNG...WEVV
Idiomarina_loihiensis 149 KEGLLVMNHEYVDPQIL.....HAS...FSPNAPRPQQEVLKEMMAHGVSVAHIIRKPDGE...WQLV
Parvularcula_bermudensis 169 RRGMLCVNHEYGTNPHALGKS.....FPESLTDVRLSQNIHGVSVLAIEKSFRGWD.....V
Mesorhizobium_loti 174 EHGLLVVNHEYTNPHLMFPG...IVSIVEKDGKKKAEVAPLSKEQVDVEMAAHGGTIVEIRKE..GGK..WQVV
Paracoccus_denitrificans 150 DRGLLVVNHEYTNEHLMFPG...IVTVE..DGK..IAVAEATEARVDIEMAAHGGTVIEIRKV..DGK..WQPV
Leptospira_interrogans 97 NAGLLWNNHETLGTLEYYVNGYDSQ.....KQGPNLRTDQQIEQYLYALGGSVIHIYKK..KGE..WILN

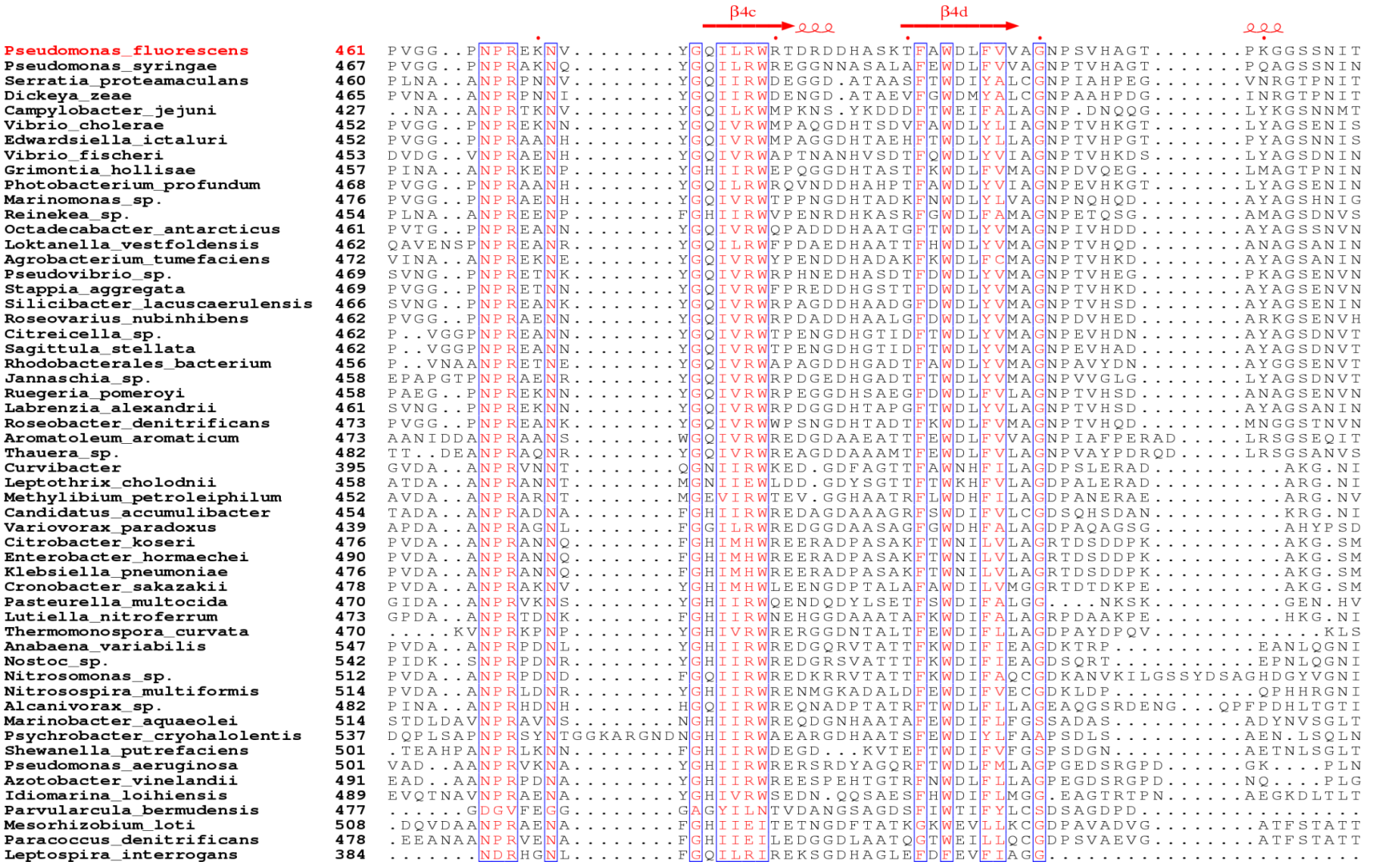

β2a
β2b
β2c
β2d
Pseudomonas_fluorescens 182 QGSRYNRRIHGNS.PLRISGPAAGHELMKT...SADKHGKKVLG.....TFQNCANGKTPWGTYLTCEEN
Pseudomonas_syringae 188 QDSRYNRRVHGNT.PIRLSGPAAGHDWLKT...DADKTGTRVLG.....TFQNCSSGQTPWGTYLTCEEN
Serratia_proteamaculans 182 RPSRYNRRITANT.EMQFSGPAAGHALLQT...AADPTGRKVLG.....TFGNCANGKTPWGTYLTCEEN
Dickeya_zeae 187 RPSRYNRRITANT.AMTFSGPAAGHALLKT...AADPQGINVSG.....TFGNCANGKTPWGTYLTCEEN
Campylobacter_jejuni 160 LDSKYNRRIDANT.KMQVSG.AAKKEVLKN...EK....FVHG.....TFANCANGQTPWRTYITCEEN
Vibrio_cholerae 177 RQGERNRRITAYT.PMMMTGPAAGHDLLKT...AEDPSGLKVLG.....TFNNCANGETPWGTYLTCEEN
Edwardsiella_ictaluri 177 RQGMRNRRITAYT.KMMMTGPAAGHDLLKT...QDDPSGLHVLG.....TCNNCANGQTPWGTYLTCEEN
Vibrio_fischeri 177 SAGKANRRITANT.EMEVTGPAAGHDLLKT...KADPTGKKVLG.....TFNNCANGETPWGTYLTCEEN
Grimontia_hollisae 181 TQGRLNRKITAYT.PMEMTGPARGSDLLKT...HADPTGTKILG.....TFNNCANGQTPWGTYLTCEEN
Photobacterium_profundum 192 IDSKYNRRITANT.EMEITGLVAGHPDVKT...AADKAGKKALG.....TFGNCANGFTPWGTYLTCEEN
Marinomonas_sp. 203 QNSDFNRRIHVQT.EMQLTGPAAGHDLMKT...RIDASGTQVFG.....TMNNCSNGATPWGTYLTCEEN
Reinekea_sp. 182 KNSDLNRRLHGNADEFVLTGPVAGHPLVQT...DMDRTGTKVIG.....TLNNCGNGQTPWGTYLTCEEN
Octadecabacter_antarcticus 194 IDSDLNRRITQNT.PMTVAGPAAGHDLLKT...DDDPTGRTVLG.....TMNNCGAGKTPWGTYLTCEEN
Loktanella_vestfoldensis 194 VDSPFNRRITQNT.PMTFSGPAAGHPLLQT...AADPDGMIVLG.....TFNNCGSGKTPWGTYLTCEEN
Agrobacterium_tumefaciens 199 LDSPFNRRIHHNT.PMKLSGPAAGSDLVKT...AADPKGVDCLG.....TFNNCGAGRTPWGTYLTCEEN
Pseudovibrio_sp. 196 KNSPFNRRITHNT.PMTLAGPAAGHLLLQT...AADPSGTQSLG.....TFNNCGSGKTPWGTYLTCEEN
Stappia_aggregata 197 VDSPFNRRIHQNT.PMKISGPAAGHDLLKT...EADPSGTSALG.....TMNNCGSGKTPWGTYLTCEEN
Silicibacter_lacuscaerulensis 195 KDSPFNRRIHHNT.PMKIAGPAAGHDLMKT...EADPTGTRSLG.....TLNNCGAGKTPWGTYLTCEEN
Roseovarius_nubinhibens 194 IDSPFNRRITHLT.PIRISGPAAGHEMLRT...EADPEGKLALG.....TFNNCGAGRTPWGTFLTCEEN
Citreicella_sp. 196 VDSPFNRRITHET.QMTMDGPAAGSALVQT...NADPEGMSPKG.....TMNNCGSGRTLWGTYLTCEEN
Sagittula_stellata 196 VDSPFNRRITHET.QMTMDGPAAGSALVQT...NADPEGMSPKG.....TMNNCGSGKTLWGTYLTCEEN
Rhodobacterales_bacterium 196 KDSPYNRRITHET.QMTMDGPAAGHALVQT...NADPEGMSPKG.....TMNNCGSGKTLWGTYLTCEEN
Jannaschia_sp. 190 VDSPLNRRIHHNT.PMTLAGPAAGHPLLQT...AADPSGTEVFG.....TMNNCGSGRTLWGTYLTCEEN
Ruegeria_pomeroyi 190 TDSPYNRRITPRT.EMALTGPAAGHDLMKT...AADPTGTVTFG.....TWNNCGSGTTPWGTYLACEEN
Labrenzia_alexandrii 192 KDSPYNRRITPQS.EMELTGPAAGHDLMKT...EADPTGTKSLG.....TWNNCGNGVTPWGTYLTCEEN
Roseobacter_denitrificans 204 KDSPYNRRFTPDE.PMDITGPARGHALMQT...NADPTGETVLG.....TFNNCGNGRTPWGTYLACEEN
Aromatoleum_aromaticum 198 RDSKYNRRITGNT.PMAISGPASGHPLMRT...SADPEGVEVLG.....TLNNCANGFTPWGTYLACEEN
Thauera_sp. 206 IGSPYNRRITGKT.PMTLTGPAAGDALLRT...TADPSGLNVLG.....TLNNCANGKTLWNTYLTCEEN
Curvibacter 127 KPSPWARRITANT.RMELSGPAAGHAMLKT...AADPSGRVVLG.....TINNCASGITPWGTYLTSEEN
Leptothrix_cholodnii 188 LPSKFARRITATT.PMTLAGPAAGHALTKT...AADPSGTKVLG.....TFNNCGSGMTPWGTYLTCEEN
Methylibium_petroleiphilum 183 RPSAYARRISART.PMTISGPAAGHALMRT...EFDPGGREVIG.....TFNNCAHGVTPWGTYLACEEN
Candidatus_accumulibacter 190 RPSAYARRITART.PMRIDGPAAGAVDMRT...RDDRRGNVVFG.....TLANCAMGLTPWGTYLSCEEN
Variovorax_paradoxus 162 RPSPYARRVHGQT.PMRIAGPAAGAVPMRT...AANPAGDEVLG.....TFANCAMGVTPWGTYLTCEEN
Citrobacter_koseri 206 RPSSFARRITVNT.PMQLSGPARHQALMKT...AADPQGEVVLG.....TMQNCANGKTPWGTYLTCEEN
Enterobacter_hormaechei 220 RPSSFARRITVNT.PMQLSGPARHQALMKT...AADPQGEVVLG.....TMQNCANGKTPWGTYLTCEEN
Klebsiella_pneumoniae 206 RPSSFARRITVNT.PMQLSGPARHQALMKT...AADPQGEVVLG.....TMQNCANGKTPWGTYLTCEEN
Cronobacter_sakazakii 208 RPSSFARRITVNT.PMKITGPALHNPLMQT...ADDSKGETILG.....TMQNCANGFTPWGTYLTCEEN
Pasteurella_multocida 197 RPSKYARRITPHT.PMRLTGPAKHNELMKT...VADPLGEFVLG.....TMQNCANGETPWRTYLTCEEN
Lutiella_nitroferrum 202 RPSKYARRIHANT.PIAIQGPARGHALMQT...EADPKGVEVLG.....TFQNCASGPTPWGTYLTCEEN
Thermomonospora_curvata 210 .DSRYNRRITGTT.PMTFSGPVGPR....H...PALKAENAPMG.....TLNNCSNGQTPWGTYLACEEN
Anabaena_variabilis 229 RNSPYGRRVTANT.QMRVSGPAAGDVLLQSKKFNITPNGSVEIGTNDGYTAYGTLNNCANGYTPWGTYLTCEEN
Nostoc_sp. 229 RNSEFGRRITANT.PMRISGPAAGSDLLKSKKFAINPDGSVEIGTNDGFTAYGTANNCAHGYTPWGTYLTCEEN
Nitrosomonas_sp. 215 RHSRLNRRITGNT.LCKISGPAAGHALMKT...AADPHGTRVIG.....TLNNCAHGYTPWGTYLTCEEN
Nitrosospira_multiformis 217 RTSAFNRRITGNS.LCKLSGPAAGHDLMKT...AADPGGMNVLG.....TLNNCSNGHTPWGTYLTCEEN
Alcanivorax_sp. 203 MPSRWNRRISGLT.EMEIRGPLAGHGLMKT...VADPDGKTVIG.....TLNNCSMGVTPWGTYLACEEN
Marinobacter_aquaeolei 213 TDSPYNRRLTVFT.EMQMQGPVAGEDFVKT...KLTPTGTTRYG.....TMNNCANGYTPWGTYLAAEEN
Psychrobacter_cryohalolentis 235 ADSKYNRRLTSSS.TAQLTGPVAGSDLVKT...KFDPTGFQTRG.....INNNCGAGLSPWGTYLTTEEN
Shewanella_putrefaciens 210 LNDSHNRRFTGAT.VMDMAGPLAYTSLLET...RYSPDGSQARG.....TLNNCGNGYTPWGTYLTCEEN
Pseudomonas_aeruginosa 222 P.SQRNRRITAAT.PMRIAGPARGHELLRT...RHSPDGTRTRG.....THNNCSNGFTPWGTYLTCEEN
Azotobacter_vinelandii 210 GDSKFGRRITAAT.PMEIRGPARGHALLRT...RYSRDGSATRG.....TQNNCSNGFTPWGTYLTCEEN
Idiomarina_loihiensis 206 EGSPYNRRITAST.PMQMSGPVAGHDKLVT...RFSPDGTKTRG.....TINNCSHGHTPWGTYLTCEEN
Parvularcula_bermudensis 221 VASPNSRRITVNS.RVAFSGPAADSPLLQN......PAGNVPLG.....TVNNCGSGYTPWGTYLTCEEN
Mesorhizobium_loti 241 RDGKRNRRITSTT.EMALSGPVAGHDRVKT...NADPSGTKVFG.....TFNNCAGGVTPWGTYVMAEEN
Paracoccus_denitrificans 213 LDGALNRRITAKT.PMEICGPATGHARMQT...AADPTGTRVLG.....TINNCAGGVTPWGTYIMAEEN
Leptospira_interrogans 158 PDSKYGRRIHGLT.QFRLTGPAAGSAAVGN...TN....LVFG.....TFANCAGGLTFWNTVLSCEEN

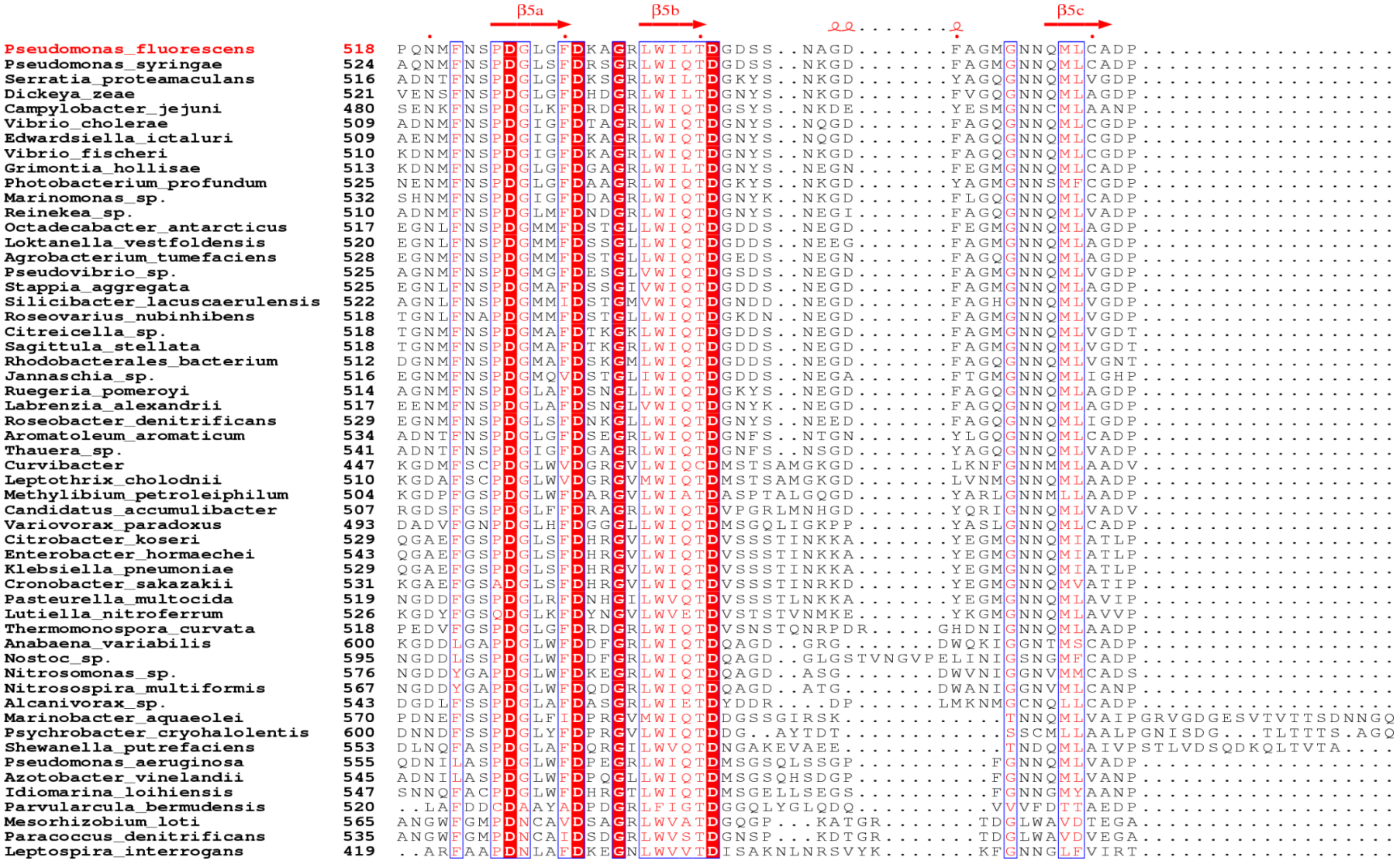

Pseudomonas_fluorescens 243 FTDCFG.....SSNAQ.....QQFDPAQKRYGVS..AASREINWHPFD.....PRFDMAKN
Pseudomonas_syringae 249 FTDCFG.....SSNAE.....QTFDAGMKRYGVV..AASKEINWHPHD.....PRFDVAKN
Serratia_proteamaculans 243 FDTYFG.....THQAD.....YKTTADQKRYTLK..VSEPERNWPDYD.....ERFDVAKN
Dickeya_zeae 248 FDTYFG.....TRQAD.....YQTTPEQKRYTLK..VHEPERNWVDFD.....ERFDVAKH
Campylobacter_jejuni 215 FDDFFG.....SSDEN.....LEFNDSLKRYGFN..KTS.LYGWEKFD.....ERFDLAKN
Vibrio_cholerae 238 FDDFFG.....ADQEG.....SVDADQKRYGIA..AEPSDYQWHKHD.....ARFDITKN
Edwardsiella_ictaluri 238 FNDFFG.....SDRQG.....SVTADHTRYGIA..AAPSIYRWHQHD.....ARFDVTQN
Vibrio_fischeri 238 FNDFFG.....SEQNA.....ELTPDDKRYGIK..AEQSDYQWHNFD.....DRFDVAKN
Grimontia_hollisae 242 FNGYFA.....TSENV.....DLGPTYKRYGIQ..DKDWGYDWWKHD.....ERFDVAKN
Photobacterium_profundum 253 FNGYFG.....SKSTV.....KVNKEQKRYGVN..EADSGYHWFQHD.....DRFDIAKN
Marinomonas_sp. 264 FNGYFG.....ADANV.....ERTEEQKMYGIK..NDEKRYQWYKQD.....DRFDFGRN
Reinekea_sp. 244 FNGYFG.....APAGT.....AFSDEQKAYGLS..ETGFDYNWWPTE.....ERFNLAKH
Octadecabacter_antarcticus 255 FNGYFG.....STDAA.....FKLPEDYARYGIS..LEGR.YNYEKFD.....ARFDVSKT
Loktanella_vestfoldensis 255 FNGYFG.....STDAA.....FARPDDYVRYGIG..LESR.YGYEKFD.....ARFDVSQN
Agrobacterium_tumefaciens 260 FNGYFG.....TTDAA.....FKLPDDYKRYGIV..AETR.YAYEKFD.....ARFDVAKN
Pseudovibrio_sp. 257 FNGYFG.....ATDAAPVERMDGVDAGYKRYGVT..DTGYNYDYHKFD.....ERFDIAKN
Stappia_aggregata 258 FNTYFG.....TTEASQDQKPIERPFGYDRYGIA..NEGR.YGYEKFD.....PRFDFSKN
Silicibacter_lacuscaerulensis 256 FNGYFG.....SSDETASYDP.KVADGYKRYGIK..AQSW.NGYELFD.....PRFDTSKN
Roseovarius_nubinhibens 255 FNGYFG.....ASDEA.....HEMPEDFKRYGIS..HDGWGYDYHKHD.....ARFDVVAT
Citreicella_sp. 257 FNGYFG.....TTG.....EMPEGEGCSRYGIG..AEGR.YAYEKYD.....ARYDLSQE
Sagittula_stellata 257 FNGYFG.....TTG.....EMPEGEGYSRYGIG..GEGR.YAYEKYD.....ARYDLSQE
Rhodobacterales_bacterium 257 FNGYFG.....ATG.....GYDETDGLSRYGIG..GEGR.YAYEKFD.....PRYDLTAE
Jannaschia_sp. 251 FNGYFG.....STDA.....DLDLPEGFDRYGID..HESR.YAYELYH.....ARYDISQE
Ruegeria_pomeroyi 251 FNGYFS.....AADEAH.....EVSAELNRYGVS..AKDWGYGWARVD.....DRFDVAKN
Labrenzia_alexandrii 253 FNGYFS.....APENPDR....EVSPELARYGVS..ATDWGYNWAVIE.....DRYDVSKH
Roseobacter_denitrificans 265 FNGYFG.....TSDADGY....EQDAGKKRYGVS..GKDRGYGWWKID.....DRFDVVKE
Aromatoleum_aromaticum 259 FQGYFG.....TTSG.....IDARDDVMRRYGVS..PLGSGYRWEEHD.....DRFDYVKE
Thauera_sp. 267 FNGYFA.....TAASP....APTRSAAFVRYGIS..AGGSGYRWHEHE.....DRFDYAKE
Curvibacter 188 FINYFS.....GGD.....SLTEHDKRWGLKK.GGG.GYNWHQHD.....ARFDATKT
Leptothrix_cholodnii 249 FINYFN.....GPA.....SPDAHGKRWGLKK.VDDTLYRWHEFD.....ERFDAVKN
Methylibium_petroleiphilum 244 FVGYFH.....GPA.....QPDAHAKRWGLT..PGGYGLRWHEHD.....ARFNATLH
Candidatus_accumulibacter 251 FNGYFN.....ASG.....TPTAAQRRYGIR..EGGSGLRWHEHD.....SRFDVASE
Variovorax_paradoxus 223 FHGYFG.....GPKEAA....KDATPAQRRYGTV..PGSQWVEYWRFE.....ERFDMSRH
Citrobacter_koseri 267 WSDIFV.....KK.....APRNVLEKRYGIS..DSDESYRWNEVD.....ERFSVDKT
Enterobacter_hormaechei 281 WSDIFV.....KK.....APRNVLEKRYGIS..DSDESYRWNEVD.....ERFSVDKT
Klebsiella_pneumoniae 267 WSDIFV.....KK.....APRNVLEKRYGIS..DSDESYRWNEVD.....ERFSVDKT
Cronobacter_sakazakii 269 WSDIFV.....KK.....AEMNPLEKRYGIS..DSDDSYRWNEVD.....KRFSVDAT
Pasteurella_multocida 258 WSDIFV.....RES.....GDFTKLDKRYGIMKKEKEDKYRWNEFD.....ERFNTDKH
Lutiella_nitroferrum 263 FDGVFV.....NDS.....GKLTELEKRYGIS..NKDAGYKWAWED.....FRFDAQLN
Thermomonospora_curvata 266 WNSYFG.....TTDSS.....WKPTPQQARYGVS...ASSSYRWHEAD.....PRFDVAVN
Anabaena_variabilis 302 WNGYFANPTLAANSTSDVESIPGIDKS.DILVGQRRYGIP..SQSS.YRWPDVD.....PRFNAQTN
Nostoc_sp. 302 WNGYFG.....NNGEPLASTPGINDS.EVLAVQNRYGLS..AGGFGYRWHEVD.....PRFNNRTN
Nitrosomonas_sp. 276 WNGYFS.....NETGEVEGESDSERTLKIIQNQLRYGVV..KGGFGYRWHEVD.....ERFRADLH
Nitrosospira_multiformis 278 WNGYFS.....NETGDVAGANDPEQKRRILNGQARYGIG..KGGFGYRWHEMD.....ARFRADLN
Alcanivorax_sp. 264 WHDYFV.....NRDSE.....DQRVSHRRYGIGEGERADKYAWNTAE.....PRFDATFD
Marinobacter_aquaeolei 274 WYAYFRVG.....NVDALSEKQKGWLARYGISGN..GWAYR.....QWDEVPGDQYQRFGVNTS
Psychrobacter_cryohalolentis 296 FLGVFARGQ.....DASQLSAGHNYGRERYGATENFPGWEYLWHTPDAKDAKISDEFSRWDMTAA
Shewanella_putrefaciens 271 WPEYFANR.....GTLKPDQS..RIGISKTSTRYGWDHLAG....HDEERLDEFARFDITPK
Pseudomonas_aeruginosa 282 WVGYFSTQ.....DSELPRELTRYGIR.NTSRFGWETLAG.....DEFERFDATRK
Azotobacter_vinelandii 271 WANYFATQ.....DSELPRELQRYGIG.KTGRFGWNTLEG.....DEFQRFDATRK
Idiomarina_loihiensis 267 WAGCFMND.....DDTHPREHERYGVNRENSRYYWETATPG.....LDETLRFNASST
Parvularcula_bermudensis 279 FNGYFGANG.....AFDPSARQARYGLS..ANGFGYGWHVFD.....KRFDLTDP
Mesorhizobium_loti 302 IHGYFSG.....ELPEGHKEAANYKRLGIPEGAYEWGAHY.....DRFNLAKE
Paracoccus_denitrificans 274 IHGYFTG.....ALPEGHREAENYKRMGIPEDSYAWGRFH.....DRFDIGKE
Leptospira_interrogans 214 VEWIIEP.....CKLT.....

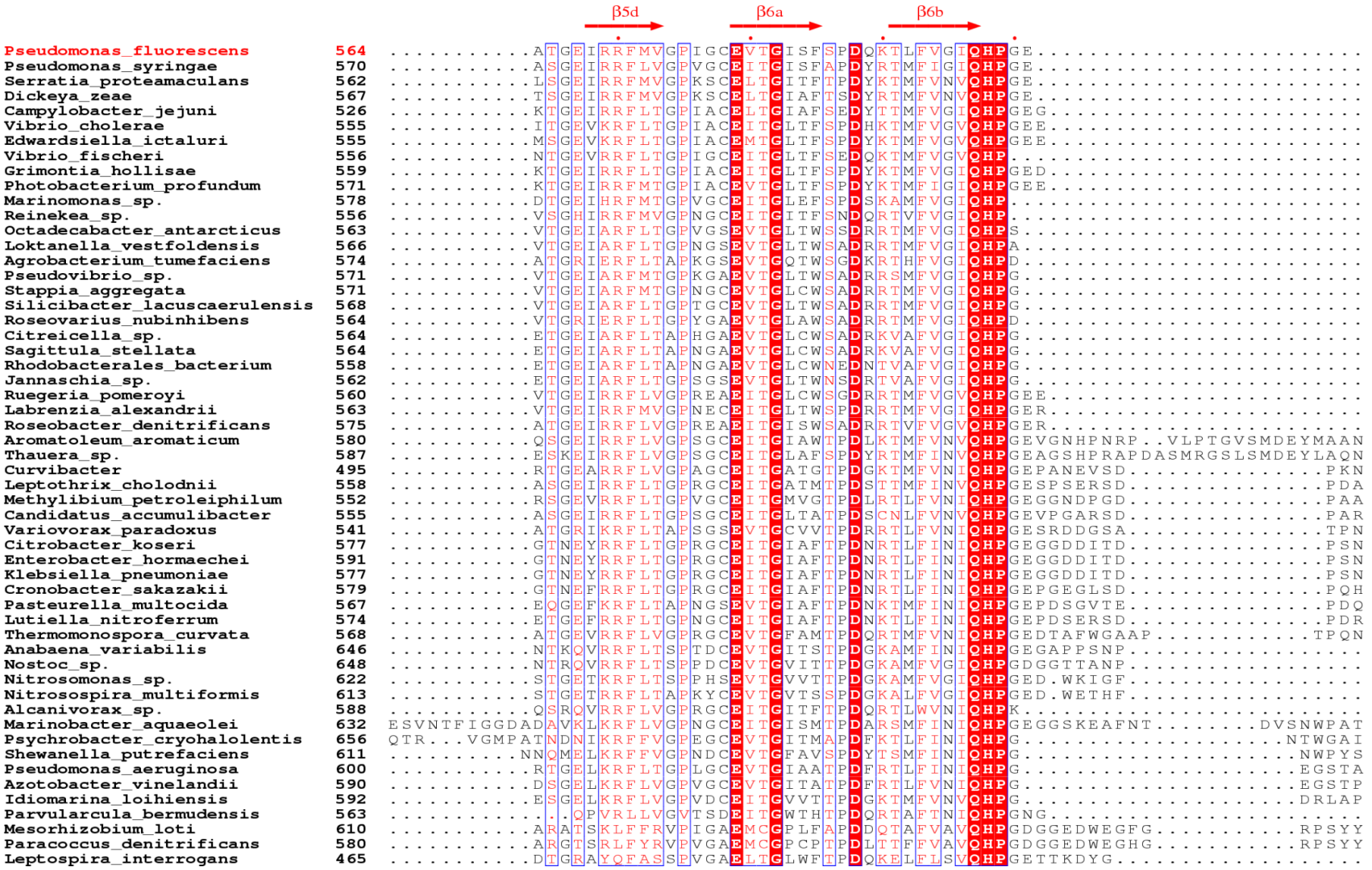

β2e
β2f
β3a
β3b
β3c
Pseudomonas_fluorescens 287 P.....NELNRHGWVVEIDPFDPQSTPVKRTALGRFKHENAALAETDDGR.AVVYMGDDERGEFIYKFVSR
Pseudomonas_syringae 293 P.....NEVNRHGWVVEIDPFDPDSTPVKRTALGRFKHENAAFTQTRDGR.AVVYMGDDERGEFIYKFISR
Serratia_proteamaculans 287 P.....NEFNRHGWIVEIDPMNPTSTPIKHTALGRFKHENAAVTVAKDGR.LVVYMGDDERGEHIYKYVSK
Dickeya_zeae 292 P.....NEFNRHGWIVEIDPLNPTSTPVKRTALGRFKHENAAVTVAKSGH.LVVYMGDDERGEFIYKYVSN
Campylobacter_jejuni 258 I.....NEANRFGWIVEINPFDAKSTPVKRTSLGRFKHENAEIIVEKDGS.VIVYMGDDEMNEFIYKFVSK
Vibrio_cholerae 281 P.....KEPNRFGWVVEIDPHNPNSTPLKRTALGRFKHENAALVINNDGH.VVVYLGDDERGEHLYKFVSK
Edwardsiella_ictaluri 281 P.....QEPNRFGWVVEIDPHNPQSVPQKRTALGRFKHENAALTLNKDGR.VVIYLGDDERGEHLYRFVSH
Vibrio_fischeri 281 P.....NEPNRFGWVVEIDPNDPTSTPLKRTGLGRFKHENAALVVNDDGH.VVVYLGDDERGEHLYKFVSK
Grimontia_hollisae 285 P.....HEPHRFGWVVEIDPMDPTSTPKKRSALGRFKHENAALTINKDGH.VVVYLGDDERGEHLYKFVSK
Photobacterium_profundum 296 P.....NEANRHGWVVEIDPMDPTSTPMKRTAMGRFKHENVAIVVNKDGH.VVAYMGDDERGEHLYRFVSK
Marinomonas_sp. 307 P.....NEPHKFGWIVEIDPKDPKSLPLKRTAMGRFKHENVAIHVNHDGS.VVAYMGDDERGAHIYKFVSE
Reinekea_sp. 287 P.....NEANRFGWIVEIDPYNPSSKAMKRTAMGRFKHENAALTLSKDGH.AVAYMGDDERGEFIYKFISR
Octadecabacter_antarcticus 298 P.....NEPHRFGYIVEIDPANPLAAPIKRTALGRFKHENAAATLAPNGQ.VVVYLGDDERGEFLYKYVSN
Loktanella_vestfoldensis 298 V.....NEPRRQGYVVEIDPSDATSTPVKHTALGRFKHENAAVVVARDGR.VVVYLGDDERGEFLYKFVSA
Agrobacterium_tumefaciens 303 P.....NEPRRAGYVVEIDPSDAKSTPIKRTALGRIKHENAAVVIARDGR.VVVYMGDDERGEFLYKFVSN
Pseudovibrio_sp. 306 P.....NEPHRAGWVVEIDPTDPNSTPVKRTALGRFKHENAEAVLAPDGR.VVVYMGDDERGEFMYRFVSN
Stappia_aggregata 306 P.....NEPNRFGWIVEIDPTDPNNTPVKRTALGRFKHENAEVVLSRDGR.VVVYLGDDERGEFLYKFVSN
Silicibacter_lacuscaerulensis 303 P.....NEPHRAGWVVEIDPTDPESTPVKRTGLGRFKHENAAVELARDGR.VVVYMGDDERGEYLYRFVSN
Roseovarius_nubinhibens 299 P.....NEPRRAGYVVEIDPADPNSVPVKRTGLGRFKHENAAVVLAADGR.VVVYMGDDERGEFLYKFVSD
Citreicella_sp. 299 P.....NEPHRHGWITEIDPSDASSTPVKHTALGRFKHENAEMVQAADGR.VVVYMGDDERGEFIYKYVSN
Sagittula_stellata 299 P.....NEPHRHGWITEIDPSDASSTPVKHTALGRFKHENAEMVQAADGR.VVVYMGDDERGEFIYKYVSN
Rhodobacterales_bacterium 299 P.....NEPNRHGWITEIDPTDPESTPVKHTALGRFKHENAEMVQAADGR.VVVYMGDDERGEFIYRYVSN
Jannaschia_sp. 294 P.....NEAHRFGYIVEIDPADPTSQPVKRTALGRFKHENAAMTLAEDGR.VVVYMGDDERGEFMYKYVSN
Ruegeria_pomeroyi 295 P.....NEPNRAGYVVEIDPRDPNSTPKKRTALGRFKHENAECVVNRDGR.LVIYMGDDERGEFLYRYVSN
Labrenzia_alexandrii 298 P.....NEPNRAGYVVEIDPTDPTSTPKKRTALGRFKHENAEAVVNNDGR.IVVYMGDDERGEFVYRYVSN
Roseobacter_denitrificans 310 P.....NEANRHGYITEVDPYDPTSKPKKRTALGRFKHENAEVVVNFDGR.IAIYMGDDERGEFLYRYVSD
Aromatoleum_aromaticum 303 P.....NESNCFGWIVEIDPFDPDSVPVKHTALGRIKHENAALTVAGTGR.VVVYMGDDQTNDYIYKYVSD
Thauera_sp. 312 P.....NEANRFGWVVEINPFEPGSTPKKRTALGRFKHENAEMRLAADKR.VVVYMGDDQANDYIYKFVSD
Curvibacter 229 P.....NEPNRFGWIVEIDPYNPSSTPMKRTAMGRAAHEGATVAVTKDNR.AVVYMGEDSRFEYIYKFVSR
Leptothrix_cholodnii 291 P.....NEPNRHGWVVEIDPAKPGSTPVKRTALGRAAHEGATVAVTKDQR.AVVYMGEDARFEYIYKFVSR
Methylibium_petroleiphilum 285 P.....NEFNRFGWVVEIDPLDPRSTPVKRTALGRAAHEGAFVTQAADGR.AVVYMGEDARFEYLYKFVSR
Candidatus_accumulibacter 292 P.....NEANRFGWVVEVDPWQPGRPPVKHTALGRFKHEGATVTLASDGR.VVVYMGDDEAFEYIYKYVSR
Variovorax_paradoxus 268 P.....NEPHRFGWVVEIDPFDPAAKPVKRTALGRKRQESATCTLAKDGR.VVVYMGDDARFEYIYKFVSH
Citrobacter_koseri 308 P.....NEPNRFGWVVEIDPYDPTSTPRKHTALGRFKHEGAAVTLAGDNR.VVVYMGDDQKFEYIYKFISE
Enterobacter_hormaechei 322 P.....NEPNRFGWVVEIDPYDPTSTPRKHTALGRFKHEGAAVTLAGDNR.VVVYMGDDQKFEYIYKFISE
Klebsiella_pneumoniae 308 P.....NEPNRFGWVVEIDPYDPTSTPRKHTALGRFKHEGAAVTLAGDNR.VVVYMGDDQKFEYIYKFISE
Cronobacter_sakazakii 310 P.....NEPNRFGWVVEIDPYDPHSVPRKHTALGRIKHEGAAVTLAPDNR.VVVYMGDDQKFEYIYKFVSE
Pasteurella_multocida 302 P.....NEPHRFGWVVEIDPFDPNSTPVKHTALGRFKHEGAMLVLSKEGH.AVVYMGDDQRFEYIYKFVSK
Lutiella_nitroferrum 305 P.....NEPNRFGWVVEIDPFDPKSKPVKRTALGRFKHEGATVAIAPSGH.AVVYMGDDQRFEYIYKFVSK
Thermomonospora_curvata 309 P.....NELNRFGWVVEIDPFDPKSTPVKRTALGRIKHEGATVTESR.GR.IVVYTGDDQDGEYIYKFVGS
Anabaena_variabilis 360 P.....LEPHLFGWVVEIDPYDPQSTPVKRTALGRFKHESAQVVIDDNNR.AAFYMGDDERNEYIYKFVCA
Nostoc_sp. 355 P.....LEPHTFGWVVEIDPYTPRSTPVKRTALGRFKHESAQVVVDDNNR.VAFYQGDDERNDYIYKFVCA
Nitrosomonas_sp. 330 P.....NEPNRFGWVVEIDPWDPKSTPIKRTAMGRMKHESAQVVIDHKDR.VAFYMGDDERNEYIYKFVCD
Nitrosospira_multiformis 332 P.....NESHRFGWVVEIDPWDPKSTPVKRTALGRFKHENASCVVDPDNT.VVIYMGDDERNEYLYKFVCA
Alcanivorax_sp. 309 ETQAFGGYLQEPHRYGWVVEIDPFDPESTPIKRTAFGRFCRECCTASLGEDGR.LAFYSGDDSRGEYLYKFVPE
Marinobacter_aquaeolei 326 GSTATEDFRYEPNLHGYVTEVDPFRPTQKPRVRTGLGRFSHEGAWVAPAKEGAPIVIYSGDDSRREYMYKFVSK
Psychrobacter_cryohalolentis 356 SASAADDYRNGFNTFGYITEIDPFDQNSIPQKRTALGRFAHENCAYAPVEQGKPVVFYMGDDARGEYIYKFVSK
Shewanella_putrefaciens 322 GANAQDDYRNEANGHGYIVEIDPFNPASRAVKRTALGRFRHEGCTFGKLTEGKPVVFYSGHDARFEYLYKFVSA
Pseudomonas_aeruginosa 327 GKQAQDDYRNEPNNFGWIVEIDPFDPQSVPVKRTALGRFAHEGLVFAPVKPGRQVVCYSGDDSQNEYIYKYVSR
Azotobacter_vinelandii 316 AETASGDYRNEPNHFGWIVEIDPFDPQSVPVKHTALGRFAHEGVIFAPTGPGKRLVCYSGDDSRNEYIYKYVSR
Idiomarina_loihiensis 315 GSTAADDYRNEPNHFGWIVEIDPFNAKSTPVKRSMLGRFAHEGIVFAPAVEGEPLVAYSGDDSRFEYIYKFVSA
Parvularcula_bermudensis 322 AY.....VNENNRFGWIVEINPFDGTELPTKRTALGRFKHEAIAIAETPDGR.VAAYMGDDQSNDYCYKFLSA
Mesorhizobium_loti 345 P.....NEPNRFGWIVEVDVNDPNSTPRKRTAMGRFKHEGAESIVAKDGR.VVFYLGDDERFDYVYKFVTK
Paracoccus_denitrificans 317 P.....NEPNRFGWIVEVDVMDPTSTPRKRTALGRFKHEGAASALAKDGR.VVFYLGDDERFDYVYKFVTA
Leptospira_interrogans 225 .....HETHYGWVIEVDPFDPKSVPVKHTALGRFAHENAALVLSPSGK.LVVYMGDDARDEFVYKFVSK


β3d
β3e
Pseudomonas_fluorescens 352 DKINHRN.....AKANRDILDHGTLYVARFDAGDGNPDHPKGQGQWIELTHGKN....GIDASSG.....
Pseudomonas_syringae 358 DRINRKN.....PKANRDLLDHGTLYVAQFDAGNGNPDHPKGNGQWIELKHGKN....GLDAAAG.....
Serratia_proteamaculans 352 GVVDVAN.....PANNRSLLDEGTLYVAQFDG.DAGGTPLKGTGRWIALEFGKN....GLTPENG.....
Dickeya_zeae 357 GKVDHAN.....PSANSRLLDDGVLYVAQFDG.DASGTPLKGKGRWVALEFGKN....GLTPENG.....
Campylobacter_jejuni 323 HKYKKG.....ADTSKILDEGTLYVGQFNG...NVGDFKGQGKWIALEYGKN....SLDEKNG.....
Vibrio_cholerae 346 HRYQAGN.....DQQNRNLLEEGTLYVAKFDINE...NELKGSGRWMELSFGKN....GLTPENG.....
Edwardsiella_ictaluri 346 QRYQPGN.....DAANRNLLDHGTLYVAQFDMDE...NELKGNGRWIALTFGQN....GLTPENG.....
Vibrio_fischeri 346 DKYQEGN.....NRANRNLLEEGTLYVAKFAMKD...DKLEGDGQWLELTYGKN....GLTRENG.....
Grimontia_hollisae 350 NKYKPGV.....HASNRDLLEEGTLYVAKFEGKD...GELKGKGEWLELTFGKN....GLTPDNG.....
Photobacterium_profundum 361 NKYVEGN.....DTENRKLLEEGTIYVAKFSTEE...GKIGGTGQWIDLTHGKN....GLTKENG.....
Marinomonas_sp. 372 HKYNPNN.....PDANRNLLVDGTLYVAKFSMTQ...D..KGAGEWVELTFGKN....GLTPENG.....
Reinekea_sp. 352 DTYREGD.....RAHNMTLLENGTLYAAKFN.....DNGSGEWVELTYGKN....GLTQANG.....
Octadecabacter_antarcticus 363 GIYTPGG.....D..TSTLLDDGQLYVAKFND.....DMSGEWIALTEDAT....DMS.....
Loktanella_vestfoldensis 363 GVYTPGG.....E..TSSLLSDGQLYVARFAT.....DMTGEWVALTPEAT....GMD.....
Agrobacterium_tumefaciens 368 GIYVPGG.....D..TSRLLDEGTLHVAKFAD.....DGAGEWVALTPETT....GMK.....
Pseudovibrio_sp. 371 GVYAPGQ.....P..TDGLLDDGTLYVAKFND.....DQTGEWLALTPDTT....GMD.....
Stappia_aggregata 371 GSYVPGG.....P..TEGLLDDGTLYVAKFND.....DQSGEWLALTPDTT....GMP.....
Silicibacter_lacuscaerulensis 368 GIYTPGG.....E..TEGLLDDGTLYVARFND.....DMTGQWVPLTPETT....GME.....
Roseovarius_nubinhibens 364 GVYAPGA.....D..TEALLDEGTLYAARFLD.....DGRGEWLPLTEETT....GMT.....
Citreicella_sp. 364 GTWAEGQ.....P..TDGLLSDGTLYVAKFNP.....DQTGEWLALTPEST....GMS.....
Sagittula_stellata 364 GTWAEGQ.....P..TDGLLSDGTLYVAKFNP.....DQTGEWLALTPEST....SMS.....
Rhodobacterales_bacterium 364 GTYAEGG.....S..TEGLLSDGQLYVAKFND.....DMTGEWIALTPEAT....GMS.....
Jannaschia_sp. 359 GTYTDGG.....P..TDGLLDDGTLYAAYFSA.....DGTGEWRALTPDTT....QMD.....
Ruegeria_pomeroyi 360 GVYAAGV.....D..TDELMENGTLYAAKFDA.....SGKGIWLALTPEAT....GMS.....
Labrenzia_alexandrii 363 GVYAPGA.....P..TDELLNDGILYVAKFSD.....NGTGAWVALTPETT....GMD.....
Roseobacter_denitrificans 375 GYYAEGG.....N..TDALMENGTIYAAKFEA.....DGTGKWLALTPEAT....GMD.....
Aromatoleum_aromaticum 368 DRFIPGR.....KTANSELLDHGRLYVAKFSDGAVS.GDFIGTGEWILLDKHANP...ILAADPR.....
Thauera_sp. 377 GVFDASR.....GLANGNLLDAGKLYVAKFDAGAAS.GDFMGVGEWLLLDKAANP...TLAADAR.....
Curvibacter 294 DAIKPGG.....AAANATLLDHGTLYVAKFNA.....DGKGQWIALTHGQG....PLTAANG.....
Leptothrix_cholodnii 356 DAIKAGG.....AAANATLLDHGTLYVARFDA.....EGKGTWLPLVHGQG....PLTAANG.....
Methylibium_petroleiphilum 350 DPVRTGG.....YANNREVLDHGTLYVARFDS.....DGTGEWLALVHGQN....GLDAQAG.....
Candidatus_accumulibacter 357 DRFIAGS.....SATSGRWLEVGTLYVARFEA.....DGRGQWVALQHGQN....NLNEANG.....
Variovorax_paradoxus 333 GKVSMAGDGAR..ASANSRLLDEGTLYAARFDA.....DGRGQWLELVHGRN....GLDADSG.....
Citrobacter_koseri 373 NKYDPGD.....RKANMQLLESGTLYVARFND.....DGSGDWLPLIFGEN....GLDQSKG.....
Enterobacter_hormaechei 387 NKYDPGD.....RKANMQLLESGTLYVARFND.....DGSGDWLPLIFGEN....GLDQSKG.....
Klebsiella_pneumoniae 373 NKYDPGD.....RKANMQLLESGTLYVARFND.....DGSGDWLPLIFGEN....GLDQSKG.....
Cronobacter_sakazakii 375 GKFNPQD.....RKANMHLLEKGTLYVAKFND.....DGKGEWLPLVFGQN....GLDASKG.....
Pasteurella_multocida 367 GKYNPAD.....RAANMYLLSEGTLYVAHFNE.....DNTGEWRPLVHNQN....GLTAENG.....
Lutiella_nitroferrum 370 NTFDPKN.....RKANMQLLDEGTLYVARFDA.....DGSGQWLPLVAGQN....GLSADKG.....
Thermomonospora_curvata 373 APWRLLR.....ALRKSPLDHGTLYVAKFND.....DGTGRWLPLKHGTG....PLTKAGG.....
Anabaena_variabilis 425 QPYNPGN.....RAANRDLLDNGILYVAKFND.....NGTGQWIPLVYGQN....GLTPENG.....
Nostoc_sp. 420 QPYNPRN.....REANRDLLDTGVLYVAKFND.....NGTGQWLPLVYGQN....GLTPENG.....
Nitrosomonas_sp. 395 KKFNREG.....HAENRNMLDHGTLYVAKFNA.....DGSGQWLPLIWGKN....GLTRKNG.....
Nitrosospira_multiformis 397 NKYNPRN.....RAANRDLLDSGTLFVARFNA.....DGGGKWLPLVWNQN....GLTPANG.....
Alcanivorax_sp. 382 GRHIEGD.....REHNRHLLDTGTLYVARFDG.....DGSGVWLPLVHGEN....GLTEANG.....
Marinobacter_aquaeolei 400 TAWTESDRNGG..LAAGDKYLDEGTLYVAVFND.....DGTGAWKALTISNPEL.NGDHTYNDDGTTYQYS
Psychrobacter_cryohalolentis 430 ATWSNADIGGG..LKAGDKYLNEGTLYVAIFNE.....NGRGEWKALIHGQNGL.DVSNTVLP.....
Shewanella_putrefaciens 396 AAWDAKDANSSNRIQTGNKYMDEGTLYVAKFNE.....NGGGVWLPLTLDSQTS.DGKTLGDT.....
Pseudomonas_aeruginosa 401 DKFRPGRS.NT.....RLLDEGTLYVARFNA.....DGSGQWLPLDIADGNF.RAACRKAG.....VS
Azotobacter_vinelandii 390 DKFRPGRPDNS.....RLLDEGTLYVARFNA.....DGTGDWLPLDIQDATF.QAACAAKG.....IG
Idiomarina_loihiensis 389 KPYYKATA.SG.....ELLNEGTLYVAQFND.....DGTGRWLPLDIENSEF.QAKASEAG.....VG
Parvularcula_bermudensis 389 LPWREAIA.....QGVSPLDEGTLYVARFND.....DGTGDWLPLTIDNP....LLAAEFA.....
Mesorhizobium_loti 410 GAFNAND.....HAANKDLLDEGTLHVAKFAEDG.....TVDWLPIVFGQ....GPLTAENG.....
Paracoccus_denitrificans 382 GRYNPDD.....RAANLDLLDEGTLYVARFDEDG.....SFVWLPLVHGQ....GPLTAENG.....
Leptospira_interrogans 288 KSYDPSL.....GAKNSTLLEEGILYVADFDKG.....TWIPLDLETNETLKNYKTENGN.....RQ


β4a
β4b
Pseudomonas_fluorescens 408 FADQAEVLIHARLAASVVGATTRMDRPEWIVVSPKDG.....QVYCTLTNNAKR.....GEDGQ
Pseudomonas_syringae 414 FANQAEVLIHARLAASVVKATTRMDRPEWIVVSPKDG.....QVYCTLTNNSKR.....GDEGQ
Serratia_proteamaculans 407 FRDEAEVLIFARKAAAQVGATTKMDRPEWIAVNPHDG.....RAYCTLTNNSKR.....GEEGM
Dickeya_zeae 412 FNSQAEVLIFARKAASQVNATTRMDRPEWIAVNPHDG.....RAYCTLTNNSKR.....GEEGM
Campylobacter_jejuni 374 FKSQAQVLINTRLAASIVGATTPMDRCEWIASHKQSGSK...EVFATLTNNKNR.....QEP..
Vibrio_cholerae 399 FKDQAEVLIFARRAATQVGATTMDRPEWVAVHPDKK.....HVFCTLTNNKNRG.....KEGQ
Edwardsiella_ictaluri 399 FNDQGEVLIFARRAATQVGATTMDRPEWVAIHPNNQ.....HVFCTLTNNKNRG.....KEGQ
Vibrio_fischeri 399 FNSQAEVMIFARRAATQVGATTMDRPEWVAVHPDKK.....HVFCTLTNNKYRGV.....KEGQ
Grimontia_hollisae 403 FNNQAEVLIFARRAATQVGATTMDRPEWVAIHPDKK.....HVFCTLTNNKYRGV.....RENQ
Photobacterium_profundum 414 FPNQEYIHLYTRLAATQVGATTMDRPEWIAVHPNQE.....YVFCTLTNNKNRGV.....KEGQ
Marinomonas_sp. 423 FADHAEVLIYTRKAAKHVGATTMDRPEWIAVHPDGH.....QVYCTLTNNKNRG.....KEGQ
Reinekea_sp. 400 FDSQEYILVYARLAARFVGATTMDRPEWVACHPSSP.....MVFCTLTNNKYRGV.....RDNQ
Octadecabacter_antarcticus 405 ...AAEIAIFTRQAGSRVGATTMDRPEWVAVNPVSV.....ECYCALTNNKNRGLKP.....NAGGDET
Loktanella_vestfoldensis 405 ...AAEIAVYTRVAASRVGATTMDRPEWVAVNPVAI.....EGYCALTNNSRRGADAT.....NAGGDAM
Agrobacterium_tumefaciens 410 ...IDEICVFTRQAASKVGATTMDRPEWVAINPVAI.....EAYCALTNNSRRGEMKDGKLRA....NAGGDAM
Pseudovibrio_sp. 413 ...AASIVIFSRIAGSAVGATTMDRPEWISANPNAV.....EAYCCLTNNKNRGVKP.....NAGGDDT
Stappia_aggregata 413 ...LAEILVYSRIAASKVGATTMDRPEWVAANPHAV.....EAYCCLTNNKNRGVKT.....NAGGDET
Silicibacter_lacuscaerulensis 410 ...PAEILIFSRMAASKAGGTTMDRPEWVATSPVAS.....EAYCCLTNNKNRGVKP.....NAGGDDT
Roseovarius_nubinhibens 406 ...KAEIAILTRMAGSNAGATTMDRPEWVAVNPNAV.....EAYCALTNNKNRGIKP.....NAGGDET
Citreicella_sp. 406 ...LEEILVFSRMAGSKVRATTMDRPEWIACNPLKA.....EAYCALTNNKNRGL.KT.....NAGGDET
Sagittula_stellata 406 ...LEEILVFSRMAGSKVGATTMDRPEWIACNPLKA.....EAYCALTNNKNRGL.KT.....NAGGDET
Rhodobacterales_bacterium 406 ...LEETLVFSRIAASAVGATSMDRPEWIAANPLKA.....EAYCALTNNSKR.....GGDAM
Jannaschia_sp. 401 ...AAMICIHTRHAGSSVGATTMDRPEWVAVSTTLA.....EAYCCLTNNSRRAEGFT.....NAGGDPA
Ruegeria_pomeroyi 402 ...QAEICIHTRIAASKLGATTMDRPEWVTANPNAD.....EVYCALTNNKNRGVKP.....NRGGDDT
Labrenzia_alexandrii 405 ...QAEIHIHTRQAGSKVGATTMDRPEWISANPKAP.....EIYCCLTNNKKRGVKP.....NAGGDDT
Roseobacter_denitrificans 417 ...KAMICIQTRMAASAVGATTMDRPEWVTANPLKA.....EVYCALTNNKNRGIKP.....NAGGDET
Aromatoleum_aromaticum 424 FVDQGELLIKTRLAADAVGATKMDRPEWITVHPGSG.....EVYCTLTNNSGR.....A
Thauera_sp. 433 FATQAEVLIHARLAADAVGATKMDRPEWITTHPQTG.....EVYCALTNNSGR.....T
Curvibacter 342 FADQGEVLIKARQASDLLGATKMDRPEWIAVD..KQG.....WVYTTLTNNSNRGG.....DKQP
Leptothrix_cholodnii 404 FADQGEVLIKSRQASDALRATKMDRPEWTAVDP.KTR.....KVYCTLTNNSNRGR.....DKQP
Methylibium_petroleiphilum 398 FADQGELLVKSRQASDKVGATKMDRPEWIAGDP.RTA.....TLYCSLTNNSRRGT.....DGYP
Candidatus_accumulibacter 405 FAGQAEVLIRARQAADLVGATKMDRPEWIAVQP.QSG.....EVYCTLTNNRQR.....V
Variovorax_paradoxus 385 FADQAEVLIHARLAADAVGATRMDRPEWIAVHP.QSG.....EVYVTLTNNSQRGE.....PGKP
Citrobacter_koseri 421 FDNQGDLLIKTRLAADTVGATKMDRPEWIAVDTHAKG.....SVYCTLTNNSDRGK.....EGKA
Enterobacter_hormaechei 435 FDNQGDLLIKTRLAADTVGATKMDRPEWIAVDTHAKG.....SVYCTLTNNSDRGK.....EGKA
Klebsiella_pneumoniae 421 FDNQGDLLIKTRLAADTVGATKMDRPEWIAVDTHAKG.....SVYCTLTNNSDRGK.....EGKA
Cronobacter_sakazakii 423 FENQGDLLIKTRLAADAVGATKMDRPEWIAVDPYHTG.....SVYCTLTNNSDRGK.....EGKA
Pasteurella_multocida 415 FLNQGDVTIKARMAADVVGGTKMDRPEWIAVDPYQTG.....SVYCTLTNNSQRGT.....EGKA
Lutiella_nitroferrum 418 FASQGDVVIKARMAADVMGATKMDRPEWIAVHSHKPG.....EVYCTLTNNSERGK.....DGKP
Thermomonospora_curvata 420 WADQADVLLRTREAADAVGATKLDRPEWVAVHPRTK.....DVYLTLTNGSGWNN.....
Anabaena_variabilis 473 FRSQAEVLVKTRQAADRVGATMMDRPEWTAVRPRIGGYKEIEVYCTLTNNRRGSTPPSSNNPNGTTTAASARP
Nostoc_sp. 468 FRNQAEVLVKTRQAADRVGATMMDRPEWVAVRPRIGGFREIEVYCTLTNNRRGTGTESVNQADGSTTAGSARP
Nitrosomonas_sp. 443 FADQAEVLIKTRQAADRLGATMMDRPEWGAVHPVTG.....DIYMTLTNNSRRGNTPVSGNNIDGTGSAGSANP
Nitrosospira_multiformis 445 FADQAEVLIKARQAADRAGATMMDRPEWIAAHPASR.....EVYMTLTNNRRGSNPPSGNSIDGSTPAGSARP
Alcanivorax_sp. 430 FADQGELVLVNARAAADRLGATTMDRPEWVAVDPNSR.....EVYATLTNNHKRG.....HAH
Marinobacter_aquaeolei 463 FASEAEVLASARLAADVLGATPMDRPEWAAVVHPTNG.....DVYLTLTNGNPS.N.....RP
Psychrobacter_cryohalolentis 485 FHGQDEVLVFARAAAADVVGATTKMDRPEWVSVSPITG.....DVYVTLTNNKYRGV.....RD
Shewanella_putrefaciens 453 FASLADIILNTAGAADLVGATTPMDRPEWCAVDPFTG.....SVYLTLTNNSKR.....
Pseudomonas_aeruginosa 452 FADQGEVLINTRLAADVLGATKMDRPEWGAVNPDNG.....DVYFTLTNNTSR.....T
Azotobacter_vinelandii 442 FADQGEVLINTRTAADAVGATKMDRPEWGAVHPRSG.....EVYFTLTNNTNR.....T
Idiomarina_loihiensis 440 FHSQADVLLNTRLAADVMGATKMDRPEWGAVHPTTK.....EVYFSLTNNTRR.....T
Parvularcula_bermudensis 436 ..DQAEILIFARAAADILGATPMARPEWTTIAKDGS.....TYWTMTD.....
Mesorhizobium_loti 458 FGSQADVLIETRRAADLLGATKMDRPEDIQPNGVNG.....KVYVMLTNNSKRKA.....
Paracoccus_denitrificans 430 FESQADVLIETRRAADLLGATPMDRPEDIEPNAETG.....RAYVMLTNNTKR.....
Leptospira_interrogans 340 FETQADVLVHCRAAAKICGATPMDRPEDIFIHPLDG.....TIFIAMTN.....


β4c
β4d
Pseudomonas_fluorescens 461 PVGG..PNPREKNV.....YGQILRWRTDRDDHASKTFAWDLFVVAGNPSVHAGT.....PKGGSSNIT
Pseudomonas_syringae 467 PVGG..PNPRAKNQ.....YGQILRWREGGNNASALAFEWDLFVVAGNPTVHAGT.....PQAGSSNIN
Serratia_proteamaculans 460 PLNA..ANPRPNNV.....YGQIIRWDEGGD.ATAASFTWDIYALCGNPIAHPEG.....VNRGTPNIT
Dickeya_zeae 465 PVNA..ANPRPNNI.....YGQIIRWDENGD.ATAEVFGWDMYALCGNPAAHPDG.....INRGTPNIT
Campylobacter_jejuni 427 ..NA..ANPRTKNV.....YGQILKWMPKNS.YKDDDFTWEIFALAGNP.DNQQG.....LYKGSNNMT
Vibrio_cholerae 452 PVGG..PNPREKNN.....YGQIVRWMPAQGDHTSDVFAWDLYLIAGNPTVHKGT.....LYAGSENIS
Edwardsiella_ictaluri 452 PVGG..PNPRAANH.....YGQIVRWMPAGGDHTAEHFTWDLYLLAGNPTVHPGT.....PYAGSNNIS
Vibrio_fischeri 453 DVDG..VNPRAENH.....YGQIVRWAPTNANHVSDTFQWDLYVIAGNPTVHKDS.....LYAGSDNIN
Grimontia_hollisae 457 PINA..ANPRKENP.....YGHIIRWEPQGGDHTASTFKWDLFVMAGNPDVQEG.....LMAGTPNIN
Photobacterium_profundum 468 PVGG..PNPRAANH.....YGQILRWRQVNDDHAHPTFAWDLYVIAGNPEVHKGT.....LYAGSENIN
Marinomonas_sp. 476 PVGG..PNPRAENH.....YGQIVRWTPPNGDHTADKFNWDLYLVAGNPNQHQD.....AYAGSHNIG
Reinekea_sp. 454 PLNA..ANPREENP.....FGHIIRWVPENRDHKASRFGWDLFAMAGNPETQSG.....AMAGSDNVS
Octadecabacter_antarcticus 461 PVTG..PNPREANN.....YGQIVRWQPADDDHAATGFTWDLYVMAGNPIVHDD.....AYAGSSNVN
Loktanella_vestfoldensis 462 QAVENSPNPREANR.....YGQILRWFPDAEDHAATTFHWDLYVMAGNPTVHQD.....ANAGSANIN
Agrobacterium_tumefaciens 472 VINA..ANPREKNE.....YGQIVRWYPENDDHADAKFKWDLFCMAGNPTVHKD.....AYAGSANIN
Pseudovibrio_sp. 469 SVNG..PNPRETNK.....YGQIVRWRPHNEDHASDTFDWDLYVMAGNPTVHEG.....PKAGSENVN
Stappia_aggregata 469 PVGG..PNPRETNN.....YGQIVRWFPREDDHGSTTFDWDLYVMAGNPTVHKD.....AYAGSENVN
Silicibacter_lacuscaerulensis 466 SVNG..PNPREANK.....YGQIVRWRPAGDDHAADGFDWDLYVMAGNPTVHSD.....AYAGSENIN
Roseovarius_nubinhibens 462 PVGG..PNPRAENN.....YGQIVRWRPDADDHAALGFDWDLYVMAGNPDVHED.....ARKGSENVH
Citreicella_sp. 462 P..VGGPNPREANN.....YGQIVRWTPENGDHGTIDFTWDLYVMAGNPEVHDN.....AYAGSDNVT
Sagittula_stellata 462 P..VGGPNPREANN.....YGQIVRWTPENGDHGTIDFTWDLYVMAGNPEVHAD.....AYAGSDNVT
Rhodobacterales_bacterium 456 P..VNAANPRETNE.....YGQIVRWAPAGDDHGADTFAWDLYVMAGNPAVYDN.....AYGGSENVT
Jannaschia_sp. 458 EPAPGTPNPRAENR.....YGQIVRWRPDGEDHGADTFTWDLYVMAGNPVVGLG.....LYAGSDNVT
Ruegeria_pomeroyi 458 PAEG..PNPREKNN.....YGQIVRWRPEGGDHSAEGFDWDLFVLAGNPTVHSD.....ANAGSENVN
Labrenzia_alexandrii 461 SVNG..PNPREKNN.....YGQIVRWRPDGGDHTAPGFTWDLYVLAGNPTVHSD.....AYAGSANIN
Roseobacter_denitrificans 473 PVGG..PNPREANK.....YGQIVRWWPSNGDHTADTFKWDLFVMAGNPTVHQD.....MNGGSTNVN
Aromatoleum_aromaticum 473 AANIDDANPRAANS.....WGQIVRWREDGDAAEATTFEWDLFVVAGNPIAFPERAD.....LRSGSEQIT
Thauera_sp. 482 TT..DEANPRAQNR.....YGQIVRWREAGDDAAAMTFEWDLFVLAGNPVAYPDRQD.....LRSGSANVS
Curvibacter 395 GVDA..ANPRVNNT.....QGNIIRWKED.GDFAGTTFAWNHFILAGDPSLERAD.....AKG.NI
Leptothrix_cholodnii 458 ATDA..ANPRANNT.....MGNIIEWLDD.GDYSGTTFTWKHFVLAGDPALERAD.....ARG.NI
Methylibium_petroleiphilum 452 AVDA..ANPRARNT.....MGEVIRWTEV.GGHAATRFLWDHFILAGDPANERAE.....ARG.NV
Candidatus_accumulibacter 454 TADA..ANPRADNA.....FGHIIRWREDAGDAAAGRFSWDIFVLCGDSQHSDAN.....KRG.NI
Variovorax_paradoxus 439 APDA..ANPRAGNL.....FGGILRWREDGGDAASAGFGWDHFALAGDPAQAGSG.....AHYPSD
Citrobacter_koseri 476 PVDA..ANPRANNQ.....FGHIMHWREERADPASAKFTWNILVLAGRTDSDDPK.....AKG.SM
Enterobacter_hormaechei 490 PVDA..ANPRANNQ.....FGHIMHWREERADPASAKFTWNILVLAGRTDSDDPK.....AKG.SM
Klebsiella_pneumoniae 476 PVDA..ANPRANNQ.....FGHIMHWREERADPASAKFTWNILVLAGRTDSDDPK.....AKG.SM
Cronobacter_sakazakii 478 PVDA..ANPRAKNV.....YGHIMHWLEENGDPTALAFAWDILVMGGRTDTDKPE.....AKG.SM
Pasteurella_multocida 470 GIDA..ANPRVKNS.....YGHIIRWQENDQDYLSETFSWDIFALGG....NKSK.....GEN.HV
Lutiella_nitroferrum 473 GPDA..ANPRTDNK.....FGHIIRWNEHGGDAAATAFKWDIFALAGRPDAAKPE.....HKG.NI
Thermomonospora_curvata 470 .....KVNPRKPNP.....YGHIVRWRERGGDNTALTFEWDIFLLAGDPAYDPQV.....KLS
Anabaena_variabilis 547 PVDA..ANPRPDNL.....YGHIIRWREDGQRVTATTFKWDIFIEAGDKTRP.....EANLQGNI
Nostoc_sp. 542 PIDK..SNPRPDNR.....YGHIIRWREDGRSVATTTFKWDIFIEAGDSQRT.....EPNLQGNI
Nitrosomonas_sp. 512 PVDA..ANPRPDND.....FGQIIRWREDKRRVTATTFKWDIFAQCGDKANVKILGSSYDSAGHDGYVGNI
Nitrosospira_multiformis 514 PVDA..ANPRLDNR.....YGHIIRWRENMGKADALDFEWDIFVECGDKLDP.....QPHHRGNI
Alcanivorax_sp. 482 PINA..ANPRHDNH.....HGQIIRWREQNADPTATRFTWDLFLLAGEAQGSRDENG...QPFPDHLTGTI
Marinobacter_aquaeolei 514 STDLDAVNPRAVNS.....NGHIIRWREQDGNHAATAFEWDIFLFGSSADAS.....ADYNVSGLT
Psychrobacter_cryohalolentis 537 DQPLSAPNPRSYNTGGKARGNDNGHIIRWAEARGDHAATSFEWDIYLFAAPSDLS.....AEN.LSQLN
Shewanella_putrefaciens 501 .TEAHPANPRLKNN.....FGHIIRWDEGD...KVTEFTWDIFVFGSPSDGN.....AETNLSGLT
Pseudomonas_aeruginosa 501 VAD..AANPRVKNA.....YGHIIRWRERSRDYAGQRFTWDLFMLAGPGEDSRGPD.....GK....PLN
Azotobacter_vinelandii 491 EAD..AANPRPDNA.....YGHIIRWREESPEHTGTRFNWDLFLLAGPEGDSRGPD.....NQ....PLG
Idiomarina_loihiensis 489 EVQTNAVNPRAENA.....YGHIVRWSEDN.QQSAESFHWDIFLMGG.EAGTRTPN.....AEGKDLTLT
Parvularcula_bermudensis 477 ......GDGVFEGG.....GAGYILNTVDANGSAGDSFIWTIFYLCSDSAGDPD.....
Mesorhizobium_loti 508 .DQVDAANPRAENA.....FGHIIEITETNGDFTATKGKWEVLLKCGDPAVADVG.....ATFSTATT
Paracoccus_denitrificans 478 .EEANAANPRVENA.....FGHIIEILEDGGDLAATQGTWEILLQCGDPSVAEVG.....ATFSTATT
Leptospira_interrogans 384 .....NDRHGNL.....FGQILRIREKSGDHAGLEFDFEVFIAGG.....


β5a
β5b
β5c
Pseudomonas_fluorescens 518 PQNMFNSPDGLGFDKAGRLWILTDGDSS..NAGD.....FAGMGNNQMLCADP
Pseudomonas_syringae 524 AQNMFNSPDGLSFDRSGRLWIQTDGDSS..NKGD.....FAGMGNNQMLCADP
Serratia_proteamaculans 516 ADNTFNSPDGLGFDKSGRLWILTDGKYS..NKGD.....YAGQGNNQMLVGDP
Dickeya_zeae 521 VENSFNSPDGLGFDHDGRLWILTDGNYS..NKGD.....FVGQGNNQMLAGDP
Campylobacter_jejuni 480 SENKFNSPDGLKFDRDGRLWIQTDGSYS..NKDE.....YESMGNNCMLAANP
Vibrio_cholerae 509 ADNMFNSPDGIGFDTAGRLWIQTDGNYS..NQGD.....FAGQGNNQMLCGDP
Edwardsiella_ictaluri 509 AENMFNSPDGIGFDKAGRLWIQTDGNYS..NQGD.....FAGQGNNQMLCGDP
Vibrio_fischeri 510 KDNMFNSPDGIGFDKAGRLWIQTDGNYS..NKGD.....FAGQGNNQMLCGDP
Grimontia_hollisae 513 KDNMFNSPDGLGFDGAGRLWILTDGNYS..NEGN.....FEGMGNNQMLCADP
Photobacterium_profundum 525 NENMFNSPDGLGFDAAGRLWIQTDGKYS..NKGD.....YAGMGNNSMFCGDP
Marinomonas_sp. 532 SHNMFNSPDGIGFDDAGRLWIQTDGNYK..NKGD.....FLGQGNNQMLCADP
Reinekea_sp. 510 ADNMFNSPDGLMFDNDGRLWIQTDGNYS..NEGI.....FAGQGNNQMLVADP
Octadecabacter_antarcticus 517 EGNLFNSPDGMMFDSTGLLWIQTDGDDS..NEGD.....FEGMGNNQMLAGDP
Loktanella_vestfoldensis 520 EGNLFNSPDGMMFDSSGLLWIQTDGDDS..NEEG.....FAGMGNNQMLAGDP
Agrobacterium_tumefaciens 528 EGNMFNSPDGMMFDSTGLLWIQTDGEDS..NEGN.....FAGQGNNQMLAGDP
Pseudovibrio_sp. 525 AGNMFNSPDGMGFDESGLVWIQTDGSDS..NEGD.....FAGMGNNQMLVGDP
Stappia_aggregata 525 EGNLFNSPDGMAFDSSGIVWIQTDGDDS..NEGD.....FAGMGNNQMLAGDP
Silicibacter_lacuscaerulensis 522 AGNLFNSPDGMMIDSTGMVWIQTDGNDD..NEGD.....FAGHGNNQMLVGDP
Roseovarius_nubinhibens 518 TGNLFNAPDGMMFDSTGLLWIQTDGKDN..NEGD.....FAGMGNNQMLVGDP
Citreicella_sp. 518 TGNMFNSPDGMAFDTKGKLWIQTDGDDS..NEGD.....FAGMGNNQMLVGDT
Sagittula_stellata 518 TGNMFNSPDGMAFDTKGRLWIQTDGDDS..NEGD.....FAGMGNNQMLVGDT
Rhodobacterales_bacterium 512 DGNMFNSPDGMAFDSKGMLWIQTDGDDS..NEGD.....FAGQGNNQMLVGNT
Jannaschia_sp. 516 EGNMFNSPDGMQVDSTGLIWIQTDGDDS..NEGA.....FTGMGNNQMLIGHP
Ruegeria_pomeroyi 514 AGNMFNSPDGLAFDSNGLLWIQTDGKYS..NEGD.....FAGQGNNQMLAGDP
Labrenzia_alexandrii 517 EENMFNSPDGLAFDSNGLVWIQTDGNYK..NEGD.....FAGQGNNQMLAGDP
Roseobacter_denitrificans 529 EGNMFNSPDGLSFDNKGLLWIQTDGNYS..NEED.....FAGQGNNQMLIGDP
Aromatoleum_aromaticum 534 ADNTFNSPDGLGFDSEGRLWIQTDGNFS..NTGN.....YLGQGNNQMLCADP
Thauera_sp. 541 ADNTFNSPDGIGFDGAGRLWIQTDGNFS..NSGD.....YAGQGNNQMLVADP
Curvibacter 447 KGDMFSCPDGLWVDGRGVLWIQCDMSTSAMGKGD.....LKNFGNNMMLAADV
Leptothrix_cholodnii 510 KGDAFSCPDGLWVDGRGVMWIQTDMSTSAMGKGD.....LVNMGNNQMLAADP
Methylibium_petroleiphilum 504 KGDPFGSPDGLWFDARGVLWIATDASPTALGQGD.....YARLGNNMLLAADP
Candidatus_accumulibacter 507 RGDSFGSPDGLFFDRAGRLWIQTDVPGRLMNHGD.....YQRIGNNQMLVADV
Variovorax_paradoxus 493 DADVFGNPDGLHFDGGGLLWIQTDMSGQLIGKPP.....YASLGNNQMLCADP
Citrobacter_koseri 529 QGAEFGSPDGLSFDHRGVLWIQTDVSSSTINKKA.....YEGMGNNQMIATLP
Enterobacter_hormaechei 543 QGAEFGSPDGLSFDHRGVLWIQTDVSSSTINKKA.....YEGMGNNQMIATLP
Klebsiella_pneumoniae 529 QGAEFGSPDGLSFDHRGVLWIQTDVSSSTINKKA.....YEGMGNNQMIATLP
Cronobacter_sakazakii 531 KGAEFGSADGLSFDHRGVLWIQTDVSSSTINRKD.....YEGMGNNQMVATIP
Pasteurella_multocida 519 NGDDFGSPDGLRFDNHGILWVQTDVSSSTLNKKA.....YEGMGNNQMLAVIP
Lutiella_nitroferrum 526 KGDYFGSQDGLKFDYNGVLWVETDVSTSTVNMKE.....YKGMGNNQMLAVVP
Thermomonospora_curvata 518 PEDVFGSPDGLGFDRDGRLWIQTDVSNSTQNRPDR.....GHDNIGNNQMLAADP
Anabaena_variabilis 600 KGDDLGAPDGLWFDDFGRLWIQTDQAGD..GRG.....DWQKIGGNTMSCADP
Nostoc_sp. 595 NGDDLSSPDGLWFDDFGRLWIQTDQAGD..GLGSTVNGVPELINIGSNGMFCADP
Nitrosomonas_sp. 576 NGDDYGAPDGLWFDKEGRLWIQTDQAGD..ASG.....DWVNIGGNVMMCADS
Nitrosospira_multiformis 567 NGDDYGAPDGLWFDQDGRLWIQTDQAGD..ATG.....DWANIGGNVMLCANP
Alcanivorax_sp. 543 DGDLFSSPDGLAFDASGRLWIETDYDDR....DP.....LMKNMGCNQLLCADP
Marinobacter_aquaeolei 570 PDNEFSSPDGLFIDPRGVMWIQTDDGSSGIRSK.....TNNQMLVAIPGRVGDGESVTVTTSDNNGQ
Psychrobacter_cryohalolentis 600 DNNDFSSPDGLYFDPRGVLWIQTDDG..AYTDT.....SSCMLLAALPGNISDG...TLTTTS.AGQ
Shewanella_putrefaciens 553 DLNQFASPDGLAFDQRGILWVQTDNGAKEVAEE.....TNDQMLAIVPSTLVDSQDKQLTVTA....
Pseudomonas_aeruginosa 555 QDNILASPDGLWFDPEGRLWIQTDMSGSQLSSGP.....FGNNQMLVADP
Azotobacter_vinelandii 545 ADNILASPDGLWFDPQGLLWIQTDMSGSQHSDGP.....FGNNQMLVANP
Idiomarina_loihiensis 547 SNNQFACPDGLWFDHRGTLWIQTDMSGELLSEGS.....FGNNGMYAANP
Parvularcula_bermudensis 520 ..LAFDDCDAAYADPDGRLFIGTDGGQLYGLQDQ.....VVVFDTTAEDP
Mesorhizobium_loti 565 ANGWFGMPDNCAVDSAGRLWVATDGQGP...KATGR.....TDGLWAVDTEGA
Paracoccus_denitrificans 535 ANGWFGMPDNCAIDSDGRLWVSTDGNSP...KDTGR.....TDGLWAVDVEGA
Leptospira_interrogans 419 ..ARFAAPDNLAFDKEGNLWVVTDISAKNLNRSVYK.....KFGNNGLFVIRT


β5d
β6a
β6b
Pseudomonas_fluorescens 564
Pseudomonas_syringae 570
Serratia_proteamaculans 562
Dickeya_zeae 567
Campylobacter_jejuni 526
Vibrio_cholerae 555
Edwardsiella_ictaluri 555
Vibrio_fischeri 556
Grimontia_hollisae 559
Photobacterium_profundum 571
Marinomonas_sp. 578
Reinekea_sp. 556
Octadecabacter_antarcticus 563
Loktanella_vestfoldensis 566
Agrobacterium_tumefaciens 574
Pseudovibrio_sp. 571
Stappia_aggregata 571
Silicibacter_lacuscaerulensis 568
Roseovarius_nubinhibens 564
Citreicella_sp. 564
Sagittula_stellata 564
Rhodobacterales_bacterium 558
Jannaschia_sp. 562
Ruegeria_pomeroyi 560
Labrenzia_alexandrii 563
Roseobacter_denitrificans 575
Aromatoleum_aromaticum 580
Thauera_sp. 587
Curvibacter 495
Leptothrix_cholodnii 558
Methylibium_petroleiphilum 552
Candidatus_accumulibacter 555
Variovorax_paradoxus 541
Citrobacter_koseri 577
Enterobacter_hormaechei 591
Klebsiella_pneumoniae 577
Cronobacter_sakazakii 579
Pasteurella_multocida 567
Lutiella_nitroferrum 574
Thermomonospora_curvata 568
Anabaena_variabilis 646
Nostoc_sp. 648
Nitrosomonas_sp. 622
Nitrosospira_multiformis 613
Alcanivorax_sp. 588
Marinobacter_aquaeolei 632
Psychrobacter_cryohalolentis 656
Shewanella_putrefaciens 611
Pseudomonas_aeruginosa 600
Azotobacter_vinelandii 590
Idiomarina_loihiensis 592
Parvularcula_bermudensis 563
Mesorhizobium_loti 610
Paracoccus_denitrificans 580
Leptospira_interrogans 465

β6c

```
Pseudomonas_fluorescens                  602  .NGGSTFPEHLP.........NGKPRSSVMAITREDGGIVGA....
Pseudomonas_syringae                     608  .NGGSTFPEHLP.........NGKPRSSVMVITREDGGVIGA....
Serratia_proteamaculans                  600  .EGDSHFPNNSP.............RPRSSVLMITREDGGVIGA....
Dickeya_zeae                             605  .KGDSHFPNNSV.............RPRSSVVMITREDGGVIGA....
Campylobacter_jejuni                     565  .LKGSTFPYGKT.............PRSSVMMIRKLDGGVIGS....
Vibrio_cholerae                          594  .AAPSHFPYGG..........TSKPRSTIMMITREDGGVIGA....
Edwardsiella_ictaluri                    594  .GAPSHFPDGG..........NSLPRSTIMMITREDGGVIGA....
Vibrio_fischeri                          592  ...SGHFPQGG..........NSKPRSTIVMITKDNGGVIGS....
Grimontia_hollisae                       598  .LAPSHFPDGG..........DSIPRSTVMRITRKDGGVIGT....
Photobacterium_profundum                 610  .LEPSHWPDGG..........NSTPRSSIVMITKKDGSVIGS....
Marinomonas_sp.                          614  ...DAHFPEGG..........NSKPRSTVVVITRKDGGVIGA....
Reinekea_sp.                             592  ...SNAWPNAEK..........DGVPRSSVIAIRREDGGIVGA....
Octadecabacter_antarcticus               600  ....APFPDGE..........GKLARSSIIAVKRDDNAPVG.....
Loktanella_vestfoldensis                 603  ....APFPDGD..........GKLARSTVIAVKKADGAVIG.....
Agrobacterium_tumefaciens                611  ....APFPDGE..........GKLPRSAVIAIKRDDNAKIG.....
Pseudovibrio_sp.                         608  ....GNWPDG...........GKLPRSSVIVIKRDDNGVIG.....
Stappia_aggregata                        608  ....SNWPDG...........GKLTRSSVIAIKRSDNGLIG.....
Silicibacter_lacuscaerulensis            605  ....KNWPDG...........G.TPRSAIIAIKRNDNEVLG.....
Roseovarius_nubinhibens                  601  ....APFPDGE..........GRLPRSAVIAVWRDDGAMVG.....
Citreicella_sp.                          601  ....GSWPDG...........EGKPRSSVIAVWREDGDTIG.....
Sagittula_stellata                       601  ....GSWPDG...........EGKPRSSVISVWREDGETVG.....
Rhodobacterales_bacterium                595  ....GSWPDG...........NGKPRSSIIQVWRTDGAPVG.....
Jannaschia_sp.                           599  ....GTWPDG...........DSLPRSSVIAVWRDDGAAIG.....
Ruegeria_pomeroyi                        599  ..GDSHWPDGG..........DSVPRSAIIAVTREDGGVIG.....
Labrenzia_alexandrii                     602  ..GNSHWPAGG..........ASVPRSAIIAVTREDGGLVG.....
Roseobacter_denitrificans                614  ..GDSNWPDGG..........DKTPRSSIIAVRRDDGAVIG.....
Aromatoleum_aromaticum                   641  PLAFSQWPEAA..........GGRPRSATVVITKEDGGVIGT.....
Thauera_sp.                              650  PLAFSQWPEAG..........GGRPRSATVVITKDDGGVVGS.....
Curvibacter                              543  PRAVSNWPEKK..........ANGRPRSATVVIRKADGGVIGT.....
Leptothrix_cholodnii                     606  PGKISNWPDHK..........QGGRPRSATVVIRKLDGGVIGS.....
Methylibium_petroleiphilum               600  ARKRSTWPG............GGGRPRSATVAIRRIDGGAIGT.....
Candidatus_accumulibacter                603  PLAVSGWPANQFA......EVTGGRPRSATLVIRRRDGGVVGT.....
Variovorax_paradoxus                     589  ....SAWPDGTVP.........GSARPRSATLAIRRRDGAIVGT.....
Citrobacter_koseri                       625  PRAISNWPDSR..........SDGRPRSSTVVITKSNGGIIGT.....
Enterobacter_hormaechei                  639  PRAISNWPDSR..........SDGRPRSSTVVITKSNGGIIGT.....
Klebsiella_pneumoniae                    625  PRAISNWPDSR..........SDGRPRSSTVVITKSNGGIIGT.....
Cronobacter_sakazakii                    627  PTAVSSWPDGD..........KAGRPRSSTVVIVKADGGMIGT.....
Pasteurella_multocida                    615  VTAISTWPDRQ..........GKTRPRSATVVIQKEDGGVISS.....
Lutiella_nitroferrum                     622  PLAISSWPEGP..........KAGRPRAATVVIRKLDGGVIGS.....
Thermomonospora_curvata                  619  PNAVSSWPDHDP.........AGRPRSATVVIRKKDGGVIGT.....
Anabaena_variabilis                      691  .TQTSNWPYSQGYGPSG.......RPRSSTVVITRNDGGVIGGL...
Nostoc_sp.                               693  .TQNSHWPYSQGYGLPG.......RPRSATVVITRNDGGVIGGL...
Nitrosomonas_sp.                         666  .TDNSTWPDNGYNGATTYTGTSVSKPRSSVIVITKDDGGVIGS....
Nitrosospira_multiformis                 657  .TQNSTWPDSGQNGPTT.AGGSPSKPRSAVVVITKDDGGVIGT....
Alcanivorax_sp.                          625  ....ISYPASDG..........TTRPRSTTVVISRDDGGVVGT....
Marinobacter_aquaeolei                   698  SRDATAFGE.A..........GNRPRSATIVIYREDGGELAL.....
Psychrobacter_cryohalolentis             707  AG.....-.............GSTPRSATVMITREDGDVILAESFV
Shewanella_putrefaciens                  654  ANAAEQTPAGT..........TLRPRSATVLIRKQDGKQVGV....
Pseudomonas_aeruginosa                   642  DNLLSTWPDGP..........GRRPRSATVVITREDGRRLL.....
Azotobacter_vinelandii                   632  DNLLSTWPDGP..........GKRPRSSTVVITREGWRRVL.....
Idiomarina_loihiensis                    634  GKFSSSWPAGN..........GARPRSATVIITREDGGVVGA....
Parvularcula_bermudensis                 599  NPALTNFPLPYDG.........VTIPRDATLIIEKKGPGKVGS....
Mesorhizobium_loti                       662  EDLSTRWPDFKP..........EMPVRPSVVAITKQGGGKIAV....
Paracoccus_denitrificans                 632  EDLSTRWPDFRD..........DMPVRPAVVAITKQGGGRIG.....
Leptospira_interrogans                   509  .KPTSHWPHGGE..........SVPRSGVVAIYLK.....-......
```

Appendix 1: Multiple sequence alignment of PhoX from 50 different species

Identical sequences are in red boxes with white character. Similar residues, having a 80% threshold determined with the RISLER substitution matrix are in red characters. The secondary structure of PhoX with β strands labelled is above the sequence alignments