Structural biochemistry of the Tol-Pal inner membrane protein assembly

By
Cara Press
Linacre College



A thesis presented for the degree of Doctor of Philosophy at the University of Oxford

Michaelmas Term 2022

Supervised by
Professor Colin Kleanthous
Professor Simon Newstead

Declaration

I confirm that the work presented in this thesis, submitted for the degree of Doctor of Philosophy, is my own original work. Work that is not my own is referenced. This thesis has not been previously submitted at this or any other university.

Cara Press

December 2022

Word count including abstract, tables, figure legends, appendices, and references: 44,520.

Acknowledgements

I am incredibly grateful for the many wonderful people in my life who have supported me throughout this endeavour. I want to thank Professor Colin Kleanthous for the opportunity to work in his lab and for his supervision and encouragement throughout this project. I also want to thank my secondary supervisor Professor Simon Newstead for his guidance. I am grateful to the Wellcome Trust and the Structural Biology Programme organisers for funding my research and facilitating my DPhil training. I am thankful for the opportunity I had to work with an amazing group of people in the Kleanthous lab. I can't thank Melissa Webby enough, for her scientific contribution and mentorship throughout my project, and for her friendship. I'd like to thank Jo Szczepaniak for her microscopy contribution to my thesis, for her Tol-Pal expertise, and for her constant support. I also want to thank Renata Kaminska for her assistance in cloning and protein purification, Nick Housden for his advice, proof-reading and help with experimental work, and Danny Williams-Jones for his editing and scientific discussion. Thank you to all members of the Kleanthous lab past and present for being wonderful colleagues to work with: Patrick, Vivian, Shreya, Jonathan, Iva, Gideon, Nathalie, Ruth, Sandip, and all others that have passed through the lab during my time.

I was fortunate to work with some great collaborators during my project and would like to thank them for their contribution to my work: Professor Edward Lowe for his help and advice in tackling a challenging crystallography project. Professor Christina Redfield for her work and guidance on the NMR project. Dr Tarick El-Baba for his enthusiasm and support in the native mass spectrometry work. Dr Dheeraj Praakesh for conducting the molecular dynamic simulations, and Professor Syma Khalid for facilitating this collaboration. Professor David Staunton for his guidance on the biophysical experiments conducted. Members of the Higgins lab for their guidance on crystallography and for allowing me to steal their reagents. The Lea Group for providing the pT12 plasmid used in my work.

Outside of the lab, I am grateful to OUWPC for providing an excellent escape from work, and to the members that have been a source of friendship and support throughout my

DPhil. I'd like to thank my Linacre friends, and some honorary Linacrites, for so many memories and for making my Oxford experience. I'd also like to thank my 'cirque-le' of friends from undergrad for their support and friendship. Particularly, Alice Burleigh and Hannah Stephens for always being there. I also want to thank Mari Johnson and Hannah Sharpe for being wonderful friends and pandemic housemates. I would like to thank my family, especially my brother Kieran, for their support. Finally, I am incredibly grateful to my parents, Breda and Steve, for their constant support, encouragement, and for believing in me even when I didn't.

Abstract

Understanding fundamental processes within bacteria is crucial to building new approaches to tackle the problem of antimicrobial resistance. In particular, the cell envelope is a key focus as it is a target for current and future antibiotics. The Gram-negative cell envelope typically comprises of an outer membrane (OM), cell wall and inner membrane (IM). The OM acts as a barrier to the outside world and can confer intrinsic resistance to antibiotics. However, the OM is non-energised and so active processes that are carried out at the OM must harness an energy source from elsewhere. One system that carries out such a process is the Tol-Pal system that works to actively maintain the integrity of the OM. Tol-Pal utilises the proton motive force (PMF) across the inner membrane to carry out this role.

There are five core components of the Tol-Pal system that span the cell envelope. Pal is a lipoprotein anchored into the OM that binds the peptidoglycan (PG) cell wall. TolB is a periplasmic protein that binds Pal with high affinity. When in complex with TolB, Pal can freely diffuse within the plane of the OM. The Tol-Pal system works to modulate the binding state of Pal between PG and TolB. During cell division, Pal is recruited to the septum to ensure the OM is tethered to the PG as the cell divides. The IM components TolQ-TolR are responsible for transducing the PMF to the IM effector protein TolA. It is thought that upon interaction and protonation, TolA extends across the periplasm to bind the TolB-Pal complex at the OM. TolA then relaxes back to its ground state, dissociating TolB from Pal, allowing for the deposition of Pal onto PG.

The work described in this study aimed to better our understanding of the IM components of Tol-Pal by integrating a range of structural, biochemical, computational, and biophysical techniques. The first focus was to elucidate the role of the second domain of TolA in extending across the periplasm to bind TolB-Pal. A combination of Far UV circular dichroism, size exclusion chromatography multi-angle light scattering, and nuclear magnetic resonance determined that the domain is helical, monomeric and has flexible N- and C-termini.

Analysis of the AlphaFold2 model of the domain reveals an extended helical hairpin structure in this domain that is the length of the periplasm. Molecular dynamic simulations demonstrated the helices of this domain can interact, and provided a basis for a new mechanism by which TolA reaches the OM.

The assembly of the IM stator component of Tol-Pal, TolQ-TolR, was also investigated. First, extensive optimisation was carried out to identify conditions where TolQ-TolR could be purified. Following this, the stoichiometry of the complex was investigated through biochemical methods, native mass spectrometry, and cryo-EM. In addition, the assembly of TolQ-TolR was modelled through computational means and the model evaluated in relation to experimental data. The final project of the study investigated the IM interactions of the TolQ-TolR-TolA assembly through a photoactivatable crosslinking approach. The TolAHis22^{BPA} mutation was demonstrated to be functional, but crosslinks between TolA and TolQ at this site could not be detected. Modelling of AlphaFold2 TolA into the TolQ-TolR model showed TolA binds to a groove between two transmembrane helices of TolQ. In combination, the study furthers our understanding of how the Tol-Pal IM components assemble, and suggests a new 'helical hinge' mechanism for how TolA operates, allowing it to bridge the periplasm and bind to TolB-Pal at the OM.

Table of Contents

Abstract **5**

1 Introduction **17**

1.1 General introduction **18**

1.2 The Gram-negative cell envelope **18**

1.2.1 The outer membrane 18

1.2.2 The periplasm and cell wall 21

1.2.3 The inner membrane 22

1.3 Proton pathways and the proton motive force **22**

1.3.1 Definition of the proton motive force 22

1.3.2 Generating a proton gradient across the IM 23

1.3.3 Proton wires 23

1.4 Trans-envelope machinery in Gram-negative bacteria **24**

1.4.1 Powering motility 25

1.4.2 Powering nutrient import 27

1.5 The Tol-Pal system **31**

1.5.1 The *tol-pal* operon 31

1.5.2 The proteins of the Tol-Pal system 33

1.5.3 Functions of the Tol-Pal system 41

1.6 Aims and Objectives **46**

2 Materials and Methods **48**

2.1 Microbiology **49**

2.1.1 Strains, media, and growth conditions 49

2.1.2 Plasmids 50

2.2 Molecular Biology **52**

2.2.1 Polymerase Chain Reaction (PCR) 52

2.2.2 Restriction digests and ligation 55

2.2.3 InFusion cloning 55

2.2.4 Preparation of chemically competent cells 55

2.2.5 Bacterial heat shock transformation 56

2.2.6 DNA gel electrophoresis 56

2.2.7 Plasmid preparation and sequencing of DNA 56

2.3 Protein Expression and Purification **57**

2.3.1 Expression and purification of TolA-II and TolA-II+III from *Escherichia coli, Proteus mirabilis, Yersinia enterocolitica* and *Salmonella* Typhimurium 57
2.3.2 Purification of truncated *E. coli* TolA-II constructs 58
2.3.3 ^{15}N labelling of TolA constructs 59
2.3.4 Purification of TolA with BPA mutations 59
2.3.5 Expression and purification of TolQ-TolR 61
2.3.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 62
2.3.7 Blue Native-PAGE 63
2.3.8 LPS staining 63
2.3.9 Protein quantification 64
2.3.10 Peptide mass fingerprinting 65
2.4 Biophysical Techniques 65
2.4.1 Circular Dichroism 65
2.4.2 Nanoscale Differential Scanning Fluorimetry 66
2.4.3 Size Exclusion Chromatography Multi Angle Light Scattering (SEC-MALS) 66
2.5 Crystallography 67
2.5.1 Crystallisation 67
2.5.2 Crystal screening 68
2.6 Nuclear Magnetic Resonance Spectroscopy 68
2.7 Electron Microscopy 69
2.8 Computational Methods 69
2.8.1 Model Building 69
2.8.2 Model Analysis 70
2.8.3 Molecular Dynamics Simulations 71
2.9 Native Mass Spectrometry 71
2.10 Assays to investigate Tol-Pal function 72
2.10.1 Colicin toxicity assay 72
2.10.2 Detergent sensitivity assay 73
2.10.3 Western blot 74
2.10.4 Brightfield microscopy 75
3 Characterisation of domain II of TolA 77
3.1 Introduction 78
3.2 Results 79
3.2.1 Domain II of TolA is helical in nature 79
3.2.2 The N- and C-termini of domain II are flexible 83
3.2.3 *E. coli* TolA domain II crystallises at high concentrations 87

3.2.4 Structural predictions of TolA domain II reveal an extended helical hairpin 89
3.2.5 Molecular dynamic simulations of AlphaFold2 *E. coli* TolA reveal interactions between the two helices in domain II 95
3.3 Discussion 98
4 Investigation into the assembly of the TolQ-TolR stator 102
4.1 Introduction 103
4.2 Results 104
4.2.1 The TolQ-TolR complex can be purified using the detergent Cymal-6 104
4.2.2 Biochemical approaches to validate the TolQ-TolR complex 107
4.2.3 The TolQ pentamer is observable using cryo-EM 110
4.2.4 Modelling TolQ-TolR highlights a proton channel conserved with ExbB-ExbD and MotA-MotB 113
4.3 Discussion 118
5 Using photoactivatable crosslinking to probe the interaction between TolQ-TolR and TolA 121
5.1 Introduction 122
5.2 Results 124
5.2.1 The TolA His22^{BPA} and Ile15^{BPA} mutants are functional 124
5.2.2 No crosslinks between TolQ and TolAHis22^{BPA} were detected 132
5.2.3 Modelling the TolQ-TolR-TolA assembly shows TolA TMH docks into groove composed of TolQ TMHs 135
5.3 Discussion 137
6 General Discussion 141
6.1 Updating the TolQ-TolR-TolA interaction mechanism 142
6.1.1 The TolQ-TolR rotary mechanism 142
6.1.2 The TolA length paradox 145
6.1.3 The 'helical hinge' mechanism underpinning the role of TolQ-TolR-TolA in the Tol-Pal system 146
6.2 Future work 148
6.3 Conclusions 150
7 Appendices 152
7.1 Appendix 1. Protein sequences 152
7.2 Appendix 2. Supplementary Figures 157

Bibliography **164**

List of Figures

Figure 1.1 The tripartite Gram-negative cell envelope consists of an outer membrane (OM), peptidoglycan layer (PG) and inner membrane (IM). 19

Figure 1.2 The Mot, Ton and Tol systems use the PMF to drive force transduction at the OM 24

Figure 1.3 The MotA-MotB stator. 26

Figure 1.4 Structures of the Ton system. 28

Figure 1.5 The Tol-Pal system. 32

Figure 1.6 TolR's periplasmic domain forms two structural conformations. 34

Figure 1.7 TolA-III binds to the N-termini of TolB through β-augmentation. 36

Figure 1.8 TolB and Pal bind to one another, and competitively Pal binds to PG. 38

Figure 1.9 The structures of CpoB and YbgC. 40

Figure 1.10 Mobilisation and capture mechanism of the Tol-Pal system 43

Figure 3.1 GFP-tagged TolA can diffuse freely in the IM and exhibits restricted diffusion upon addition of translocation incompetent colicin E9. 78

Figure 3.2 The design and purification of TolA-II and TolA-II+III constructs from four bacterial species. 80

Figure 3.3 Biophysical characterisation confirmed the TolA-II and TolA-II+III constructs are helical and monomeric. 83

Figure 3.4 NMR analysis reveals TolA-II has unstructured N and C-termini. 85

Figure 3.5 *E. coli* TolA-II structured domain begins at Ser69 and ends at Glu293. 86

Figure 3.6 *E. coli* $TolA_{72-298}$ crystallises at high concentrations and when crosslinked with 0.1% glutaraldehyde. 88

Figure 3.7 AlphaFold2 predicts *E. coli* TolA domain II contains a long helical hairpin 89

Figure 3.8 Alignment of primary sequences of TolA-II from 9 different bacterial species. 92

Figure 3.9 Comparison of AlphaFold2 predicted structures of TolA from eight γ-proteobacteria. 93

Figure 3.10 Comparison of electrostatic surfaces of TolA from nine γ-proteobacteria. 94
Figure 3.11 The AlphaFold2 predicted structures of TolA from representative species from different classes of proteobacteria. 96
Figure 3.12 Molecular dynamics simulations of the AlphaFold2 TolA model demonstrate an intramolecular interaction between α2 and α3, and interactions with PG and the IM. 97
Figure 4.1 Optimisation of expression of the TolQ-TolR 105
Figure 4.2 The detergent Cymal-6 best solubilised TolQ. 106
Figure 4.3 Immunoblotting and cleavage with TEV protease identified the presence of both TolQ and TolR. 108
Figure 4.4 Techniques unable to validate the purification of the TolQ-TolR complex. 109
Figure 4.5 The presence of TolQ oligomers was confirmed by native MS 111
Figure 4.6 Preliminary cryo-EM shows TolQ is a pentamer. 112
Figure 4.7 TolQ-TolR complex model resembles ExbB-ExbD and MotA-MotB. 114
Figure 4.8 Monomers of TolQ and TolR predicted by AlphaFold2 coloured by pLDDT score. 115
Figure 4.9 Docking of TolQ-TolR model into 8 Å EM map demonstrates the pentamer of the TolQ model fits in the pentameric map. 116
Figure 4.10 The TolQ-TolR model has conserved protonatable residues consistent with a proton channel. 117
Figure 5.1 The BPA crosslinking method. 123
Figure 5.2 The positions of sites chosen for BPA crosslinking in the TMH of TolA visualised on the AlphaFold2 prediction of TolA-I. 125
Figure 5.3 Choosing a concentration of IPTG to use in detergent sensitivity assays. 126
Figure 5.4 The TolAHis22^{BPA} mutant is partially functional, and the TolAIle15^{BPA} and TolASer32^{BPA} mutants are non-functional. 127
Figure 5.5 Full length TolA is expressed in the TolAHis22^{BPA} and TolAIle15^{BPA} mutants in the presence of BPA. 130

Figure 5.6 *ΔtolA::kan* expressing TolAHis22^{BPA} has the same colony morphology as *ΔtolA::kan* expressing TolA and the synthetase pair on pVol. 131
Figure 5.7 No crosslinks could be detected for TolAHis22^{BPA}. 134
Figure 5.8 Docking TolQ-TolR and TolA models shows interaction of TolA TMH within TolQ's TMH groove 136
Figure 5.9 TolA His22 fits into a pocket formed by TolQ TMH1 and TMH2 138
Figure 6.1 A model for TolQ-TolR proton transduction 143
Figure 6.2 Proline 37 of *E. coli* TolR's position in different structural models implies piston action of TolR TPHs. 144
Figure 6.3 The 'helical hinge' mechanism. 147

List of Tables

Table 2.1 *E. coli* strains used in this study 49

Table 2.2 Plasmids used in this study 50

Table 2.3 Typical PCR reaction composition 53

Table 2.4 Typical PCR reaction incubations 53

Table 2.5 Primers 53

Table 2.6 Typical InFusion cloning reaction 55

Table 2.7 Molecular weights and extinction coefficients of proteins purified in this study determined by ExPasy ProtParam 65

Table 2.8 Protein concentrations used for CD and nanoDSF 66

Table 2.9 TolA sequences and AlphaFold2 models analysed. 70

Table 2.10 Strains and plasmids used in colicin killing assays 72

Table 2.11 Plate contents for detergent sensitivity assay to determine concentration of IPTG to use in future assays. 73

Table 2.12 Strains and plasmids used for detergent sensitivity assay 74

Table 2.13 Plate contents for detergent sensitivity assays. Dark blue filled box indicates the component is included. 74

Table 3.1 Comparison of the constructs of TolA-II+III and TolA-II from four different bacterial species designed in this study. 81

Table 3.2 Truncated constructs of TolA-II. 85

Table 3.3 Comparison of AlphaFold2 predicted structures of TolA from different bacterial species. 91

Supplementary Table 7.1 Commercial crystallographic screens 157

List of Abbreviations

ABC	dimethylamine-borane complex
BPA	p-benzoyl-L-phenylalanine
CD	circular dichroism
CGMD	coarse grained molecular dynamics simulations
cryo-EM	cryo electron microscopy
DEER	double electron-electron resonance
DNA	deoxyribonucleic acid
DSF	differential scanning fluorimetry
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionisation mass spectrometry
ETC	electron transport chain
FRAP	fluorescence recovery after photobleaching
GTase	glycosyltransferase
HetNOE	heteronuclear Overhauser effect
HSQC	heteronuclear single quantum coherence spectroscopy
IM	inner membrane
IPTG	isopropyl β- d-1-thiogalactopyranoside
IUTD	intrinsically unstructured translocation domain
Kdo	3-deoxy-D-manno-oct-ulosonic acid
LB	Luria Bertani/lysogeny broth
Lpp	Braun's lipoprotein
LPS	lipopolysaccharide
mDAP	meso-diaminopimelic acid
MS	mass spectrometry
NAG	N-acetylglucosamine
NAM	N-acetylmuramic acid
NEB	New England Biolaboratories
NMR	nuclear magnetic resonance
OD	optical density
OM	outer membrane
OMP	outer membrane protein
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PG	peptidoglycan
pLDDT	local distance difference test
PMF	proton motive force
PVDF	polyvinylidene difluoride
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography

SEC-MALS	size exclusion chromatography multi-angle light scattering
SPT	single particle tracking
TBDT	TonB-dependent transporter
TBE	tris borate EDTA
TBST	tris-buffered saline Tween-20
TCEP	tris(2-carboxyethyl)phosphine
TMH	transmembrane helix
TOCSY	total correlation spectroscopy
TPase	transpeptidase
TPH	trans-pore helix
TPR	tetratricopeptide
TROSY	transverse relaxation optimised spectroscopy

1 Introduction

1.1 General introduction

Antibiotic resistant Gram-negative bacteria pose one of the greatest challenges to modern medicine at present. As the threat of a post-antibiotic age approaches, alternatives to the traditional antibiotics we use to treat bacterial infections must be explored. Furthermore, to develop and apply new treatments to target bacteria, the fundamental processes within bacteria must be better understood. The work described in this study on the Tol-Pal system is an example of how furthering our understanding of its basic function can help with the development of traditional and new techniques to target bacteria. The Tol-Pal proteins themselves are a potential novel antibiotic target as they are required for virulence and essential for survival in several pathogens. In addition, the development of so-called 'protein antibiotics' relies upon import through the Tol-Pal system. It is hoped that advancing knowledge of how Tol-Pal works will open doors to new therapeutic potential in Gram-negative bacteria.

1.2 The Gram-negative cell envelope

The cell envelope of bacteria defines the structures found outside the cytoplasm that encase the bacterium. They are classified by their retention of the crystal violet stain used in Gram staining. Gram-negative bacteria have a cell envelope of at least three layers: an asymmetric outer membrane; an aqueous periplasm containing the cell wall; and an inner membrane (Figure 1.1). The Tol-Pal system's role is to maintain the outer membrane of this cell envelope (Szczepaniak *et al.*, 2020), so understanding the characteristics of the molecular components within each layer is fundamental to investigating the mechanism of the system.

1.2.1 The outer membrane

The outer membrane (OM) of the Gram-negative cell envelope is the first point of contact between the bacterium and the environment. It contains an asymmetric bilayer with a phospholipid inner leaflet and outer leaflet made of lipopolysaccharide (LPS) (Nikaido, 2003). LPS is a glycolipid with a structure comprised of three sections: lipid A, core polysaccharide,

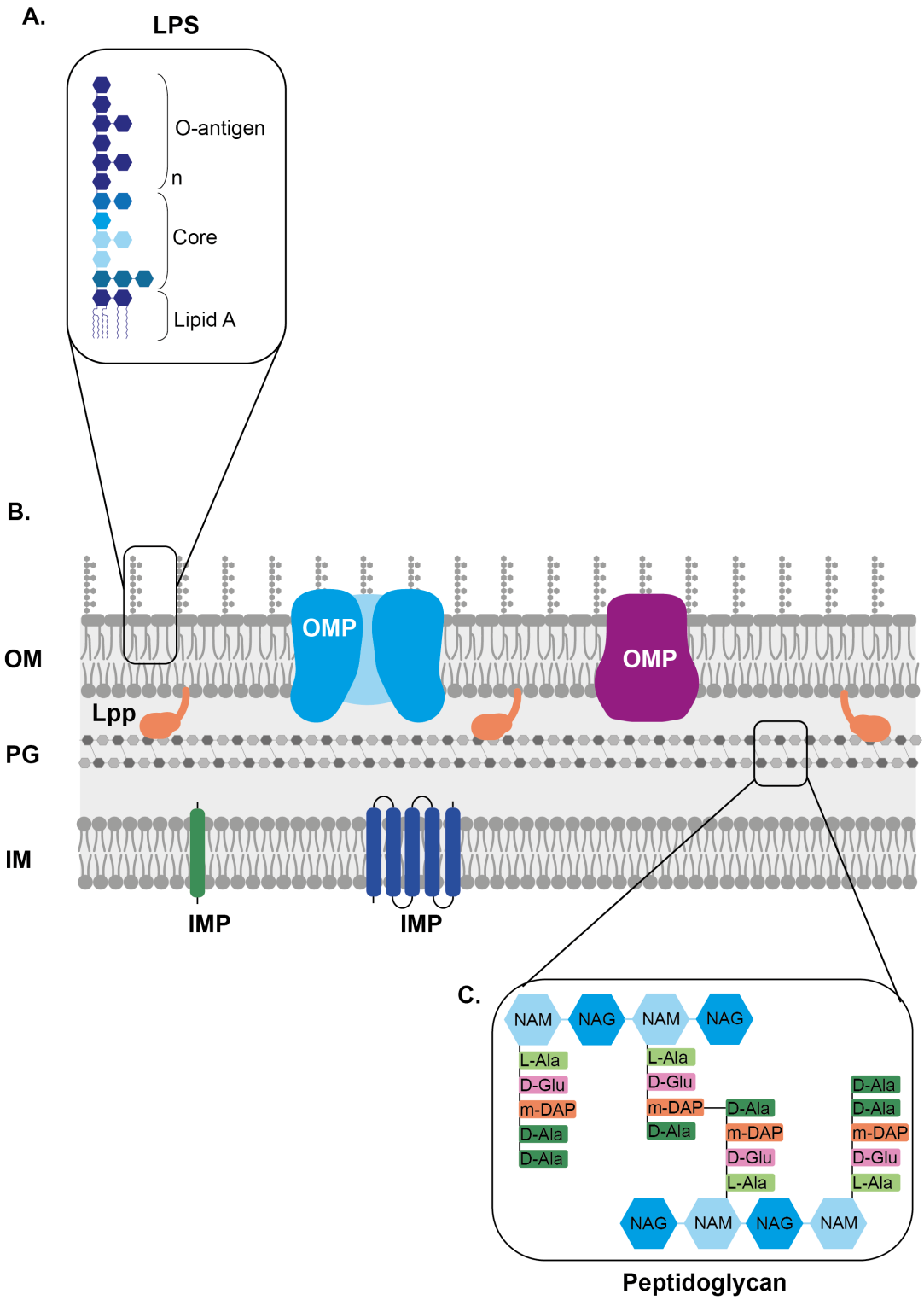


Figure 1.1 The tripartite Gram-negative cell envelope consists of an outer membrane (OM), peptidoglycan layer (PG) and inner membrane (IM).
A. Representative structure of LPS found on the outer leaflet of the OM. **B.** The *E. coli* Gram-negative cell envelope, including representations of the proteins found in the IM and OM. The two classes of OMP are shown with representative porins in blue and purple, and Braun's lipoprotein (Lpp) in orange. Two representative IMPs are shown in the IM, one with a single transmembrane helix and one with multiple transmembrane helices. **C.** Diagram of un-crosslinked and crosslinked *E. coli* peptidoglycan. The NAG-NAM disaccharide repeating units form the glycan strands that are linked by 3-4 crosslinks between the peptides extending from NAM.

and the O-antigen (Figure 1.1.A). The LPS outer leaflet is what provides Gram-negative bacteria with an effective diffusion barrier against antibiotics and other small molecules.

Lipid A acts as the hydrophobic anchor of LPS. It was first characterised as 'endotoxin', named after its role in endotoxic shock during septicaemia (Raetz, 1990), and purified lipid A is recognised by the innate immune system at picomolar levels (Raetz and Whitfield, 2002). *E. coli*'s lipid A structure is formed of glucosamine disaccharides, with six acyl chains (Figure 1.1.A) (Steimle, Autenrieth and Frick, 2016).

The core oligosaccharide begins with the scaffold of 3-deoxy-D-manno-oct-ulosonic acid (Kdo) that is linked to the lipid A disaccharide. Kdo is found in all Gram-negative core oligosaccharides, commonly alongside L-glycero-D-mannoheptose (heptose), and the inner components of the core are well conserved within each genus. The outer core contains a series of branched hexose sugars such as glucose, mannose and rhamnose, and varies more substantially between species (Raetz and Whitfield, 2002). Divalent cations such as Mg^{2+} or Ca^{2+} are present in the lipid A and core oligosaccharide layers to neutralise the phosphate charges, stabilising the OM and preventing the permeation of antibiotics (Clifton *et al.*, 2015). At the outermost part of LPS is the O-antigen that extends from the core oligosaccharide and interacts with the environment. O-antigen is highly variable, and strains can be classified based upon the length of the polysaccharides.

The LPS-containing bilayer encasing Gram-negative bacteria provides an efficient barrier to the environment, but to allow the selective movement of nutrients into the cell outer membrane proteins (OMPs) are required. There are two classes of proteins found in the OM of Gram-negative bacteria (although some, such as TonB-dependent-transporters fall outside of these categories): porins and lipoproteins. Lipoproteins are embedded into the inner leaflet of the OM by lipid moieties attached to an N-terminal cysteine residue (Sankaran and Wu, 1994). The most abundant of these is Braun's lipoprotein (Lpp), which is inserted in the OM and covalently attached to the cell wall (Braun, 1975). Porins are responsible for the passive diffusion of molecules into the cell and are typically β-barrel pores, for example OmpF and OmpC in *E. coli* (Silhavy, Kahne and Walker, 2010). OMPs cluster into regions called OMP

islands within the OM outer leaflet (Rassam *et al.*, 2015). The interactions between the OMPs are mediated by LPS and lipids, forming a network of interactions thought to explain the load-bearing capacity of the OM (Webby *et al.*, 2022a).

1.2.2 The periplasm and cell wall

The periplasm defines the area between the two membranes in the Gram-negative cell envelope and contains the cell wall. The periplasm itself is a viscous, aqueous compartment, home to an array of proteins, including chaperones and enzymes involved in cell envelope formation. It is also of note that there is no ATP present within the periplasm nor a proton gradient across the OM. Therefore, energy requiring processes carried out in the periplasm must use the IM as their energy source, either through coupling to ATP hydrolysis in the cytoplasm or utilising electrochemical gradients across the IM.

The peptidoglycan (PG) cell wall is an essential structure in the cell envelope of bacteria. It is responsible for retaining the rigidity of the cell against the turgor pressure from the cytoplasm, and for defining the cell shape (Rojas *et al.*, 2018). The overall structure of the Gram-negative PG is formed from chains of repeating disaccharides and crosslinked short peptides. The glycan strands in *E. coli*'s cell wall have a mean length of 20-40 disaccharide units that run parallel to the membrane in a layered manner, with a thickness of 2-6 nm (Matias *et al.*, 2003; Vollmer and Seligman, 2010). The disaccharide unit contains N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) linked by a β 1-4 glycosidic bond. A short peptide extends from the lactoyl group bound to position 3 of NAM. The model Gram-negative PG has a pentapeptide of L-Ala, D-Glu, meso-diaminopimelic acid (mDAP), D-Ala, D-Ala (Figure 1.1.C).

Crosslinking between glycan strands most commonly occurs between the mDAP at position 3 and D-Ala at position 4 and releases the D-Ala at position 5. This crosslink can be direct or indirect through a peptide bridge, though direct is most common in Gram-negative PG (Vollmer, Blanot and De Pedro, 2008). Pores and variation in the degree of thickness of

the PG in individual Gram-negative bacteria has been observed, and it is at these pores that biosynthesis of new PG is thought to occur (Turner *et al.*, 2013).

1.2.3 The inner membrane

The inner membrane (IM) in Gram-negative bacteria is a phospholipid bilayer composed of 75-85% phosphatidylethanolamine, 10-20% phosphatidyl glycerol and 5-15% cardiolipin (Raetz and Dowhan, 1990). IM proteins are responsible for many cellular processes including energy production, cell envelope biosynthesis, nutrient transport and protein transport (Silhavy, Kahne and Walker, 2010).

One crucial distinction between the inner and outer membranes of Gram-negative bacteria is the existence of a proton gradient across the IM that is not present in the OM. This, coupled with the lack of ATP in the periplasm, means that mechanical work carried out at the OM must be driven by energy generated from ion gradients across the IM.

1.3 Proton pathways and the proton motive force

Tol-Pal uses the energy derived from the proton motive force (PMF) to carry out its role in stabilising the OM of Gram-negative bacteria. Here, the PMF is introduced as a concept and the fundamental understanding of how protons travel through proteins is explored.

1.3.1 Definition of the proton motive force

Across the IM of Gram-negative bacteria there is a gradient of hydrogen ions (H^+), with a higher concentration of H^+ in the periplasm. The movement of ions across a membrane is defined by two terms: 1) The chemical gradient where molecules tend to diffuse from higher to lower concentration. 2) The electrical gradient where cations move from the positive side of the membrane to the negative side of the membrane (and vice versa for anions). The IM is impermeable to H^+. This means that the gradient of protons across the membrane essentially acts as an energy store that can be harnessed to carry out energetically unfavourable processes. This energy store is called the proton motive force (PMF) (Δp). It is worth noting that ‘force’ in this instance does not refer to the vector force, but the potential for motion.

The PMF is specifically described by two terms: the change in concentration of H^+ (ΔpH), and the change in electrical gradient ($\Delta\Psi$). The change in pH is defined as $[H^+]$ inside the cell minus $[H^+]$ outside the cell and change in electrical gradient Ψ is defined as Ψ inside minus Ψ outside. The Equation 1.1 demonstrates the relationship between ΔpH and $\Delta\Psi$, and shows that $\Delta\Psi$ has a greater contribution to Δp than ΔpH.

Equation 1.1 The definition of the proton motive force

$$\Delta \boldsymbol{p}(\boldsymbol{V}) = \Delta \boldsymbol{\Psi} - \mathbf{0.06}\Delta \boldsymbol{pH}$$

1.3.2 Generating a proton gradient across the IM

The PMF is generated by the electron transport chain (ETC) in a process called chemiosmosis, a concept first put forward by Mitchell (Mitchell, 1966). The ETC couples the transfer of electrons from substrates to the movement of protons across the membrane. In brief, a series of protein complexes facilitate redox reactions with an electron donor such as NADH. The energy released from this reaction is used to transfer protons across the membrane, building a higher concentration of protons in the periplasm.

1.3.3 Proton wires

To travel through proteins, protons move through a network of hydrogen bonds that may also include amino acid side chains of polar or charged amino acids. The network of hydrogen bonded water and side chains is known as a ‘proton wire’ and creates a path for protons through protein, across the membrane (Nagle and Tristram-Nagle, 1983). The side chains of polar and charged amino acids introduce new mechanisms for proton transfer (DeCoursey and Hosler, 2014). Polar amino acids with hydroxyl groups (Ser, Thr, Tyr) simultaneously bind to an incoming proton and release the proton currently bound, carrying the proton downstream without ionisation. Acidic (Asp, Glu) or basic (Arg, His, Lys) amino acids acquire and release the proton and can allow for rapid chain movements (Cukierman,

2003). In summary, organised networks of amino acid side chains and water molecules form proton wires through proteins to allow protons to be transduced.

1.4 Trans-envelope machinery in Gram-negative bacteria

Proteins containing proton wires that reside within the IM of Gram-negative bacteria can use the proton electrochemical gradient to power actions carried out in the periplasm and OM. This study focuses on the Tol-Pal IM machinery that converts the PMF to mechanical force for the purpose of OM maintenance. Two sister systems to Tol-Pal that also use the PMF to carry out work at the OM are first introduced as similarity among the three systems allows for inferences between them and helps to further understand Tol-Pal (Figure 1.2). More detailed comparison between the stator components and their mechanisms is reserved for later discussion.

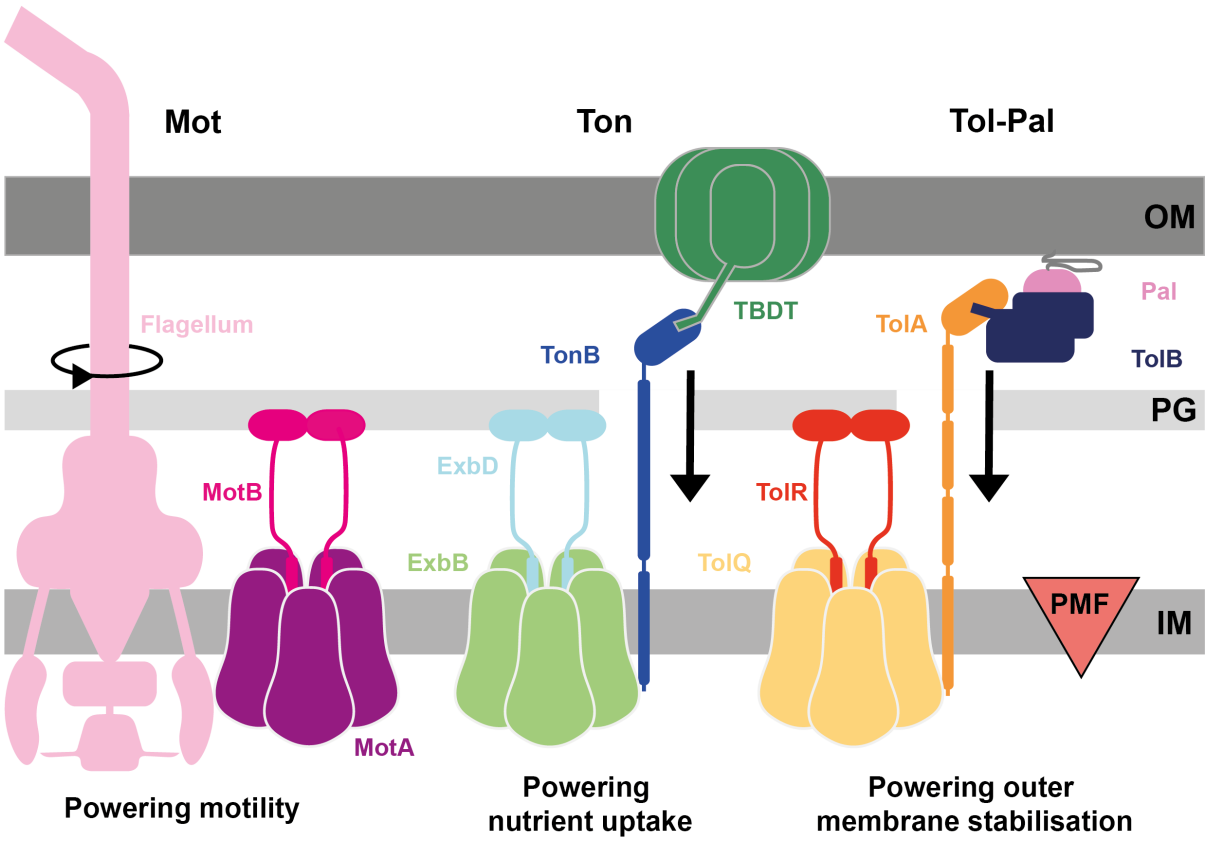


Figure 1.2 The Mot, Ton and Tol systems use the PMF to drive force transduction at the OM
Mot, Ton, and Tol-Pal all contain a stator unit that uses the PMF to drive force transduction. Black arrows indicate the direction of movement.

1.4.1 Powering motility

Moving a bacterium through its environment is a process that requires energy. In the case of the *E. coli* flagellum, it is the stator units of the system that drive the rotation of the flagellum, propelling the bacterium forwards. The stator is comprised of the proteins MotA and MotB that use the PMF as the energy source to generate torque that in turn powers the rotor. Up to 11 MotA-MotB units associate with the rotor through interactions with the 'torque helix' in the C-terminal domain of the flagellar C-ring protein FliG (Zhou, Lloyd and Blair, 1998; Reid *et al.*, 2006).

The MotA-MotB stator unit has a stoichiometry of 5:2, and the structures of this complex from three different species were solved in 2020 (Deme *et al.*, 2020; Santiveri *et al.*, 2020). The transmembrane helices (TMHs) 3 and 4 of MotA form the central pore, surrounded by TMHs 1 and 2 and an amphipathic helix that creates a belt at the cytoplasm/membrane interface (Figure 1.3.A-B). TMHs 3 and 4 extend 30 Å into the cytoplasm where a series of 4 shorter helices make up the cytoplasmic domain (Figure 1.3.A-B). It is within this region that residues Arg89 and Glu97 (*E. coli* numbering) that are known to be involved in torque generation and interaction with FliG are found (Zhou and Blair, 1997).

The N-terminal helices of MotB, previously described as transmembrane, here described as trans-pore helices (TPHs), reside within the hydrophobic pore made by MotA. They are followed by a plug helix that interacts with the tops of MotA's TMHs 3 and 4 (Santiveri *et al.*, 2020). The periplasmic domain of MotB could not be resolved in any of the available cryo-EM complex structures, however crystal structures of the isolated periplasmic domain have been determined (Roujeinikova, 2008; Kojima *et al.*, 2009; O'Neill *et al.*, 2011). The structure of *Helicobacter pylori* (*Hp*) MotB has four pairs of alternating β strands and α helices, with β strands 1-3 forming a parallel sheet, and β4 an antiparallel sheet. Helices α1, α2 and α4 pack against one face of the β-sheet, and α3 extends from β4 (Figure 1.3.D) (Roujeinikova, 2008). The MotB periplasmic domain was found to dimerise through an antiparallel interaction of β3 from each subunit, forming an extended β-sheet.

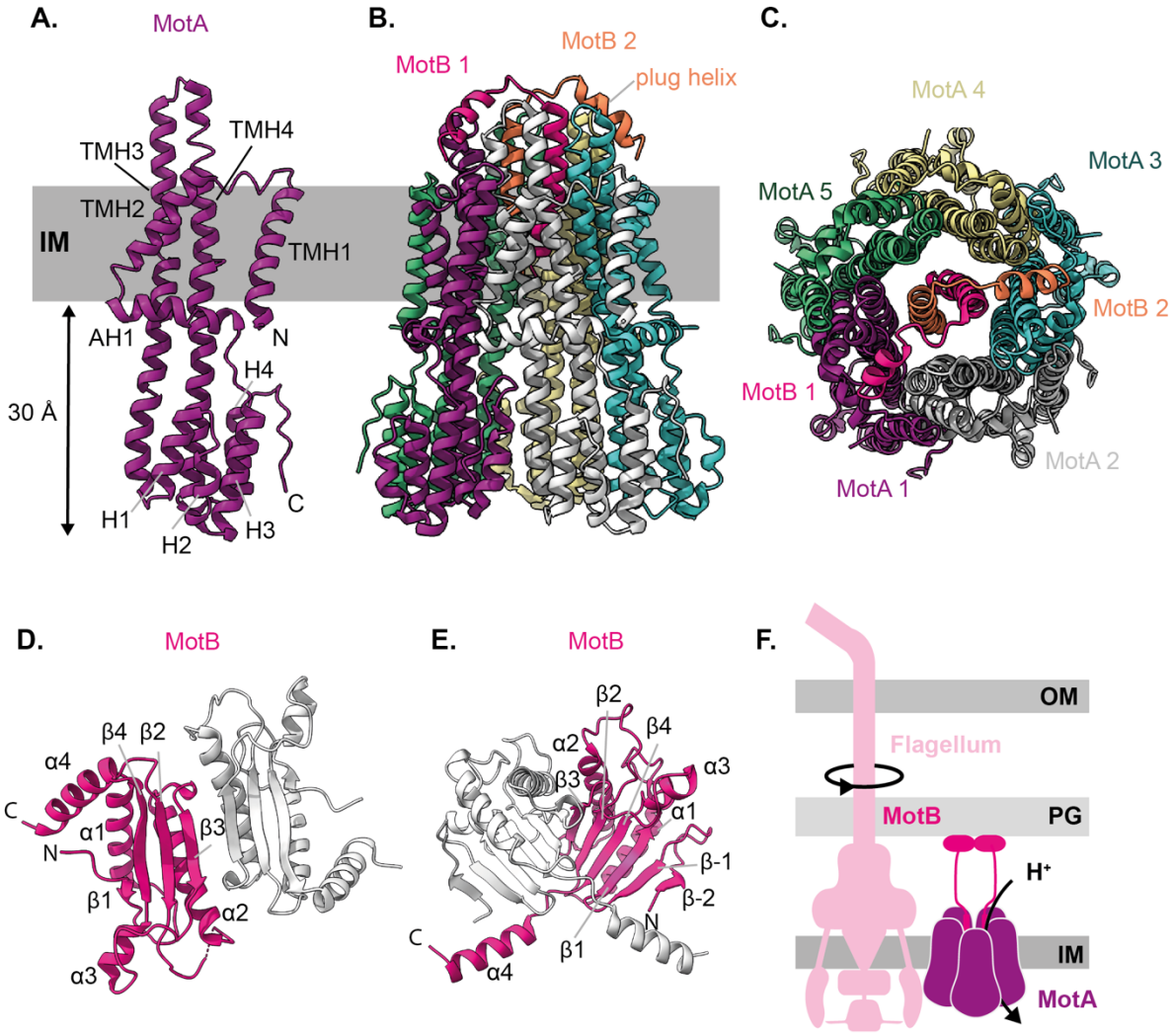


Figure 1.3 The MotA-MotB stator.
A. The structure of monomeric MotA from *Campylobacter jejuni* (PBD: 6YKM) (Santiveri *et al.*, 2020). The amphipathic helix (AH1), TMHs 1-4 and cytoplasmic helices (H1-4) are labelled. The cytoplasmic domain extends 30 Å from the IM. **B.** The structure of the pentameric MotA with MotB's TPHs occupying the pore formed by MotA. The MotB plug helices block the channel in this structure. **C.** The top-down view of the MotA pore with MotB's TPHs and plugging by the plug helices. **D.** The structure of *H. pylori* MotB's truncated periplasmic domain, residues 125-256 (PDB: 3CYP) (Roujeinikova, 2008) forms a dimer through an extended β-sheet interaction. **E.** The structure of *H. pylori* MotB residues 78-256 (PDB: 3S0W) (O'Neill *et al.*, 2011) forms a dimer in which the linker region, previously truncated, occludes the PG binding site. **F.** Schematic representation of the MotA-MotB stator driving flagellum rotation as protons are transduced through the stator.

MotB's periplasmic domain was proposed to bind to PG (De Mot and Vanderleyden, 1994) and the fold of *Hp*MotB was found to have high structural similarity with PG-binding OmpA-like fold proteins RmpM (Grizot and Buchanan, 2004) and Pal (Parsons, Lin and Orban, 2006). Co-crystallisation with NAM identified a glycan binding groove and residues involved in the interaction (Roujeinikova, 2008). Later structures of the *Salmonella enterica* serovar

Typhimurium MotB periplasmic domain identified an alternative dimerisation interaction (Kojima *et al.*, 2009). Analysis of truncations of *Hp*MotB showed that the N-terminal linker is responsible for inhibiting binding to PG by occluding the PG binding site (O'Neill *et al.*, 2011). The dimerisation of MotB through its periplasmic domain was also shown to be essential for motility (Roujeinikova, 2008; Kojima *et al.*, 2009).

It is proposed that MotA's interaction with the FliG torque helix of the rotor triggers the unplugging of MotA through the movement of MotB's plug domains, allowing for proton flux through the stator (Santiveri *et al.*, 2020). The linker between MotB's domains unfolds, exposing the PG-binding region and allowing it to extend to bind to the PG (O'Neill *et al.*, 2011). Once MotB is bound to the PG it provides an anchor for MotA to rotate around, in turn driving the rotation of the rotor through interactions with FliG in the C-ring (Santiveri *et al.*, 2020).

1.4.2 Powering nutrient import

Another process within Gram-negative bacteria that requires energy for function at the OM is the import of nutrients. The Ton system uses the PMF to drive the active transport of nutrients into the periplasm. It catalyses the import of sugars, siderophores, vitamin B12 and other non-ferrous cations (Lim, 2010). It can also be parasitised by bacteriophages and bacteriocins, which hijack the trans-envelope machinery for import into the periplasm. Like the homologous Mot system (Marmon, 2013), Ton has a stator unit comprised of ExbB and ExbD. However, unlike its flagellar counterpart, the effector unit coupled to ExbB-ExbD is a single protein, TonB, that spans the periplasm to interact with a TonB dependent transporter (TBDT) in the OM. When the TBDT is loaded with a ligand, the ExbB-ExbD-TonB motor pulls the plug domain of the TBDT to open the pore and allow for import (Figure 1.4.A).

Initially, a range of stoichiometries of ExbB:ExbD, from 4:1 to 6:2, were observed in cryo-EM structures (Sverzhinsky *et al.*, 2015; Celia *et al.*, 2016; Maki-Yonekura *et al.*, 2018). However, more recent cryo-EM data (Celia *et al.*, 2019; Deme *et al.*, 2020; Biou *et al.*, 2022)

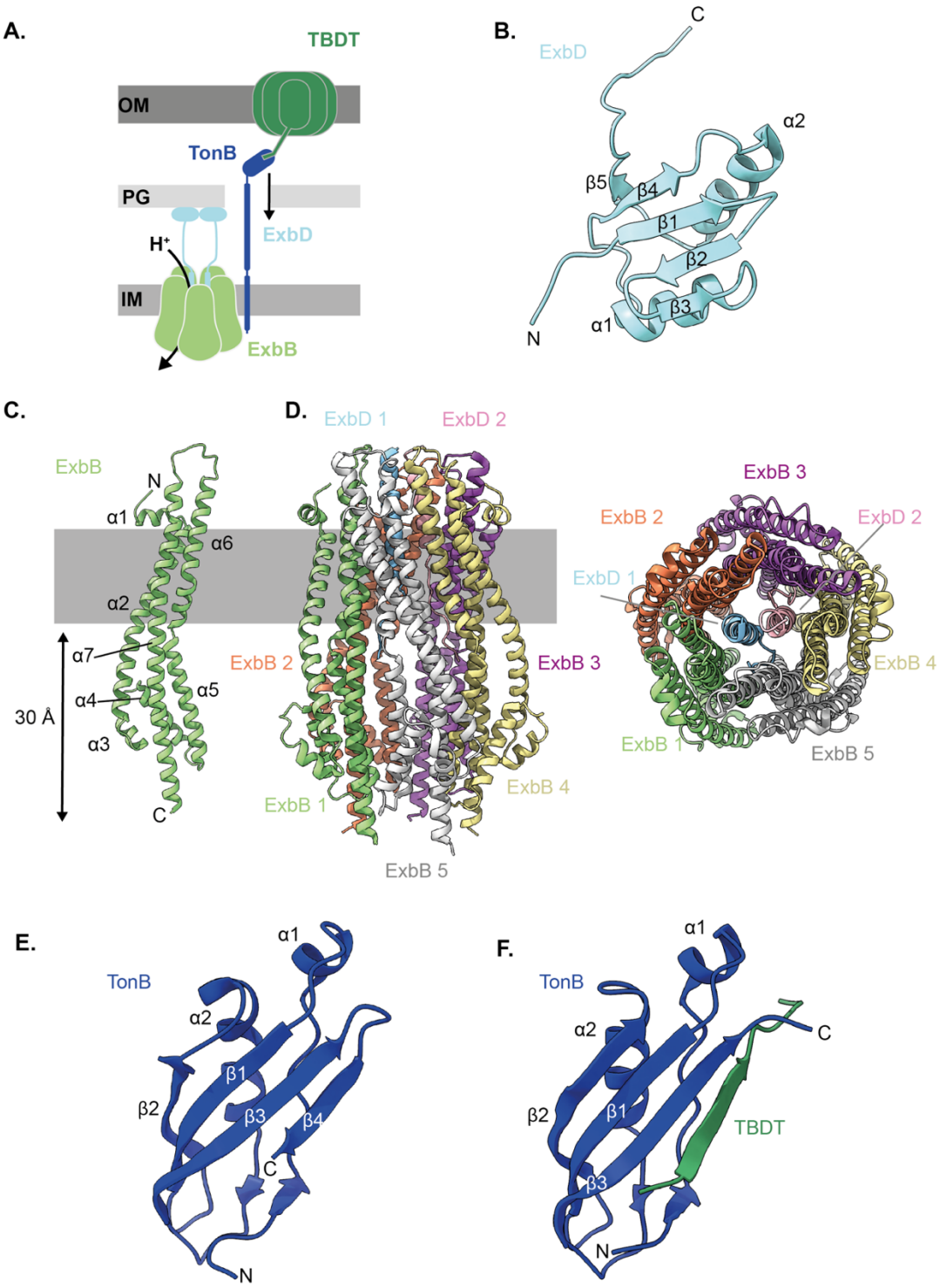


Figure 1.4 Structures of the Ton system.
A. Schematic representation of the Ton system. Flux of protons drives TonB to dislodge the plug domain of the TBDT. **B.** Solution structure of the periplasmic domain, residues 61-141 of *E. coli* ExbD as a monomer (PDB: 2PFU) (Garcia-Herrero *et al*., 2007). Residues 43-60 are flexible so for clarity are not displayed here. **C.** The monomer of *E. coli* ExbB has seven helices, including three transmembrane regions, and extends 30 Å into the cytoplasm. (PDB: 6TYI) (Celia *et al.*, 2019). **D.** ExbB forms a pentameric pore that ExbD's TPHs reside in. Side and top-down views shown. **E.** Apo form of *E. coli* TonB-III (PDB: 1XX3) (Peacock *et al.*, 2005). **F.** *E. coli* TonB-III (blue) bound to the TonB box of BtuB (green) (PDB: 2GSK) (Shultis et al., 2006). The β4 strand is displaced by the TBDT TonB box to allow for parallel β-sheet formation. **D.** ExbB forms a pentameric pore that ExbD's TPHs reside in. Side and top-down views shown. **E.** Apo form of *E. coli* TonB-III (PDB: 1XX3) (Peacock *et al.*, 2005). **F.** *E. coli* TonB-III (blue) bound to the TonB box of BtuB (green) (PDB: 2GSK) (Shultis *et al.*, 2006). The β4 strand is displaced by the TBDT TonB box to allow for parallel β-sheet formation.

and *in-vivo* native mass spectrometry (Chorev *et al.*, 2018) confirmed the stoichiometry of the ExbB:ExbD complex as 5:2. The monomer of ExbB is formed of seven α helices, where α2 and α7 form extended helices of up to 100 Å in length, and α2, α6 and α7 contain the transmembrane regions (Figure 1.4.C) (Celia *et al.*, 2016; Celia *et al.*, 2019). The pentamer of ExbB forms a pore in which the TPHs of the ExbD dimer sit within a hydrophobic environment (Celia *et al.*, 2019; Deme *et al.*, 2020; Biou *et al.*, 2022). The periplasmic domain of ExbD was not resolved in the cryo-EM structures, but a nuclear magnetic resonance (NMR) solution structure shows it to have two α helices and a five-stranded β-sheet, with a flexible N-terminal linker to the TPH (Figure 1.4.B) (Garcia-Herrero *et al.*, 2007). The cryo-EM structures, double electron-electron resonance (DEER) spectroscopic analysis and crosslinking data point to ExbD being a homodimer (Ollis *et al.*, 2009; Celia *et al.*, 2016). Combined, data on the ExbB-ExbD stator suggests it too has a stoichiometry of 5:2, in agreement with the MotA-MotB stator.

The effector protein for the Ton system that couples the IM stator unit to the TBDT in the OM is TonB. It has three distinct domains; the transmembrane helix TonB-I, a proline rich central domain TonB-II, and a C-terminal domain, TonB-III that interacts with the TBDT (Figure 1.4.A) (Postle and Skare, 1988; Hannavy *et al.*, 1990; Clarke, Tari and Vogel, 2005). TonB-III is responsible for interacting with the TBDT via a conserved 5/6 residue sequence at the N terminus of the TBDT called the TonB box. The NMR solution structure of *E. coli* TonB-III revealed it to have a 4-stranded antiparallel β-sheet with two α helices and an extended structured region (Figure 1.4.E) (Peacock *et al.*, 2005). Crystal structures of *E. coli* TonB in complex with TBDTs FhuA and BtuB demonstrated that the TonB box of the TBDT displaces the β4 strand and forms an interprotein β-sheet with TonB-III (Figure 1.4.F) (Shultis *et al.*, 2006; Pawelek *et al.*, 2006). Dimeric crystal structures of shorter constructs of TonB-III observed an altered β-sheet conformation (Chang *et al.*, 2001; Ködding *et al.*, 2005). However, the crystal structure dimers could not be observed *in vivo* (Postle *et al.*, 2010) and the

functional relevance of TonB dimers is debated (Sauter, Howard and Braun, 2003; Ködding *et al.*, 2005).

The central domain of TonB, TonB-II, is less well characterised. It contains an unusual poly-proline sequence, consisting of $(Glu\text{-}Pro)_n…(Lys\text{-}Pro)_n$ repeats (Postle and Good, 1983), and deletion of this region does not impact function of the Ton system (Larsen and Wood, 1993). NMR analysis of a 33 residue synthetic peptide based upon the *E. coli* TonB polyproline region demonstrated rigidity (Evans *et al.*, 1986). In addition, DEER spectroscopy analysis of *E. coli* TonB 56-126 also showed the region to be rigid and have the potential to span the length of the periplasm (Köhler *et al.*, 2010). Finally, circular dichroism spectra of the domain was consistent with a polyproline helix (Köhler *et al.*, 2010). At present little is known about the structure of TonB-II, and even less about the role it plays in the function of TonB in transducing the PMF to TonB-III interacting with TBDTs at the OM.

The ExbB-ExbD stator and TonB assemble through interactions between the TMHs of TonB and ExbB. A conserved $SX_3HX_6LX_3S$ sequence, where X denotes any other residue, named the 'SHLS' motif, was observed in the TMHs of both TonB and TolA (Koebnik, 1993). Mutational analysis of this motif found that when residue His20 (*E. coli* numbering) is mutated to Ala or Arg, function of the Ton system was abolished (Traub, Gaisser and Braun, 1993). Mutation of the other residues within the motif, however, had little impact on Ton function (Larsen *et al.*, 2007). More recently, the cryo-EM structure of *Pseudomonas aeruginosa* ExbB-ExbD contained density adjacent to ExbB TMH 1 that was thought to be TonB (Deme *et al.*, 2020). This provided evidence that the assembly of the ExbB-ExbD-TonB complex is dependent upon interactions between the TMHs of ExbB and TonB.

One of the major questions regarding the mechanism of the Ton system is how a stator with a seemingly rotary motion translates to a pulling force at the OM. Atomic force microscopy and molecular dynamic simulations demonstrated that the TonB-TBDT-TonB-box interaction is capable of withstanding force (Hickman *et al.*, 2017). Steered molecular dynamic simulations showed using a pulling force perpendicular to the membrane to unfold the plug domain of the TBDT is the least energetically demanding mechanism (Gumbart, Wiener and

Tajkhorshid, 2007). However, fluorescence anisotropy data for GFP-tagged TonB suggested that TonB undergoes PMF-dependent movement that may be rotary in motion (Jordan *et al.*, 2013). Coupled with the evidence for MotA-MotB's rotary mechanism, this points towards a rotary movement of ExbB around ExbD.

1.5 The Tol-Pal system

The focus of this study is the Tol-Pal system that is found across all proteobacteria subclasses and frequently present in the phyla *Chlorobi, Chlamydiae* and *Acidobacteria* (Krachler *et al.*, 2010). Tol-Pal is essential in some species such as *Pseudomonas aeruginosa* and *Caulobacter crescentus* (Dennis, Lafontaine and Sokol, 1996; Yeh *et al.*, 2010) but is not essential in *E. coli* K-12. Here the system's genes, component proteins, and overall function is introduced.

1.5.1 The *tol-pal* operon

The genes within the *tol-pal* operon were first identified by mutations that conferred resistance to bacteriocins and bacteriophages, giving rise to the name 'tol' for 'tolerant' (Gratia, 1964; Reeves, 1966; Nagel de Zwaig and Luria, 1967). The operon contains seven genes in the following order: *ybgC, tolQ, tolR, tolA, tolB, pal, cpoB*, and the order of the genes within the operon is highly conserved (Figure 1.5) (Sturgis, 2001). The operon contains two open reading frames (ORFs) and two promoters: P1, the promoter for the whole operon or ORF1 (*ybgC, tolQ, tolR, tolA*), and PB, the promoter for ORF2 (*tolB, pal, cpoB*) (Figure 1.5.B) (Vianney *et al.*, 1996). The capsule regulatory protein RcsC has been linked to regulation of the operon (Clavel *et al.*, 1996). More recently the two-component system CenKR in *Rhodobacter sphaeroides*, was shown to regulate Tol-Pal expression, though the signal sensed by the system remains to be identified (Lakey *et al.*, 2022).

Mutation of any one of the five core members of the operon - *tolQ, tolR, tolA, tolB* and *pal* – results in the classical *tol* phenotype. Cells become sensitive to detergents and antibiotics they were previously resistant to, such as vancomycin (Nagel de Zwaig and Luria,

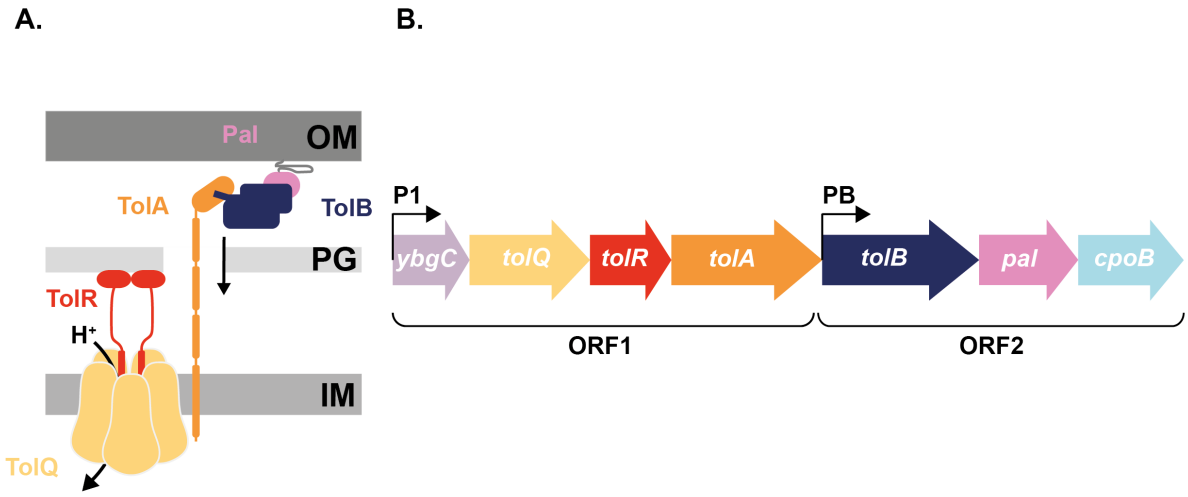


Figure 1.5 The Tol-Pal system.
A. Schematic representation of the Tol-Pal system including the five core proteins: TolQ, TolR, TolA, TolB and Pal. **B.** The *E. coli* Tol-Pal operon contains seven genes organised into two ORFs with two promoters (Vianney et al., 1996).

1967; Bernstein, Rolfe and Onodera, 1972; Paterson *et al.*, 2009). Cell division defects occur, resulting in blebbing and cell chaining (Weigand and Rothfield, 1976; Meury and Devilliers, 1999), and periplasmic proteins are released (Lazdunski *et al.*, 1998). The *tol* phenotype also affects the ratio of phospholipids-to-LPS in the OM (Shrivastava, Jiang and Chng, 2017; Masilamani, Cian and Dalebroux, 2018) and results in aberrant PG processing (Gray *et al.*, 2015; Yakhnina and Bernhardt, 2020). Finally, the cells become resistant to colicins and bacteriophages that parasitize the system, though this attribute of the phenotype does not occur for *Δpal*.

Loss of Tol-Pal function has also been shown to impact pathogenesis and virulence in some species: *E. coli* (Hellman *et al.*, 2002), *Edwardsiella ictaluri* (Abdelhamed *et al.*, 2016), *Dickeya dadantii (Erwinia chrysanthemi)* (Dubuisson *et al.*, 2005), *Haemophilus ducreyi* (Fortney *et al.*, 2000), *S. enterica* serovar Typhimurium (Bowe *et al.*, 1998) and uropathogenic *E. coli* all exhibit loss of virulence in *tol* mutants. In addition, mutations or deletions of *tol-pal* genes have been shown to impact biofilm formation in *Pseudomonas putida* and *E. coli* (López-Sánchez *et al.*, 2016; Ranjith *et al.*, 2019). The mechanistic understanding of how loss of Tol-Pal function results in reduced virulence and biofilm formation is limited, and the impact

of membrane destabilisation on pathogenesis compared with Tol-Pal having a direct role in virulence has not been well explored.

1.5.2 The proteins of the Tol-Pal system

The five core proteins of Tol-Pal span the cell envelope, from Pal anchored in the OM through the periplasmic TolB to the IM proteins TolQ, TolR and TolA. The protein-protein interactions between each component of the system have been extensively studied using a range of biophysical techniques, and structures of domains of Tol-Pal have been determined through both X-ray crystallography and NMR.

1.5.2.1 The TolQ-TolR stator

Like the MotA-MotB and ExbB-ExbD stators discussed above, Tol-Pal includes an IM molecular machine that converts the PMF to force transduction, consisting of the proteins TolQ and TolR. As there are no experimentally determined structures of TolQ or the TolQ-TolR complex available, most information on TolQ comes from crosslinking studies and comparison to MotA-MotB and ExbB-ExbD. TolQ, 25 kDa, is predicted to have three transmembrane helices and a large cytoplasmic loop (Kampfenkel and Braun, 1993; Vianney *et al.*, 1994). Formaldehyde and cysteine crosslinking studies showed TolQ to interact with TolR, TolA and to form oligomers (Derouiche *et al.*, 1995; Lazzaroni *et al.*, 1995; Zhang *et al.*, 2011). TolQ-TolR is expected to be organised with five copies of TolQ forming a pore and two copies of TolR residing within this pore, as in the MotA-MotB and ExbB-ExbD structures (Celia *et al.*, 2019; Deme *et al.*, 2020; Santiveri *et al.*, 2020). In this study detailed comparison between the stators is reserved for discussion in Chapter 4.

TolR, 16 kDa, can be separated into two domains, the TPH and a periplasmic domain, and has been found to form a dimer *in vivo* (Journet *et al.*, 1999; Goemaere *et al.*, 2007; Zhang *et al.*, 2009) and *in vitro* (Parsons, Grishaev and Bax, 2008; Wojdyla *et al.*, 2015). The two structures of the periplasmic domain of TolR that have been determined suggest that TolR oscillates between two different dimeric states. The solution structure of *H. influenzae* TolR

residues 59-130 showed each TolR monomer contains a β-β-β-α-β-α organisation where the first three β-sheets are antiparallel, β4 is parallel to the first, and the α helices sit on one face of the β-sheet (Figure 1.6.A) (Parsons, Grishaev and Bax, 2008). The β4 from each monomer interact to form an extended β-sheet where the dimer interface forms a hydrophobic groove between each α2 (Figure 1.6.A). The crystal structure of *E. coli* TolR 36-142 revealed a strand-swapped dimer (Wojdyla *et al.*, 2015). The fold of the domains remains similar, but the two copies of TolR are rotated 180° relative to one another. The additional residues, compared with the *H. influenzae* structure, form two β strands and two α helices, and the hydrophobic cleft is disrupted (Figure 1.6.B). The validity of the strand-swapped dimer is supported by previous *in vivo* cysteine crosslinking data that identified Tyr117Cys could form a spontaneous disulphide bond (Goemaere *et al.*, 2007). It is only in the strand-swapped structure that the two copies of Tyr117 are close enough to form a disulphide when mutated to Cys. The TolR dimer crosslinked with Tyr117Cys was found to inactivate Tol-Pal function, demonstrating that the cycling between two forms of TolR is likely to have a functional relevance (Goemaere *et al.*, 2007; Szczepaniak, Press and Kleanthous, 2020).

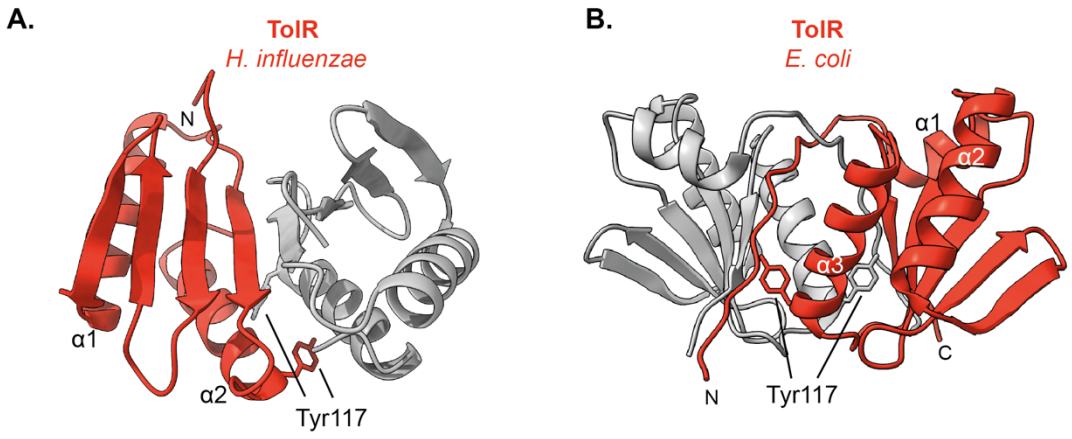


Figure 1.6 TolR's periplasmic domain forms two structural conformations.
A. The structure of *H. influenzae* TolR truncated periplasmic domain (PDB: 2JWK) (Parsons, Grishaev and Bax, 2008) forms a dimer where the β-strands form a continuous inter-molecular β-sheet. The Tyr residues (*E. coli* numbering) are too far away to interact. **B.** *E. coli* TolR's periplasmic domain including N-terminal linker and C-termini (PDB: 5BY4) (Wojdyla et al., 2015) forms a strand-swapped dimer. In this structure the Tyr117 residues are close enough to form a disulfide crosslink when substituted for Cys (Wojdyla et al., 2015).

A truncated version of *E. coli* TolR equivalent to the *H. influenzae* structure was found to bind PG *in vivo*, however no binding was observed for the intact *E. coli* periplasmic domain (Wojdyla *et al.*, 2015). These data were later supported by molecular dynamic simulations investigating the interaction between TolR and PG (Boags, Samsudin and Khalid, 2019). This is consistent with the PG binding region identified and characterised in MotB (Roujeinikova, 2008; O'Neill *et al.*, 2011). Combined, the structures and *in vivo* studies suggest that TolR alternates between two conformational states in a functionally relevant manner.

1.5.2.2 TolA

TolA is the effector protein that associates with the TolQ-TolR stator and spans the periplasm to interact with TolB-Pal at the OM. It is a 40 kDa monomeric protein (Witty *et al.*, 2002) with three domains: an N-terminal transmembrane region (TolA-I), a 25 kDa long helical domain rich in alanine and charged residues (TolA-II) and a 12 kDa periplasmic C-terminal domain responsible for interacting with its binding partners (TolA-III) (Levengood, Beyer and Webster, 1991). The interactions of TolA with other Tol-Pal proteins were first characterised by extensive crosslinking studies and yeast two-hybrid analysis. Interactions were identified with TolQ and TolR in the inner membrane, and TolB and CpoB in the periplasm (Germon *et al.*, 1998; Dubuisson, Vianney and Lazzaroni, 2002; Walburger, Lazdunski and Corda, 2002). An interaction between TolA and Pal in the OM has also been reported (Cascales *et al.*, 2000), but could not be validated using *in vitro* biophysical methods (Bonsor *et al.*, 2009).

The crystal structure of *P. aeruginosa* TolA-III revealed an $\alpha\beta$ domain with three β strands and four α helices. The β strands form an antiparallel β-sheet, and the α helices all sit on the same side of the β-sheet (Figure 1.7.A) (Witty *et al.*, 2002). TolA-III interacts with TolB, the bacteriophage Ff, colicin A and colicin N through β-augmentation interactions similar to the interactions between TonB and TBDT's discussed above (Lubkowski *et al.*, 1999; Li *et al.*, 2012; Ridley and Lakey, 2015; Szczepaniak *et al.*, 2020). The solution structure of *P. aeruginosa* TolA-III in complex with a peptide of TolB's N-terminal 12 residues demonstrated

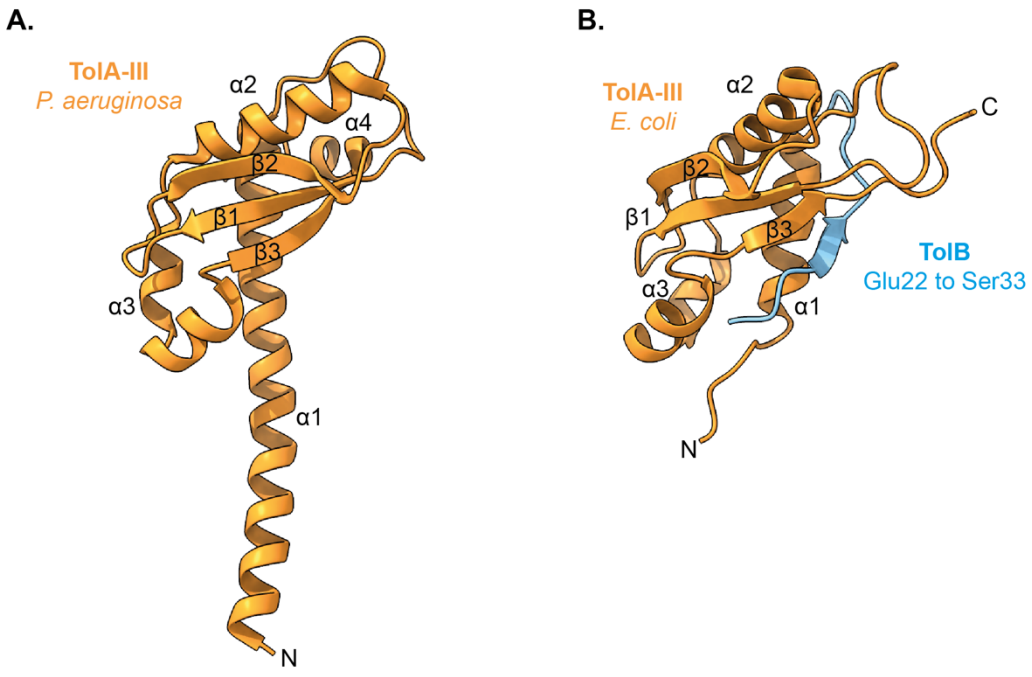


Figure 1.7 TolA-III binds to the N-termini of TolB through β-augmentation.
A. Crystal structure of *P. aeruginosa* TolA-III (PDB: 1LR0) (Witty et al., 2002). **B.** Solution structure of *E. coli* TolA-III (orange) bound to N terminal residues of *P. aeruginosa* TolB 22-33 (blue) (PDB: 6S3W) (Szczepaniak et al., 2020). The α4 helix is displaced to allow a β augmentation inter-protein interaction.

that the α4 helix is displaced to allow the TolB box residues to form a parallel β-sheet with β3 of TolA-III (Figure 1.7.B) (Szczepaniak *et al.*, 2020).

Much less is known about the central domain of TolA, TolA-II. In *E. coli* its primary sequence is ~35% alanine residues and contains 14 repeats of the motif Lys-Glu-Ala_3-Glu/Asp (Levengood, Beyer and Webster, 1991; Schendel *et al.*, 1997). Far UV circular dichroism, analytical ultracentrifugation and solution X-ray scattering all suggest TolA-II has an elongated helical structure (Levengood, Beyer and Webster, 1991; Derouiche *et al.*, 1999; Witty *et al.*, 2002). Studies into the role of TolA-II using a series of deletions found it to be essential; deletion of the whole domain resulted in sensitivity to detergent consistent with Δ*tolA*. However, the system was still functional when half of the domain was deleted (Schendel *et al.*, 1997). This implies that the length of TolA-II is important to its function, but its precise role remains unclear.

The first domain of TolA, TolA-I, is predicted to be a transmembrane helix that anchors TolA into the IM (Levengood, Beyer and Webster, 1991). It is this transmembrane domain that

has been implicated in interaction with TolQ (Koebnik *et al.*, 1993). A TonB-equivalent SHLS motif is present in TolA-I (Koebnik, 1993). Mutations of Ser18 and His22 impact binding to TolQ, and suppressor mutations identified TolQ as the site of interaction (Germon *et al.*, 1998).

1.5.2.3 TolB and Pal

The outermost protein of the Tol-Pal system is the OM-anchored, PG interacting protein Pal. Pal's primary role is to act as a tether between the OM and PG. It is attached to the OM through a lipid linkage at its N-terminal cysteine residue (Mizuno, 1979) and non-covalently binds to PG (Lazzaroni and Portalier, 1992). TolB is a soluble 45 kDa periplasmic protein that competitively binds to Pal's PG-binding site (Bouveret *et al.*, 1995; Clavel *et al.*, 1998). Its role is to modulate the state of Pal-PG interaction, manipulating whether Pal is static (PG anchored), or diffusing within the inner leaflet of the OM (TolB bound) (Szczepaniak *et al.*, 2020). Pal has also been shown to interact with OmpA and TolA (Cascales *et al.*, 2000; Cascales and Lloubès, 2004), however the interaction with TolA is debated as it could not be replicated using *in vitro* methods (Bonsor *et al.*, 2009).

The apo form of *E. coli* Pal has an α/β sandwich structure containing four α helices and four β-sheets, plus 15 unstructured N-terminal residues (Figure 1.8.B). The strands $\beta1$, $\beta2$ and $\beta4$ form an antiparallel sheet, and $\beta3$ and $\beta2$ a parallel sheet (Abergel *et al.*, 2001). The structure of *H. influenzae* Pal was found to bind the PG precursor PG-P's pentapeptides at the $\beta1$-$\alpha2$, $\beta2$-$\alpha3$ loops (Figure 1.8.C) (Parsons, Lin and Orban, 2006). The crystal structure of Pal in complex with TolB demonstrated that the same loops are used to bind to TolB, therefore Pal cannot bind both TolB and PG at the same time (Bonsor *et al.*, 2007).

TolB's structure has two distinct domains: an N-terminal α/β domain and a C-terminal β propeller, linked by a 9-residue linker sequence (Figure 1.8.A). The N-terminal α/β domain consists of two β-sheets, one parallel and one antiparallel, and two α helices (Abergel *et al.*, 1999; Carr *et al.*, 2000). The β propeller domain has pseudo-six-fold symmetry and is made

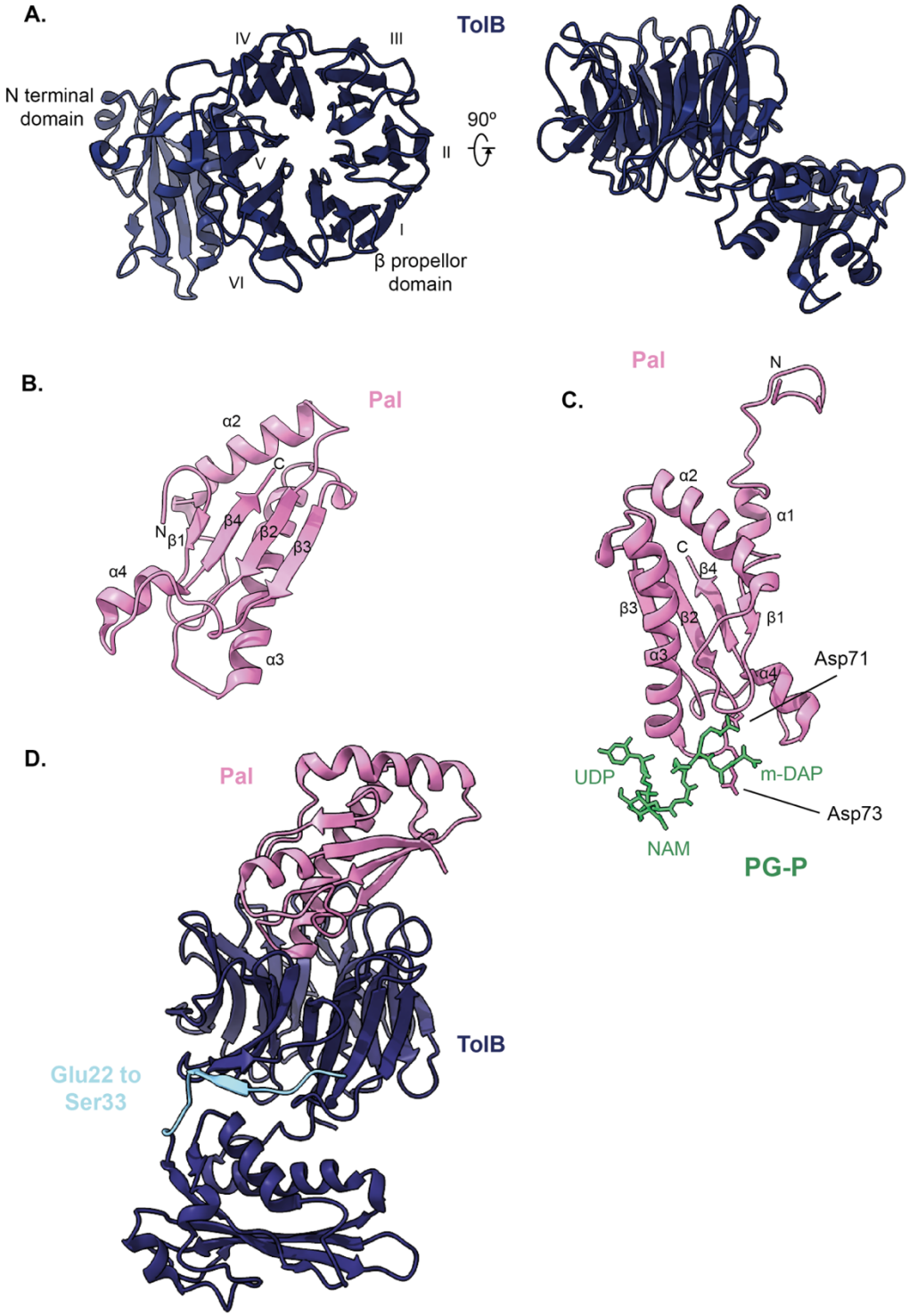


Figure 1.8 TolB and Pal bind to one another, and competitively Pal binds to PG.
A. The structure of *E. coli* TolB (PDB: 1CRZ) (Abergel *et al.*, 1999) shows its two distinct domains: the N-terminal domain and C-terminal β-propeller domain comprised of 6 repeating 4 β strand motifs (I-VI). Top down and side views shown. **B.** The crystal structure of apo *E. coli* Pal (PDB: 1OAP). **C.** The solution structure of *H. influenzae* Pal bound to PG-P, a PG precursor (PDB: 2AIZ) (Parsons, Lin and Orban, 2006). Residues Asp71 and Asp73 are important in binding to the m-DAP moiety of the PG-P. **D.** Pal (pink) binds to TolB's β-propeller domain (navy blue) (PDB: 2W8B) (Bonsor *et al.*, 2009). TolB's N terminal 12 residues Glu22-Ser33 (light blue) involved in binding TolA-III are sequestered between the β-propeller and N-terminal domains.

up of six repeated motifs of 4-stranded antiparallel β-sheets. The sixth blade closes the propeller using a β-sheet from the C terminus that forms an antiparallel sheet with three strands from the N-terminal end of the domain. The propeller creates a funnel shaped central cavity where the diameter ranges from 7 Å to 40 Å. At the entrance to the funnel there is a ring of prolines followed by a stabilising ring of aromatic residues (Abergel *et al.*, 1999; Carr *et al.*, 2000).

When bound to TolB, Pal blocks the central channel of the funnel and TolB undergoes conformational changes. The loops in the N-terminal domain of TolB move towards the C-terminal β-propeller by 1.5-5.0 Å, and the linker between domains forms a 2-stranded antiparallel sheet (Figure 1.8.D). The strand in the sixth blade of the β-propeller, responsible for closing the propeller, moves, opening a proline gate and creating a groove in which the N-terminal residues of TolB bind (Bonsor *et al.*, 2009). The N-terminal 12 residues of TolB that are sequestered in the groove are also the residues involved in the β-augmentation interaction with TolA, thereby preventing the TolB-TolA interaction (Szczepaniak *et al.*, 2020). However despite this, a ternary complex of TolA-TolB-Pal was observed through *in vitro* crosslinking (Bonsor *et al.*, 2009). This demonstrated the existence of a third, low abundance conformational state that allows binding of both Pal to the TolB β-propeller and TolA to the TolA N-terminal 12 residues (Bonsor *et al.*, 2009).

1.5.2.4 CpoB and YbgC

Deletion of the peripheral genes of the Tol-Pal operon *cpoB* and *ybgC* does not result in the classical *tol* phenotype, and therefore these ancillary genes have garnered less attention than the core components of the system. However, CpoB's role in the coordination of PG biosynthesis and OM constriction was identified in recent years (Gray *et al.*, 2015). The *ΔcpoB* phenotype of increased sensitivity to β-lactams, decreased cell envelope stability and OM blebbing, all to a lesser extent than in *tol*, and mirrors that of the bifunctional PG synthase PBP1B deletion *ΔmrcB*, inferring its role in directing PG biosynthesis (Gray *et al.*, 2015).

CpoB has two domains: an N-terminal coiled coil, and a C-terminal tetratricopeptide (TPR) domain. The crystal structure of the *E. coli* N-terminal domain revealed a trimeric coiled coil containing a hydrophobic core and 'trigger motif' of polar residues in the core that define the oligomeric state of the domain (Figure 1.9.A) (Krachler *et al.*, 2010). The structures of the *Xanthomonas campestris* and *E. coli* C-terminal TPR domains were both monomeric and contained three TPRs each with a helix-turn-helix motif, followed by final capping helix (Figure 1.9.B) (Krachler *et al.*, 2010; Egan *et al.*, 2018). Characterisation of the interaction of CpoB and TolA found that TolA disrupts the trimerisation of CpoB and binds with a 1:1 stoichiometry (Krachler *et al.*, 2010). The interaction site of TolA-CpoB was mapped to between residues 280 and 313 of TolA, and within the TPR domain of CpoB (Krachler *et al.*, 2010). The TPR

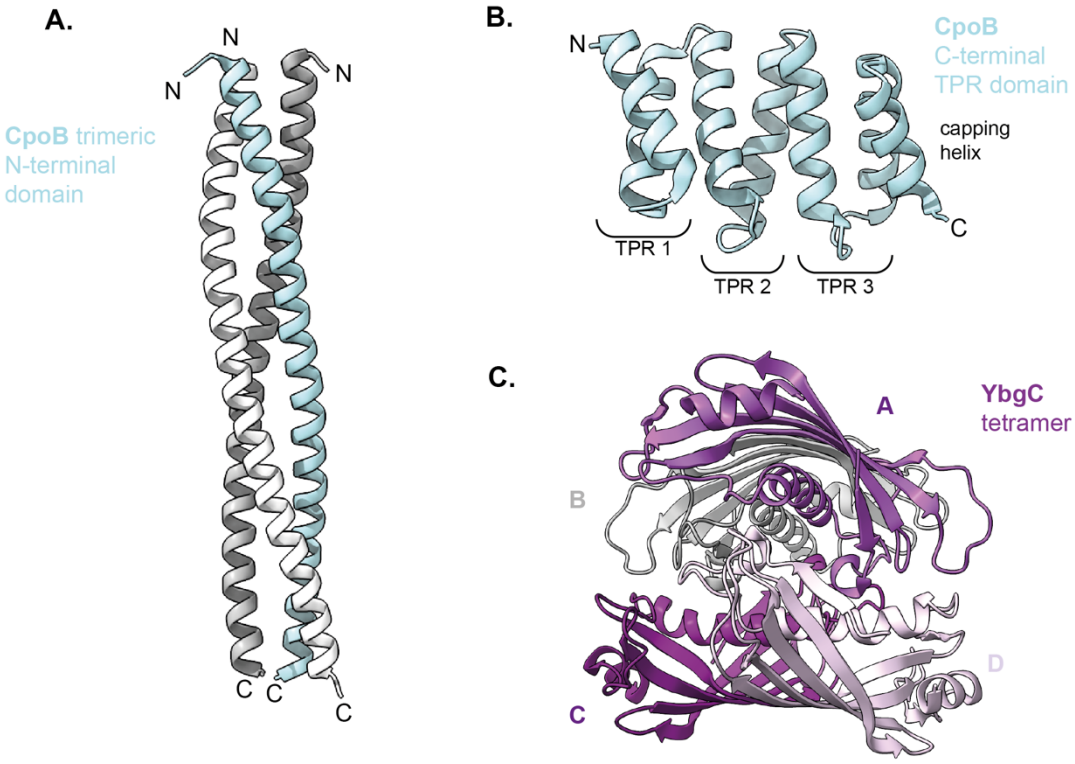


Figure 1.9 The structures of CpoB and YbgC.
A. The crystal structure of trimeric *E. coli* CpoB N-terminal domain forms a coiled coil (2XDJ) (Krachler *et al.*, 2010), each copy of CpoB is a different colour (white/grey/light blue). **B.** The crystal structure of the C-terminal TPR domain of CpoB (PDB: 2XEV) (Krachler *et al.*, 2010). Each copy of the helix-turn-helix motif is labelled, alongside the capping helix. **C.** The tetramer of *H. pylori* YbgC (2PZH) (Angelini *et al.*, 2008). Each monomer (A-D) forms a hot-dog fold characteristic of thioesterases.

domain is also responsible for interaction with the UB2H domain of PBP1B, inhibiting PBP1B's transpeptidase function (Egan *et al.*, 2018).

Deletion of the gene encoding *ybgC* does not give rise to the *tol* phenotype or similar (Sun and Webster, 1987; Gully and Bouveret, 2006) and *S.* Typhimurium *ΔybgC* has cell morphology typical of wild-type (Masilamani, Cian and Dalebroux, 2018). However, the deletion does give rise to the accumulation of phospholipids in the OM (Masilamani, Cian and Dalebroux, 2018). YbgC was identified as a cytoplasmic acyl-CoA thioesterase linked to phospholipid biogenesis (Zhuang *et al.*, 2002; Angelini *et al.*, 2008). Crystal structures of *E. coli* and *H. pylori* YbgC saw a tetrameric hotdog fold conserved across thioesterases (Figure 1.9.C) (Angelini *et al.*, 2008; Kim *et al.*, 2004). The activity of YbgC with long-chain acyl-CoA in *H. pylori* or short chain acyl-CoA in *H. influenzae*, and its interaction with acyl carrier protein (ACP) have all been demonstrated (Zhuang *et al.*, 2002; Gully and Bouveret, 2006; Angelini *et al.*, 2008). However, no direct link between Tol-Pal and YbgC has been uncovered, other than the location of the gene within the Tol-Pal operon.

1.5.3 Functions of the Tol-Pal system

1.5.3.1 Maintaining the outer membrane

The primary function of the Tol-Pal system is to maintain the integrity of the OM during cell division. The pleiotropic nature of the *tol* mutant phenotype made the identification of its role particularly challenging, and it is only recently that its specific mode of action was elucidated. Overexpression of Lpp in *Δpal* does not recover the wild-type phenotype, which suggested that Pal and the Tol-Pal system must have a more active role in OM maintenance than purely acting as a tether between the OM and PG (Cascales *et al.*, 2002). Fluorescent labelling of Tol-Pal proteins demonstrated that the components of the system are recruited to the septum during cell division (Gerding *et al.*, 2007). This, coupled with the *tol* phenotypic effects of OM blebbing from the septum and cell chaining, all provided evidence that Tol-Pal likely plays a role in actively coordinating cell division.

The dynamic recruitment of Tol-Pal to the cell division site was demonstrated by live cell fluorescence microscopy of endogenously labelled Tol-Pal proteins (Petiti *et al.*, 2019), confirming previous observations. Analysis of these labelled proteins in *tol* mutants identified that TolQ and TolA are recruited to the septum independently, and Pal requires functional TolQ-TolR-TolA and TolB to be localised at the septum (Petiti *et al.*, 2019). The observation that Pal is not recruited in the TolR-Asp23Ala mutant supported the idea that Pal recruitment to the septum is a PMF dependent process (Petiti *et al.*, 2019).

Assessment of the diffusion of Pal fused to photoactivatable mCherry using fluorescence recovery after photobleaching (FRAP) and single particle tracking (SPT) confirmed the dynamic recruitment of Pal to the septum (Szczepaniak *et al.*, 2020). Application of the mathematical modelling SpatialFRAP technique allowed for the determination of effective diffusion coefficients for Pal in both dividing and non-dividing cells. Pal was found to exist in two different populations: one where it is immobile when bound to PG, and a second where it is bound to TolB and able to diffuse laterally within the OM (Szczepaniak *et al.*, 2020).

The information described here is amalgamated into a 'mobilisation-and-capture' mechanism first proposed by Szczepaniak *et al.,* 2020 (Figure 1.10) that forms the foundation for the work described in this study (Szczepaniak *et al.*, 2020). Pal exists in one of two states: static when bound to PG, or mobile when bound to TolB. The affinity of Pal binding to TolB is much higher than for binding PG, therefore any TolB in the outer periplasm will be bound to Pal (Bonsor *et al.*, 2009). The interaction between TolQ-TolR and TolA results in proton flux through the stator, and TolA undergoes a conformational change allowing it to reach the OM and bind TolB-Pal. TolQ-TolR-TolA pulls TolB from Pal, freeing Pal and enabling it to bind PG. TolB is now spatially separated from Pal by the PG layer, and able to diffuse within the periplasm, eventually diffusing back up to the outer region and interacting with Pal again. In a non-dividing cell, TolQ-TolR and TolA independently diffuse within the IM. The result of this action is that TolQ-TolR-TolA ensures that Pal is bound to PG, not sequestered by TolB.

In a dividing cell, the Tol-Pal IM machinery is recruited to the septum. This means that Pal will only be released from TolB, allowing it to bind to PG, at the septum. The net effect of

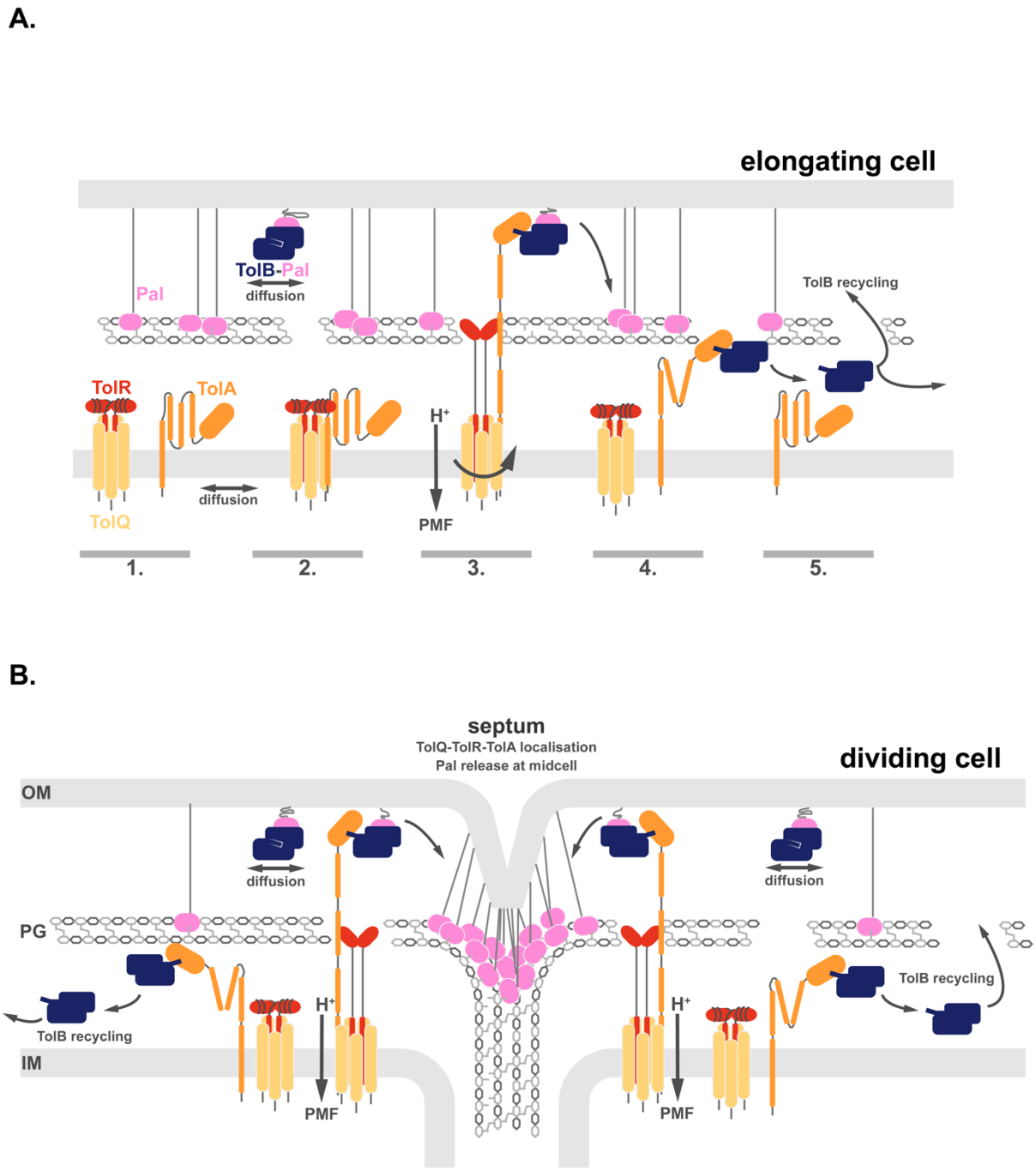


Figure 1.10 Mobilisation and capture mechanism of the Tol-Pal system
A. The Tol-Pal mechanism in a dividing cell. 1. In an elongating cell TolQ-TolR and TolA diffuse freely within the IM. 2. TolQ-TolR interacts with TolA in the IM. 3. Proton flux through TolQ-TolR is transduced to TolA, causing the extension of TolA to bind to TolB-Pal, forming a ternary complex. 4. Loss of protonation results in the relaxation of TolA pulling TolB from Pal, allowing Pal to bind to PG. 5. TolB is now spatially separated from Pal by the PG and is able to diffuse away freely.
B. In an elongating cell the TolQ-TolR and TolA IM machinery is recruited to the septum. As TolB is recycled away to mobilise more Pal it is only captured by the TolQ-TolR-TolA complex at the septum, resulting in the accumulation of Pal at the division site. Adapted from (Szczepaniak, Press and Kleanthous, 2020).

this is the deposition of Pal at the septum, actively creating a tether between PG and the OM as the OM invaginates and new PG is synthesised. TolB is again spatially separated from Pal, allowing it to be recycled to pick up another Pal and effectively bring it to the septum. This mechanism provides an explanation of how the invagination of a non-energised membrane is coordinated with septation in cell division (Figure 1.10).

1.5.3.2 Coordinating peptidoglycan biosynthesis and lipid trafficking

The Tol-Pal system is also involved in the mediation of PG biosynthesis. During cell division, two PG synthase activities coordinate the synthesis of new PG: glycosyltransferases (GTases) polymerise glycan strands and transpeptidases (TPases) crosslink peptides between glycan strands. PBP1B is a penicillin binding protein that has both GTase and TPase activity and is involved in synthesis of new PG at the septum during cell division. LpoB is a regulatory protein that activates both TPase and GTase activity of PBP1B (Egan *et al.*, 2018). CpoB, on the other hand, inhibits only the TPase activity. The binding of TolA to PBP1B reverses inhibition by CpoB, thereby increasing the TPase activity and the degree of crosslinking of PG at the division site (Gray *et al.*, 2015). The relationship between the trimer and monomer of CpoB in these interactions, particularly in relation to the known action of TolA coordinating the oligomeric state of CpoB, is unknown. The inhibitory activity of TolA requires the PMF. Therefore, the full Tol-Pal system is involved in modulating the crosslinking of PG during cell division (Egan *et al.*, 2018).

Along with the role of TolA and CpoB in PG synthesis, Tol-Pal has also been implicated in the process of separating glycan strands in late-stage cell division to allow for the splitting of daughter cells. The cell chaining characteristic of the *tol* phenotype was previously attributed to OM defects. However, it was recently identified that the PG sacculi of *tol* mutants are chained, implying the PG has been incorrectly processed to allow for cell separation (Yakhnina and Bernhardt, 2020). Suppressor mutations that recover the cell chaining phenotype were identified in the protease Prc and its partner MlpI, an enzyme that hydrolyses MepS, which is involved in cell elongation. In addition, the overexpression of PG glycan

cleaving enzymes such as DigH and MltB was found to suppress the *tol* cell chaining phenotype (Yakhnina and Bernhardt, 2020). The exact mechanism of how Tol-Pal mediates processing of PG in late-stage cell division, however, remains to be understood.

Another process Tol-Pal has been linked to is lipid trafficking. *tol* mutants have an accumulation of phospholipids in the outer leaflet of their OM, resulting from impaired retrograde phospholipid transport (Shrivastava, Jiang and Chng, 2017). Overexpression of the Mla system, a lipid transport system between the IM and OM, can partially recover OM asymmetry in *tol* mutants (Shrivastava, Jiang and Chng, 2017). In addition, YbgC is a thioesterase and *⊿ybgC* mutants also have accumulation of phospholipids in the OM (Masilamani, Cian and Dalebroux, 2018). Whether Tol-Pal has a direct role, or the phospholipid accumulation is a result of the destabilised membrane in the *tol* mutant is unclear.

The identification of Tol-Pal's functions in maintaining the OM as the cell invaginates, mediating the synthesis of new PG, and inferred role in lipid biogenesis, all point towards Tol-Pal as an active coordinator of cell division at the septum. Spatially and functionally organising multiple processes across the cell envelope to ensure successful division is carried out.

1.5.3.3 Bacteriocins hijack the Tol-Pal system

The Tol-Pal and Ton systems are both parasitised by protein toxins called bacteriocins. These 'protein antibiotics' are secreted by bacteria to kill one another when in competition for resources (Behrens *et al.*, 2017). Group A bacteriocins target Tol-Pal and group B target the Ton system. The bacterium secreting the bacteriocin is protected from its effects by a complimentary immunity protein. In-depth examination of the import of bacteriocins via the Tol-Pal and Ton systems falls outside the scope of this study, see (Kleanthous, 2010; Atanaskovic and Kleanthous, 2019), which focuses on the native role of Tol-Pal. However, it is important to have an appreciation of these toxins as they can provide a powerful tool in building our understanding of Tol-Pal function.

The group A bacteriocins are first localised to the cell surface by binding to a protein receptor or LPS itself (Kleanthous, 2010). Next, the bacteriocin is translocated through the

OM via the same or a different protein (translocator). The first step of the import process is the insertion of an intrinsically unstructured translocation domain (IUTD) through the translocator to bind to either TolA directly or indirectly via TolB. In the case of colicin E9, the TolB box in the IUTD binds to TolB, which recruits TolA, allowing the colicin to access the PMF (Housden *et al.*, 2013). Mediated by TolB-TolA, the PMF drives the import of the bacteriocin into the periplasm, unfolding it through the pore of the translocator (Vankemmelbeke *et al.*, 2009; Housden and Kleanthous, 2012). The role of the PMF in bacteriocin import is, however, still debated (Cramer, Sharma and Zakharov, 2018), particularly for specific colicins such as colicin E1 shown to kill *ΔtolR E. coli* that cannot access the PMF through Tol-Pal (Lazdunski *et al.*, 1998). Once in the periplasm the onward journey of the bacteriocin depends on its mechanism of killing. Nuclease bacteriocins such as colicin E9 then cross the IM to access their DNA or RNA target. Pore forming bacteriocins including colicin N and colicin A kill by depolarising the IM.

Bacteriocins can also provide a useful tool for studying the biology of the Tol-Pal system. They provide a method of assaying the function of the system in a given mutant as *tol* is resistant to killing by bacteriocins. In addition, engineering of bacteriocins such as colicin E9 to add fluorescent tags or disulfide crosslinks allows for the capture of the assembly machinery, and for investigation into the organisation of Tol-Pal and OMP islands (Rassam *et al.*, 2015; Rassam *et al.*, 2018).

1.6 Aims and Objectives

Tol-Pal's role in maintaining OM integrity and orchestrating cell division machinery at the septum has recently been identified. However, our understanding of the molecular detail of this mechanism is limited. The 'mobilisation-and-capture' mechanism put forward by Szczepaniak *et al.*, 2020 explains how Tol-Pal actively modulates the binding state of Pal through TolB, to accumulate Pal at the septum during cell division (Szczepaniak *et al.*, 2020). Questions remain over the role of the PMF in this mechanism, and how the TolQ-TolR rotary motor translates to the pulling force required to disassociate TolB from Pal. The structures of

the IM components of the system have yet to be determined, and this study set out to characterise the IM proteins through a combination of structural and biophysical approaches.

The first aim was to investigate the role of the elusive helical domain II of TolA using a range of structural, biophysical, and computational methods. The second aim was to purify the complex of TolQ-TolR and confirm the stoichiometry of the complex. The final aim was to gain molecular detail on the interaction between TolA and TolQ-TolR in the IM using a photoactivatable crosslinking approach. The work described in this thesis integrates different structural methodologies to probe our understanding of this challenging system and provides a framework for future studies of the IM components of the Tol-Pal system.

2 Materials and Methods

All reagents purchased from Merck or Fisher Scientific unless stated otherwise. All chromatography columns were purchased from Cytivia.

2.1 Microbiology

2.1.1 Strains, media, and growth conditions

Table 2.1 *E. coli* strains used in this study

Strain	Genotype	Source
DH5α	fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	New England Biosciences (NEB) C2987H
BL21 (DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5	NEB
BW25113	F-, Δ(araD-araB), Δlac(::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB), hsdR514	(Datsenko and Wanner, 2000)
JW0729-3	BW25113 *ΔtolA*	(Baba *et al.*, 2006)

All bacteria were cultured in Luria Bertani/Lysogeny broth (LB) (Tryptone 10 g/l, Yeast Extract 5 g/l, NaCl 10 g/l, pH 7.2) unless stated otherwise. Solid phase cultures were grown on LB broth with 1.5% agar (Melford Biolaboratories) at 37°C in the presence of antibiotic where appropriate. Liquid cultures were inoculated from a single colony and grown aerobically at 37°C, 190 rpm, supplemented with appropriate antibiotics where required, and are referred to as ‘overnight cultures’. Glycerol stocks were maintained at -80°C in 50% LB 25% glycerol (Scientific Laboratory Supplies). Antibiotic selection was carried out using the following antibiotic concentrations; ampicillin (amp) dissolved in MilliQ water, 100 μg/ml, kanamycin (kan) dissolved in MilliQ water, 50 μg/ml, chloramphenicol (cam) dissolved in ethanol, 30 μg/ml.

2.1.2 Plasmids

Table 2.2 Plasmids used in this study

Name	Protein expressed	Description	Resistance	Source
pREN20	*Ec*TolA-His_6	*E. coli* MG1655 TolA cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	Dr Renata Kaminska (unpublished)
pCEP18	*Ec*TolA-II-His_6	*E. coli* MG1655 $TolA_{48-310}$ derivative of pREN20	amp	This study
pCEP19	*Ec*TolA-II+III-His_6	*E. coli* MG1655 $TolA_{48-421}$ derivative of pREN20	amp	This study
pREN155	*Ye*TolA-II-His_6	*Y. enterocolitica* subspecies palearctica 13030 $TolA_{45-291}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	This study Dr Renata Kaminska
pCEP26	*Ye*TolA-II+III-His_6	*Y. enterocolitica* subspecies palearctica 13030 $TolA_{45-362}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	This study
pCEP24	*Pm*TolA-II	*P. mirabilis* HI4320 $TolA_{48-260}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	This study
pCEP25	*Pm*TolA-II+III-His_6	*P. mirabilis* HI4320 $TolA_{48-357}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	This study
pREN156	*ST*TolA-II-His_6	*S. enterica* subspecies *enterica* serovar Typhimurium str. LT2 $TolA_{49-296}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	This study Dr Renata Kaminska
pREN157	*ST*TolA-II+III-His_6	*S. enterica* subspecies *enterica* serovar Typhimurium str. LT2 $TolA_{49-407}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal His_6 tag	amp	This study Dr Renata Kaminska
pCEP40	$TolA_{82-298}$-TEV-His_6	*E. coli* MG1655 $TolA_{82-298}$ cloned into pET21a(+) XhoI and NdeI sites with C-terminal TEV cleavage site and His_6 tag	amp	This study
pCEP41	$TolA_{72-298}$-TEV-His_6	*E. coli* MG1655 $TolA_{72-298}$ cloned into pET21a(+) XhoI and NdeI sites, with C-terminal TEV cleavage site and His_6 tag.	amp	This study
pCEP42	$TolA_{62-298}$-TEV-His_6	*E. coli* MG1655 $TolA_{62-298}$ cloned into pET21a(+) XhoI and NdeI sites, with C-	amp	This study

		terminal TEV cleavage site and His_6 tag.		
pCEP43	$TolA_{92-298}$-TEV-His_6	*E. coli* MG1655 $TolA_{92-298}$ cloned into pET21a(+) XhoI and NdeI sites, with C-terminal TEV cleavage site and His_6 tag.	amp	This study
pT12_*Ec*ExbBD	*Ec*ExbB-ExbD-TEV-strep-strep	*E. coli exbBD*$_{TEV2xstrep}$	kan	(Deme *et al.*, 2020)
pCEP00	*Ec*ExbB-ExbD-TEV-strep-strep	Derivative of pT12_*Ec*ExbBD, XhoI site added at 3' end of *exbD* in ORF	kan	This study
pCEP09	*Ec*TolQ-TolR-TolA-strep-strep	*E. coli* MG1655 TolQ-TolR-TolA ORF cloned into pCEP00 at EcoRI and XhoI sites. C-terminal TEV cleavage sites followed by twin strep tags.	kan	This study
pCEP04	*Ye*TolQ-TolR-TolA-strep-strep	*Y. enterocolitica* subspecies palearctica 13030 TolQ-TolR-TolA ORF cloned into pCEP00 at EcoRI and XhoI sites. C-terminal TEV cleavage sites followed by twin strep tags.	Kan	This study
pCEP08	*Pm*TolQ-TolR-TolA-strep-strep	*P. mirabilis* HI4320 TolQ-TolR-TolA ORF cloned into pCEP00 at EcoRI and XhoI sites. C-terminal TEV cleavage sites followed by twin strep tags.	kan	This study
pCEP05	*ST*TolQ-TolR-TolA-strep-strep	*S. enterica* subspecies *enterica* serovar Typhimurium str. LT2 TolQ-TolR-TolA ORF cloned into pCEP00 at EcoRI and XhoI sites. C terminal TEV cleavage sites followed by twin strep tags.	kan	This study
pCEP11	*Ec*TolQ-TolR-strep-strep	*E. coli* MG1655 TolQ-TolR ORF cloned into pCEP00 at EcoRI and XhoI sites. C-terminal TEV cleavage sites followed by twin strep tags.	kan	This study
pCEP10	*Ye*TolQ-TolR-strep-strep	*Y. enterocolitica* subspecies palearctica 13030 TolQ-TolR ORF cloned into pCEP00 at EcoRI and XhoI sites. C -terminal TEV cleavage sites followed by twin strep tags.	kan	This study
pCEP13	*Ec*TolQ-TolR-strep-strep-His_6	Derivative of pCEP11 with C-terminal His_6 tag inserted through megaprimer mutagenesis	kan	This study
pNP4	GFP-*Ec*TolA	BR322 *E. coli* GFP-TolA	amp	(Gerding *et al.*, 2007)

pEVOL-pBpF		Expression of *Methanococcus jannaschii* tyrosyl-tRNA synthetase pair	cam	(Chin *et al.*, 2002)
pCEP44	*Ec*TolA-His_6	*E. coli* MG1655 TolA with C-terminal His_6 in pNP4 cloned into XbaI and SalI sites	amp	This study
pCEP37	*Ec*TolAI15^{BPA}-His_6	*E. coli* MG1655 TolAI15^{BPA} with C-terminal His_6 in pNP4 cloned into XbaI and SalI sites	amp	This study
pCEP38	*Ec*TolAH22^{BPA}-His_6	*E. coli* MG1655 TolAH22^{BPA} with C-terminal His_6 in pNP4 cloned into XbaI and SalI sites	amp	This study
pCEP39	*Ec*TolAS32^{BPA}-His_6	*E. coli* MG1655 TolAS32^{BPA} with C-terminal His_6 in pNP4 cloned into XbaI and SalI sites	amp	This study

2.2 Molecular Biology

2.2.1 Polymerase Chain Reaction (PCR)

All PCR reagents were purchased from New England Biolaboratories (NEB) unless stated otherwise. PCR was carried out to amplify genes of interest for insertion into a new vector, amplify a megaprimer or for whole plasmid mutagenesis. The typical PCR reaction composition is described in Table 2.3. The cycle conditions used for each reaction are described in Table 2.4. PCR reactions were carried out in an Eppendorf Mastercycler Nexus. Primers used are detailed in Table 2.5. Following PCR, methylated DNA template was removed by digestion with 1 μl DpnI incubated at 37°C for 1-3 hours. PCR amplimers were purified using an NEB PCR clean-up kit as per manufacturer's instructions, eluted with MilliQ water. PCR amplimers were analysed using gel electrophoresis, described in section 2.2.6.

Table 2.3 Typical PCR reaction composition

Reagent	Final volume/amount	Volume (μl)
5 X Phusion HF Buffer	1X	10
dNTPs	0.2 mM	1
Vector	100 μg	
Forward Primer	2 μM	1
Reverse Primer	2 μM	1
Phusion DNA Polymerase	2 units in 50 μl	1
DMSO	10%	5
MilliQ water to total volume 50 μl		

Table 2.4 Typical PCR reaction incubations

Step	Temperature (°C)	Length	Cycle
Denaturation	98	3 min	
Denaturation	98	10s	X 30
Annealing	55	30s	
Extension	72	30 s per kb template	
Extension	72	3-10 min	

Table 2.5 Primers

Number	**Sequence**	**Direction**	**Use**
CP1	GTAACGATCCGGTCACCGATG AAAC	Fwd	Site directed mutagenesis to add XhoI site to pT12*exbBD* to generate **pCEP00**
CP2	GTACAGGTTTTCCTCGAGCTTC GCTTTGGC	Rev	
CP3	GAAATTCAGGAGGAATCCAT GGTGACTGACATGAATATCC	Fwd	InFusion cloning of *E. coli* TolQ-TolR-TolA into pCEP00 to generate **pCEP09**
CP15	TGGAAGTACAGGTTTTCCTCGAG CTCGAGCGGTTTGAAGT	Rev	
CP5	CTGGAAGTACAGGTTTTCCTCGAG TTAGATAGGCTGCGTC	Rev	InFusion cloning of *E. coli* TolQ-TolR into pCEP00 to generate pCEP02
CP6	GAAATTCAGGAGGAATCCGTGGCT GACATGAATGTTATCG	Fwd	InFusion cloning of *P. mirabilis* TolQ-TolR-TolA into pCEP00 to generate **pCEP08**
CP7	CTGGAAGTACAGGTTTTCCTCGAG ATCTTGGCTTAAACTC	Rev	
CP9	GAAATTCAGGAGGAATCCGTGGCT GACATGAACATTCTGG	Fwd	InFusion cloning of *Y. enterocolitica* TolQ-TolR-TolA into pCEP00 to generate **pCEP04**
CP10	CTGGAAGTACAGGTTTTCCTCGAGT TGAGGCTTAAATACA	Rev	
CP11	CTGGAAGTACAGGTTTTCCTCGAGT ATCGGTTGAGTCATT	Rev	InFusion cloning of *Y. enterocolitica* TolQ-TolR into pCEP00 to generate **pCEP10**
CP12	GAAATTCAGGAGGAATCCGTGACT GACATGAATATCCTTG	Fwd	InFusion cloning of *S.* Typhimurium TolQ-TolR-TolA into pCEP00 to generate **pCEP05**
CP13	CTGGAAGTACAGGTTTTCCTCGAGCAG TTTAAAGTCTAGT	Rev	
CP23	AGTACACCGTGGTGGTTGAGA	Fwd	Site directed mutagenesis to remove stop codon from
CP24	TTCCTCGAGGATAGGCTGCGT	Rev	

			pCEP02 to generate **pCEP11**
CP28	AGTCGGTTTTTCTGTTTTAGTG ATGGTGATGGTGATGGT GATGGTGATGTTTTTCGAACTGCG	Rev	Site directed mutagenesis to add His_6 to pCEP11 to generate **pCEP13**
CP49	AAGGAGATATACATATGTCTGAT GGGCAAGTTATTGATGC	Fwd	InFusion cloning of *P. mirabilis* TolA-II and TolA-II+III into pET21a(+) to generate **pCEP24** and **pCEP25**
CP50	GGTGGTGGTGCTCGAGTTGACCT GCTGCAGAAGCCC	Rev	
CP51	GGTGGTGGTGCTCGAGTCTTGGCT TAAACTCAATTGTTGG	Rev	
CP52	AAGGAGATATACATATGGCTGGCGG TGAAGTTATTGACG	Fwd	InFusion cloning of *Y. enterocolitica* TolA-II+III into pET21a(+) to generate **pCEP26**
CP54	GGTGGTGGTGCTCGAGTCTTTGAGG CTTAAATACAAGCGG	Rev	
NdeTolA2Y	GGAATTCCATATGGCTGGCGGTGA AGTTATTGACG	Fwd	PCR amplification of *Y. enterocolitica* TolA-II+III for ligation into pET21a(+) to generate **pREN155**
XhoTolA2Y	GCTCTCGAGTTTACCGGCGGCGGC GGCGCCTGCAC	Rev	
NdeTolA2S	GGAATTCCATATGTCCGCTATCGA CGCGGTGATGGTC	Fwd	PCR amplification of *S.* Typhimurium TolA-II for ligation into pET21a(+) to generate **pREN156**
XhoTolA2S	GCTCTCGAGGGTTTTCGGCGCATT CTTAC	Rev	
NdeTolA2S	GGAATTCCATATGTCCGCTATCGAC GCGGTGATGGTC	Fwd	PCR amplification of *S.* Typhimurium TolA-II+III for ligation into pET21a(+) to generate **pREN157**
XhoTolA3S	CCGCTCGAGCAGTTTAAAGTCTAGTT TGGCATCC	Rev	
CP63	GCCATTGGACTTCAAACCGCTCGAGCA CCACCACCACCACCACTGAGGCTGTTT TGGCGGATGAG	Fwd	Site directed mutagenesis of pBAD*tolA* to add His_6 and generate pCEP28
CP64	CTACCATCGGCGCTACGGCGTTTCA	Rev	
CP33	TTCTCACGCAGTTCTTCAGCAGCCTGCTG	Rev	Site directed mutagenesis to add I15TAG to pCEP28 to make pCEP29.
CP34	GCTCAAGCGGGCGTAGATTATTTCAGCAGT	Fwd	
CP65	GCGGATAACAATTCCCCTCTAGAATG GTGTCAAAGGCAACCGAAC	Fwd	InFusion cloning of *E. coli* TolA-His_6 or TolAI15^{BPA}-His_6 from pCEP28 or pCEP29 into pNP4 to generate **pCEP44** and **pCEP37**
CP66	GTGCCAAGCTTGTCGACTCAGTGGTG GTGGTGGTGGTGCTCGA	Rev	
JSH22XF	TTATTTCAGCAGTGCTGTAGGTCATCTTATTT GCGGCGCT	Fwd	Site directed mutagenesis of pCEP44 to add H22TAG mutation to generate **pCEP38**. pCP33 as Rev primer
CP37	CGCTGATCTGGTAGTCGTTCGATGA	Fwd	Site directed mutagenesis of pCEP44 to add S32TAG mutation to generate **pCEP39**. pCP33 as Rev primer

2.2.2 Restriction digests and ligation

Restriction digestion was performed to linearise vectors. A typical digestion reaction digests 1 μg of vector with 1 μl of restriction enzyme in 1X CutSmart Buffer made up to 50 μl with MilliQ water. All digestion reagents were purchased from NEB. Reactions were incubated at 37°C for 1-3 hours. Digestion products were analysed by gel electrophoresis as described in section 2.2.6 and extracted using the NEB gel extraction kit as per the manufacturer's instructions, eluted with MilliQ water.

Plasmids pREN155, pREN156 and pREN157 were constructed through ligation of restriction digested pET21a(+) and the amplified gene of interest using T4 DNA ligase for 1 hour at room temperature.

2.2.3 InFusion cloning

Genes synthesised by Thermo Fisher Scientific or amplified from a plasmid were cloned into the new vector using the InFusion cloning technique (Takara Biosciences). A typical InFusion reaction is described in Table 2.6. InFusion reactions were incubated at 50°C for 15 min before transformation into competent DH5α cells as described in section 2.2.5.

Table 2.6 Typical InFusion cloning reaction

Reagent	**Volume (μl)**
Vector	To achieve 100 ng
Fragment	To achieve 200 ng
5X InFusion HD Premix	1μl
MilliQ water to 5 μl	

2.2.4 Preparation of chemically competent cells

Overnight cultures of *ΔtolA::kan*, 5 ml, were sub-cultured into LB with kanamycin and incubated at 37°C, 190 rpm, to an optical density at 600 nm (OD_{600}) of 0.3. Cultures were pelleted by centrifugation at 3783 x g, 4°C, 10 min. The pellet was resuspended in cold 20 mM Tris/HCl pH 8.0, 200 mM $CaCl_2$ and incubated on ice for 1 hour. Cells were pelleted by centrifugation at 3783 x g, 4°C, 10 min. The pellet was resuspended in cold 20 mM Tris/HCl pH 8.0, 200 mM $CaCl_2$, 15% (v/v) glycerol, aliquoted, and stored at -80°C.

2.2.5 Bacterial heat shock transformation

For commercial competent cells, a 25-50 μl aliquot was thawed on ice. Plasmid, 1-5 μl, was added to the cells and incubated on ice for 30 min. Cells were heat-shocked at 42°C for 30 s and incubated on ice for 2 min. SOC media (NEB), 400 μl, was added and cells were incubated at 37°C, 190 rpm for 1 hour. The culture, 100 μl, was spread on LB agar with appropriate antibiotic. For transformations following InFusion reactions, after incubation cells were pelleted by centrifugation at 4,000 x g, resuspended in 100 μl SOC media, and spread on LB agar with appropriate antibiotic.

For *ΔtolA::kan* competent cells, a 100-200 μl aliquot was thawed on ice. Plasmid, 1-5 μl was added to the cells and incubated on ice for 30 min. Cells were heat shocked at 42°C for 1 min and incubated on ice for 5 min. SOC media, 1 ml, was added and cells were incubated at 37°C, 190 rpm for 1 hour. Cells were pelleted as described above and plated on LB agar containing 1/10th the working stock of appropriate antibiotic because *tol* cells are more susceptible to antibiotics.

2.2.6 DNA gel electrophoresis

DNA gel electrophoresis was used to assess PCR amplimers and digestion of plasmids. 1% agarose gels (w/v) were prepared in 1 X Tris Borate EDTA (TBE) (89 mM Tris pH 8.3, 89 mM boric acid, 0.5 mM ethylenediaminetetraacetic acid (EDTA)) with 1 X SYBR Safe. 1 kb ladder (NEB) and samples mixed 5:1 with Gel Loading Dye (NEB) were loaded into gel wells and electrophoresis carried out at a constant voltage of 80 V in 1 X TBE running buffer. Gels were imaged in GBOXCHEMI-XRQ using GeneSys software.

2.2.7 Plasmid preparation and sequencing of DNA

Plasmids were purified from DH5α cells grown overnight at 37°C in appropriate antibiotics using the NEB plasmid miniprep kit or Promega midiprep kit as per the manufacturer's instructions, using MilliQ water for elution. Plasmids were sequenced by Genewiz.

2.3 Protein Expression and Purification

2.3.1 Expression and purification of TolA-II and TolA-II+III from *Escherichia coli, Proteus mirabilis, Yersinia enterocolitica* and *Salmonella* Typhimurium

BL21 (DE3) cells containing pCEP18 (*E. coli* $TolA_{48-310}$-His_6), pCEP19 (*E. coli* $TolA_{48-421}$-His_6), pCEP24 (*P. mirabilis* $TolA_{48-260}$-His_6), pCEP25 (*P. mirabilis* $TolA_{48-357}$-His_6), pREN155 (*Y. enterocolitica* $TolA_{45-291}$-His_6), pCEP26 (*Y. enterocolitica* $TolA_{45-362}$-His_6), pREN156 (*Salmonella* Typhimurium $TolA_{48-296}$-His_6) or pREN156 (*Salmonella* Typhimurium $TolA_{48-407}$-His_6) were grown overnight at 37°C, 190 rpm. The overnight cultures, 8 ml, were used to inoculate 800 ml LB supplemented with appropriate antibiotics and grown to an OD_{600} of approximately 0.6. Cultures were induced with 1 mM isopropyl ß-D-1-thiogalactopyranoside (IPTG) and incubated overnight at 20°C. Cells were harvested by centrifugation at 3,378 x g (Beckman High Speed Avanti JXN-26), 20°C, 12 min. Pellets were stored at -20°C.

Cells were thawed and resuspended in Lysis Buffer (20 mM Tris/HCl pH 8.0, 100 mM NaCl, 5 mM imidazole, 1 mM phenylmethanesulfonylfluoride (PMSF)). Cell lysis was carried out by sonication, using a Misonix S-4000 ultrasonic processor with a 3/8-inch stud probe and a programme of 70% maximum amplitude for 3 s and 7 s off, for 3 min total sonication. The programme was run twice per sample and sonication carried out on ice.

Lysed cells were centrifuged at 12,500 x g, 4°C, 30 min. The soluble fraction was filtered through a 0.44 μm filter and hand-loaded onto a HisTrap FF 5 ml column that had been equilibrated in 2 column volumes (CV) of MilliQ water followed by 2 CV Lysis Buffer. The column was washed with 4 CV Wash Buffer (20 mM Tris/HCl pH 8.0, 100 mM NaCl, 50 mM imidazole) and protein step eluted with 100% Elution Buffer (20 mM Tris/HCl pH 8.0, 100 mM NaCl, 500 mM imidazole) collecting 1 ml fractions of eluent and using absorbance at 280 nm (A280) to monitor the elution. Fractions corresponding to absorbance peaks were analysed by SDS-PAGE, see section 2.3.6, to ascertain where TolA had eluted. Fractions containing the relevant protein were pooled and dialysed into Dialysis Buffer (20 mM Tris/HCl pH 8.0,

100 mM NaCl) overnight at 4°C using dialysis tubing (Enzo Life Sciences) with molecular weight cut-off (MWCO) 12-14 kDa. Any precipitated material was pelleted by centrifugation at 16,000 x g, 18°C, 10 min, and sample was applied to a HiLoad 26/600 Superdex 200 PG or HiLoad 26/600 Superdex 75 PG size exclusion chromatography (SEC) column equilibrated in SEC Buffer (20 mM Tris/HCl pH 8.0, 100 mM NaCl). SEC buffer, 1.5 CV, was used to elute proteins present and elution was monitored by A280 and A220. Fractions were analysed by SDS-PAGE and fractions containing TolA-II or TolA-II+III were pooled and centrifuged at 15,000 x g, 18°C to pellet any precipitated material. The sample was concentrated using VivaSpin 20 with MWCO of 10 kDa, then either immediately used for further analysis or snap frozen in liquid nitrogen and stored at -20°C. For *Y. enterocolitica* TolA-II and *S.* Typhimurium TolA-II and TolA-II+III the expression and purification steps were carried out by Dr Renata Kaminska.

2.3.2 Purification of truncated *E.* coli TolA-II constructs

The constructs of *E. coli* $TolA_{62-298}$, $TolA_{72-298}$, $TolA_{82-298}$, and $TolA_{92-298}$ were expressed in BL21 (DE3) cells as described above. Purification of the truncated versions of TolA was carried out as described above, in buffers at pH 7.5, with the following additional steps:

1. During dialysis, the His_6 tag was removed by digestion with 300 μg Tobacco Etch Virus (TEV) protease, purified in-house. A second nickel affinity chromatography step was carried out to remove the TEV protease, His_6 tag and any non-cleaved TolA construct. The sample was loaded onto a HisTrap FF 5 ml column, washed with 2.5 CV SEC Buffer, collecting 1 ml fractions of the flow-through. Bound TEV and un-cleaved TolA was eluted with 2 CV Elution Buffer. The presence of the relevant protein in the flow through was identified using analysis by SDS-PAGE. The SEC purification step was carried out as described above. 2. Following SEC and analysis by SDS-PAGE, the sample was pooled into a weighed 50 ml centrifuge tube and snap frozen in liquid nitrogen. Sample was freeze-dried in a VirTis Bechtop K lyophiliser until only powder present, then stored at -20°C until required.

2.3.3 ^{15}N labelling of TolA constructs

Constructs of *E. coli* TolA-II, $TolA_{62-298}$, $TolA_{72-298}$, $TolA_{82-298}$, and $TolA_{92-298}$ were labelled with ^{15}N for use in Nuclear Magnetic Resonance Spectroscopy (NMR). Overnight cultures of BL21 (DE3) with the appropriate plasmid were inoculated 8 ml into 800 ml LB with amp and grown at 37°C, 150 rpm to an OD_{600} of 1. Cells were pelleted by centrifugation at 3,378 x g, 20°C, 12 min and resuspended in 20 ml M9 media (42.3 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.5 mM NaCl, 1 g/l NH_4Cl, 0.4% (w/v) glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, pH 7.0). Cells were pelleted by centrifugation at 4,000 x g, 18°C, 10 min and resuspended in 800 ml ^{15}N labelled M9 media, where $^{15}NH_4Cl$ replaces NH_4Cl. Cells were then grown for 1 hour at 37°C, 150 rpm before induction with 1 mM IPTG and expression overnight at 20°C. Cells were harvested by centrifugation at 3,378 x g, 20°C, 12 min.

Purification of ^{15}N *E. coli* TolA-II, $TolA_{62-298}$, $TolA_{72-298}$, $TolA_{82-298}$, and $TolA_{92-298}$ carried out as described in section 2.3.2. For $TolA_{62-298}$, $TolA_{72-298}$, $TolA_{82-298}$ and $TolA_{92-298}$, SEC buffer was substituted with NMR Buffer (20 mM Sodium Phosphate Buffer pH 7.0, 100 mM NaCl). TolA-II was dialysed into NMR Buffer in dialysis tubing 12-14 kDa MWCO, overnight at 4°C following purification.

2.3.4 Purification of TolA with BPA mutations

Pull-downs of TolAHis22^{BPA} were performed with the aim of identifying crosslinked adducts. Overnight cultures, 10 ml, of *ΔtolA::kan* with pCEP38 (pNP4*tolA*His22^{BPA}-His_6) and pEVOL-pBpF or the *ΔtolA::kan* with pCEP44 (pNP4*tolA*-His_6) control were inoculated 8 ml into 800 ml LB with 0.5 mM p-benzoyl-L-phenylalanine (BPA) dissolved in 1.3 M NaOH, 1.3 M HCl to match pH, amp and cam and grown at 37°C to OD_{600} 0.6. IPTG, 1 mM, and 0.125% arabinose were added to induced expression and incubated overnight at 20°C. Cells were harvested by centrifugation at 3,378 x g, 20°C, 12 min and pellets resuspended in 6-8 ml Lysis Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 5 mM imidazole). Resuspended cells were poured into 10 x 10 cm square petri dishes and exposed to UV at 365 nm for 90 min in a CL-1000 UV crosslinker, with mixing after 45 mins. Following UV exposure cells were stored at -

20°C. Samples of cell lysates from before and after induction for each condition were taken, heated at 98°C for 20 min and analysed by SDS-PAGE.

Cells were thawed and a cOmplete EDTA free protease inhibitor tablet added. Lysis was carried out using a Misonix S-4000 ultrasonic processor with a 3/8-inch stud probe and a programme of 70% maximum amplitude for 3 s and 7 s off, for 3 min total sonication. Lysed cells were centrifuged at 7,500 x g for 15 min and the supernatant containing soluble material ultracentrifuged at 200,000 x g, 4°C, 45 min. The pelleted membrane fraction was homogenised in Solubilisation Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 5 mM imidazole, 2% β-OG) and incubated for 2 hours at 4°C. Following this, the solubilised material was clarified by ultracentrifugation 200,000 x g, 45 min, 4°C. The solubilised inner membrane fraction was stored at 4°C.

Nickel affinity pulldowns were performed of both the soluble material from the first ultracentrifugation step, and the solubilised material from the second ultracentrifugation step, to identify whether TolAHis22^{BPA} was present in the IM. Samples were filtered through a 0.44 μm filter and applied to a gravity column containing cOmplete nickel resin with a bed volume of 0.9 ml, equilibrated in 5 x CV MilliQ and 5 x CV Lysis Buffer. The flow-through was collected and 5 CV Lysis Buffer, then 5 CV Wash Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 50 mM imidazole) applied to wash the column. Sample was eluted with 5 CV Elution Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 500 mM imidazole). Samples of the flow-through and resin were taken for analysis by SDS-PAGE and western blot at each step.

Further purification steps were also carried out comparing UV-crosslinked TolAHis22^{BPA} with non-UV TolAHis22^{BPA} and UV-TolA controls. Nickel affinity chromatography was carried out using a 5 ml HisTrap FF column and Wash Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 50 mM imidazole, 1% β-OG) and gradient elution with Elution Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 500 mM imidazole, 1% β-OG) over 15 ml. Elution peaks were pooled and desalted using a 5 ml HiTrap desalting column in SEC Buffer (20 mM Tris/HCl pH 7.5, 100 mM NaCl, 1%β-OG). Sample following desalting was pooled and loaded onto a

HiLoad 16/600 Superdex 200 and eluted over 1.5 CV in SEC Buffer. Elution peak fractions were pooled and concentrated in a Millipore concentrator with a MWCO of 10 kDa. Samples were taken following each purification step and of each elution peak for analysis by SDS-PAGE and western blot.

2.3.5 Expression and purification of TolQ-TolR

TolQ-TolR-2xstrep-His_6 was expressed in BL21 (DE3) cells on the pCEP13 plasmid. Overnight cultures, 5 ml, of BL21 (DE3) pCEP13 were inoculated into 50 ml LB with appropriate antibiotic and grown to OD_{600} 1. BL21 (DE3) pCEP13 was sub-cultured 20 ml into 1 L Terrific Broth (24 g/l Yeast Extract, 20 g/l Tryptone, 4 ml glycerol, 10 mM phosphate buffer pH 7.4) with 1 g/l rhamnose and appropriate antibiotic. Cultures were induced overnight at 20°C. Cells were harvested by centrifugation at 3,378 x g, 20°C, 12 min and pellets stored at -20°C.

Cell pellets were thawed and resuspended in Lysis Buffer (50 mM Tris/HCl pH 7.4, 300 mM NaCl, 5 mM imidazole) with a cOmplete EDTA free protease inhibitor tablet. Lysis was carried out by sonication using a Misonix S-4000 ultrasonic processor with a 3/8-inch stud probe and a programme of 70% maximum amplitude for 3 s and 7 s off, for 3 min total sonication. Lysed cells were centrifuged at 7,500 x g, 4°C, 15 min and the supernatant containing the soluble material ultracentrifuged 200,000 x g, 4°C, 45 min.

The pelleted whole cell membrane fraction was homogenised in Solubilisation Buffer (50 mM Tris/HCl pH 7.4, 300 mM NaCl, 5 mM imidazole, 0.5% Cymal-6) and incubated for 2 hours at 4°C. For the solubilisation detergent screen Solubilisation Buffer contained 0.5% Cymal-6, 1% DDM, 1% DDM with 0.25% cholesterol hemisuccinate or 0.1% LMNG. The solubilised material was ultracentrifuged at 200,000 x g, 4°C, 45 min. The solubilised inner membrane fraction was stored at 4°C.

The inner membrane fraction was filtered through a 0.44 μm filter and loaded onto a HisTrap FF 5 ml column equilibrated in MilliQ water then Solubilisation Buffer. The column was washed with 2 CV Wash Buffer (50 mM Tris/HCl pH 7.4, 300 mM NaCl, 50 mM imidazole,

0.1% Cymal-6) and eluted by gradient elution with Elution Buffer (50 mM Tris/HCl pH 7.4, 300 mM NaCl, 500 mM imidazole, 0.1% Cymal-6) taking 1 ml fractions of eluent. The fractions were analysed to identify TolQ-TolR using SDS-PAGE, described below.

The purified TolQ-TolR sample was desalted to remove the imidazole present. Sample was loaded onto a 5 ml HiTrap desalting column equilibrated in Desalt Buffer (50 mM Tris/HCl pH 7.4, 300 mM NaCl, 0.1% Cymal-6) and eluted over 1.5 CV. Following desalting, TolQ-TolR was incubated with 300 μg TEV protease overnight at 4°C. The sample was loaded onto a HisTrap FF 5 ml column equilibrated in Desalt Buffer and washed with 2 CV Desalt Buffer. Elution fractions, 1 ml, were collected from the flow through and wash. Fractions containing TolQ-TolR were assessed by SDS-PAGE. Fractions containing cleaved TolQ-TolR were evaluated by SDS-PAGE then pooled and concentrated using a VivaSpin concentrator MWCO 50 kDa.

The TolQ-TolR was loaded onto a HiLoad 16/600 Superdex 200 PG column equilibrated in Desalt Buffer and eluted over 1.5 CV. Fractions containing TolQ-TolR were assessed by SDS-PAGE then pooled and TolQ-TolR was concentrated using a VivaSpin concentrator MWCO 100 kDa. Purifications of TolQ-TolR using LMNG for the purpose of cryo-EM substituted Cymal-6 for 1% LMNG in Solubilisation Buffer, and 0.01% LMNG in all other relevant buffers.

2.3.6 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to identify proteins during purification and assess the purity of proteins. Samples were diluted 1:3 with SDS Sample Buffer (200 mM Tris/HCl pH 8.0, 8% w/v SDS, 0.4 w/v bromophenol blue, 40% glycerol, 400 mM β-mercaptoethanol), heated at 98°C for 2 min, and loaded onto 12% Bis-Tris gels in Hoefer Mighty Small SE250 gel tanks. Electrophoresis was carried out at a constant amperage of 30 mA in running buffer (25 mM Tris, 192 mM glycine, 3.5 mM SDS). Gels were stained with 0.2% (w/v) Coomassie Brilliant

Blue R-250, de-stained with 10% acetic acid and 10% ethanol and imaged in GBOXCHEMI-XRQ using GeneSys software.

SDS-PAGE with TolQ-TolR was carried out using tricine gels (reagents purchased from Invitrogen). Samples were diluted 1:1 with Sample Buffer (450 mM Tris/HCl pH 8.45, 12% glycerol, 4% SDS, 0.00075% Coomassie Blue G, 0.0025% phenol red), heated for 2 min at 98°C and loaded onto 10-20% Novex Tricine gels in a Mini Gel Tank. Electrophoresis was carried out at a constant voltage of 125 V for 1 hour in Tricine Running Buffer (100 mM Tris/HCl pH 8.3, 100 mM tricine, 0.1% SDS). Gels were Coomassie stained and imaged as described above, with 1 hour incubation steps.

2.3.7 Blue Native-PAGE

Blue Native-PAGE was used to look for the TolQ-TolR complex in different detergents (reagents purchased from Invitrogen, catalogue numbers detailed here). Protein sample was diluted to 10 μM in appropriate SEC buffer that matched the protein's current buffer, 5% Coomassie G-250 was added to achieve a concentration of ¼ the detergent concentration and the sample incubated for 5 min at room temperature. 4X Native-PAGE sample buffer (BN2003) added, 3 μl. The 1 X Running Buffer (400 ml) was prepared from 20 X Running Buffer stock (BN2001) with MilliQ water. The 1 X Dark Blue Cathode Buffer (100 ml) was prepared from 20 X Cathode Additive (BN2002) with 20X Running Buffer Stock using MilliQ water to dilute. Buffers were cooled to 4°C prior to use. Samples, 8 μl, were loaded onto a Native-PAGE Novex 4-16% Bis-Tris protein gel alongside native mark unstained protein ladder. Electrophoresis was carried out at a constant voltage of 150 V for 1 hour, then 250 V for 1 hour. Coomassie staining as described in 1.3.5, increased to 1 hour.

2.3.8 LPS staining

Pro-Q Emarald 300 LPS Gel Stain (Invitrogen) was used to look for LPS contamination in TolQ-TolR samples destined for native MS. LPS sample control prepared by dissolving 2.5 mg in 1 ml 50 mM Tris/HCl pH 7.4, 300 mM NaCl, 0.1% Cymal-6. Pro-Q Emerald 300 Stock Solution prepared by dissolving Pro-Q Emerald 300 in DMSO as per manufacturer's

instructions. Samples were prepared and run on 10-20% Tricine gels as described in 1.3.5. Following electrophoresis, gels were stained following manufacturer's protocol for Pro-Q Emerald 300 LPS stain, all incubation steps carried out at room temperature with agitation. The gel was fixed with Fix Solution (50% methanol, 5% acetic acid) for 45 min, followed by repetition with fresh Fix Solution. Next, it was washed with Wash Solution (3% acetic acid) for 3 X 15 min, then incubated in Oxidising Solution (1% w/v periodic acid) for 30 min. This was followed by 2 X washes with Wash Solution for 15 min. Staining was then carried out with Staining Solution (500 μl Pro-Q Emerald 300 stock solution in 25 ml Staining Buffer from kit) for 2 hours in the dark. Two final washes were performed with Wash Solution each for 15 min, followed by imaging in GBOXCHEMI-XRQ using GeneSys software.

2.3.9 Protein quantification

Protein concentration was calculated from absorbance at 280 nm and theoretical molecular weights and extinction coefficients estimated through ExPasy ProtParam (Table 2.7). For calculations regarding TolQ-TolR, only the extinction coefficient of TolQ was used. For proteins concentrated through lyophilisation, the protein present was calculated by weighing its container before addition of protein and after freeze-drying.

The protein mass of purified $TolA_{72-298}$ was validated by denaturing electrospray ionisation mass spectrometry (ESI-MS) by Dr David Staunton (Molecular Biophysics Suite, Department of Biochemistry, University of Oxford).

Table 2.7 Molecular weights and extinction coefficients of proteins purified in this study determined by ExPasy ProtParam

Protein	Molecular Weight (Da)	Extinction Coefficient (M^{-1} cm^{-1})
*Ec*TolA-II-His6	28190	1490
*Ec*TolA-II+III-His6	39440	7450
*Pm*TolA-II-His6	23844	1490
*Pm*TolA-II+III-His6	34135	5960
*Ye*TolA-II-His6	26606	1490
*Ye*TolA-II+III-His6	36952	7450
*ST*TolA-II-His6	26737	1490
*ST*TolA-II+III-His6	38083	7450
$EcTolA_{62-298}$	25481	2980
$EcTolA_{72-298}$	24110	1490
$EcTolA_{82-298}$	22964	1490
$EcTolA_{92-298}$	21791	1490
*Ec*TolQ	25696	23950
*Ec*TolR	15382	2980

2.3.10 Peptide mass fingerprinting

The identity of proteins within SDS-PAGE bands of interest was confirmed using peptide mass fingerprinting. The protocol was as described in Webby *et al.* 2022 and carried out by Dr Melissa Webby (Webby *et al.*, 2022b). Briefly, excised gel bands were destained with ammonium bicarbonate/acetonitrile, dehydrated with acetonitrile, reduced with tris(2-carboxyethyl)phosphine (TCEP) and cysteines alkylated with chloroacetamide. Protein was digested overnight with 5 μg sequencing grade modified trypsin (Promega) at 37°C. A Q Exactive mass spectrophotometer was used to analyse the peptides, and the protein ID determined through searching the *E. coli* database in MaxQuant (Cox and Mann, 2008).

2.4 Biophysical Techniques

2.4.1 Circular Dichroism

Protein samples analysed by Far UV Circular Dichroism (CD) were dialysed into 20 mM potassium phosphate buffer pH 7.5, 100 mM NaF. The concentration of each protein analysed is detailed in Table 2.8. CD was measured on a Jasco J-815 Spectropolarimeter between 250 nm and 190 nm at 0.5 nm intervals, taking an average from 9 accumulations at a scanning speed of 100 nm/min. Mean residue ellipticity was calculated by dividing the

baseline corrected spectra (mdeg) by the molar concentration multiplied by the path length (1 mm) and number of residues.

2.4.2 Nanoscale Differential Scanning Fluorimetry

The melting temperature of protein samples was measured in a Prometheus NT.48 (NanoTemper Technologies). Protein samples at the concentrations described in Table 2.8. Samples were excited at 280 nm between 20°C and 95°C at 1 min/1°C intervals. The emission spectra at 330 nm and 350 nm were measured and the first derivative of 330 nm/350 nm plotted against temperature to identify the melting temperature.

Table 2.8 Protein concentrations used for CD and nanoDSF

Protein	Concentration for CD (μM)	Concentration for nanoDSF (μM)
E. coli TolA-II	99.3	85
E. coli TolA-II+III	12.7	25
P. mirabilis TolA-II	55.7	50
P. mirabilis TolA-II+III	2.2	527
Y. enterocolitica TolA-II	8.5	285
S. Typhimurium TolA-II	14.5	50
S. Typhimurium TolA-II+III	10.0	50
E. coli $TolA_{62-298}$	31.2	314
E. coli $TolA_{72-298}$	24.8	200
E. coli $TolA_{82-298}$	10.9	50
E. coli $TolA_{92-298}$	39.6	1239

2.4.3 Size Exclusion Chromatography Multi Angle Light Scattering (SEC-MALS)

The molecular mass and therefore oligomeric state of TolA constructs were determined by SEC-MALS. Samples were separated in Running Buffer (20 mM Tris/HCl pH 8.0, 100 mM NaCl) on a Superdex S200 HR10/30 column. Light scattering was detected by a Wyatt Dawn HELEOS-II 8-angle light scattering detector and Wyatt Optilab rEX refractive index monitor linked to a Shimadzu HPLC system. SEC-MALS was performed by Prof David Staunton. The mass averaged molar mass was determined from refractive index and light scattering using ASTRA software.

2.5 Crystallography

2.5.1 Crystallisation

Table of crystal plates set up detailed in Supplementary Figure 1. Crystallisation trials using the JCSG+, Morpheus, Morpheus2, MIDASplus, PACT, MemGold2, Proplex and STURA screens (Molecular Dimensions) were set up in a 96-well MRC 2-drop vapour diffusion plate (SWISSCI) using a Mosquito robot (TTP Labtech). Drops contained a 1:1 ratio of protein to reservoir, or when seed stock added, 10:2:10 protein:seed stock:reservoir. Pre-crystallisation tests (Hampton) were conducted according to the manufacturer's instructions to examine the ideal concentration of protein required for crystallography. Seed stock was prepared by transferring crystallisation drop and reservoir solution to a seed bead 1.5 ml tube (Hampton) and vortexing. Plates were stored and imaged at 18°C or 4°C in a JANSi UVEX PS-640 crystallisation hotel. UV and brightfield images were taken using Crystal Detect software viewed through Webview.

Reductive methylation was used to aid crystallisation of $TolA_{72-298}$. $TolA_{72-298}$ was diluted to 8.7 mg/ml and dialysed into 20 mM HEPES pH 7.5, 100 mM NaCl. 1 M dimethylamine-borane complex (ABC), 20 μl, and 40 μl 1 M formaldehyde were added to 500 μl $TolA_{72-298}$ 8.7 mg/ml and incubated in the dark at 4°C for 2 hours. The addition of ABC and formaldehyde and incubation steps were repeated twice, with the final incubation overnight. 50 mM Tris/HCl pH 8.0 was added 1:1 to the methylated $TolA_{72-298}$ to quench the reaction then centrifuged 16,000 x g 5 min to pellet precipitated material. Methylated $TolA_{72-298}$ was dialysed into MilliQ water, lyophilised, and resuspended in SEC buffer to achieve 137 mg./ml as previously described.

Following identification and screening of crystals in STURA condition H8 (0.1 M sodium citrate pH 5.5, 20% PEG 4000, 20% 2-propanol) optimisation plates were set up to improve crystal resolution (Supplementary Figure 1). The first screen varied pH between 4.5 and 6.25, PEG 4000 18-24% or isopropanol 18-24%, using stock reagents, 1 M sodium citrate pH 4.5, 1 M sodium citrate pH 6.5, 50% PEG 4000 (Molecular Dimensions), 99.9% 2-propanol.

The second added the Hampton Silver Bullet Additive Screen to sitting drops at the ratio 10:5:10 protein:additive:reservoir. The third added condition H8 (Molecular Dimensions) to 150 mg/ml $TolA_{72-298}$ resuspended in MilliQ water, 20 mM HEPES pH 7.5, 20 mM HEPES pH 7.5 100 mM NaCl, or resuspended in MilliQ water and crosslinked with 0.1% glutaraldehyde for 15 min before quenching with 20 mM Tris/HCl pH 8.0.

2.5.2 Crystal screening

Crystals were harvested and flash cooled in liquid nitrogen prior to screening. Crystals were screened at SOLEIL beamline Proxima 2A. Screening for further crystals was carried out at Diamond Light Source on the beamlines I03 and I04. Diffraction data was visualised using dials.image_viewer. Crystal harvesting and screening carried out by Dr Edward Lowe.

2.6 Nuclear Magnetic Resonance Spectroscopy

The ^{15}N-labelled TolA constructs analysed by NMR spectroscopy were first dialysed into NMR Buffer (20 mM Phosphate Buffer pH 7.0, 100 mM NaCl) at 4°C overnight. The following concentrations of each ^{15}N-labelled TolA construct was used: TolA-II 110 μM, $TolA_{72-298}$ 18 μM, $TolA_{82-298}$ 54 μM, $TolA_{62-298}$ and $TolA_{92-298}$ 110 μM. D_2O was added to achieve 5% (v/v) in the sample, and the sample was then centrifuged at 10,000 x g for 10 min to pellet any precipitated material.

NMR experiments were collected using a Bruker 750 Hz spectrometer at 20°C. For TolA-II, $TolA_{62-298}$, $TolA_{82-298}$, and $TolA_{92-298}$, standard 2D and 3D double resonance data was collected, including 2D BEST-TROSY, 3D TOCSY-HSQC, as well as ^{15}N HetNOE to examine protein dynamics. For $TolA_{72-298}$ only 2D BEST-TROSY data was collected due to low concentration of material. NMR data processing was conducted using NMRPipe (Delaglio *et al.*, 1995) and processed spectra visualised and assigned using CCPN software suite (Vranken *et al.*, 2005). NMR experiments and processing conducted by Prof Christina Redfield.

2.7 Electron Microscopy

TolQ-TolR in 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 2 mM EDTA, 0.01% LMNG was applied to 1.3/1.2, 300 mesh Cu grids prepped with either glow discharge at 20 mAu for 45 s or plasma cleaning. Vitrification was carried out using Vitrobot Mark VI at 100% humidity for 3 s blot with -10 force and a wait time of 3 s, at 8°C.

Data were collected on a 200 kV Talos Arctica microscope (Thermo Fisher) at magnification of 190 K, pixel size 0.74 Å and a dose rate of 10.48 e/Å/s. A collection of 1468 movies were aligned and CTF estimation carried out in SIMPLE before further processing in CryoSparc to pick and extract particles used to generate 2D classes and 3D volumes (Punjani *et al.*, 2017; Caesar *et al.*, 2020; Punjani, Zhang and Fleet, 2020). Grid preparation and analysis carried out by Dr Melissa Webby.

2.8 Computational Methods

2.8.1 Model Building

A model of TolQ-TolR-TolR was generated by Dr Melissa Webby using AlphaFold and the Rosetta suite (Maguire *et al.*, 2021; Khatib *et al.*, 2011). A pentamer of *E. coli* TolQ (P0ABU9) was made from the AlphaFold2 model of the monomer (Jumper et al., 2021) through symmetry expansion in Rosetta. The dimer of full-length TolR was made from structural alignment of the *E. coli* AlphaFold monomer with the *H. influenzae* dimer (PDB: 2JWL) (Parsons, Grishaev and Bax, 2008). The TolR dimer was then aligned to the trans-pore helices of ExbD in the ExbB-ExbD complex (PDB: 6TYI) (Celia *et al.*, 2019). Manual rigid body movements of flexible linker regions were carried out to reduce obvious steric clashes. Rosetta Relax was run for refinement and the best scoring model taken forward (Nivón, Moretti and Baker, 2013). The first 100 residues of *E. coli* TolA from the AlphaFold2 model were taken and the transmembrane helix placed within 10 Å of TolQ-TolR to run Rosetta Protein-Protein Docking (Marze *et al.*, 2018). The best scoring models were refined in Rosetta Relax and the

best model following this was taken as the TolQ-TolR-TolA model and analysed in UCSF ChimeraX (Pettersen *et al.*, 2021).

2.8.2 Model Analysis

The electrostatic surface of the *E. coli* AlphaFold2 model (P19334) was calculated, and measurements of the length of its helices were made in ChimeraX (Pettersen *et al.*, 2021). The AlphaFold2 models of TolA from 13 different species (Table 2.9) and *E. coli* TonB (P02929) were taken from the AlphaFold database, https://alphafold.ebi.ac.uk (Jumper *et al.*, 2021) and coloured by pLDDT score and electrostatic surfaces calculated in ChimeraX. Secondary structure alignments were carried out in Pymol (Delano, 2010). Primary sequence alignment was conducted using T-COFFEE, https://www.ebi.ac.uk/Tools/msa/tcoffee/ (Notredame, Higgins and Heringa, 2000).

The MOLE 2.5 online tool, https://mole.upol.cz/, was used to predict channels in the TolQ-TolR model (Pravda *et al.*, 2018). The bottleneck radius for channel prediction was set to 1 Å and Leu175 omitted from calculations. The channel best fitting the mutational data was visualised using UCSF Chimera (Pettersen *et al.*, 2004).

Table 2.9 TolA sequences and AlphaFold2 models analysed.

Class	Species	Uniprot Code
γ-proteobacteria	*Escherichia coli*	P19934
	Haemophilus influenzae	P44678
	Klebsiella pneumonia	A0A0H3GU6
	Pseudomonas aeruginosa	P50600
	Proteus mirabilis	B4ESS3
	Salmonella Typhimurium	Q8ZQT6
	Shigella dysentriae	Q32II2
	Serretia marcescens	A0A2S4XFN9
	Yersinia enterocolitica	A0A0H3NMM3
β-proteobacteria	*Burkholder ambifaria*	Q0BHY6
α-proteobacteria	*Caulobacter vibrioides*	A0A258D9E8
δ-proteobacteria	*Desulfovibrio vulgaris*	A0A0H3A4A9
ε-proteobacteria	*Campylobacter ureolyticus*	A0A0K1HEI9

2.8.3 Molecular Dynamics Simulations

The *E. coli* AlphaFold2 model (P19934) was first manipulated to angle TolA-I α1 away from α2 using MODELLER 9.16. The loop residues IEASAGGGGGSSIDAV were cut and remodelled. This model was then coarse-grained using the *martinize.py* tool. Coarse-grained molecular dynamics (CGMD) simulations were conducted using Gromacs 2021.3 with the Martini 2.2 forcefield. TolA was embedded in the IM (90% POPE, 5% POPG, 5% cardiolipin) using the *insane* tool (Wassenaar et al 2015). A coarse-grained form of PG layer, as constructed by Boags *et al.* 2019 was positioned such that TolA penetrates through one if its pores but not contacting it. For the simulation of TolA in the IM only, the α3 helix of TolA-II was positioned 85° from α2 and the simulation run for a total of 2000 ns. For the simulation of TolA in IM with PG, PG was positioned 15 nm from the IM, α3 positioned 40° from α2, and the simulation run for a total of 80 ns. Each simulation was solvated with the Martini water model and neutralised with 0.15 M NaCl at 313 K. Three independent simulations were run for each system (with or without PG). CGMD simulations conducted by Dr Dheeraj Praakesh.

2.9 Native Mass Spectrometry

Native mass spectrometry (native MS) carried out by Dr Tarick El-Baba. In initial native MS experiments TolQ-TolR (20-100 μM) in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 0.1% Cymal-6 was buffer exchanged into 200 mM ammonium acetate pH 7.4, 0.06% Cymal-6 (MS buffer). Further detergent screening altered the MS buffer, replacing 0.06% Cymal-6 with either 0.02% DDM, 0.5% tetraethylene glycol monooctyl ether or 0.04% custom detergent G1 (Urner *et al.*, 2020).

Native mass spectrometry was carried out on a Thermo Q-Exactive UHMR Quadrupole-Oritrap platform. Sample, 3 μl, was loaded into a gold-coated borosilicate electrospray ionisation capillary and mounted on the nano electrospray ionisation source. Typical instrument setting based upon Housden *et al.*, 2021: Capillary voltage 1.0-1.2 kV, resolution 12,500 (at *m/z* 200), S-lens RF 150%, 200 ms fixed injection time, 200-250 V HCD

energy at pressure setting 8.0. Data was visualised and processed using OriginPro v. 9.5.1.195 (Housden *et al.*, 2021).

2.10 Assays to investigate Tol-Pal function

2.10.1 Colicin toxicity assay

Overnight cultures, 5 ml, of strains detailed in Table 2.10 were sub-cultured into LB, 5 ml, with appropriate antibiotics. Two sub-cultures per strain were set-up, one containing 0.5 mM BPA and (dissolved in 1.3 M NaOH) with 1.3 M HCl to match pH, and a second without BPA. Sub-cultures were grown to OD_{600} 0.6, induced with 0.125% arabinose and 1 mM IPTG, then grown for a further 1 hour.

Subculture, 2.5 ml, was added to soft agar (50% LB, 50% LB agar) containing the required antibiotics, 0.125% arabinose, 1 mM IPTG, and 0.5 mM BPA with 1.3 M HCl if required. The soft agar containing the culture was mixed gently and poured onto agar plates containing the same additives. Once set, colicin E9, purified in-house (Housden *et al.*, 2005; Housden *et al.*, 2021), was serially diluted from 10 μM to 0.2 μM, spotted, 2 μl, onto the plate and dried. Plates were incubated overnight at 37°C and imaged using white light in GBOXCHEMI-XRQ using GeneSys software. Cytotoxicity was determined by examination of the zones of killing where colicin E9 was spotted.

Table 2.10 Strains and plasmids used in colicin killing assays

Number	Strain	Plasmids
1	BW25113 (WT Keio)	
2	JW0729-3 (*ΔtolA::kan*)	
3	JW0729-3 (*ΔtolA::kan*)	pCEP44 (pNP4tolA-His_6)
		pEVOL-pBbF
4	JW0729-3 (*ΔtolA::kan*)	pCEP37 (pNP4tolAI15^{BPA}-His_6)
		pEVOL-pBbF
5	JW0729-3 (*ΔtolA::kan*)	pCEP38 (pNP4tolAH22^{BPA}-His_6)
		pEVOL-pBbF
6	JW0729-3 (*ΔtolA::kan*)	pCEP39 (pNP4tolAS32^{BPA}-His_6)
		pEVOL-pBbF

2.10.2 Detergent sensitivity assay

Overnight cultures, 5 ml, were sub-cultured into LB, 3 ml, with 0.125% arabinose and/or 0.5 mM BPA (dissolved in 1.3 M NaOH), with 1.3 M HCl added to match pH, and/or 1 mM IPTG. Cultures were grown to OD_{600} 0.4, then diluted in LB to OD_{600} 0.2 and serially diluted 10-fold in LB to OD_{600} 2 $x10^{-4}$. Each serial dilution was spotted, 1 μl, onto plates with corresponding culture conditions. The plates were incubated overnight at 37°C and imaged in GBOXCHEMI-XRQ using GeneSys software.

For preliminary experiments to test the concentration of IPTG required the following strains were used: BW25113 (WT Keio), JW0729-3 (*ΔtolA::kan*), *ΔtolA::kan* with pCEP44 (pNP4tolA-His_6) and pEVOL-pBbF, and *ΔtolA::kan* with pCEP37 (pNP4tolAI15^{BPA}-His_6) and pEVOL-pBbF. Plate contents detailed in Table 2.11. Two biological repeats of the assay were carried out. For detergent sensitivity assays testing the TolA^{BPA} mutants, strains used are detailed in Table 2.12. Plate contents detailed in Table 2.13. Three biological repeats of the assay were conducted.

Table 2.11 Plate contents for detergent sensitivity assay to determine concentration of IPTG to use in future assays.

Plate	2% SDS	0.5 mM BPA	0.125% arabinose	IPTG
1				
2				
3				
4				100 μM
5				1 mM
6				10 mM

Table 2.12 Strains and plasmids used for detergent sensitivity assay

Strain	Plasmids
BW25113 (WT Keio)	
JW0729-3 (*ΔtolA::kan*)	
JW0729-3 (*ΔtolA::kan*)	pCEP44 (pNP4tolA-His_6)
JW0729-3 (*ΔtolA::kan*)	pCEP44 (pNP4tolA-His_6)
	pEVOL-pBbF
JW0729-3 (*ΔtolA::kan*)	pCEP37 (pNP4tolAI15^{BPA}-His_6)
JW0729-3 (*ΔtolA::kan*)	pCEP37 (pNP4tolAI15^{BPA}-His_6)
	pEVOL-pBbF
JW0729-3 (*ΔtolA::kan*)	pCEP38 (pNP4tolAH22^{BPA}-His_6)
JW0729-3 (*ΔtolA::kan*)	pCEP38 (pNP4tolAH22^{BPA}-His_6)
	pEVOL-pBbF
JW0729-3 (*ΔtolA::kan*)	pCEP39 (pNP4tolAS32^{BPA}-His_6)
JW0729-3 (*ΔtolA::kan*)	pCEP39 (pNP4tolAS32^{BPA}-His_6)
	pEVOL-pBbF

Table 2.13 Plate contents for detergent sensitivity assays. Dark blue filled box indicates the component is included.

Plate	2% SDS	0.5 mM BPA	0.125% arabinose	1 mM IPTG
1				
2		■		
3	■			
4	■	■		
5	■	■	■	
6	■	■	■	■

2.10.3 Western blot

For western blots to test for expression of full-length TolA in the TolAI15^{BPA} and TolAH22^{BPA} mutants, cells were grown as described for the colicin toxicity assay (2.10.1). Samples were taken before induction and after induction for 3 hours at 37°C. The OD_{600} was measured, cells pelleted by centrifugation at 4,000 x g for 3 min and resuspended in SDS sample buffer calculated relative to the OD_{600}. Cells were lysed by heating at 98°C for 20 mins. Samples, 7.5 μl, were loaded onto a 12% Bis-Tris gel, with the Amersham rainbow ladder (Cytivia), and electrophoresis was carried out at as described in 2.3.5. For western blots analysing the nickel affinity pull-downs and purifications of UV-exposed TolAHis22^{BPA}

samples, 7.5 μl, were also run on 12% or 10% Bis-Acrylamide gels with the Amersham rainbow ladder.

Protein was transferred onto a sequi-blot polyvinylidene (PVDF) membrane through semi-dry blotting at a constant current of 45 mA for 45 mins. The membrane was washed with transfer buffer (25 mM Tris/HCl pH 8.3, 150 mM glycine, 20% (v/v) methanol) and blocked with Tris-buffered saline Tween-20 (TBST) (20 mM Tris/HCl pH 7.5, 500 mM NaCl, 0.05% Tween-20) with 8% (w/v) Marvel dried skimmed milk overnight at room temperature. The membrane was washed with TBST (five times, 1 min per wash), then incubated for 1 hour at room temperature in 4% (w/v) milk in TBST containing anti-TolA primary antibody diluted 1:1000 (Bonsor *et al.*, 2009). The wash step was repeated, followed by incubation for 1 hour at room temperature in 4% (w/v) milk in TBST with the secondary anti-rabbit IgG conjugated to horseradish peroxidase diluted 1:1000 (#A6154). The five wash steps were again repeated.

The ECL Select Western Blotting Detection Kit (Cytivia) was used for detection through chemiluminescence, as per the manufacturer's instruction, incubating in the dark for 3 min after addition of reagents. The membrane was imaged in a GBOX-CHEMI-RXQ using GeneSys software. Detection images were taken after 30 s, 40 s, 1 min and 2 mins.

2.10.4 Brightfield microscopy

The following strains were cultured overnight at 37°C with appropriate antibiotics: 1, 2,3 and 5 in Table 2.10. Sub-cultures with appropriate antibiotics and 0.5 mM BPA (dissolved in 1.3 M NaOH), with 1.3 M HCl added to match pH were grown at 37°C to reach OD_{600} 0.6. Pre-induction samples of each culture were taken, centrifuged at 4,000 x g 3 min to pellet, and resuspended in 500 μl 4% formaldehyde, stored at 4°C. Cultures were induced with 0.125% arabinose and 1 mM IPTG and grown overnight at 20°C. Post-induction samples were prepared in formaldehyde in the same manner.

A 5 μl aliquot of each sample was pipetted onto an M9 agar pad and imaged using Oxford Nanoimager (ONI) super-resolution microscope equipped with a 100X oil immersion objective (Olympus 1.49 NA). NimOS software associated with the ONI microscope was used

to take images by scanning a 50 μm x 80 μm area with exposure set at 100 ms to give a 512 x 1024 pixel image. Images were acquired as a 100-frame stack for a brightfield channel. Composite images were made through stacking using average intensity as a projection type in ImageJ software. Binary images were generated by manual thresholding of brightfield images and cell length measured by MicrobeJ (Ducret, Quardokus and Brun, 2016). Violin plots were generated using Prism software. Microscopy and analysis carried out by Dr Joanna Szczepaniak.

3 Characterisation of domain II of TolA

Contributions: Cloning and protein expression assistance – Dr Renata Kaminska. SEC-MALS, ESI Mass Spectrometry – Prof David Staunton, Biochemistry Department Biophysics Services. NMR – Prof Christina Redfield. Crystallography X-Ray screening – Prof Edward Lowe. Molecular Dynamic Simulations – Dr Dheeraj Praakesh.

3.1 Introduction

All current experimentally determined structural information for TolA relates to the third domain, TolA-III, which is responsible for interaction with TolB (Witty *et al.*, 2002; Szczepaniak *et al.*, 2020). However, it is the second domain (TolA-II), which spans the periplasm and links the TMH in domain I (TolA-I) to the TolB binding domain (TolA-III), that is likely to hold answers as to how the PMF is translated into force transduction at the OM. In the mechanism described in 1.5.3.1, it is proposed that PMF dependent conformational changes in TolA-II allow for the extension of TolA through the periplasm to reach TolB-Pal at the OM (Figure 1.8). This hypothesis arises from the following data: First, mutational studies have shown that deletion of regions of TolA-II in *E. coli* resulted in increased detergent sensitivity and reduced susceptibility to colicins (Schendel *et al.*, 1997), implying the length of the domain is important for its function. Second, the proteolysis pattern of TolA changes in response to removal of the PMF, through addition of CCCP (Germon *et al.*, 2001). However, this evidence does little to explain the nature of the conformational change in TolA.

More recently, Rassam *et al*. showed that when TolA is N-terminally tagged with GFP, it can freely diffuse within the inner membrane (Rassam *et al.*, 2018). Addition of a fluorescently-tagged TolA-dependent colicin recruited the Tol-Pal machinery for import, and resulted in restriction in GFP-TolA movement (Figure 3.1) (Rassam *et al.*, 2018). If TolA were to be constantly extended across the 300 Å periplasm to the OM, it would not be capable of

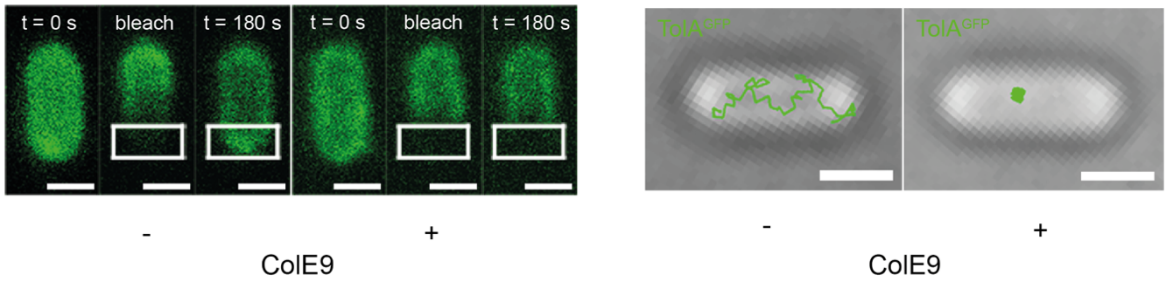


Figure 3.1 GFP-tagged TolA can diffuse freely in the IM and exhibits restricted diffusion upon addition of translocation incompetent colicin E9.
Taken from Rassam *et al.* 2018. Scale bars 1 μm. **A.** FRAP of GFP-TolA shows unrestricted diffusion as fluorescence is recovered after 180 s. Upon addition of a disulfide locked ColE9 incapable of translocation GFP-TolA diffusion is restricted. **B.** SPT of GFP-TolA reaffirms the unrestricted diffusion of GFP-TolA and restriction upon addition of ColE9.

free diffusion within the IM as it would become trapped within the holes in PG. The data from Rassam *et al*. suggests that TolA must exist in more than one conformation: an extended form that spans the periplasm to bind TolB, and one that is retracted and thus able to diffuse freely within the IM without interference from the PG (Rassam *et al.*, 2018). Coupled with the data described above, this suggests that it is TolA-II that is key to understanding how TolA adopts different conformations in a PMF dependent manner.

To begin to understand the mechanism of TolA, the following chapter aimed to characterise and determine the structure of TolA-II. Biophysical techniques including Far UV Circular Dichroism (CD), nano-scale Differential Scanning Fluorimetry (nanoDSF) and size exclusion chromatography multiangle light scattering (SEC-MALS) were used to validate the secondary structure composition, stability, and oligomeric state of versions of TolA-II. NMR was then used to identify flexible regions within the domain and aid in the design of constructs more suitable for X-ray crystallography. Crystal screens were carried out with the view to experimentally determine the structure of TolA-II. Finally, molecular dynamics simulations were used to investigate the flexibility of the AlphaFold2 TolA model and its interaction with the cell envelope.

3.2 Results

3.2.1 Domain II of TolA is helical in nature

To begin studying domain II of TolA, first an expression system was chosen, and the protein purified. In previous studies the definition of the boundaries between the domains of TolA have not always been consistent. For this study, the domain boundaries were defined by the poly-glycine regions that flank both the predicted transmembrane helix and globular C-terminal domain (Figure 3.2.A). A series of plasmid constructs were designed to express truncations of TolA from four different bacterial species. For each species a construct of TolA-II and TolA-II+III was made, excluding the poly-glycine region at the N-termini of TolA-II+III, and at the N- and C-termini of TolA-II. A C-terminal His_6 tag was also added to each construct to allow for affinity purification (Table 3.1, Figure 3.2.A).

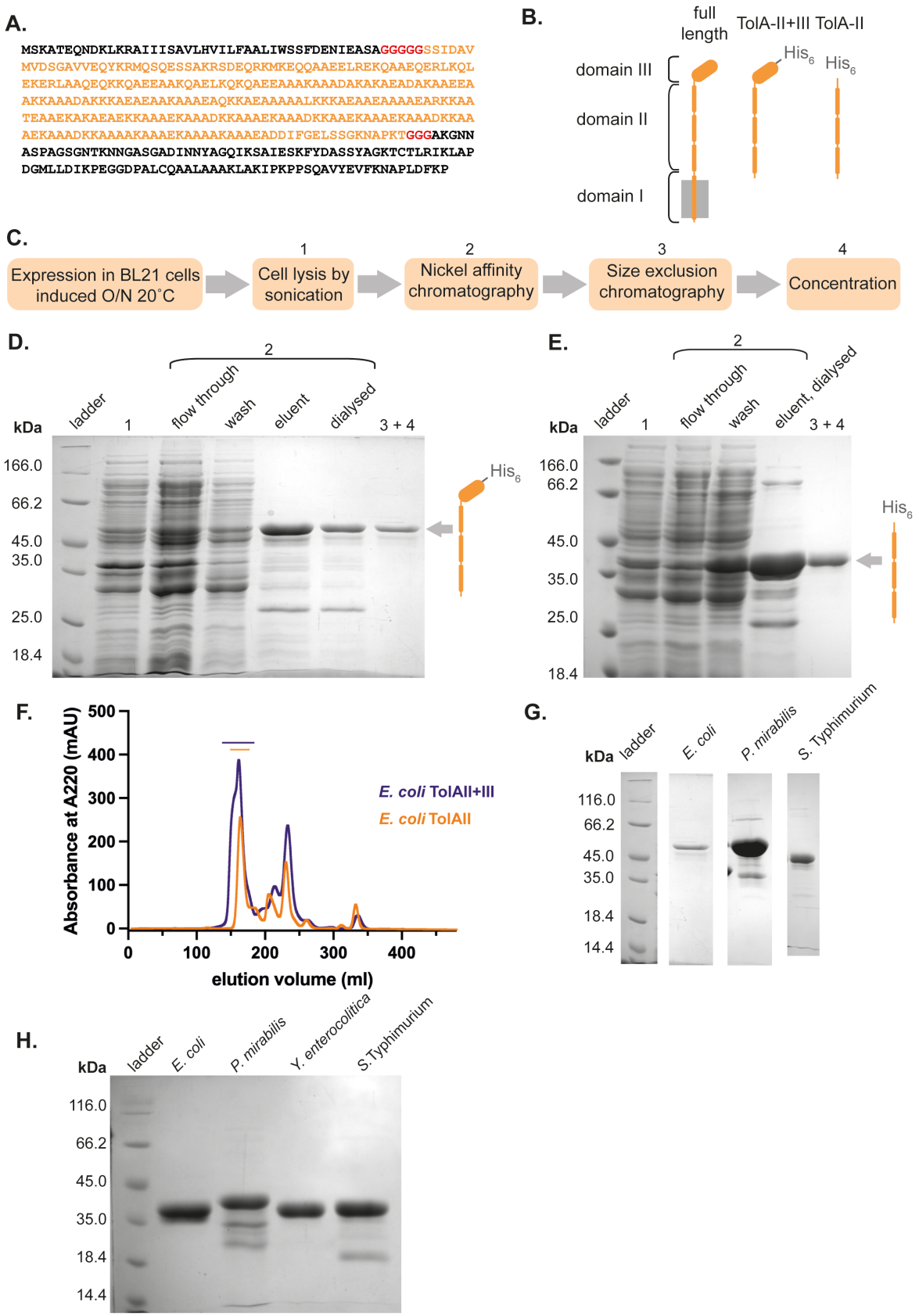


Figure 3.2 The design and purification of TolA-II and TolA-II+III constructs from four bacterial species.

Figure legend continued p81

Figure 3.2 continued A. Poly-glycine stretches (red) in the primary amino acid sequence of *E. coli* TolA were used to delineate the boundaries between domains, domain II constructs (orange) excluded the poly-glycine region. **B.** Representative schematic diagrams of the TolA-II+III and TolA-II constructs designed, including a C-terminal His_6 tag. **C.** Flowchart representing the expression and purification steps involved in purifying TolA-II+III and TolA-II. **D.** Coomassie stained 12% Bis-Acrylamide SDS-PAGE gel was used to visualise the purity of *E. coli* TolA-II+III throughout the purification process. Each number corresponds to the purification step in the flowchart (C). **E.** Coomassie stained 12% Bis-Acrylamide SDS-PAGE gel was also used to visualise the purity of *E. coli* TolA-II throughout the purification process. **F.** SEC of *E. coli* TolA-II+III and TolA-II on a HiLoad 26/60 Superdex 200 in 20 mM Tris/HCl pH 7.5, 150 mM NaCl. Purple and orange bars indicate the peak corresponding to TolA-II+III and TolA-II, respectively. **G.** Pure samples of TolA-II+III were obtained from three different species, visualised on Coomassie stained 12% Bis-Acrylamide SDS-PAGE gel. Concentrations of *E. coli* (1.6 mg/ml), *P. mirabilis* (1.9 mg/ml) and *S.* Typhimurium (5.5 mg/ml) loaded. **H.** Pure samples of TolA-II were obtained from four different species, visualised on Coomassie stained 12% Bis-Acrylamide SDS-PAGE gel. Concentrations of *E. coli* (1.4 mg/ml), *P. mirabilis* (1.2 mg/ml), *Y. enterocolitica* (1.3 mg/ml) and *S.* Typhimurium (1.3 mg/ml) loaded.

All constructs of TolA-II and TolA-II+III were expressed on a pET21a(+) plasmid in BL21 (DE3) cells, induced by IPTG. Once expression of TolA-II or TolA-II+III was established through comparison to no-plasmid control cells, the purification itself was optimised (Figure 3.2.C-H). Cells were lysed through sonication and TolA-II or TolA-II+III isolated through nickel affinity chromatography. Following this, SEC was used to obtain pure protein and spin concentration used to concentrate the protein. One attribute of the TolA-II constructs that had to be considered was the low extinction coefficient due to the absence of Trp residues. The absorbance at 220 nm was used to follow elution of protein during SEC instead of the standard

Table 3.1 Comparison of the constructs of TolA-II+III and TolA-II from four different bacterial species designed in this study.

Species:	Construct	Domain boundaries	No. of amino acids	pI	MW (kDa)	% Identity	Extinction Coefficient ($M^{-1}cm^{-1}$)	Yield (mg/L)
Escherichia coli	TolA-II-His_6	48-310	263	9.0	28.2		1490	1.4
	TolA-II+III-His_6	48-421	383	9.1	39.4		7450	1
Proteus mirabilis	TolA-II- His_6	48-260	222	6.6	23.8	47.4	1490	0.8
	TolA-II+III-His_6	48-357	319	8.7	34.1	50.3	5960	8.9
Yersinia enterocolitica	TolA-II- His_6	45-291	256	7.0	26.7	63.5	1490	0.5
	TolA-II+III-His_6	45-362	355	8.7	37.0	65.8	7450	
Salmonella enterica Typhimurium	TolA-II- His_6	49-296	257	9.1	26.7	70.5	1490	3.0
	TolA-II+III-His_6	49-407	369	9.2	38.1	68.8	7450	5.5

280 nm. The purity of purified TolA-II and TolA-II+III was assessed by SDS-PAGE, where in almost all constructs the protein of interest was the dominant species present. Some contaminants could still be observed in *S* Typhimurium TolA-II+III, *P. mirabilis* TolA-II and *S.* Typhimurium TolA-II. *Y. enterocolitica* TolA-II+III was not well expressed and could not be successfully purified. The low absorbance at 280 nm also meant measurements of the concentrations of purified TolA-II were more variable and therefore had higher error. The yields of purified TolA-II+III and TolA-II's ranged from 0.5 mg/l to 5.5 mg/l (Table 3.1) with the highest yields observed in *S.* Typhimurium and lowest in *Y. enterocolitica*.

Following the purification of the TolA-II and TolA-II+III constructs, biophysical analysis was carried out to assess the stability and fold of the proteins before moving to structural studies. The secondary structure compositions of TolA-II and TolA-II+III constructs were found to be predominantly α-helical by CD (Figure 3.3.A-B). Minima at 208 and 221 nm, and a peak at 195 nm characteristic of α-helical structures were observed for all constructs. This agrees with previously published CD spectra for constructs of *E. coli* $TolA_{43-398}$ and $TolA_{43-421}$ (Levengood, Beyer and Webster, 1991; Derouiche *et al.*, 1999) and shows the proteins to be folded in solution. The helical signal is variable between the constructs, with *P. mirabilis* TolA-II+II having the strongest α-helical signal, and *E. coli* TolA-II the weakest. This could be due to differences in helical composition of the domain, or instability of the protein in solution.

Further biophysical analysis was carried out on the *E. coli* constructs. The oligomeric states of *E. coli* TolA-II and TolA-II+II were analysed using SEC-MALS (Figure 3.3.C). TolA-II was found to be monomeric, with an experimentally determined molecular weight of 28.0 kDa compared to the theoretical molecular weight of 28.1 kDa. TolA-II+III was also found to be monomeric, with an experimentally determined molecular weight of 40.8 kDa compared to the theoretical molecular weight of 39.4 kDa. Finally, *E. coli* TolA-II's melting temperature was determined to be 47.6°C and TolA-II+III 67.1°C, by nanoDSF, suggesting the construct including domain III is more stable in solution.

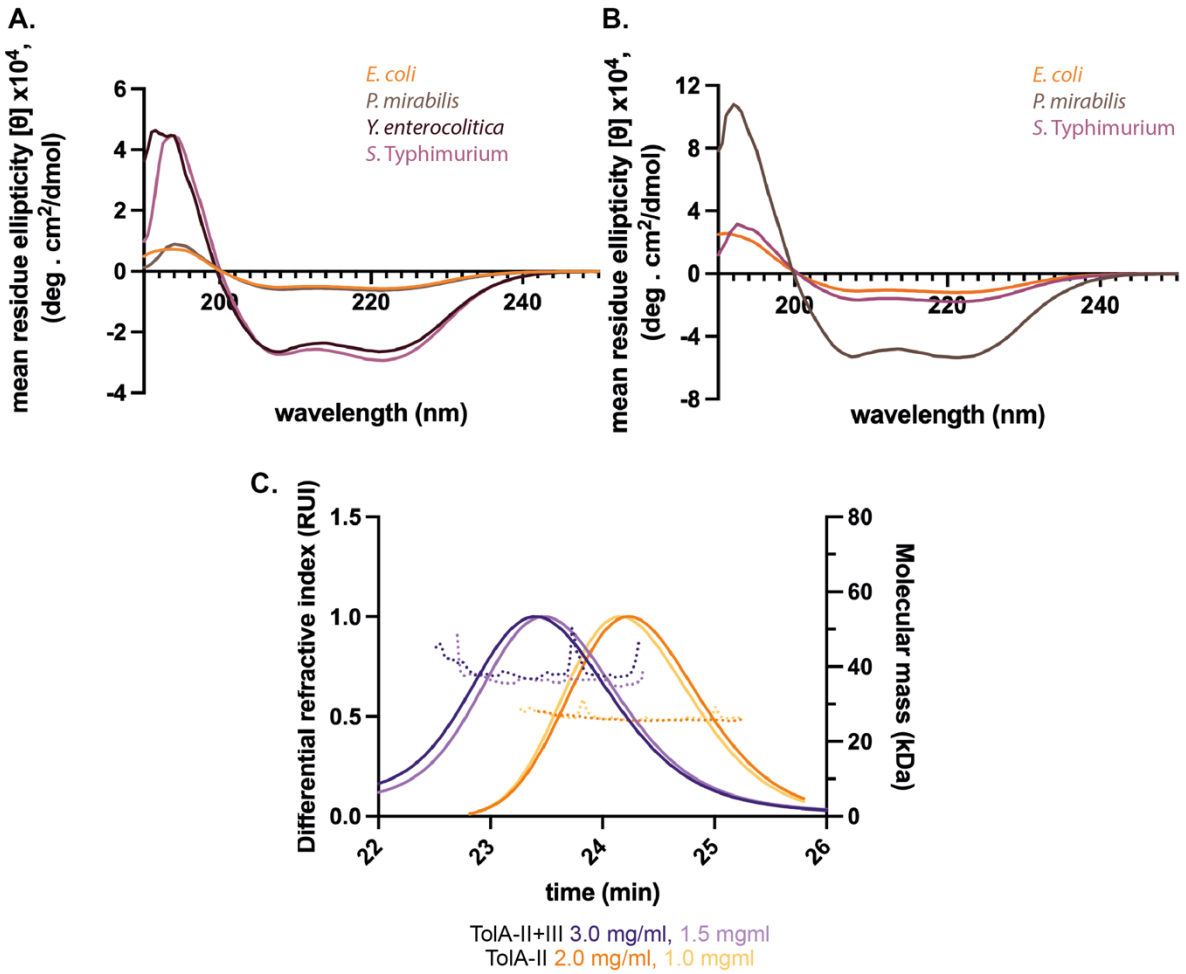


Figure 3.3 Biophysical characterisation confirmed the TolA-II and TolA-II+III constructs are helical and monomeric.
A. Far UV Circular Dichroism of TolA-II constructs from *E. coli* (orange, 2.8 mg/ml), *P. mirabilis* (beige, 1.3 mg/ml), *Y. enterocolitica* (purple, 0.2 mg/ml) and *S.* Typhimurium (pink, 0.4 mg/ml). **B.** Far UV Circular Dichroism of TolA-II+III constructs from *E. coli* (orange, 0.5 mg/ml), *P. mirabilis* (beige, 0.1 mg/ml), and *S.* Typhimurium (pink, 0.4 mg/ml). CD experiments were performed in 20 mM Potassium phosphate buffer pH 7.5, 150 mM NaF. All constructs displayed characteristic predominantly helical spectra with minima at 208 and 221 nm. **C.** SEC-MALS of *E. coli* TolA-II-III (purple/lilac, 3.0/1.5 mg/ml) and TolA-II (orange/yellow, 2.0/1.0 mg/ml) determined both to be monomeric in solution. Refractive index shown in solid lines, molecular mass in dotted lines.

3.2.2 The N- and C-termini of domain II are flexible

After establishing that TolA-II is a helical monomeric protein, the next step was to investigate if domain II adopted a helical bundle as predicted by Derouiche *et al.* (Derouiche *et al.*, 1999). To determine the structure of TolA-II, crystal trials were set up with all purified TolA-II and TolA-II+III from the bacterial species above at concentrations ranging from 4 to 16 mg/ml, with the exception of *Y. enterocolitica* TolA-II+III due to its poor purity (Supplementary Figure 1).

No crystals and little precipitation were observed across the variety of commercial screens used, which included a range of precipitants. To understand whether flexibility may be the cause of the challenges in crystallography, and whether improved constructs could be designed to remove regions of flexibility, NMR was used. *E. coli* TolA-II was grown in ^{15}N-enriched media and purified as described previously. A combination of ^{15}N 2D BEST TROSY, 3D TOCSY-HSQC and HetNOE experiments were used to identify peaks for flexible residues in TolA-II (Figure 3.4). Interestingly, the TolA-II tumbling characteristics were consistent with an elongated protein. Approximately 17 residues at the N terminus and 25 residues at the C terminus had peaks consistent with that of flexible residues. To pinpoint the start and end of the structured regions more accurately, new constructs were designed to truncate flexible residues in both the N- and C-termini (Figure 3.5.A, Table 3.2).

The new constructs were purified with the additional TEV cleavage and reverse nickel affinity chromatography steps to remove the C-terminal His_6 tag. Furthermore, lyophilisation was used as an alternative concentration method to avoid the loss of protein that had previously been observed during spin concentration (Figure 3.5.B-C). Far UV CD was carried out to confirm the constructs were folded and helical. All four constructs gave typical helical spectra, however $TolA_{92-298}$ had weaker signal than the other constructs, suggesting the truncation may have disrupted some secondary structure (Figure 3.5.D). Each TolA construct was ^{15}N labelled and purified, and ^{15}N 2D BEST TROSY, 3D TOCSY HSQC and HetNOE were carried out to assign flexible residues. For $TolA_{72-298}$ only 2D BEST TROSY was conducted due to low yield of the ^{15}N labelled construct. Comparison of peaks present in each truncation allowed for the assignment of individual residues, identifying Ser69 as the first structured residue at the N-terminus, and Glu293 as the final structured residue at the C-terminus of TolA-II (Figure 3.5.E). The repetitive nature of the primary sequence of TolA-II would make using NMR to obtain a solution structure challenging, so it was decided to continue with crystallography using the truncated versions of *E. coli* TolA-II.

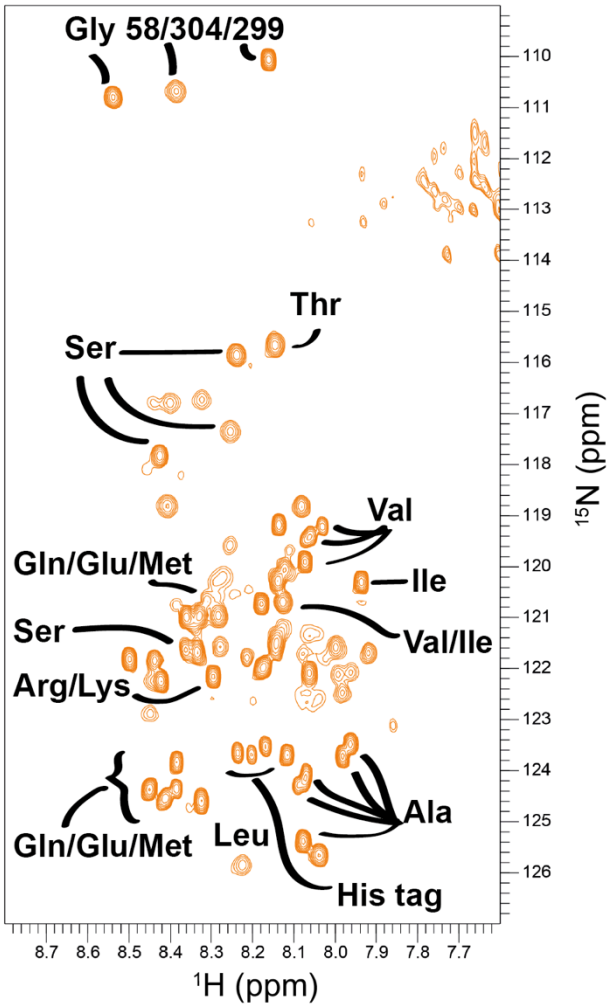


Figure 3.4 NMR analysis reveals TolA-II has unstructured N and C-termini.
^{15}N 2D best TROSY in combination with 3D TOCSY-HSQC and HetNOE allowed for the identification of approximately 17 flexible residues at the N-terminus and 25 residues at the C-terminus.

Table 3.2 Truncated constructs of TolA-II.
Molecular weight (MW) following cleavage with TEV protease.

Construct	pI	MW (kDa)	extinction coefficient (M^{-1}/cm^{-1})	Yield (mg/l)
62-298	9.0	26.5/25.4	2980	0.6
72-298	9.1	25.2/24.1	1490	6.1
82-298	9.1	24.1/23.0	1490	3.5
92-298	9.2	22.9/21.8	1490	5.4

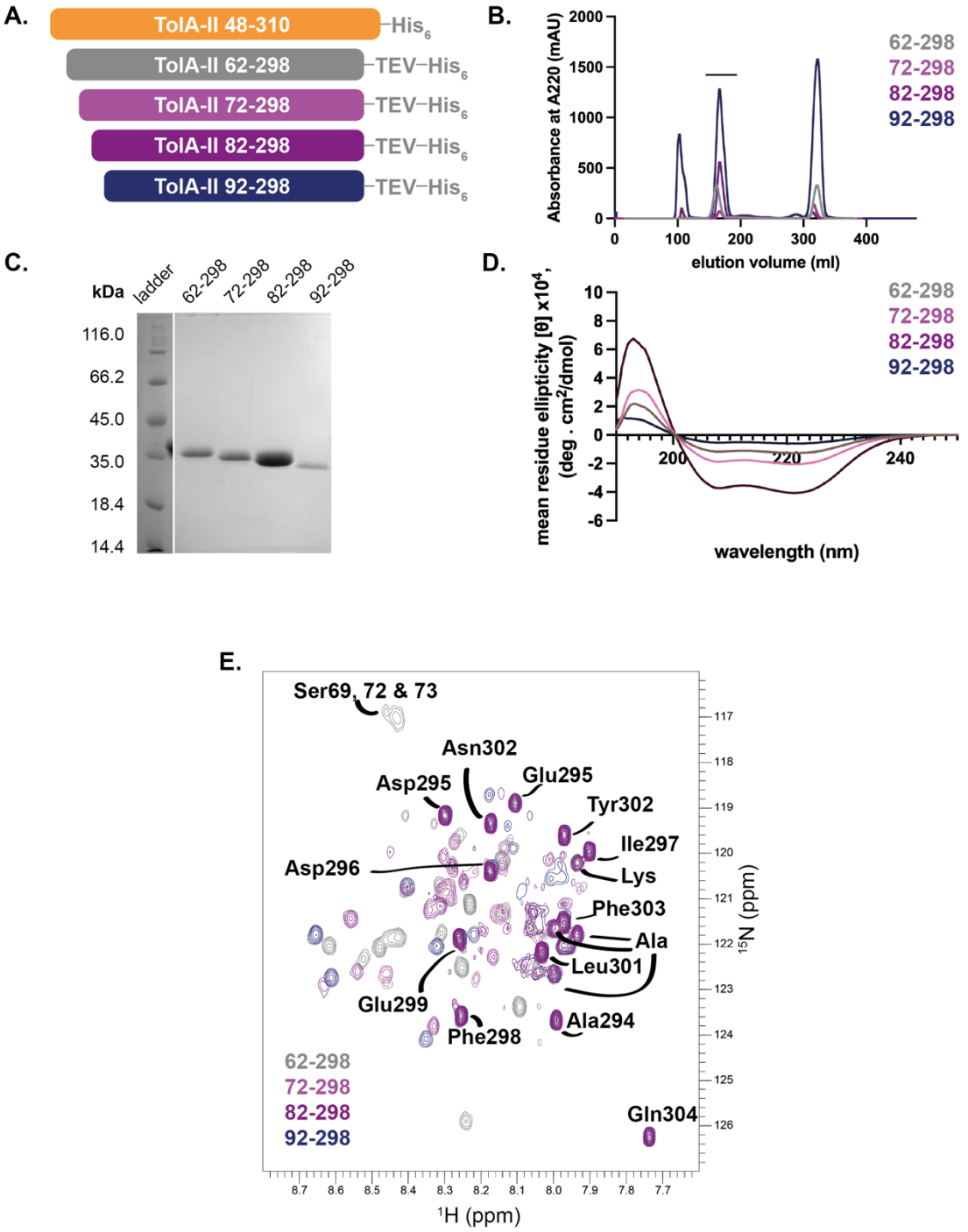


Figure 3.5 *E. coli* TolA-II structured domain begins at Ser69 and ends at Glu293.
A. A series of truncated versions of TolA-II were designed, with cleavable C-terminal His_6 tag. **B.** SEC purification elution profile of each truncated construct on a HiLoad 26/600 Superdex 200 PG column in 20 mM Tris/HCl pH 7.5, 100 mM NaCl. The second peak corresponded to each TolA-II construct. **C.** Pure samples of $TolA_{62\text{-}298}$, $TolA_{72\text{-}298}$, $TolA_{82\text{-}298}$ and $TolA_{92\text{-}298}$ visualised on Coomassie stained 12% Bis-Acrylamide SDS-PAGE gel. **D.** Far UV Circular Dichroism of $TolA_{62\text{-}298}$ (0.8 mg/ml), $TolA_{72\text{-}298}$ (0.6 mg/ml) $TolA_{82\text{-}298}$ (0.3 mg/ml) and $TolA_{92\text{-}298}$ (0.9 mg/ml) CD experiments were performed in 20 mM Potassium phosphate buffer pH 7.5, 150 mM NaF. All constructs displayed characteristic predominantly helical spectra with minima at 208 and 221 nm. **E.** ^{15}N 2D BEST TROSY NMR spectra for the constructs $TolA_{62\text{-}298}$ (grey), $TolA_{72\text{-}298}$ (magenta), $TolA_{82\text{-}298}$ (purple), and $TolA_{92\text{-}298}$ (navy). Strong peaks present in all spectra were assigned to the C-terminus, where structure was mapped to end at Glu293. The differences in the 2D spectra allowed for assignment of flexibility at the N-terminus, with order beginning at Ser69.

3.2.3 *E. coli* TolA domain II crystallises at high concentrations

Following identification of flexible residues at the N- and C-termini of *E. coli* TolA-II, further commercial crystal screens were set up with the truncated versions of *E. coli* TolA-II at concentration between 5 and 12 mg/ml. Again, little precipitation was observed. The construct $TolA_{72-298}$ became the focus of the crystallographic screens as it was the truncated version of TolA-II that came closest to the identified start of the structured region, Ser69, and gave a higher yield than $TolA_{62-298}$ (Table 3.2). The successful cleavage of the His_6 tag was validated by ESI mass spectrometry, determining the size of $TolA_{72-298}$ to be 23979 Da, consistent with the expected size of cleaved $TolA_{72-298}$ 24110 Da when assuming the N-terminal Met is lost.

Pre-crystallisation tests identified that a very high concentration of 145 mg/ml $TolA_{72-298}$ gave the desired level of precipitation in sitting drops and was the ideal concentration to use. Following this finding, commercial crystal screens were set up with $TolA_{72-298}$ (145 mg/ml). However, despite observation of precipitation and over-nucleation, no crystals were obtained. To promote crystallisation $TolA_{72-298}$ was modified by reductive methylation of Lys residues, a technique commonly used to aid crystal formation. Additional commercial screens were set up with $TolA_{72-298}$ (145 mg/ml) and with methylated $TolA_{72-298}$ (137 mg/ml). Crystals of $TolA_{72-298}$ were observed in the STURA screen, condition H8 (0.1 M sodium citrate pH 5.5, 20% PEG 4000, 20% 2-propanol) (Figure 3.6.A) and were screened at SOLEIL on beamline Proxima 2A. The best resolution observed from the crystals was approximately 6 Å although the diffraction was highly anisotropic, so meaningful structural information could not be obtained (Figure 3.6.B). This did, however, confirm the crystals to be protein and not salt.

To overcome the resolution and anisotropy observed, optimisation plates of the STURA H8 condition were set up, varying pH and precipitant concentrations, or screening STURA H8 with additives in the hope of improving the crystal quality. A further screen of the same condition was also set up where lyophilised $TolA_{72-298}$ was resuspended in MilliQ water,

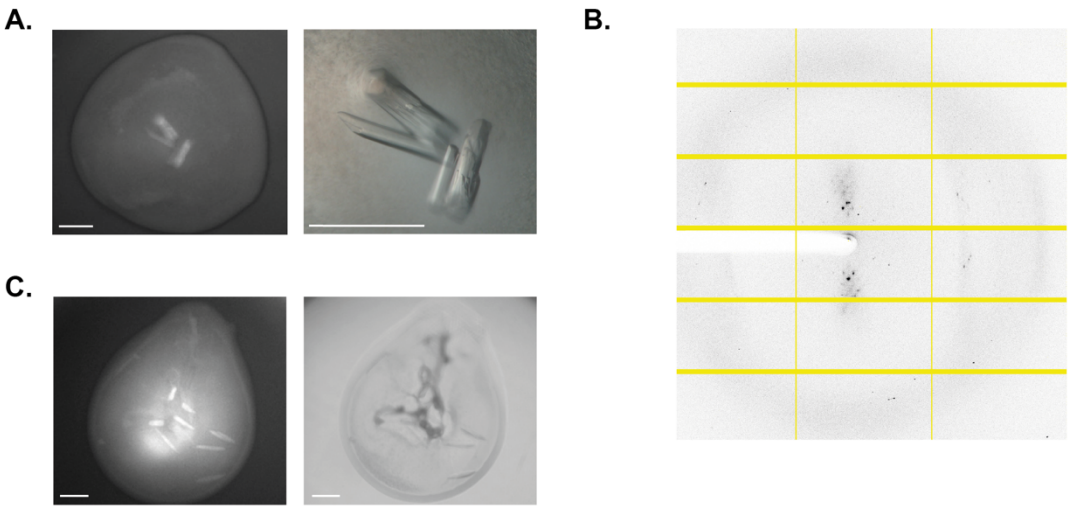


Figure 3.6 *E. coli* TolA$_{72-298}$ crystallises at high concentrations and when crosslinked with 0.1% glutaraldehyde.
A. Crystals of TolA$_{72-298}$ observed in STURA condition H8 (0.1 M sodium citrate pH 5.5, 20% PEG 4000, 20% 2-propanol). Scale bar 200 Å, UV and brightfield image. **B**. Diffraction pattern of crystal from A) at 0° revealed highly anisotropic data. Yellow bars indicate gaps in detector panels, visualised using dials.image_viewer. **C.** Crystals of glutaraldehyde crosslinked TolA$_{72-298}$ observed in STURA condition H8 (0.1 M sodium citrate pH 5.5, 20% PEG 4000, 20% 2-propanol). Scale bar 200 Å, UV and brightfield image.

20 mM HEPES pH 7.5 or 20 mM HEPES pH 7.5 100 mM NaCl, all to achieve 150 mg/ml. In addition, one of the observations from previous diffraction data was that it seemed likely that TolA$_{72-298}$ adopted a crystal packing arrangement that was impacting the achievable resolution. To attempt to overcome this, chemical crosslinking was carried out. TolA$_{72-298}$ resuspended in MilliQ water (150 mg/ml) was crosslinked with 0.1% glutaraldehyde for 15 minutes before quenching with 20 mM Tris/HCl pH 8.0. Crystals of glutaraldehyde crosslinked TolA$_{72-298}$ were harvested (Figure 3.6.C) and screened at Diamond Light Source beamline I03. The best resolution observed was ~8 Å, and again the data was highly anisotropic. Despite extensive screening and the identification of conditions that crystallise TolA$_{72-298}$, no solvable diffraction data could be gleaned from the crystals harvested in this study.

3.2.4 Structural predictions of TolA domain II reveal an extended helical hairpin

Whilst the attempts to crystallise TolA-II were ongoing, AlphaFold2 released artificial intelligence generated predictions for the structures of full-length TolA from a range of different bacterial species (Jumper *et al.*, 2021). The structural prediction of full-length *E. coli* TolA is organised into three domains as expected, and is predominantly helical, in agreement with the CD data in section 3.2.1 (Figure 3.7.A). Domain I contains a 34-residue helix with a 15-residue hydrophobic region consistent with a transmembrane helix. Domain III was also consistent with the existing structure of apo *E. coli* TolA-III, RMSD 2.6 Å in Pymol (Deprez *et al.*, 2005).

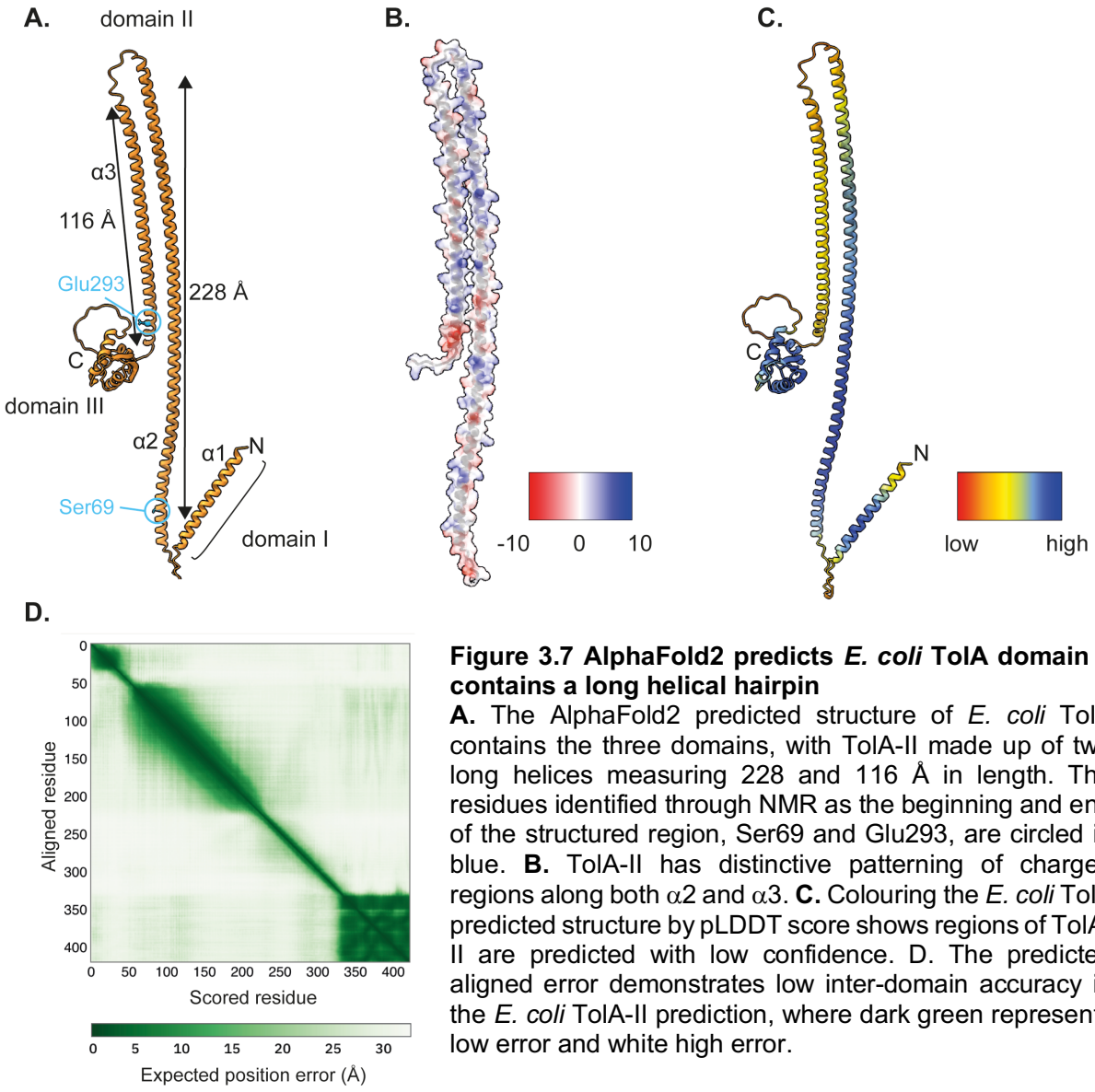


Figure 3.7 AlphaFold2 predicts *E. coli* TolA domain II contains a long helical hairpin
A. The AlphaFold2 predicted structure of *E. coli* TolA contains the three domains, with TolA-II made up of two long helices measuring 228 and 116 Å in length. The residues identified through NMR as the beginning and end of the structured region, Ser69 and Glu293, are circled in blue. **B.** TolA-II has distinctive patterning of charged regions along both α2 and α3. **C.** Colouring the *E. coli* TolA predicted structure by pLDDT score shows regions of TolA-II are predicted with low confidence. D. The predicted aligned error demonstrates low inter-domain accuracy in the *E. coli* TolA-II prediction, where dark green represents low error and white high error.

The predicted structure of domain II contains two long helices organised as a helical hairpin. The AlphaFold2 model predicts the first helix of domain II, α2, as Ser57-Lys209 and the second helix, α3, as Glu224-Leu301. These domain boundaries are reasonably close to the residues identified by NMR that showed the structured region of the domain began at Ser69 and ended at Glu293. α2, is 228 Å in length and α3 is 116 Å (Figure 3.7.A), meaning α2 alone is long enough to span most of the periplasm, assuming the periplasm width to be ~300 Å (Mandela *et al.*, 2022). The distance between the peptide backbone of α2 and α3 is on average 9.9 Å, with the closest side chain atoms 4.8 Å apart. This distance is too far to allow for inter-helical hydrogen bonding, suggesting the structure formed is in fact not a classic coiled-coil hairpin structure. Assessment of the electrostatic potential shows an alternating pattern of charge on each face along both helices. There is a combination of charged residues aligning with the poly-Ala regions, a complimentary charged residue, or a repulsive charged residue on the opposite helix (Figure 3.7.B). AlphaFold2's confidence metric pLDDT scoring assigned 27.8% of TolA-II as predicted with high confidence, but 47.9% was associated with low or very low confidence scores (Figure 3.7.C). In addition, the predicted alignment error scoring attributed low confidence in regions of α2 and α3 that are in closer contact with each other (Figure 3.7.D). Finally, the positioning of α1 relative to α2 is unrealistic given α1 should span the IM. However, this is not unexpected as AlphaFold2 does not take membrane environment into account when generating structure predictions, and the domains are connected by a flexible linker sequence that allows for multiple orientations of TolA-I relative to TolA-II.

To better understand the importance of the electrostatic pattern, helical arrangement and length of TolA-II, the sequence and AlphaFold2 predictions of TolA from nine different γ-proteobacteria species were compared (Table 3.3). Sequence homology of TolA-II between species ranged from 37-100% using T-COFFEE alignment (Notredame, Higgins and Heringa, 2000). The primary sequences of domain II showed that across the nine species the domain is predominantly made up of the residues Ala, Lys, Glu, Asp and Gln (Figure 3.8). Ala accounts

Table 3.3 Comparison of AlphaFold2 predicted structures of TolA from different bacterial species.

Species:	Uniprot code	% Identity to *E. coli*	pI	Size (kDa)	II alignment (RMSD Å)	TolA-II number of helices	α2 length	α3 length
Escherichia coli	P19934		9.1	43.16		2	228	116
Haemophilus influenzae	P44678	41.3	9.0	39.83	5.4	2	32	225
Pseudomonas aeruginosa	P50600	37.2	9.6	44.70	18.4	2	146	117
Klebsiella pneumonia	A0A0H3GU6	78.4	9.2	37.94	25.5	2	224	146
Yersinia enterocolitica	A0A0H3NMM3	63.3	8.7	37.61	2.1	1	210	
Serratia marcescens	A0A2S4XFN9	64.2	9.1	41.87	28.3	1	224	
Proteus mirabilis	B4ESS3	48.0	9.3	43.18	10.3	1	225	
Shigella dysentriae	Q32II2	99.5	9.1	41.37	1.8	2	246	117
Salmonella enterica Typhimurium	Q8ZQT6	74.8	9.1	40.23	24.6	2	194	118

for on average 35% of all residues within domain II and is present as part of poly-Ala repeating units flanked by charged residues. *E. coli* has 13 repeats of a Lys-Ala_3-Glu/Asp motif. TolA-II from *S.* Typhimurium, *S. dysentriae* and *K. pneumoniae* have between 8 to 11 copies of the same motif, whereas other species contain similar proportions of residues, but different motifs such as Lys-Ala_2-Asp in *P. aeruginosa*. The number of residues in the domain varies between 204 (in *P. aeruginosa*) and 282 (in *K. pneumoniae*) with an average of 246 residues. The regular gaps within the sequence alignment highlight that it is not specific sequence conservation that is important, but the conserved sequence composition in the residues found in the domain across all nine species.

Comparison of the AlphaFold2 predictions of TolA-II from each species reveals that the helical hairpin seen in *E. coli* is not observed across all species (Table 3.3). In *P. mirabilis, S. marcescens* and *Y. enterocolitica* only one helix is present (Figure 3.9). The α2 for each species is on average 210 Å in length, in-keeping with the length of *E. coli* TolA-II and the length of the periplasm. This is with the exception of the TolA-II of *H. influenzae* that has a short helix following the transmembrane α1. For all species the second helix, or corresponding

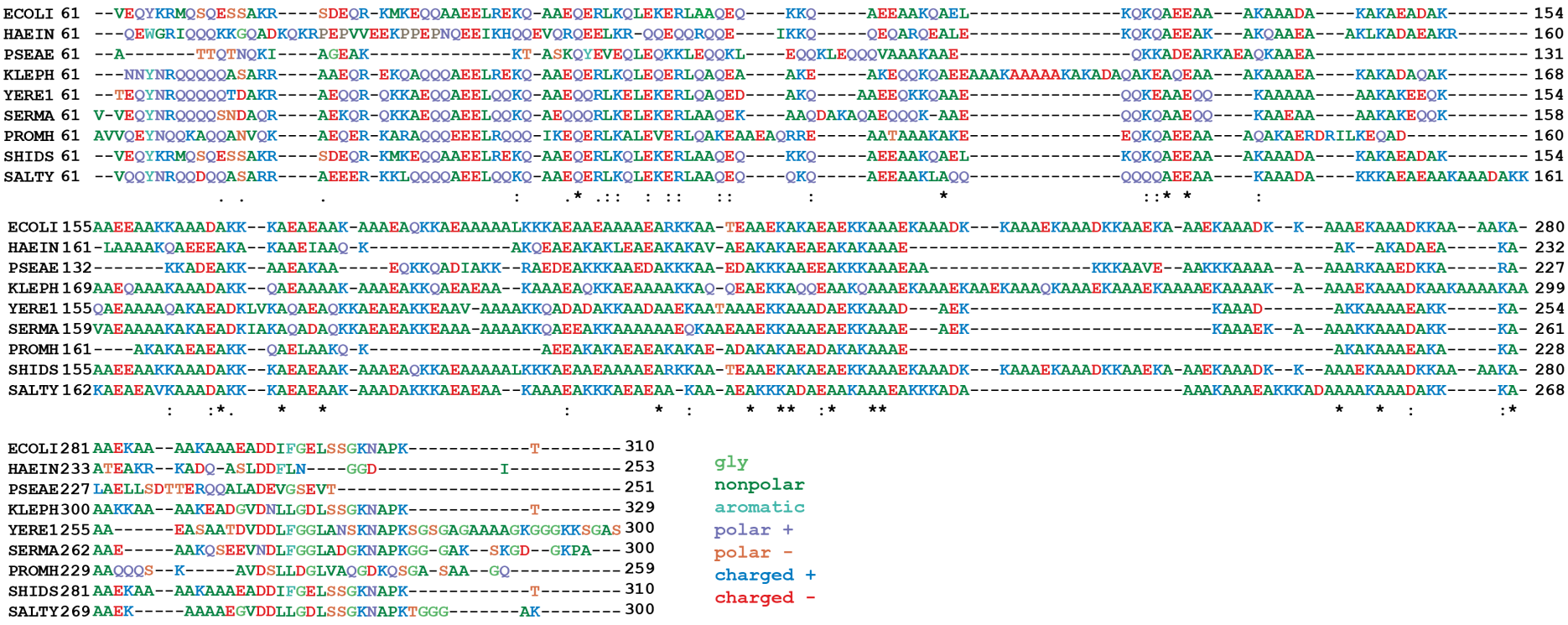


Figure 3.8 Alignment of primary sequences of TolA-II from 9 different bacterial species.
Multiple sequence alignment by T-Coffee (Notredame, Higgins and Heringa, 2000) demonstrates that the residue composition of TolA-II containing predominantly Ala, Lys, Glu and Asp is conserved across the species

region of flexibility, is the area with the lowest pLDDT score, and therefore where the prediction is least confident (Figure 3.9). Alignments of TolA-II of each prediction to *E. coli* TolA-II gave RMSD from 1.8 Å to 25.5 Å (Table 3.3). As could be expected from the nature of the primary sequence, analysis of the electrostatic surfaces of the helices in TolA-II across the nine species saw the same repeating pattern of alternating charged patches (Figure 3.10). *P. aeruginosa* is the exception in that the first half of α3 is predominantly positive, and the second

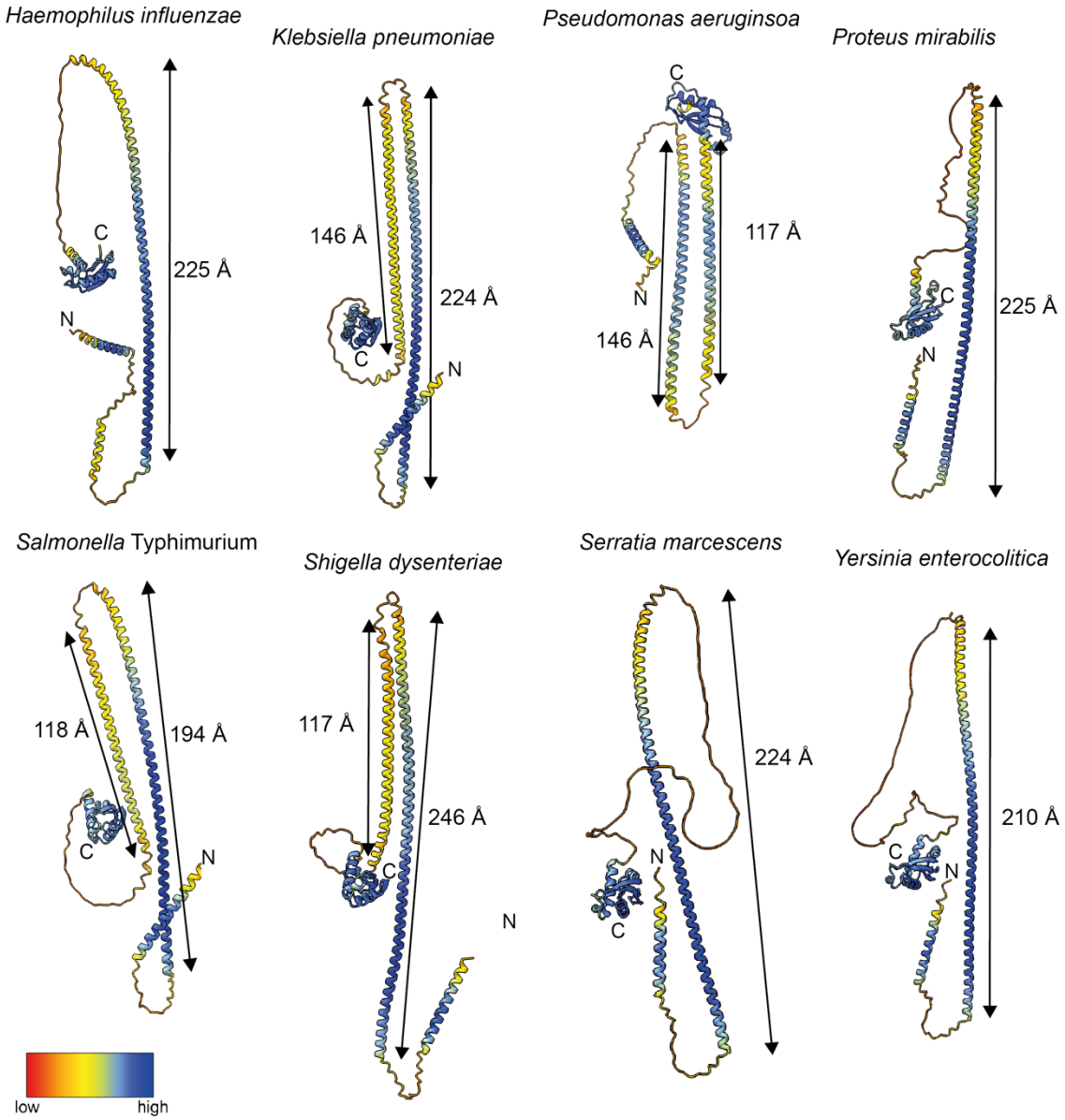


Figure 3.9 Comparison of AlphaFold2 predicted structures of TolA from eight γ-proteobacteria.
The lengths of the α helices in TolA-II are annotated and all are coloured by pLDDT score where red is low confidence and blue high confidence.

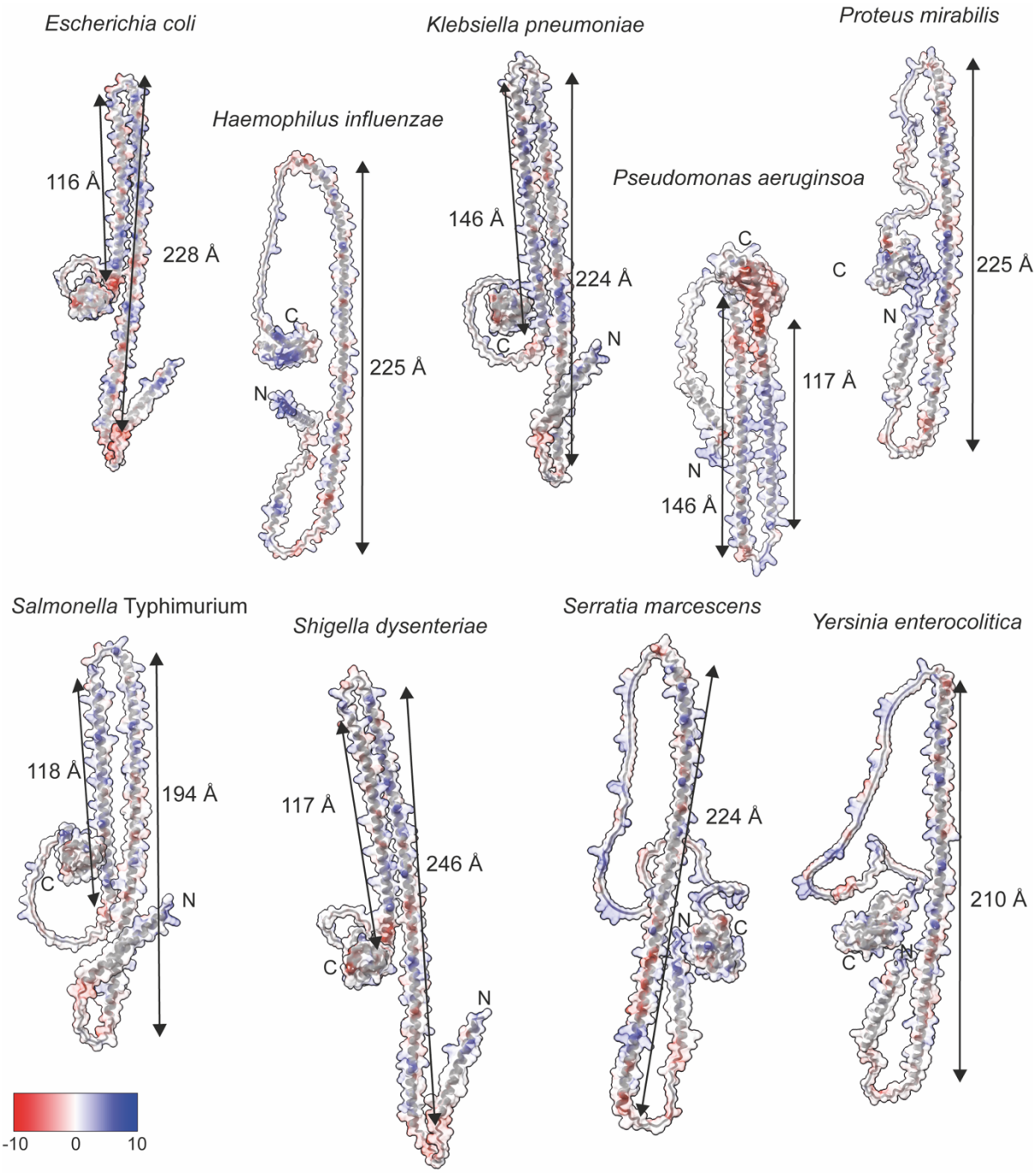


Figure 3.10 Comparison of electrostatic surfaces of TolA from nine γ-proteobacteria. The length of the helices of TolA-II are annotated and the surface coloured by electrostatic charge where red represents negative charge and blue positive charge. For TolA-II in all species there is a pattern of alternating charge along the α-helices.

half electronegative (Figure 3.10). The conservation of this charge distribution and of a long helix able to span the periplasm suggest that both could be required for and involved in the function of Tol-Pal.

Comparison of *E. coli* TolA-II to more distantly related species outside of the γ-proteobacteria shows a substantial difference in amino acid composition of between 12-20% sequence identity to *E. coli* TolA (Supplementary Figure 2) *Burkholderia ambifaria* has high Ala, Glu, Asp, and Lys content like *E. coli* TolA-II, and AlphaFold2 predicts it to have a single 220 Å helix its domain II region. *Caulobacter vibriodes, Campylobacter ureolyticus* and *Desulfovibrio vulgaris* TolA contain polyproline regions more in keeping with *E. coli* TonB's domain II (Figure 3.11).

3.2.5 Molecular dynamic simulations of AlphaFold2 *E. coli* TolA reveal interactions between the two helices in domain II

To investigate the flexibility of TolA-II, coarse-grained molecular dynamics (CGMD) simulations of the AlphaFold2 model of *E. coli* TolA embedded in the IM were conducted. First, the AlphaFold2 model was modified to angle the TolA-I helix away from α2 of TolA-II, allowing for α1 to be inserted into the IM model. Following this, α3 was angled 85° from α2 in TolA-II. The aim of this simulation was to assess whether the placement of the two helices in the AlphaFold2 prediction was indicative of an interaction surface between α2 and α3. Over a 2000 ns timescale α3 bound to α2 within 250 ns (Figure 3.12.A, Supplementary Figure 3). At the end of the simulation, the backbone carbon atoms of the two helices of TolA-II were ~5 Å from one another; close enough to form hydrogen bonds, in comparison to the ~12 Å in the original AlphaFold2 model. It therefore is likely that these helices do interact. It is also notable that embedding of the residues 45-80, which link TolA-I and TolA-II, in the polar head groups of the phospholipid bilayer occurs at the end of all three independent the simulations (Figure 3.12), suggesting this could be of mechanistic importance.

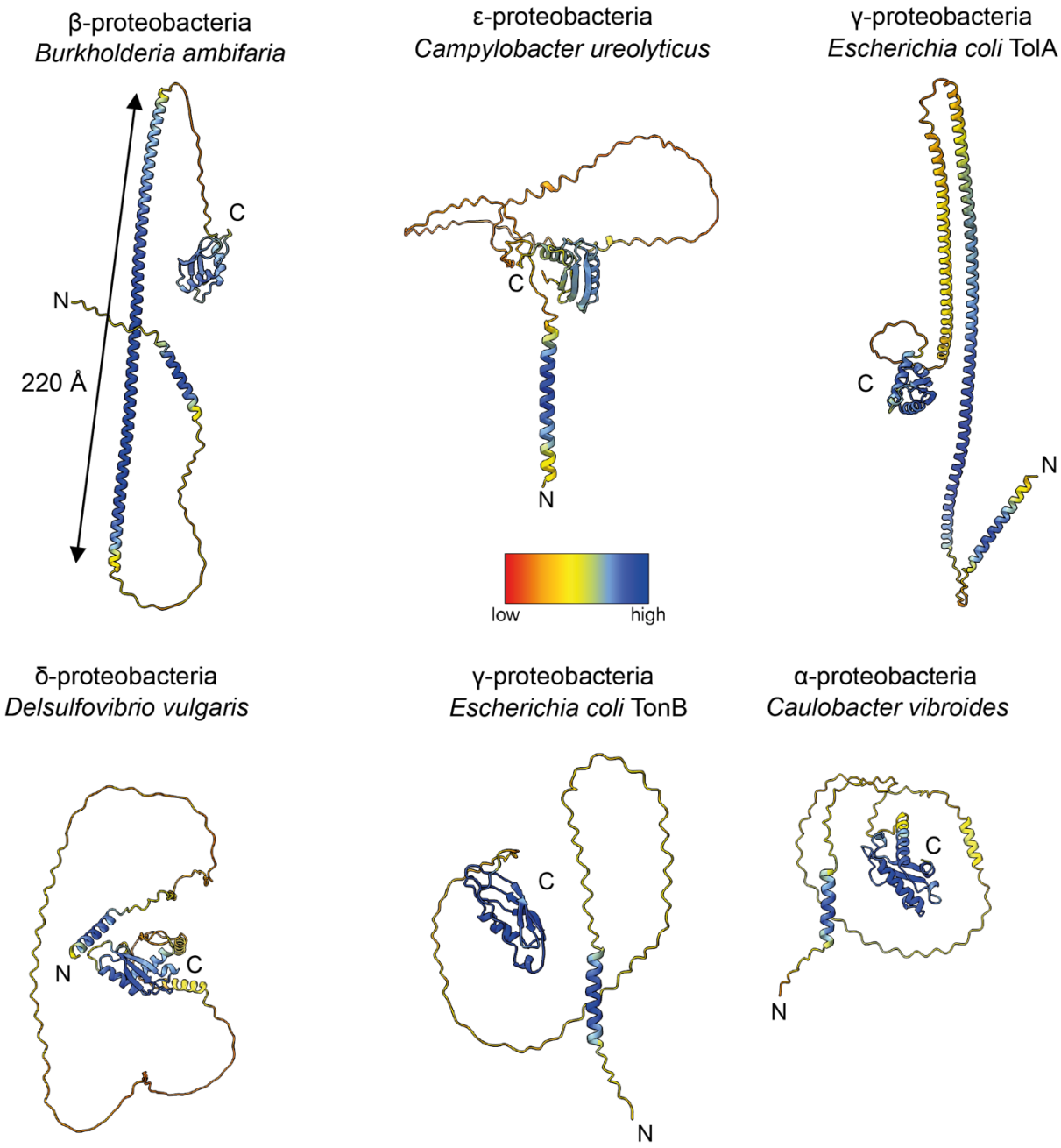


Figure 3.11 The AlphaFold2 predicted structures of TolA from representative species from different classes of proteobacteria.
The AlphaFold2 models of structures of TolA from example species for each class of proteobacteria and the lengths of the α helices if present. *E. coli* TonB is also included for comparison. All are coloured by pLDDT score where red is low confidence, blue high confidence.

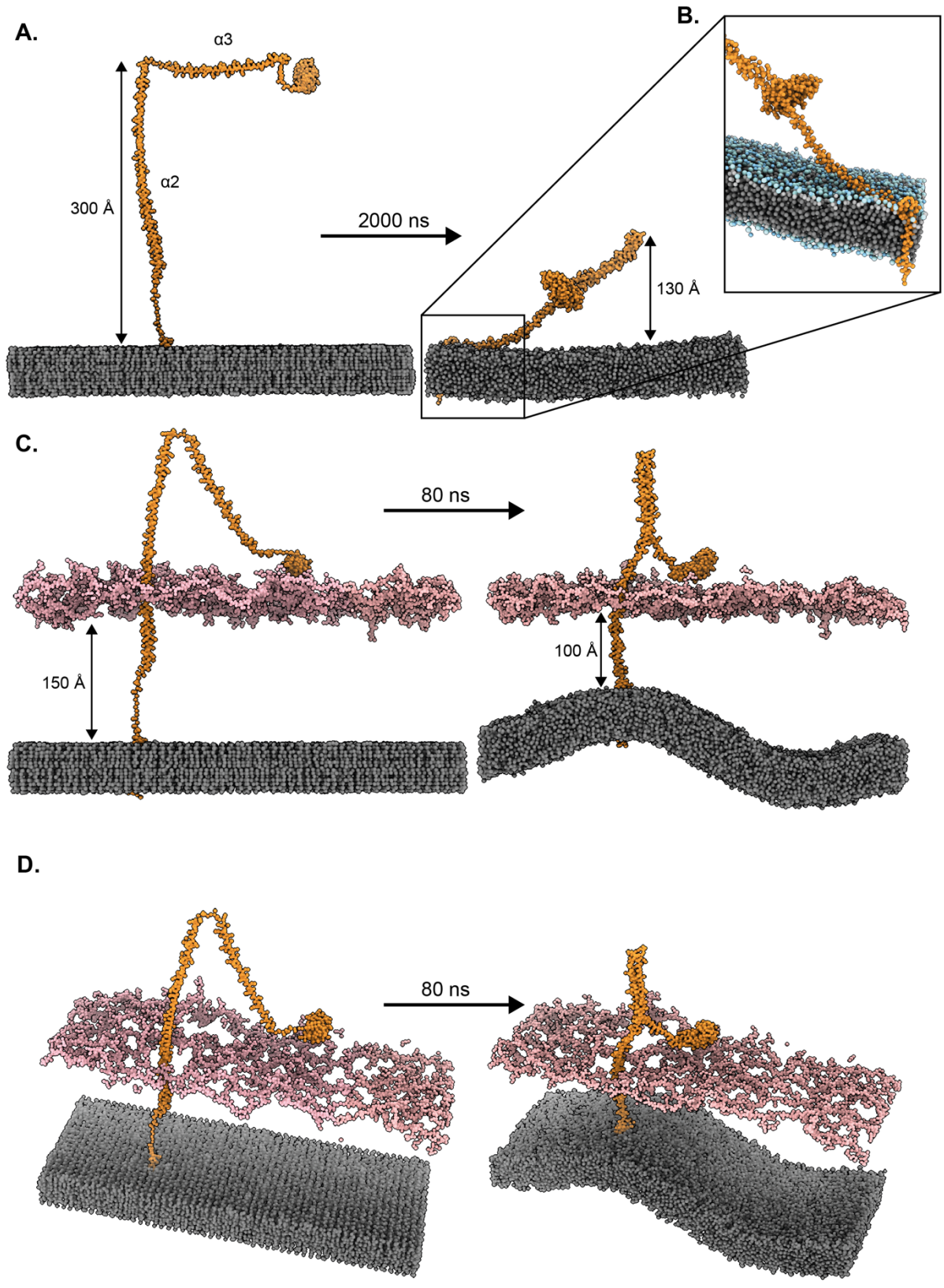


Figure 3.12 Molecular dynamics simulations of the AlphaFold2 TolA model demonstrate an intramolecular interaction between α2 and α3, and interactions with PG and the IM.
Figure Legend continued p96.

Figure 3.12 continued. A. Coarse-grained representations of before and after the 2000 ns simulation of TolA in the IM. The AlphaFold2 model of TolA was modified to straighten the flexible linker between TolA-I and TolA-II and move α3 80° from α2. After insertion into the IM (a symmetric membrane comprised of 90% POPE, 5% POPG and 5% cardiolipin) three coarse-grained simulations were run for 2000 ns. The resulting conformation of TolA showed the two helices of TolA-II interacting with one another, and supercoiling of TolA's flexible region. **B.** A zoom-in view at the end of one of the simulations shows TolA (orange) interacting with the polar head-groups of the IM (blue). **C.** Coarse-grained representations of before and after an 80 ns simulation of TolA in the presence of PG in the IM. The TolA α3 helix was placed 40° away from α3, and α2 placed through the PG mesh to insert α1 into the IM bilayer. After 80 ns, the IM bilayer was pulled 50 Å closer to the PG around TolA, and α2 and α3 of TolA-II are within 5 Å of one another. **D.** The same coarse-grained simulations as in C from an alternative angle.

The next set of simulations were conducted to investigate the impact of the presence of PG on the interaction between the two helices of TolA-II. The α3 helix was placed 40° from α2, and a layer of PG 15 nm from the IM bilayer. TolA was threaded through a hole in the PG and inserted into the IM. The 2000 ns simulations of TolA showed that the helices interact, therefore in these simulations the helices were positioned relatively close to each other, giving rise to the shorter timescale for helix-helix binding. In a simulation of 80 ns, within 55 ns the two helices of TolA-II came together in an almost zipper-like fashion, pulling TolA-III along the PG in the process (Figure 3.12.C, Supplementary Figure 3). Whether TolA-III would interact with the PG in this conformation natively, or if it is a product of the positioning implemented at the beginning of the simulation, is unclear. However, it does demonstrate that the interactions between TolA-II α2 and α3 are capable of breaking transient bonds between TolA-III and PG. The flexible linker between TolA-I and TolA-II preferred to undergo coiling whilst TolA α2 interacted with the PG, the result of this was to pull the IM bilayer up towards the PG (Figure 3.12.C-D). However, this abrupt change in the dynamics of the IM relative to the PG is expected to stabilise and evenly distribute over larger timescales. Altogether the molecular dynamics simulations demonstrate that the polar surfaces of TolA-II are capable of forming hydrogen bonding interactions with one another, and with elements of the cell envelope.

3.3 Discussion

The work described in this chapter experimentally confirmed the helical nature of TolA's central domain, evaluated the validity of the machine-learning generated model of TolA,

and investigated the flexibility of the domain through both experimental and computational approaches. The biophysical analysis of TolA-II through CD, nanoDSF and SEC-MALS confirmed that the protein purified was consistent with previously published work on the domain, identifying its helical composition and its monomeric oligomeric state (Levengood, Beyer and Webster, 1991; Derouiche *et al.*, 1999). NMR experiments indicated that there are flexible residues at the N- and C-termini of TolA-II. It also showed that the domain is predominantly structured in an aqueous environment. Further to this, the domain behaved as though it had a long, elongated shape, consistent with the biophysical analysis. Unfortunately, solving an experimentally determined structure of TolA-II through X-ray crystallography proved too challenging. As demonstrated by the diffraction data in Figure 3.6, crystals of TolA-II diffracted poorly and gave highly anisotropic data. This is consistent with TolA-II having long helices that induce heterogeneity between 2D crystal layers. TolA-II was highly soluble in aqueous solutions because of its high proportion of polar residues; this additional factor meant very high concentrations of TolA-II were required for crystallographic experiments. Obtaining a crystal structure of TolA-II is still a valuable enterprise, however the challenges described will likely persist. Re-development of the constructs through a disulphide crosslinking approach could give rise to improved crystal lattices by increasing stabilising helical interactions. Prediction of crosslinking sites using an atomistic version of the AlphaFold2 *E. coli* model at the endpoint of the CGMD simulations through Disulfide by Design (Craig and Dombkowski, 2013) found 21 potential crosslinking sites that could be used for this aim (Supplementary Figure 4). In addition, based upon the AlphaFold2 models analysed, the *P. aeruginosa* TolA-II is shorter in length (Table 3.3). Therefore, development of constructs based upon the *P. aeruginosa* TolA-II may have more success in overcoming anisotropy issues.

In the absence of an experimentally determined structure of TolA-II, the AlphaFold2 predicted structures provided a platform for further study. First, the models themselves required evaluation to determine the validity of the predictions. AlphaFold2's own metrics highlighted a lack of confidence in some regions of the predicted structures (Figure 3.10). In

future, collecting experimental SAXS data and docking in the AlphaFold2 models could validate whether the long helices of TolA-II generated in the model are consistent with the protein shape *in vitro*.

In the *E. coli* AlphaFold2 model, it was identified that the two helices α2 and α3 of domain II were placed too far from one another to form hydrogen bonding interactions. This begged the question – do these helices interact with one another or was this an arbitrary placement by AlphaFold2? Molecular dynamics simulation data using this model suggested that they do interact, and moreover, are capable of interacting with the cell envelope. This discovery opens the door for investigation into whether the interaction between α2 and α3 in TolA-II is important for system function. It is notable that of the 9 γ-proteobacteria AlphaFold2 predictions analysed, 4 did not have two helices in their domain II (Figure 3.10). However, the models were of low confidence in this region, and therefore the lack of α-helix might not be representative of the native fold or could be interpreted as an indicator of a dynamic structure. A crosslinking mass spectrometry approach could allow for identification of interactions between the two helices *in vitro*. In addition, a disulphide locking approach tested *in vivo* for Tol-Pal system function could give answers as to whether the interaction between α2 and α3 is key to TolA's mechanism of action.

The second major observation arising from the simulations was the interactions with PG and with IM head groups. TolA-II is a highly charged domain with a high pI, so it follows that it has the potential to interact with the negatively charged polar head groups of the IM phospholipid bilayer and polar PG glycan strands. The identification of these interactions puts the model of TolA in context. It highlighted that transient hydrogen bonding interactions such as those seen between TolA-III and the PG, or TolA-II and the IM, could be a part of its mechanism of action. The molecular dynamic simulations described in 3.2.5 also saw supercoiling of the flexible linker between α1 and α2 (Figure 3.12). This phenomenon has been observed in disordered regions adjacent to membrane helices previously (Prakaash *et al.*, 2021) but it is unclear whether this is physiologically or mechanistically relevant.

The process by which TolA traverses the PG is still unclear. Unfortunately, at the time of writing the PG models in use for CGMD simulations do not possess holes in the PG large enough to accommodate the diameter of TolA-II with TolA-III, however holes of this size do exist *in vivo* (Turner *et al*., 2013). Once a 'holey' model PG has been developed it may be possible to investigate whether TolA can freely work up through the PG to bring TolA-III to the OM, or whether this action may require an energy source. In addition, it would be possible to look at the process of relaxation from the extended helix through the PG down to a ground state ate the IM, and the timescale required.

Overall, the work described in this chapter makes further inroads into our understanding of the mechanism of action of TolA. The length of the helices predicted by AlphaFold2 for TolA-II would allow for the delivery of TolA-III to the OM to interact with TolB-Pal as hypothesised and observed during colicin import (Rassam *et al.*, 2018). However, how these long helices extend up to the OM, whether by a hinge or more extreme secondary structure rearrangement, remains to be understood. Further *in vivo* validation of the AlphaFold2 model of TolA could shed some light on whether the interaction between α2 and α3 is important for function. In addition, the AlphaFold2 model is of low confidence in almost 50% of TolA-II, so validating the model through an experimentally determined structure could still be valuable.

4 Investigation into the assembly of the TolQ-TolR stator

Contributions: Peptide mass fingerprinting, cryo-EM, model building – Dr Melissa Webby. Native mass spectrometry – Dr Tarick El Baba.

4.1 Introduction

Understanding how the PMF traverses the IM machinery and is translated to mechanical force is fundamental to understanding the Tol-Pal system. Compared with its sister systems Ton, Mot, and other more distant cousins such as the Gld gliding motility apparatus, very little is known about the TolQ-TolR stator complex, its assembly, and its stoichiometry.

TolQ has three transmembrane helices (TMHs) that span the IM and a structured cytoplasmic loop between the first and second TMHs (Kampfenkel and Braun, 1993; Vianney *et al.*, 1994). Suppressor mutations first identified that the TMHs of TolQ interact with TolR, and that this interaction was localised to TolR residues 20-42, originally predicted to form a helix (Lazzaroni *et al.*, 1995). Analysis of the primary sequence of the helix, referred to as a trans-pore helix (TPH) suggested that it is surrounded by protein not lipid (Cascales, Lloubès and Sturgis, 2001). Through comparison with the structures of the MotA-MotB and ExbB-ExbD stators, the stoichiometry and organisation of TolQ-TolR as a 5:2 complex can be inferred (Santiveri *et al.*, 2020; Deme *et al.*, 2020; Celia *et al.*, 2019). In this organisation, TolR sits within the pentameric pore of TolQ, supporting the previous observation that the TolR TPH is surrounded by protein.

Key to elucidating how the PMF energises the TolQ-TolR stator is the identification of a proton pathway through the complex. The TolR residue Asp23 is the most well-studied residue implicated in proton transduction through TolQ-TolR. Mutation of Asp23 to Arg confers the *tol* phenotype, and mutation to Ala also heavily impacts Tol-Pal function (Cascales, Lloubès and Sturgis, 2001). This crucial Asp residue is also conserved in ExbD and MotB (Braun and Herrmann, 2004). Residues Thr145, Thr178 and Pro187 in TolQ's third TMH have also been implicated in proton transduction through mutagenesis studies (Goemaere, Cascales and Lloubès, 2007; Zhang *et al.*, 2011) and are conserved in ExbB and MotA (Braun and Herrmann, 2004).

The work described in this chapter aimed to further our understanding of the Tol-Pal system by identifying the stoichiometry and determining the structure of the TolQ-TolR stator. To achieve this aim, the intact complex must first be purified, and therefore this was the first focus of the project. Following the purification of the complex, the stoichiometry was investigated through native mass spectrometry and cryo-EM. Simultaneously a new computational model of TolQ-TolR was generated and analysed in comparison to known ExbB-ExbD and MotA-MotB structures.

4.2 Results

4.2.1 The TolQ-TolR complex can be purified using the detergent Cymal-6

To investigate TolQ-TolR complex assembly the intact complex first had to be purified; a process that required extensive optimisation. The TolQ-TolR-TolA open reading frame from four different bacterial origins was cloned into the pT12 plasmid (C-terminal twin streptavidin tag) (Figure 4.1.A): *E. coli, Proteus mirabilis, Yersinia enterocolitica* and *Salmonella enterica* Typhimurium. The plasmid was chosen on account of its successful use in the expression of ExbB-ExbD and MotA-MotB for structural studies (Deme *et al.*, 2020). The expression of TolQ-TolR-TolA in BL21 (DE3) cells was assessed for each bacterial strain by comparing plasmid-free cells to plasmid containing cells induced with 0.1% (w/v) rhamnose (Figure 4.1.B). A band consistent with the 25 kDa TolQ was identified for all strains, with the strongest band observed in *E. coli* and *Y. enterocolitica* consistent with higher expression levels in these samples (Figure 4.1.B). It was reasoned that the TolQ-TolR-TolA complex was likely to be more transient than TolQ-TolR alone, therefore constructs of just TolQ-TolR for *E. coli* and *Y. enterocolitica* were cloned into pT12 and the expression of TolQ-TolR tested (Figure 4.1.A-D). The identity of the 25 kDa band in *E. coli* was confirmed to be TolQ by mass spectrometry, peptide mass fingerprinting.

Next, four different detergents were screened to optimise solubilisation of the *E. coli* TolQ-TolR complex (Figure 4.2.A-D). The detergent Cymal-6 was found to best solubilise TolQ

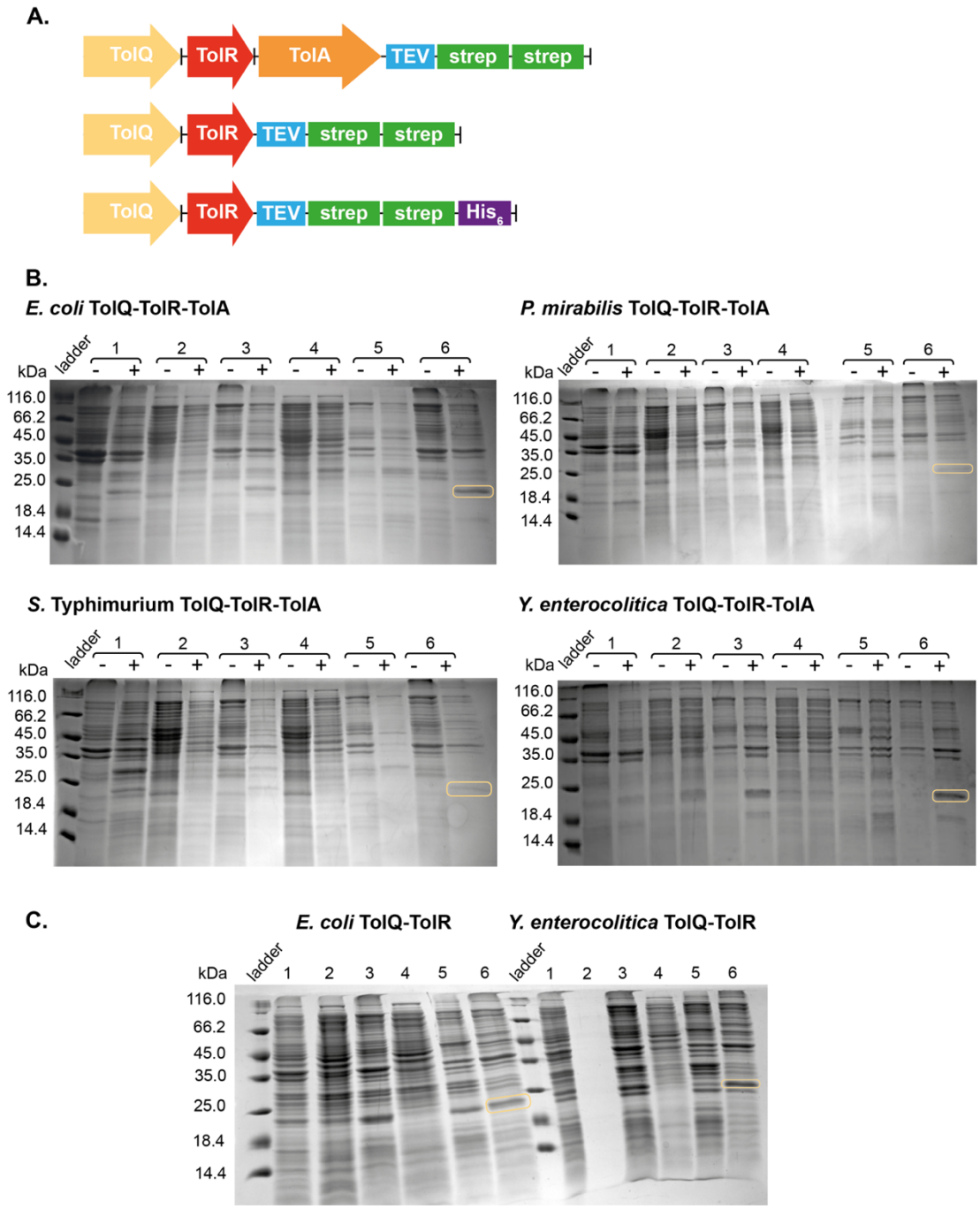


Figure 4.1 Optimisation of expression of the TolQ-TolR
A. Schematic representation of the TolQ-TolR-TolA and TolQ-TolR constructs designed. The vertical line indicates the location of a stop codon within the open reading frame. All genes were inserted into the pT12 vector with a C-terminal twin strep tag. Further cloning steps added a His_6 tag to aid in protein extraction. **B.** Coomassie stained 16% Bis-Acrylamide SDS-PAGE gels to identify expression of TolQ-TolR-TolA from four bacterial species, extracted and solubilised in 1% DDM. No plasmid control cells (-) are compared with plasmid induced expression with 1% (w/v) rhamnose (+). Yellow box indicates the band representing solubilised TolQ. Bands for TolR and TolA were not resolved. 1. Cell debris pellet. 2. Soluble material following cell lysis. 3. Total membrane fraction. 4. Total membrane soluble material. 5. Insoluble total membrane fraction. 6. Solubilised inner membrane fraction. **C.** Coomassie stained 16% Bis-Acrylamide SDS-PAGE gel to identify expression of TolQ-TolR from *E. coli* and *Y. enterocolitica*. Numbers correspond to the extraction and solubilisation fractions as above. Yellow box indicates the band representing TolQ.

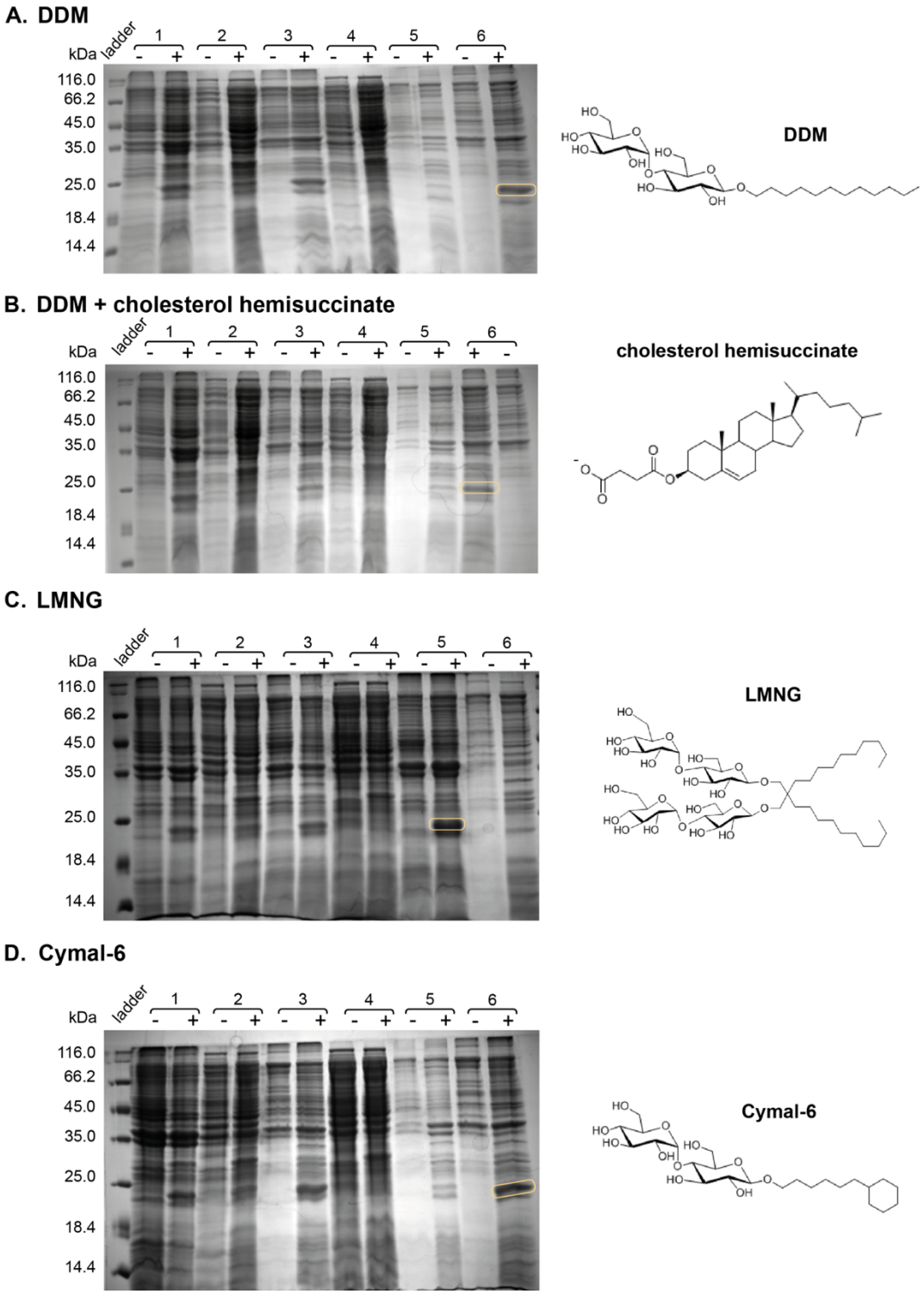


Figure 4.2 The detergent Cymal-6 best solubilised TolQ.
Coomassie stained 16% Bis-Acrylamide SDS-PAGE gels to investigate the solubilisation of *E. coli* TolQ-TolR in **A.** 1% DDM **B.** 1% DDM with 0.25% cholesterol hemisuccinate. **C.** 0.1% LMNG or **D.** 2% Cymal-6. No plasmid control cells (-) are compared with plasmid induced expression with 1% (w/v) rhamnose (+). The chemical structure of each detergent is also shown. The yellow box indicates the band representing TolQ. In all but LMNG TolQ is found in the solubilised inner membrane fraction. 1. Cell debris pellet. 2. Soluble material following cell lysis. 3. Total membrane fraction. 4. Total membrane soluble material. 5. Insoluble total membrane fraction. 6. Solubilised inner membrane fraction

and was chosen for use in future purification steps through assessment by SDS-PAGE. In addition, a hexahistidine tag (His_6) was added following the twin strep affinity tags (Figure 4.1.A). This additional tag was chosen because of issues encountered with TolQ-TolR not binding to streptavidin resin, it was reasoned that using the twin strep affinity tags as a linker and an alternative affinity purification method could improve yield.

Once a detergent was identified that could successfully solubilise TolQ, attention was turned to determining if TolR was present. Following nickel affinity chromatography and SEC, a band at 25 kDa was present when run on a 16% Bis-Acrylamide SDS-PAGE gel. However, no band at 20 kDa consistent with uncleaved TolR was identified (Figure 4.3.B). Furthermore, when run on a 10-20% Tricine SDS-PAGE gel designed to resolve lower molecular weight species, the strong band attributed to TolQ was closer to 20 kDa (Figure 4.3.A-B). Immunoblotting the TolQ-TolR with an α-strep antibody suggested the band at 25 kDa contained TolR (Figure 4.3.C). This implied both TolQ and TolR were running at the same size on SDS-PAGE. To define TolQ and TolR by SDS-PAGE, TEV protease was added to cleave TolR from the twin strep tags and His_6 tag, thereby reducing the mass of TolR from 20 kDa to 16 kDa. A new band at 16 kDa appeared on SDS-PAGE gels following cleavage with TEV protease (Figure 4.3.D). Peptide mass fingerprinting reconfirmed the identity of the band at 25 kDa as TolQ, with 53% sequence coverage, and identified the band at 16 kDa as TolR, with 37% sequence coverage. Therefore, both TolQ and TolR were purified (Figure 4.3.F). The optimised purification for TolQ-TolR gave a yield of 0.02 mg/l. The identification of both TolQ and TolR following affinity chromatography provided evidence that the TolQ-TolR complex was present as the His_6 tag was only on TolR.

4.2.2 Biochemical approaches to validate the TolQ-TolR complex

A series of biochemical techniques were employed with the aim of confirming the complex of TolQ-TolR was present and determining its stoichiometry. First, a blue native PAGE gel was run, however no band consistent with a complex was identified (Figure 4.4)

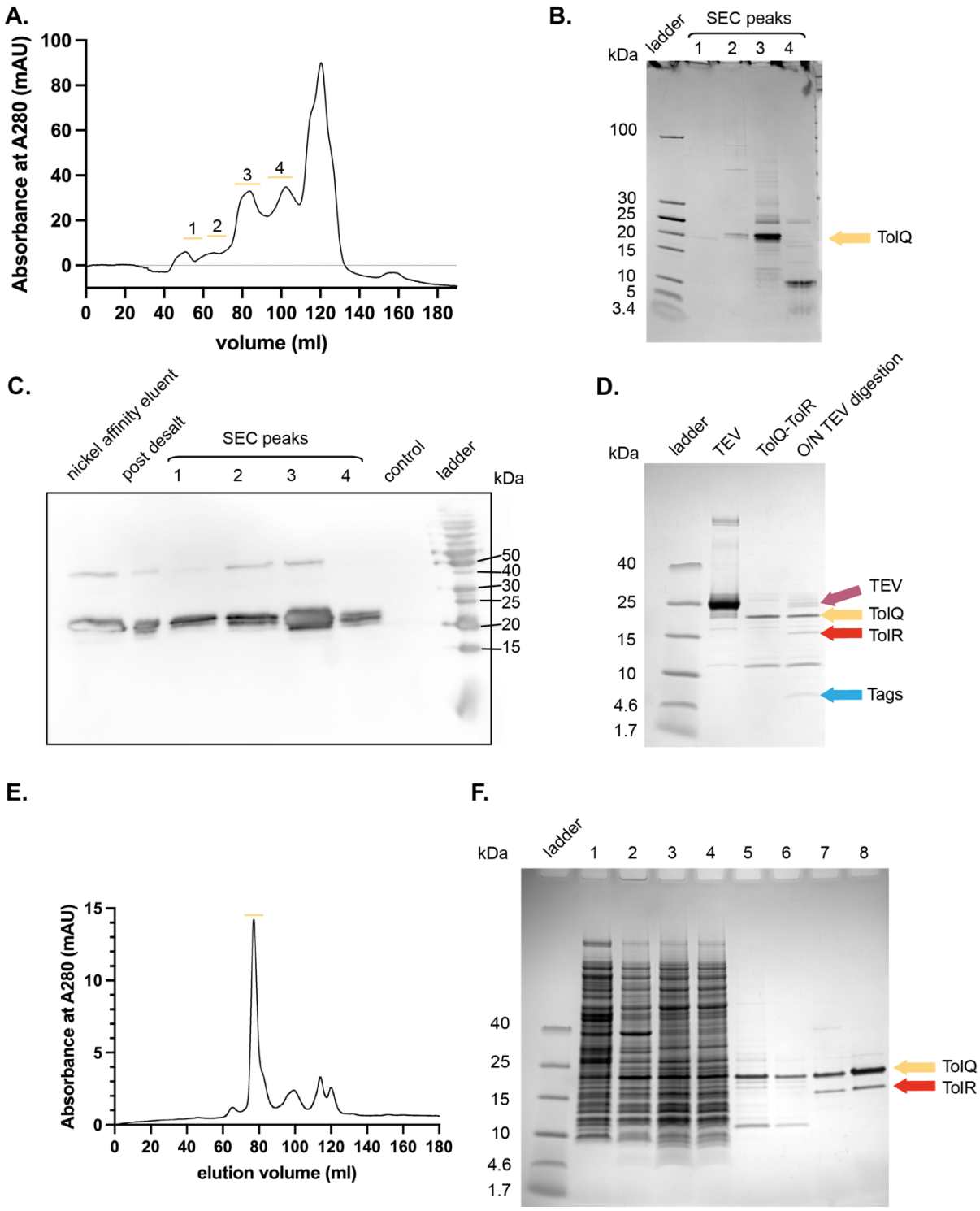


Figure 4.3 Immunoblotting and cleavage with TEV protease identified the presence of both TolQ and TolR.

Figure legend continued p106.

Figure 4.3 continued. A. SEC elution profile of TolQ-TolR on HiLoad 16/60 Superdex 200 column in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 0.1% Cymal-6. **B.** Coomassie stained 10-20% Tricine SDS-PAGE gel of SEC elution peaks 1-4. **C.** Western blot using α-strep primary antibody shows TolR is present throughout the purification and in elution peaks 1-4. The negative control is *E. coli* TolA domain II (2.4 mg/ml). Ladder used acts as positive control as proteins in ladder are strep tagged. **D.** Coomassie stained 10-20% Tricine SDS-PAGE gel of TolR cleavage with TEV protease. After incubation with TEV protease new bands were visible in the TolQ-TolR sample consistent with cleaved TolR and its tags. **E.** SEC elution profile of TolQ-TolR on HiLoad 16/60 Superdex 200 column in 50 mM Tris/HCl pH 8.0, 300 mM NaCl, 0.1% Cymal-6 following further purification optimisation. Yellow band indicates peak pursued for further work. **F.** Coomassie stained 10-20% Tricine SDS-PAGE gel. Band at 25 kDa and at 16 kDa confirmed to be TolQ and TolR respectively by peptide mass fingerprinting. 1. Soluble material following cell lysis. 2. Total membrane. 3. Solubilised inner membrane. 4. Nickel affinity chromatography flow through. 5. Nickel affinity chromatography eluent. 6. Desalted. 7. Cleaved overnight with TEV protease followed by reverse nickel affinity chromatography. 8. Concentrated SEC peak.

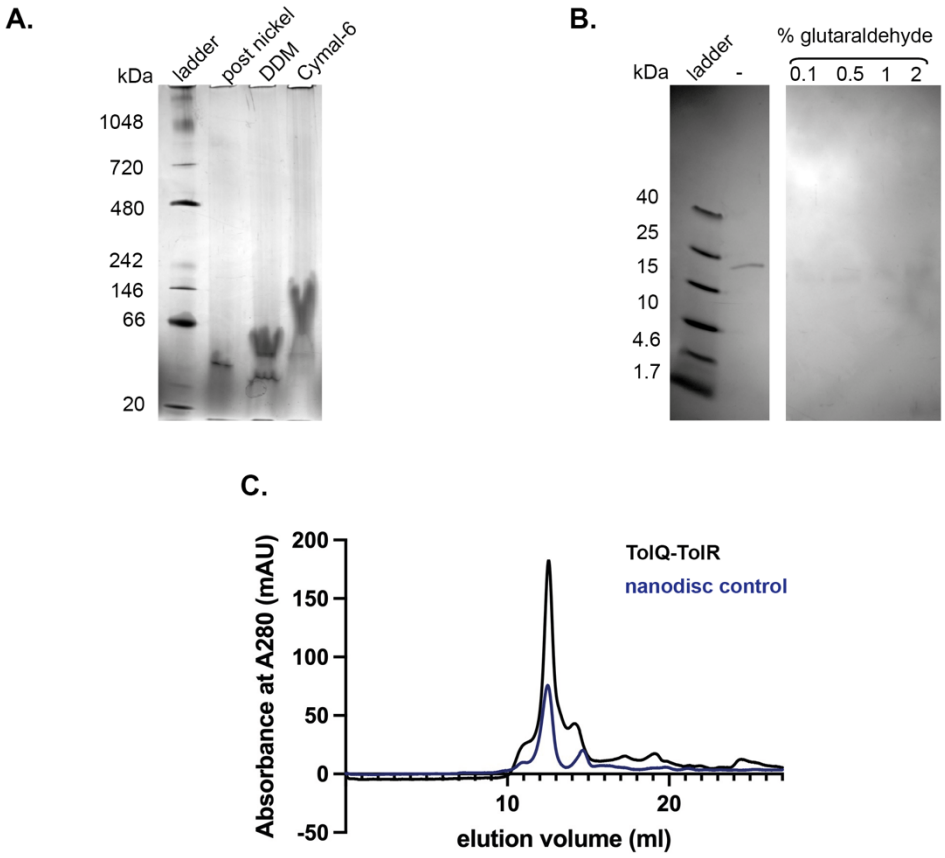


Figure 4.4 Techniques unable to validate the purification of the TolQ-TolR complex.
A. Coomassie stained Native-PAGE gel shows when TolQ-TolR nickel affinity purified in 1% Cymal-6 was desalted into 1% DDM or 1% Cymal-6, no bands consistent with a 150 kDa complex of TolQ-TolR could be observed. **B.** Coomassie stained 10-20% Tricine gel. Crosslinking of 0.1 mg/ml TolQ-TolR with a range of concentrations of formaldehyde or glutaraldehyde for 15 minutes did not reveal observable complexes of TolQ-TolR. **C.** Attempts to reconstitute TolQ-TolR into the MSP1E3D1 nanodisc with *E. coli* lipid extract were unsuccessful. SEC elution profile of reconstituted TolQ-TolR on Superdex S200 10/300 increase saw elution peak at the same volume as the protein-free control, suggesting the TolQ-TolR had precipitated during the process of reconstitution

In vitro crosslinking with glutaraldehyde was also unsuccessful in identifying a complex of TolQ-TolR or TolQ oligomers (Figure 4.4).

Next, native Mass Spectrometry (native MS) was used to investigate the stoichiometry of the putative TolQ-TolR complex. Initial experiments revealed the Cymal-6 detergent to be unsuitable for native MS and identified LPS contamination. Optimisation of the purification process through column cleaning and detergent screening following purification identified conditions more amenable to native MS (Figure 4.5.A-C). The presence of TolQ oligomers from monomers up to pentamers could be detected by native MS, however no TolR could be detected. Further work could refine the conditions required and determine the stoichiometry of the complex, but time limitations did not allow for the completion of this work within this study.

Validating the existence of the TolQ-TolR complex and experimentally determining its stoichiometry proved to be an incredibly difficult aspect of this study. The ability to pull out TolQ from the inner membrane fraction using His_6-tagged TolR provided evidence that a complex was present at this stage. However, confirming the existence of this complex following SEC was challenging.

4.2.3 The TolQ pentamer is observable using cryo-EM

Cryo-EM was used to investigate if multimeric protein complexes could be observed. Initially, an approach of reconstituting TolQ-TolR into nanodiscs was trialled, however issues with protein precipitation upon removal of detergent were encountered, and no TolQ-TolR could be observed in the assembled nanodiscs (Figure 4.4).

At this point, the earlier choice of Cymal-6 as the detergent for purifying TolQ-TolR was re-examined to find a detergent more suitable for cryo-EM. A second detergent screen was conducted to reassess the stability of TolQ-TolR solubilised and purified in either 1% Cymal-6, 1% DDM or 1% LMNG through blue native PAGE. After SEC in 0.01% LMNG higher molecular weight complexes could be observed, including a strong band between 146 and 242 kDa, consistent with a $TolQ_5$-$TolR_2$ complex (Figure 4.6.A). In addition, an improved yield

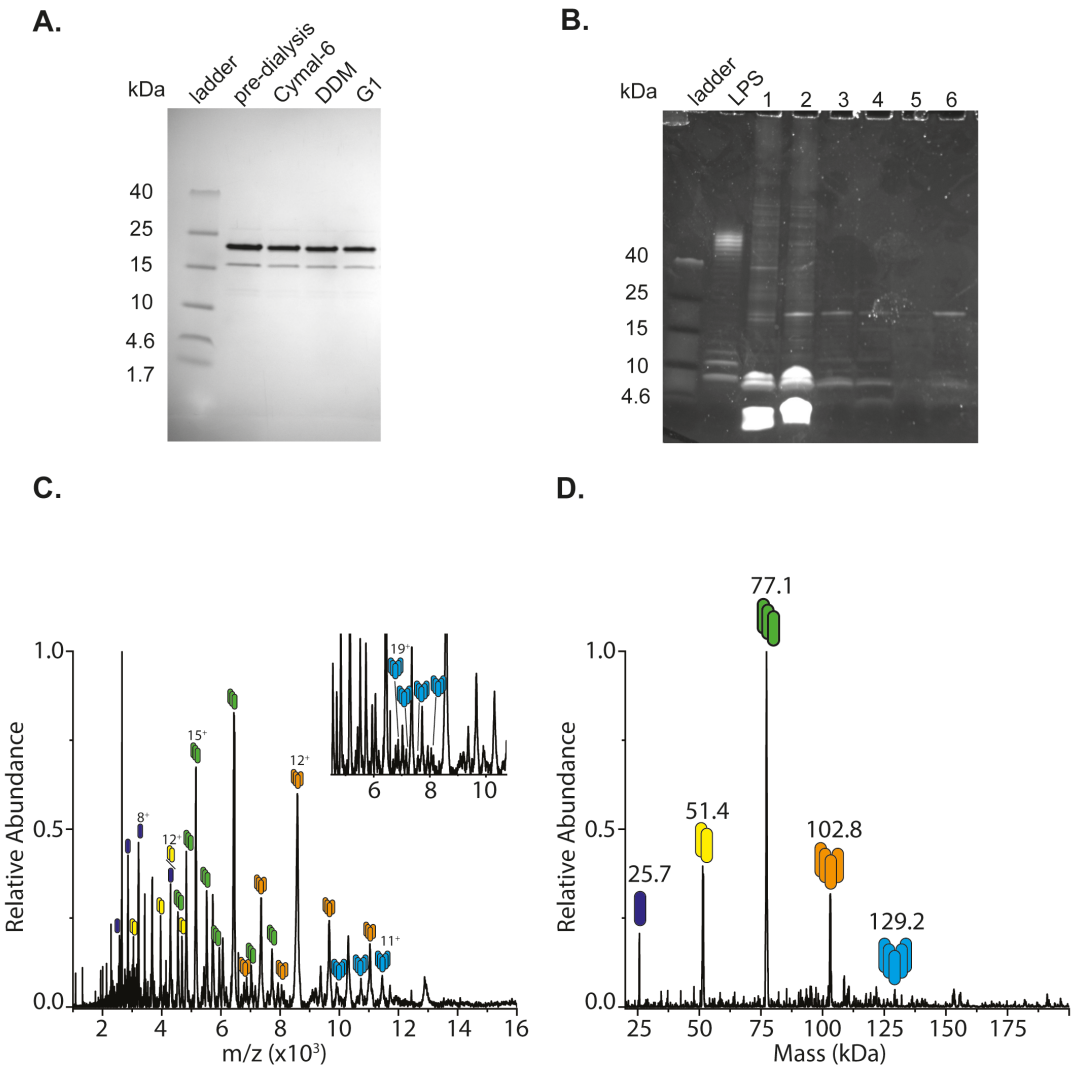


Figure 4.5 The presence of TolQ oligomers was confirmed by native MS.
A. Coomassie stained 10-20% Tricine SDS-PAGE gel to confirm the presence of TolQ-TolR in 200 mM ammonium acetate pH 7.4 with three different detergents – 0.06% Cymal-6, 0.02% DDM or 0.04% G1 compared to control of the SEC buffer (50 mM Tris/HCl pH 8.0, 300 mM NaCl, 0.1% Cymal-6). Both TolQ and TolR bands were present in all samples suggesting TolQ-TolR was stable in solution with ammonium acetate and the different detergents. **B.** ProQ Emerald Green stained 10-20% Tricine SDS-PAGE gel after optimisation of purification process showed no LPS contamination of the final TolQ-TolR sample – no characteristic LPS bands or laddering can be observed. **C.** Native MS Spectra demonstrate observable oligomers of TolQ. Peaks for monomers (dark blue), dimers (yellow), trimers (orange), tetramers (green) and pentamers (light blue) are identified. **D.** Deconvoluted native MS spectra again shows the presence of and size of oligomers of TolQ, with a weak peak for the pentamer.

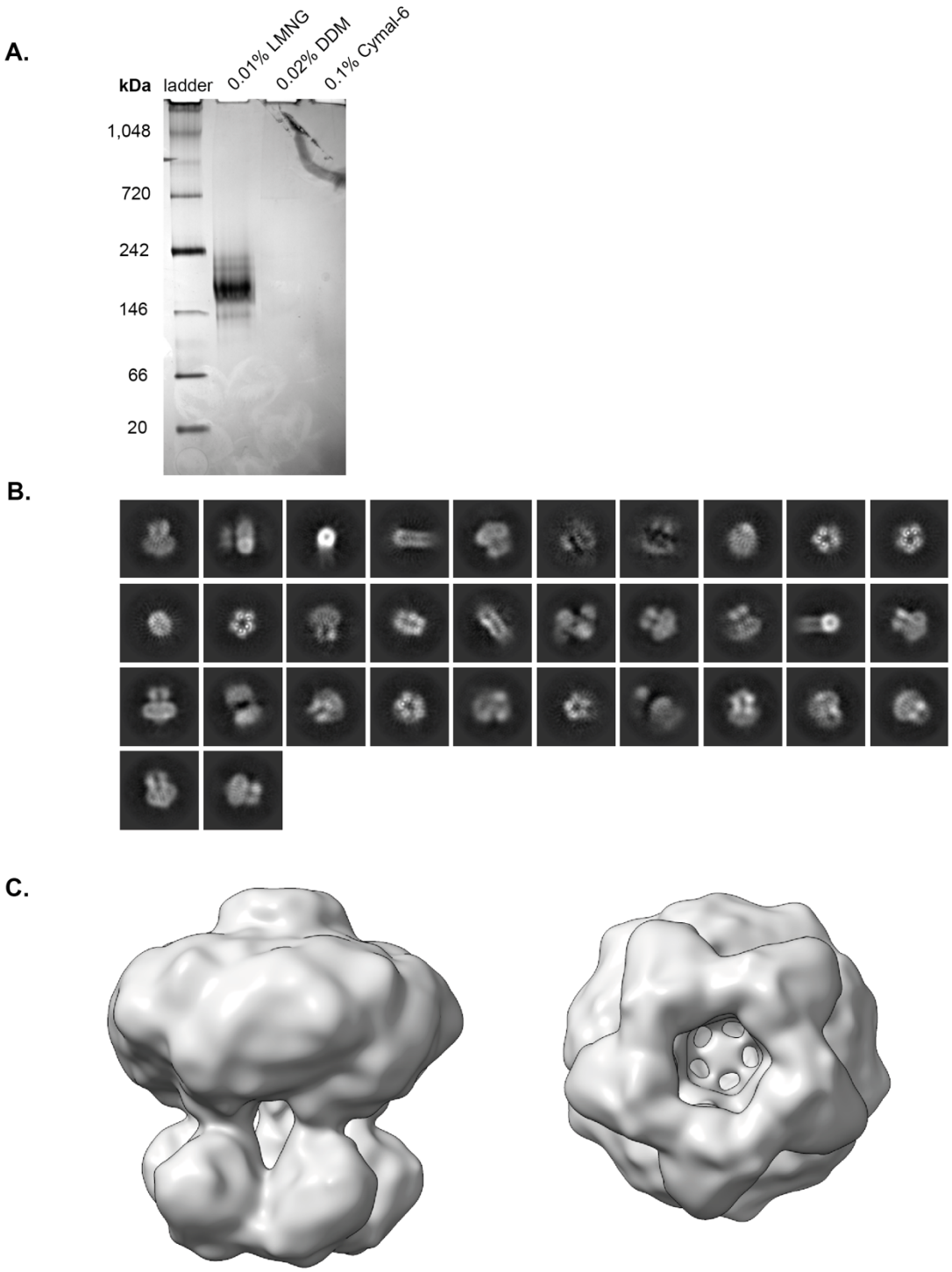


Figure 4.6 Preliminary cryo-EM shows TolQ is a pentamer.
A. Coomassie stained 4-16% Blue Native-PAGE gel comparing TolQ-TolR (20 μM) in 0.01% LMNG, 0.02% DDM or 0.1% Cymal-6. Only in 0.01% LMNG is a band between 146 and 242 kDa present, consistent with a complex of TolQ-TolR. **B.** Representative class averages from 1468 images from 50,000 particles. **C.** A low-resolution 8 Å model of TolQ-TolR was generated by processing in CryoSparc. The model demonstrates a pentamer of TolQ, with density in the centre of the pore consistent with the presence of TolR.

of 7.5 mg/l was achieved, suggesting LMNG was the appropriate detergent to use for cryo-EM going forward.

Grids with a sample of TolQ-TolR in 0.01% LMNG were prepared and screened on a Talos Arctica. 1468 images were collected from 50,000 particles and selected classes gave 18,000 particles that were aligned to generate a low resolution (8 Å) model (Figure 4.6.B-C). The preliminary low-resolution model demonstrates a pentameric pore with observable density in the pore consistent with TolR sitting within five copies of TolQ.

4.2.4 Modelling TolQ-TolR highlights a proton channel conserved with ExbB-ExbD and MotA-MotB

Alongside the experimental work to determine the stoichiometry and structure of TolQ-TolR, advancements in structure prediction and model building allowed for the generation of a computational model of TolQ-TolR. Initially a homology model of TolQ-TolR was built based upon ExbB-ExbD (PDB: 6TYI) using SWISS-Model (Waterhouse *et al.*, 2018; Szczepaniak, Press and Kleanthous, 2020). Following this an artificial intelligence-based model was built using AlphaFold2 and the Rosetta suite (Webby *et al.*, 2022b). The most recent model is discussed here.

The AlphaFold2 monomer of TolQ closely resembles the monomer of ExbB (RMSD 2.5 Å in Pymol) with both comprising three transmembrane helices and a large cytoplasmic structured loop extending 50 Å into the cytoplasm (Figure 4.7). The AlphaFold2 pLDDT scores of both TolQ and TolR were of predominantly high confidence (Figure 4.8). TolQ formed a pentamer through introduction of five-fold symmetry and TolR's TPHs reside within the TolQ pore (Figure 4.7.B). The TolQ pentamer forms a 25 Å pore, compared to ExbB's at 17 Å and smaller than MotA's at 35 Å (Figure 4.7.C). Analysis of the electrostatic surfaces of the three pores reveals discrete regions of charge: both MotA and ExbB have a positive region at the IM interface, followed by a negative region, whereas TolQ only has an electronegative band within the cytoplasmic side of the pore (Figure 4.7.D).

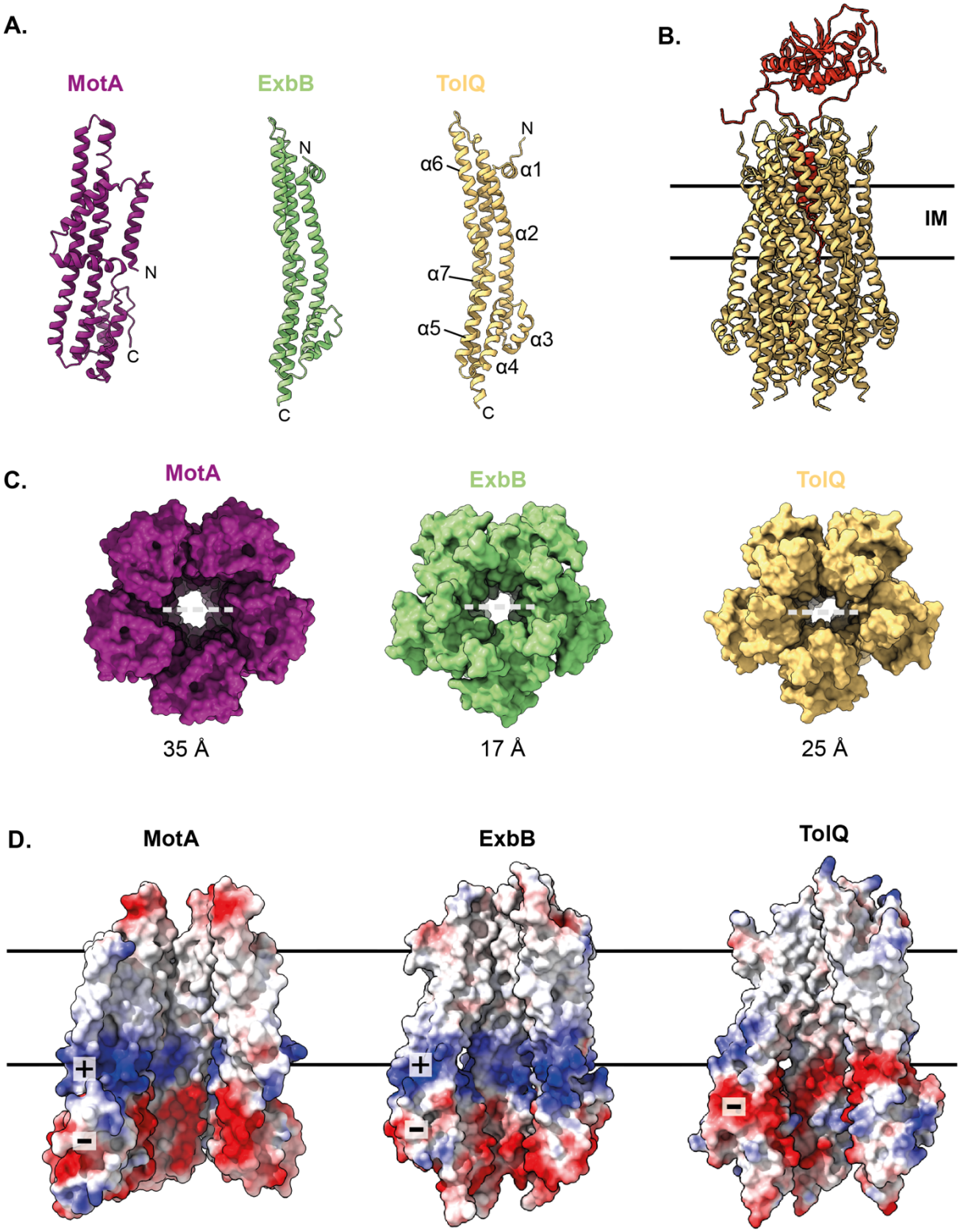


Figure 4.7 TolQ-TolR complex model resembles ExbB-ExbD and MotA-MotB.
A. Monomeric structures of *C. jejuni* MotA (PDB: 6KYM), *E. coli* ExbB (PDB: 6TYI) and the *E. coli* TolQ AlphaFold2 prediction with labelled helices. ExbB and TolQ show similar structural arrangement. **B.** The model of pentameric TolQ and dimeric full-length TolR generated using the Rosetta suite. **C.** Comparison of the pores of MotA, ExbB and TolQ, surface view of cytoplasmic face of each stator. **D.** Electrostatic analysis visualised with a cut-through show differing electrostatic regions within the three pores. All three have a band of negative charge within the cytoplasmic region, but the positive band observed in MotA and ExbB at the membrane border is not seen in TolQ. Taken from (Webby *et al.*, 2022b).

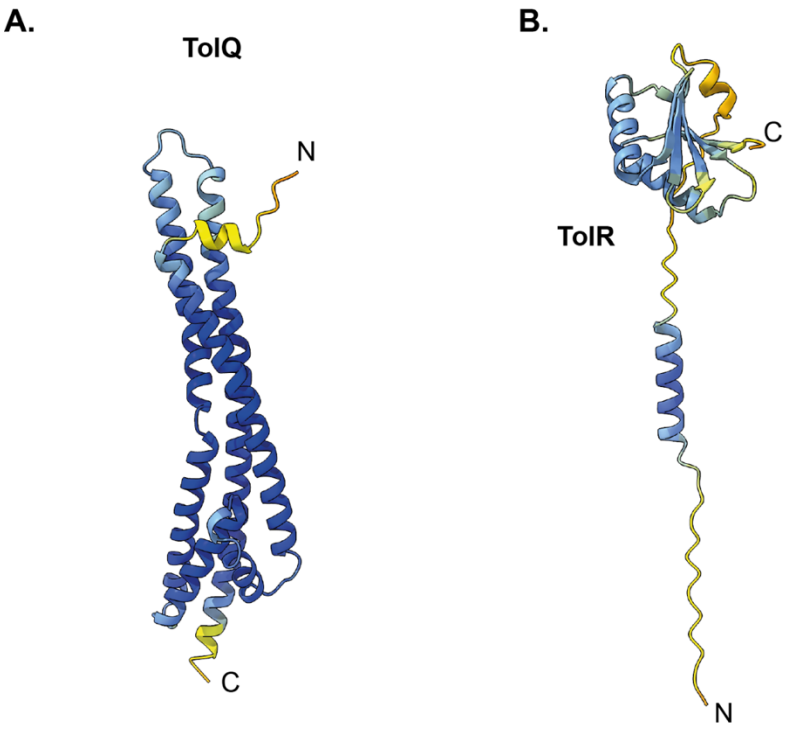


Figure 4.8 Monomers of TolQ and TolR predicted by AlphaFold2 coloured by pLDDT score.

Rigid body fitting of the TolQ-TolR model into the 8 Å EM map of TolQ-TolR was carried out to evaluate the model and EM map (Figure 4.9). The model fitted to the EM map with a correlation value of 0.813 through ChimeraX rigid body fitting, demonstrating a reasonable alignment of the model in the map. The detergent shell of the EM map aligns with the hydrophobic TMHs of TolQ in the model (Figure 4.9.A). The pentamer of TolQ fits into the pentameric EM shape and provides cross-validation of the assumption TolQ forms a pentamer when building the model (Figure 4.9.B). In future, the model could provide a useful tool for refinement of a higher resolution EM map.

To examine the potential proton pathway through TolQ-TolR conserved residues were compared between the TPH of TolR, ExbD or MotB, and the interior TMHs of TolQ, ExbB and MotA in *E. coli*. Alignment of the primary sequences highlighted highly conserved residues described previously in a consensus sequence (Figure 4.10.A) (Cascales, Lloubès and Sturgis, 2001). The positions of these residues on the stator models were identified and compared (Figure 4.10.B). In all three models the conserved Asp23 on TolR/MotB/ExbD is the

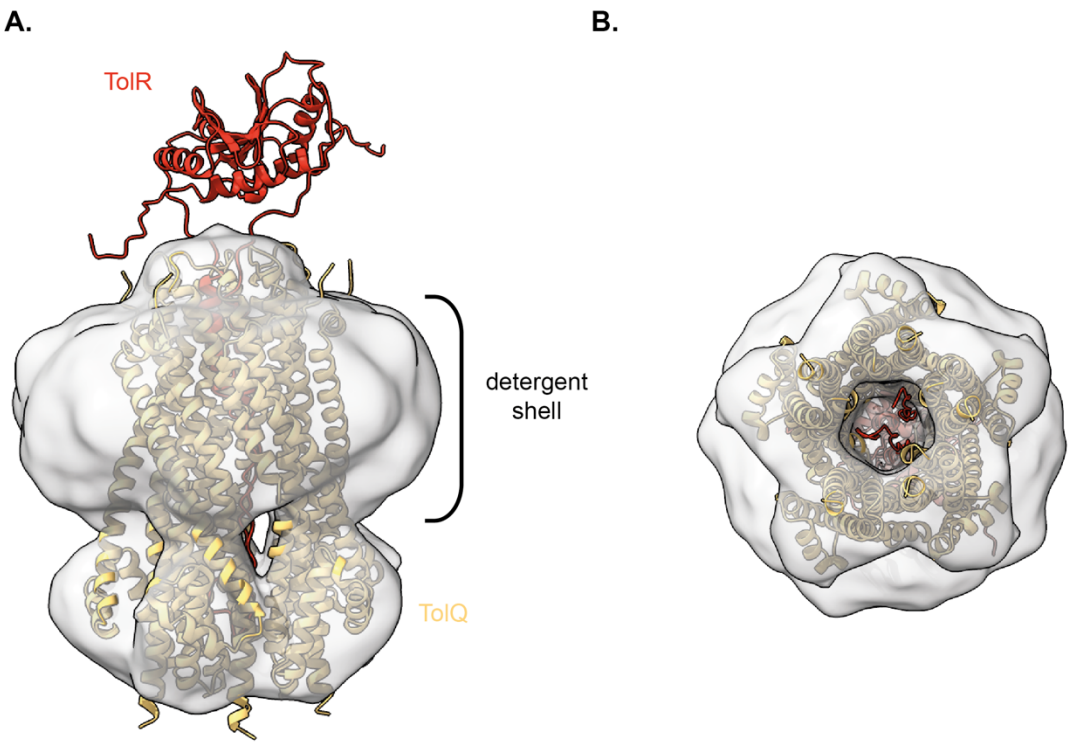


Figure 4.9 Docking of TolQ-TolR model into 8 Å EM map demonstrates the pentamer of the TolQ model fits in the pentameric map.
A. Side view of the TolQ-TolR model rigid-body fitted into the TolQ-TolR 8 Å EM map, correlation value 0.813. The detergent shell region highlighted is in line with the expected TMH of TolQ**. B.** Bottom view of the same fitting as in A. This view clearly demonstrates that the TolQ fits within the density as a pentamer**.**

closest to the cytoplasm, suggesting it is the final residue to be protonated as a proton is transduced Figure 4.10.C).

In their MotA-MotB structure, Santiveri *et al.,* 2020 identified a potential proton pathway gated by a Phe residue that moves when the MotB plug helix is removed (Santiveri *et al.*, 2020). Ratliff, Buchanan, and Celia, 2022 proposed a Leu residue, 175, in ExbB to carry out the same role in gating access to the proton channel (Ratliff, Buchanan and Celia, 2022). Leu175 in TolQ is conserved with Leu178 in ExbB, and therefore could also have the same role (Figure 4.10.C). Analysis of the TolQ-TolR stator using MOLE 2.5, software that predicts the presence of solvent channels through proteins, identified a channel consistent with mutation analysis (Figure 4.10.D). Analysis was carried out as in Santiveri *et al*. 2020 and Ratliff, Buchanan, and Celia, 2022, using a 1 Å bottleneck radius and omitting Leu175 (Santiveri *et al.*, 2020; Ratliff, Buchanan and Celia, 2022). The TolQ Thr145 and Thr178 and TolR Asp23 residues previously shown to have a role in proton transduction line the pore

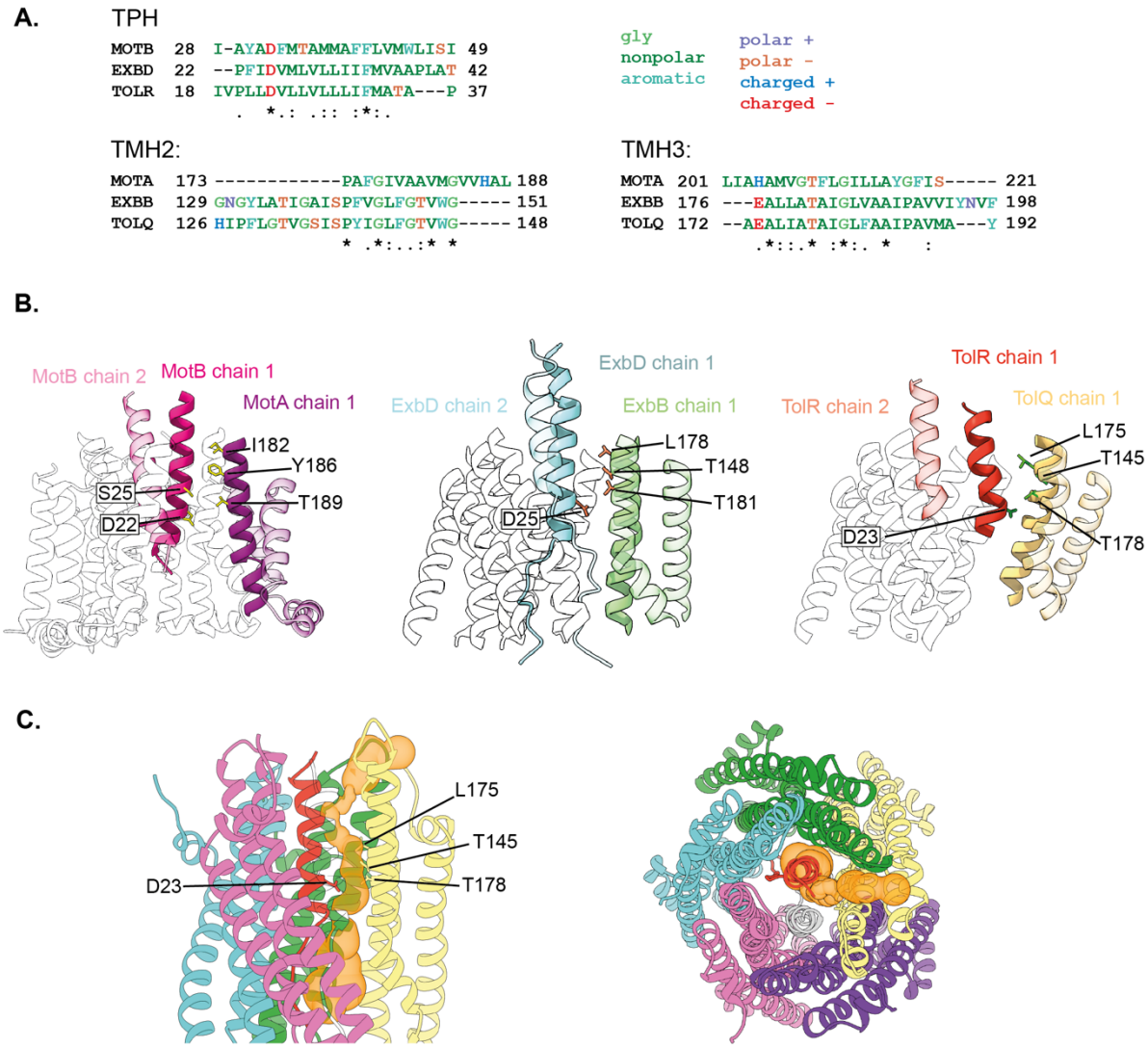


Figure 4.10 The TolQ-TolR model has conserved protonatable residues consistent with a proton channel.
A. T-Coffee alignments of MotB, ExbD and TolR TPHs, and of MotA TMH3 and TMH4 with TMH2 and TMH3 of ExbB and TolQ. Protonatable residues are conserved within the stator pore. **B.** Comparison of the positioning of the conserved residues in the pore of each stator. One copy of MotA/ExbB/TolQ is removed to allow observation of the pore interior. **C.** Predicted proton channel through TolQ-TolR represented as orange isosurface. The channel was generated using MOLE 2.5 online (https://mole.upol.cz/, Pravda et al., 2018) using the same parameters as Santiveri *et al*., 2020 and Ratliff, Buchanan, and Celia, 2022: Omitting Leu175 and a 1 Å bottleneck radius. One copy of TolQ and TolR are removed to reveal pore interior. The conserved residues Leu175, Thr145, Thr178 and Asp23 can be seen to line the channel. A top-down view also shows the location of these residues in relation to the channel.

(Goemaere, Cascales and Lloubès, 2007). Although, Leu175 does not obstruct the path of the channel in the same way as Phe186 and Leu178 in *C. jejuni* MotA and *E. coli* ExbB, respectively (Figure 4.10.D).

4.3 Discussion

The work described in this chapter details the progress made in purifying the TolQ-TolR stator component and determining the stoichiometry of this complex through a range of biochemical approaches. In the absence of an experimentally determined structure, a computational model of TolQ-TolR provided a platform for analysis and comparison to the homologues MotA-MotB and ExbB-ExbD.

To continue the work on the TolQ-TolR complex assembly described here, it is important to understand the hurdles faced and what information they offer on the TolQ-TolR complex. The first difficulty encountered regarded yield of protein. Through most of the work detailed, the best yield of TolQ-TolR obtained was 0.02 mg/l, meaning a large volume of cells had to be grown to achieve a yield with enough protein to use experimentally. In practice this meant each optimisation iteration took a long time, making progress slow.

The presence of TolQ and TolR following nickel affinity chromatography despite the His_6 tag only being present on TolR provided the first suggestion that a complex of TolQ-TolR was being purified. Following SEC and concentration in a 100 kDa concentrator bands for both TolQ and TolR could be observed. If the complex had dissociated during this step dimeric TolR would not be present with this cut-off. Combined, these two observations provided evidence that the complex of TolQ-TolR was purified.

Once purified, however, the complex proved to be unstable and highly susceptible to disassociating when the detergent or buffer conditions were altered. This was not unexpected as this complex is predicted to be highly dynamic, implying transient inter-subunit interactions. Before the complexes of ExbB-ExbD and MotA-MotB were determined to have a 5:2 ratio, the stoichiometries attributed to the two complexes ranged from 4:2 to 6:3, supporting the suggestion that obtaining a complex *in vitro* is dependent on the conditions used and is highly variable (Yonekura, Maki-Yonekura and Homma, 2011; Celia *et al.*, 2016; Maki-Yonekura *et al.*, 2018). Optimisation of the native MS conditions offered further evidence of complex

formation but requires further finessing to obtain a more definite answer regarding the stoichiometry of the complex.

Re-examining the detergent choice provided a breakthrough in identifying a detergent TolQ-TolR suitable for cryo-EM and improving yield. Preliminary cryo-EM data following this has provided the first evidence that TolQ forms a pentamer and is an encouraging indication that solving the structure is within reach. Altogether, the data points to the successful purification of the TolQ-TolR complex and provides a foundation for further work. With time and patience, it is foreseeable that optimum conditions for native MS and cryo-EM will be found, allowing for validation of the stoichiometry of the TolQ-TolR complex and the experimental determination of a high-resolution structure.

Machine learning and computational methods of structure prediction and assembly provided an avenue to explore the TolQ-TolR complex in the absence of an experimental structure. The combination of the structures of ExbB-ExbD and MotA-MotB with stoichiometry 5:2, and native MS data observing 5:2 ExbB:ExbD, point to a TolQ-TolR complex with a 5:2 stoichiometry (Chorev *et al.*, 2018; Celia *et al.*, 2019; Deme *et al.*, 2020; Santiveri *et al.*, 2020). Extensive crosslinking analysis has been carried out on TolQ-TolR in past studies, and the generation of the TolQ-TolR model allowed for analysis of the positioning of key residues within the complex pore. In particular, residues Thr145 and Thr178 of TolQ alongside Asp23 in TolR form a constellation that could feasibly form a proton pathway consistent with the pattern of conserved residues seen in MotA-MotB and ExbB-ExbD. The identification of a potential proton channel 'gating' residue, Leu175 in TolQ, stems from conservation with ExbB's Leu178 and comparison to MotA's *Cj*Phe186. In the MotA-MotB model Phe156 moves in response to the removal of the MotB plug domain, opening access to the proton channel (Santiveri *et al.*, 2020). However, neither TolR nor ExbD contain plug helices that cover the entrance to the pore. Further to this, the generation of an accessible proton channel through the TolQ-TolR complex using MOLE 2.5 allowed for investigation into the accessibility of these conserved residues. The key residues line the channel and demonstrate a potential pathway for protons through the TolQ-TolR complex. However, unlike in MotA *Cj*Phe186 and ExbB

Leu178, TolQ Leu176 does not obstruct the predicted channel. This could indicate that in the model analysed the helices of TolR are in the 'unplugged' state, or that an alternative residue plays this role (Figure 4.10.C).

The TolQ-TolR complex model also includes full-length TolR, where the single particle structures of MotA-MotB and ExbB-ExbD have been unable to resolve density for the MotB or ExbD periplasmic domains. The dimer conformation of TolR generated from the Rosetta modelling resembled the 'open' PG-binding conformation observed in the TolR *H. influenzae* structure. This was contrary to the hypothesis that the 'closed' strand-swapped dimer is the resting state. To form the strand-swapped dimer in the TolQ-TolR model starting from the open conformation, the β5-β5 interactions must be broken, the monomers must rotate 180° relative to one another, and the new strand swapped β-sheet must form using residues 81-86 from β1. To achieve this, the TPHs must move 10-15 Å out of the pore, meaning with the current positioning of the TolR TPHs within the pore, the strand-swapped dimer can't be formed. To resolve this discrepancy and understand whether the TolR 'closed' strand-swapped dimer is physiologically important, further investigation is required. Using the TolR Tyr117Cys crosslinked mutant known to impair system function could provide an avenue to explore this (Wojdyla *et al.*, 2015). Molecular dynamic simulations modelling the TolQ-TolR complex in protonated and unprotonated states could also shed light on the mechanism by which TolQ-TolR responds to the PMF.

The successful purification of the TolQ-TolR complex described provides a foundation for further structural work. In addition, the TolQ-TolR complex models generated provide clues as to how the complex transduces protons. Many questions still remain regarding the assembly and function of TolQ-TolR; obtaining structures of TolQ-TolR could provide further clues as to how the complex assembles and point to how proton transduction is translated to force transduction.

5 Using photoactivatable crosslinking to probe the interaction between TolQ-TolR and TolA

Contributions: Peptide mass fingerprinting, model building – Dr Melissa Webby. Microscopy – Dr Joanna Szczepaniak.

5.1 Introduction

TolQ-TolR and TolA are recruited independently to the septum during cell division (Gerding *et al.*, 2007), yet it is the interaction between them that is key to the translation of the PMF to force transduction. *In vivo* crosslinking using formaldehyde mapped the interaction site with TolQ to TolA's N-terminal domain (Derouiche *et al.*, 1995).

Investigation into this interaction highlighted a conserved motif named the 'SHLS' motif thought to be responsible for the interaction within the IM. First identified by Koebnik *et al.* 1993, the SHLS motif consists of a $SerX_3HisX_6LeuX_3Ser$ sequence where each named residue is found on one face of the TMH of both TolA and TonB (X denotes non-conserved residues) (Koebnik *et al.*, 1993). When mutated, crosslinking products between TolQ-TolA and TolA-TolR were no longer present, implying the mutations prevented complex formation (Germon *et al.*, 1998). In addition, suppressor mutations in TolQ could be generated to recover the function of the system (Germon *et al.*, 1998). Further study into the role of this motif in TonB through mutational analysis determined that it is the His20 and Ser16 at the start of the motif that are required for interaction with ExbB-ExbD (Larsen and Postle, 2001).

One technique that has successfully been used to examine protein-protein interactions in molecular detail is the use of photoactivatable amino acids. The method involves substituting a residue of interest with an amber stop codon, TAG. Cells expressing the mutated gene are grown in the presence of a plasmid-encoded tRNA synthetase pair capable of incorporating the photoactivatable amino acid benzoyl-L-phenylalanine (BPA) at the position of the TAG site (Figure 5.1.A) (Chin *et al.*, 2002). When the BPA is exposed to UV light at 365 nm, it will form covalent bonds with any C-H bond within 3-4 Å. This covalent crosslinking of complexes allows for the identification of what the protein of interest was interacting with *in vivo* when it was exposed to UV (Figure 5.1.B) (Dormán *et al.*, 2016). The application of LC-MS/MS of crosslinked species allows for the identification of residue-specific sites of interaction (Dietsche *et al.*, 2016).

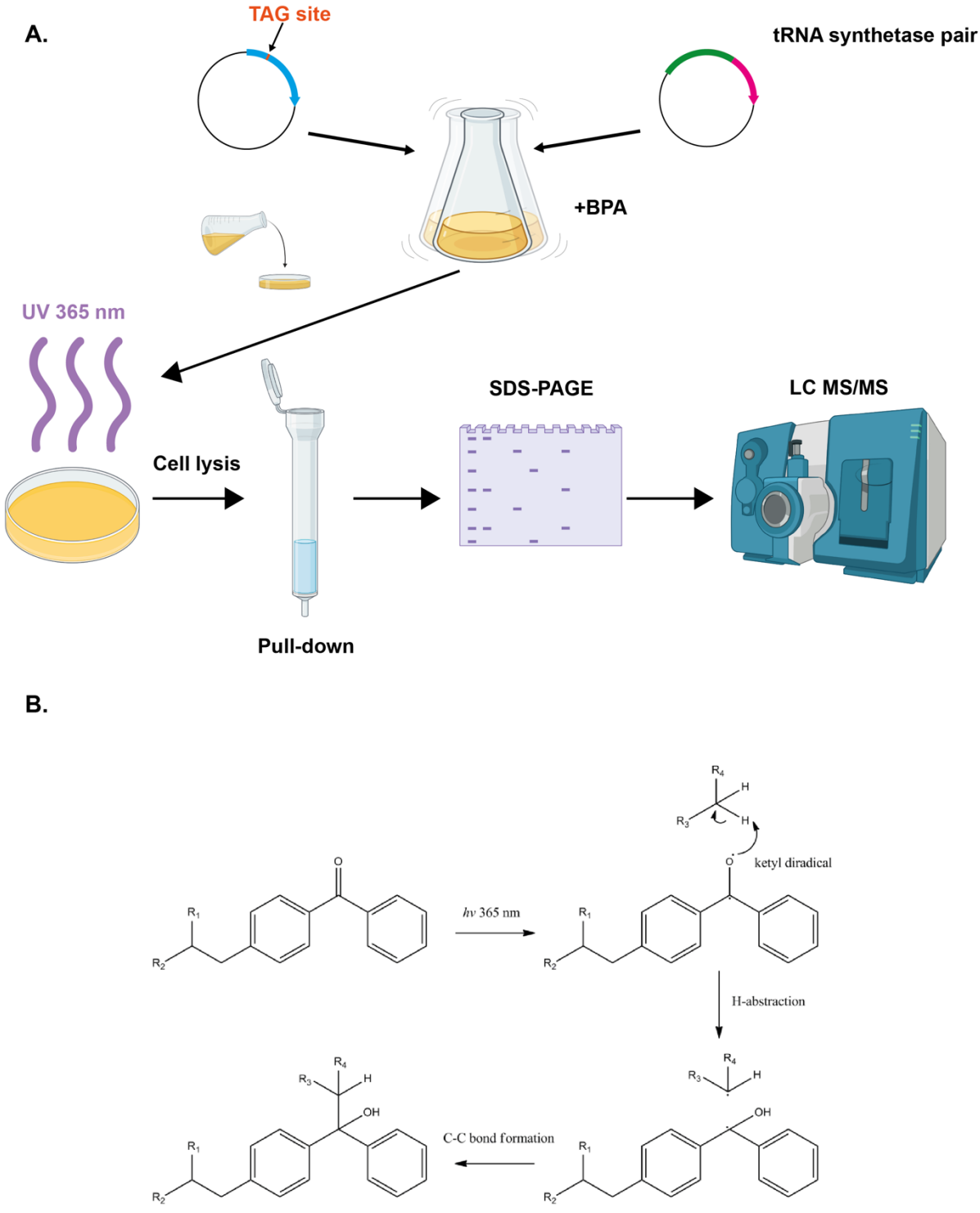


Figure 5.1 The BPA crosslinking method.
A. Schematic representation of crosslinking process. Plasmids encoding the gene of interest (blue) with an amber stop codon at the site of interest (orange), and the tRNA synthetase pair (green/pink) that incorporates BPA are expressed in cells in the presence of BPA. The cells are poured into a petri dish and exposed to UV at 365 nm. Cells are then lysed, and an affinity pulldown is used to isolate the protein of interest with any crosslinked species. The protein and any crosslinked adducts are separated by mass using SDS-PAGE, the bands excised, and the contents of the bands examined by LC MS/MS. **B.** The reaction mechanism for BPA. When irradiated with UV at 365 nm BPA forms a diradical ketyl species that reacts with a C-H bond within 3-4 Å through H-abstraction. Radical recombination with the alkyl radical then forms a crosslinked adduct (Dormán *et al.*, 2016). Figure created with Biorender.com.

To better understand the interaction between TolQ-TolR and TolA the BPA crosslinking approach was employed in this study. Firstly, sites of interest within TolA's TMH were chosen for mutation, followed by functional analysis of the Tol-Pal system and expression trials to determine if the mutants and full-length TolA were being expressed. Immunoblotting and mass spectrometry were used to search for crosslinks. Alongside the crosslinking approach, the models of TolA and TolQ-TolR described in Chapters 4 and 5 were docked computationally to examine the resulting interaction interface between the proteins.

5.2 Results

5.2.1 The TolA His22^{BPA} and Ile15^{BPA} mutants are functional

Three sites within TolA's TMH were chosen for investigation using BPA crosslinking: Ile15, His22 and Ser32 (Figure 5.2.A). The rationale behind this choice was to include sites on the same face of the helix as the SHLS motif that covered the start, middle, and end of the TMH (residues 15 to 36) (Figure 5.2.B). His22 is part of the SHLS motif and previous mutation studies showed when His22 was mutated to Pro or Arg the *tol* phenotype was observed (Germon *et al.*, 1998). The mutation of His22 to BPA was therefore expected to impact the function of the Tol-Pal system.

The TolA^{BPA} mutants were constructed in the pNP4 plasmid with a C-terminal His_6 tag, referred to as pTolAIle15^{BPA}, pTolAHis22^{BPA} and pTolASer32^{BPA}. The mutants were transformed into the *ΔtolA::kan* strain JW0729 alongside the pEVOL-pBpF plasmid (pVol/pV) encoding the tRNA synthetase/tRNA pair. The pNP4 vector was chosen over the pET21a(+) used in Chapter 3 for TolA-II expression firstly because expression in pET21 vectors requires the DE3 prophage that carries the T7 RNA polymerase gene. The *ΔtolA::kan* strain is incompatible with using pET21a(+) as it does not carry this gene nor contain a promoter compatible with *E. coli* RNA polymerase. In pNP4 the *E. coli lac* promoter is used, and so is compatible with the *E. coli* RNA polymerase. Secondly, the expression of *tolA* had to be

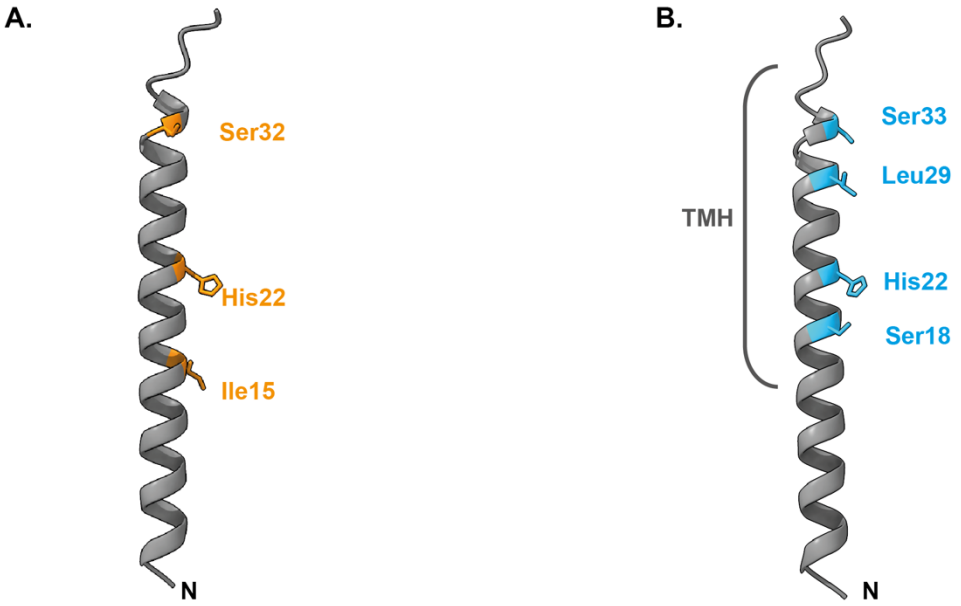


Figure 5.2 The positions of sites chosen for BPA crosslinking in the TMH of TolA visualised on the AlphaFold2 prediction of TolA-I.
A. The sites Ile15, His22 and Ser32 (orange) were chosen for BPA crosslinking studies. **B.** The residues of the SHLS motif (blue) are found on the same face as the residues chosen for BPA crosslinking studies.

induced through a different promoter than expression of the tRNA synthetase pair, which is under the arabinose inducible *araBAD* promoter. Preliminary expression trials of *tolA-His*$_6$ induced by the *araBAD* promoter in the pBAD24 plasmid alongside pVol found inducing both plasmids in the same way did not allow for optimisation of expression levels of *tolA*. Finally, pNP4 and pVol have different origins of replication and selection markers: p15A and ampicillin for pNP4, and pBR322 and chloramphenicol for pVol. Therefore, the plasmids are compatible and can be co-transformed and selected for.

The function of the Tol-Pal system in the three TolA^{BPA} mutants was first assessed by assaying detergent sensitivity. This was tested by growth and induction of protein expression in liquid culture, followed by plating on 2% SDS in solid agar. In wild-type *E. coli* strain BW25113, growth is observed on 2% SDS agar in all conditions across the dilutions plated (Figure 5.3, Figure 5.4.A). In comparison, *ΔtolA::kan* does not grow on 2% SDS, consistent with the *tol* phenotype (Figure 5.3, Figure 5.4.A). Expression of plasmid encoded TolA-His$_6$ (pTolA) recovers growth on 2% SDS to levels comparable to wild-type across all conditions

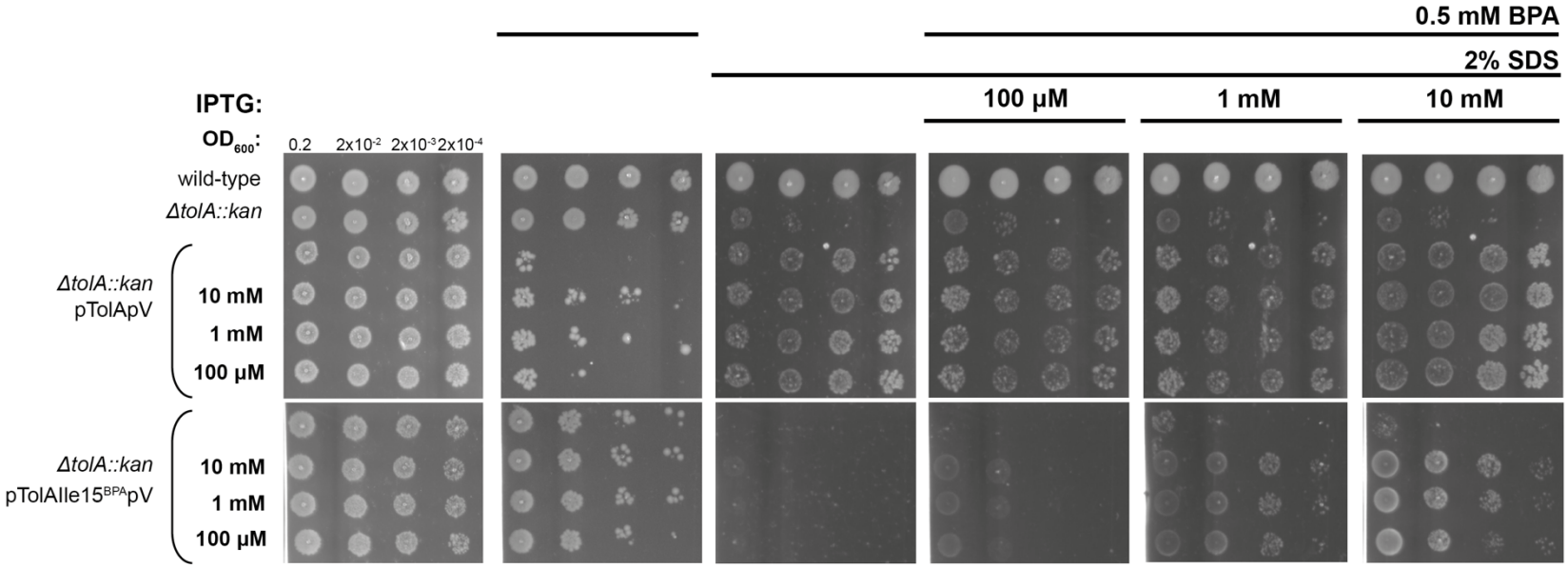


Figure 5.3 Choosing a concentration of IPTG to use in detergent sensitivity assays.
Growth of pTolApV and pTolAIle15^{BPA}pV on 2% SDS was tested with induction at three different concentrations of IPTG. All cultures were grown with 0.125% arabinose. The IPTG added to the liquid culture (y axis) and added to the plate (x axis) were assessed. The differences between expression of pTolApV with 10 mM, 1 mM and 100 µM IPTG were negligible, likely due to leaky expression levels of pTolA being enough to complement Tol-Pal function. For pTolAIle15^{BPA}pV the concentration of IPTG in the plate was more important than in the liquid culture. Growth at higher dilution than *ΔtolA::kan* was observed with pTolAIle15^{BPA}pV at 1 mM or 10 mM IPTG in the plate

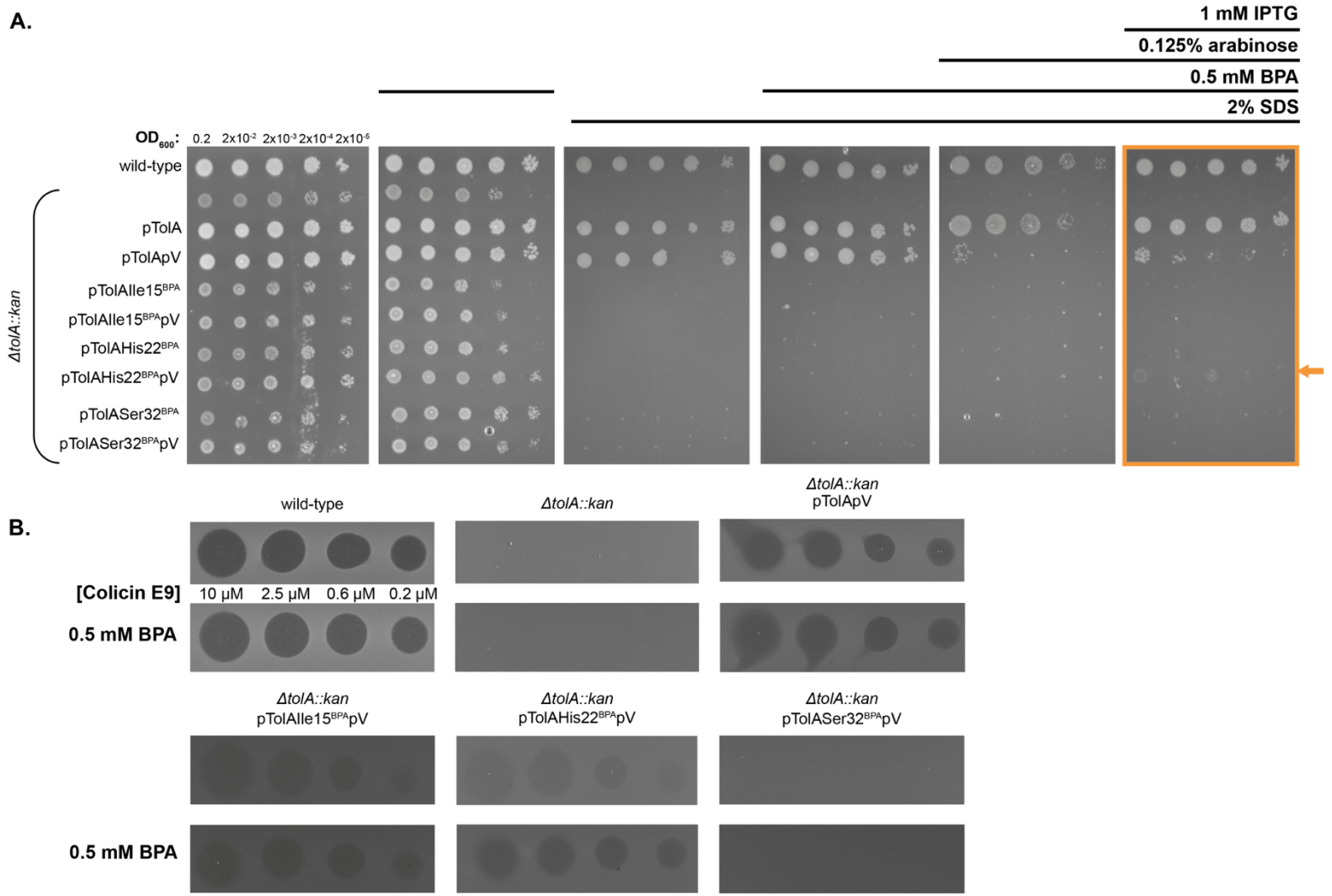


Figure 5.4 The TolAHis22^{BPA} mutant is partially functional, and the TolAIle15^{BPA} and TolASer32^{BPA} mutants are non-functional.
Figure legend continued p124

Figure 5.4 continued. A. Growth on solid agar containing 2% SDS was used to assay the functionality of the Tol-Pal system. Black lines above the plate images indicate the additives present and each plate a different condition. The background strain and plasmids present for each row are described on the left of the first plate image. The concentration of serially diluted cells plated is described using the OD_{600} shown for each spot the first plate, all other plates used the same concentrations. The orange box surrounds the key panel that contains all additives: 2% SDS, 0.5 mM BPA, 0.125% arabinose and 1 mM IPTG. The orange arrow indicates the row that demonstrates the recovery of growth of pTolAHis22^{BPA}pV and therefore partial recovery of Tol-Pal function. **B.** Plate killing assay using colicin E9 was used to assay function of the Tol-Pal system. Liquid cultures were grown with or without 0.5 mM BPA, induced with 0.125% arabinose and 1 mM IPTG, then plated in agar with the same additives. Serial dilutions of colicin E9 were spotted onto the plate and incubated overnight at 37°C. The concentrations of colicins described in the first panel applies to all plates. Killing of pTolAHis22^{BPA}pV and pTolAIle15^{BPA}pV is seen in the presence of BPA. No killing is observed for pTolASer32^{BPA}pV.

(Figure 5.3, Figure 5.4.A). It is also notable that pTolA with pVol (pTolApV) does not grow as well, with protein induction in the presence of BPA and arabinose resulting in very few colonies across the dilution series. This implies that the presence of the tRNA synthetase pair with BPA impacts detergent sensitivity. Comparison of the three mutants to these controls allowed for assessment of the function of the Tol-Pal system in the mutants.

The concentration of arabinose used for induction of the tRNA synthetase pair was chosen based upon previous optimisation work (White *et al.*, 2017; White, 2017). The concentration of IPTG used to induce *tolA* expression was chosen based upon preliminary detergent sensitivity assays conducted with pTolApV and *ΔtolA::kan* with pTolAIle15^{BPA} and pVol (pTolAIle15^{BPA}pV) (Figure 5.3). The differences between expression of pTolApV with 10 mM, 1 mM and 100 μM IPTG were negligible, likely due to leaky expression levels of pTolA being enough to complement Tol-Pal function (Figure 5.3). For pTolAIle15^{BPA}pV growth on 2% SDS was observed when 0.5 mM BPA was present, and when 1 mM or 10 mM IPTG present. The concentration of IPTG in the plate was more important than in the liquid culture, however, IPTG in the liquid culture was required for growth on 2% SDS. The difference between growth on 2% SDS with 1 mM or 10 mM IPTG was small, both grew at all dilutions plated (Figure 5.3). It was therefore decided to use the lower concentration to save wasting material. This preliminary experiment also suggested that expression of pTolAIle15^{BPA} with pVol and BPA was able to complement *ΔtolA::kan*, and therefore Tol-Pal was functional (Figure 5.3).

Once optimum conditions to assess the mutants' detergent sensitivity were identified, three biological repeats were conducted testing growth of each mutant on 2% SDS. Control plates of 2% SDS agar without 0.5 mM BPA, 0.125% arabinose or 1 mM IPTG were tested, alongside LB agar with 0.5 mM BPA. This was to determine the impact of the addition of each additive on *E. coli* growth, and how this would affect the readout of the experiment. For all pTolA^{BPA} mutants, growth was observed in the absence of SDS, confirming the cells were viable. Growth of *ΔtolA::kan* with pTolAHis22^{BPA} and pVol (pTolAHis22^{BPA}pV) was observed on 2% SDS only when 0.5 mM BPA was present and when induced with 0.125% arabinose and 1 mM IPTG. The pTolAIle15^{BPA}pV and *ΔtolA::kan* with pTolASer32^{BPA} and pVol (pTolASer32^{BPA}pV) did not grow under the same conditions (Figure 5.4.A). The lack of growth of pTolAIle15^{BPA}pV was unexpected as it had recovered growth during the preliminary work looking at IPTG concentration. The difference could be attributed to variability in the experimental conditions. The *ΔtolA::kan* in the IPTG screen grew at two dilutions on 2% SDS, whereas in the later assay it did not grow at all. The contrasting results were repeatably observed in duplicate for IPTG screening (Figure 5.3) or in triplicate for the assay with all TolA^{BPA} mutants (Figure 5.4.A). The detergent sensitivity assays suggested that TolA with the His22^{BPA} or Ile15^{BPA} mutation can partially recover function of the Tol-Pal system, but TolA with Ser32^{BPA} cannot.

To corroborate this finding and investigate the discrepancy seen with pTolAIle15^{BPA}pV detergent sensitivity, susceptibility to the group A colicin E9 was tested. Colicin E9 uses the Tol-Pal system to energise import of its cytotoxic endonuclease domain into cells. In *ΔtolA::kan* the cells are no longer susceptible to killing by colicin E9 as there is no functional Tol-Pal system to couple import with (Figure 5.4.B). Complementation with IPTG and arabinose induced pTolApV recovers colicin susceptibility (Figure 5.4.B). Plate-killing assays demonstrated that when grown in the presence of BPA and induced with IPTG and arabinose, pTolAHis22^{BPA}pV was killed by colicin E9, as indicated by the areas of clearance on plates at concentrations between 0.2 and 10 μM. Some killing was observed in the absence of BPA,

suggesting the cytotoxic domain of ColE9 is capable of translocating into cells in the absence of a functional TolA.. For pTolAIle15^{BPA}pV killing was also seen in both the presence and absence of BPA. This observation was seen across duplicate biological repeats and is inconsistent with the lack of killing observed in *ΔtolA::kan* cells (Figure 5.4.B). Only one colicin E9 molecule is required to enter the cell to induce killing, therefore if there is occasional read-through of the TAG stop codon in the absence of BPA, low level expression of TolA could account for killing in the absence of BPA. No killing was seen for pTolASer32^{BPA}pV with or without BPA, demonstrating that Tol-Pal is not functional in this mutant. Due to this no further investigation was carried out using the TolASer32^{BPA} mutation site.

For the TolAIle15^{BPA} and TolAHis22^{BPA} mutants, western blotting was used to confirm that full-length TolA was being expressed in the TolA^{BPA} mutants. Lysates of cells expressing a TolA^{BPA} mutant and the tRNA synthetase pair were grown in the presence or absence of BPA and blotted using anti-TolA primary antibody. Full length TolA was expressed in TolAIle15^{BPA} and TolAHis22^{BPA} only when BPA present (Figure 5.5). Some leaky expression was observed for TolAHis22^{BPA} when grown with BPA; however, this is consistent with the leaky expression in the plasmid encoded TolA control.

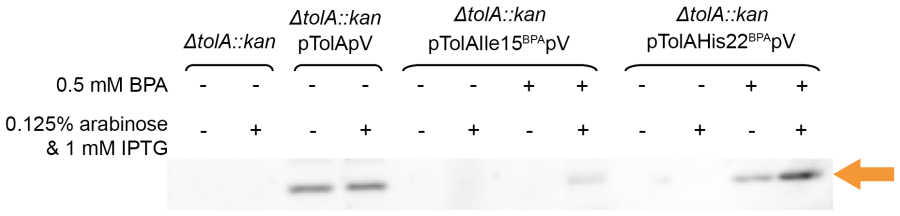


Figure 5.5 Full length TolA is expressed in the TolAHis22^{BPA} and TolAIle15^{BPA} mutants in the presence of BPA.
Cell lysate of cells grown in the presence of absence of BPA, induced with 0.125% arabinose and 1 mM IPTG, were separated by SDS-PAGE on a 12% Bis-Acrylamide gel and immunoblotted using anti-TolA primary antibody. For the controls of *ΔtolA::kan* and *ΔtolA::kan* + pTolA cultures were grown without BPA and samples taken before and after induction at 37°C for 3 hours. For the TolA^{BPA} mutants with pVol, cultures were grown with or without BPA, and samples taken before and after induction at 37°C for 3 hours. The pTolA control shows expression of full-length TolA even before induction, implying there is leaky expression. The TolAIle152^{BPA} mutant expresses full-length TolA at a low level only when BPA present and expression is induced. TolAHis22^{BPA} mutant expresses full-length TolA only in the presence of BPA and expresses best when induced.

Assessment of cell size and morphology through brightfield microscopy was used to confirm the phenotype of pTolAHis22^{BPA}pV cells was comparable to pTolA with pVol. Both prior to and following induction of protein expression (in the presence of BPA) pTolApV and pTolAHis22^{BPA}pV were of similar cell lengths (Figure 5.6). Furthermore, the attributes of the *tol* phenotype such as cell chaining were not observed (Figure 5.6). Therefore, expression of TolAH22^{BPA} complemented the *tol* phenotype.

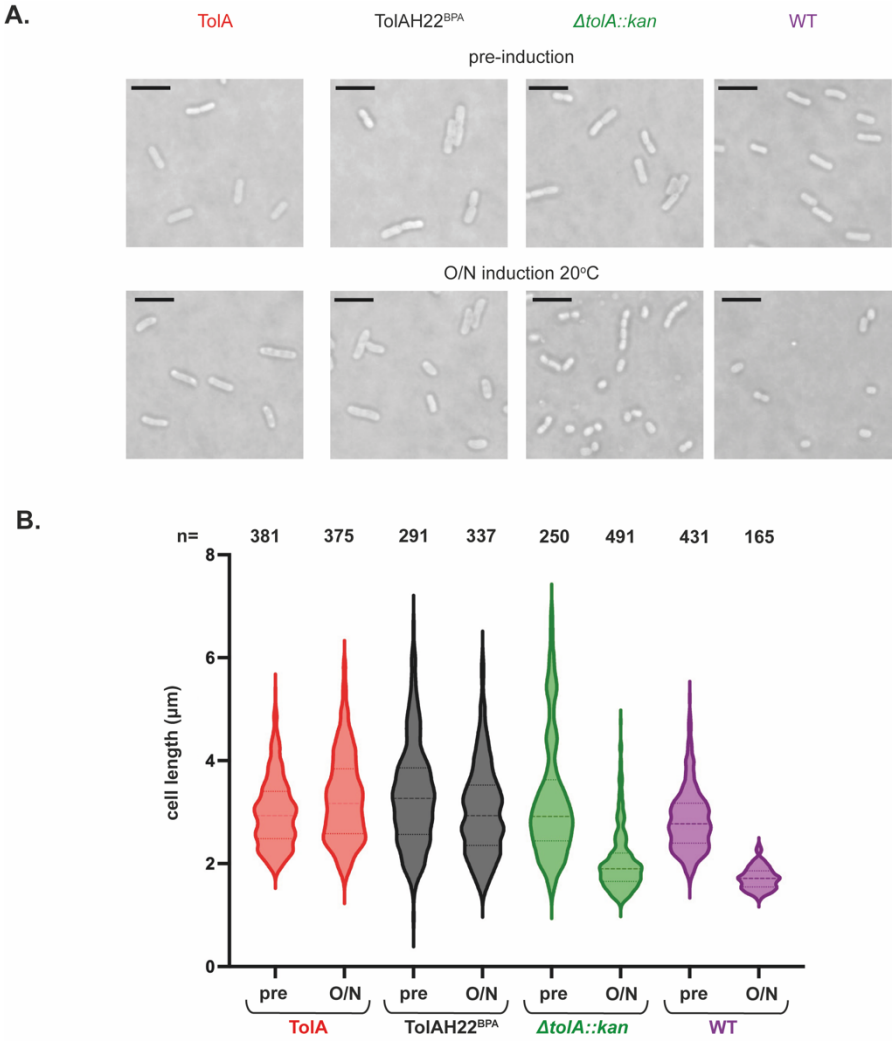


Figure 5.6 *ΔtolA::kan* expressing TolAHis22^{BPA} has the same colony morphology as *ΔtolA::kan* expressing TolA and the synthetase pair on pVol.
A. Cells grown with 0.5 mM BPA to OD_{600} 0.6 and induced with 0.125% arabinose and 1 mM IPTG overnight (O/N) at 20°C were imaged by brightfield microscopy. In *ΔtolA::kan* the cells are small and chained after overnight growth, consistent with *tol.* Comparison of representative images of pTolApV and pTolAHis22^{BPA}pV show cells of similar size pre-induction and after O/N induction. In both, the cells are larger than in *ΔtolA::kan* and are not chained. After overnight growth WT cells reach lag phase and are smaller than pre-induction. Scale bar = 5 μm. **B.** Comparison of cell sizes of *ΔtolA::kan*, pTolA with pVol, and pTolAHis22^{BPA}pV before and after O/N induction. The lengths of cells expressing TolAHis22^{BPA} are the same as the cells expressing TolA, and both have larger cells than *ΔtolA::kan*. Again, small cells are observed in WT after O/N growth, consistent with lag phase cells. Violin plot includes data from three biological repeats.

Overall, the BPA site at His22 displayed the most promise for further study with expression capable of recovering some growth on 2% SDS, inducing colicin susceptibility, and cell morphology was consistent with plasmid expression of TolA in *ΔtolA::kan*. This evidence suggested that the Tol-Pal system is partially functional under the conditions tested, therefore any crosslinks observed would be consistent with the interactions of native TolA. This site was therefore chosen for further work in optimising the crosslinking process.

5.2.2 No crosslinks between TolQ and TolAHis22^{BPA} were detected

To begin to look for the interaction partners of TolA at the His22 site, first the expression of TolAHis22^{BPA} in *ΔtolA::kan* was scaled up to 800 ml with induction overnight at 20°C with 0.125% arabinose and 1 mM IPTG. Two expression conditions of TolAHis22^{BPA}, grown with or without BPA, were compared with the plasmid encoded wild-type TolA grown with BPA and exposed to UV at 365 nm for 90 mins after harvesting the cells. Cell lysates from before and after induction for each condition were visualised by Coomassie stained SDS-PAGE. No band at 44 kDa for TolA/TolAHis22^{BPA} could be observed after induction by Coomassie stained SDS-PAGE (Figure 5.7.A). This was likely down to the expression level being too low to be observed.

Next, nickel affinity pull-downs were used to concentrate the TolA and TolAHis22^{BPA} for analysis by SDS-PAGE, allowing for confirmation of protein expression and comparison of expression levels between TolA and TolAHis22^{BPA} before looking for crosslinked adducts. TolA/TolAHis22^{BPA} were extracted and solubilised in the detergent β-OG, then applied to a nickel affinity gravity column. Samples were taken before, during and after nickel affinity for analysis by SDS-PAGE. Coomassie staining was again unable to detect the presence of TolA/TolAH22^{BPA}. However, western blotting with anti-TolA antibody revealed the presence of TolA/TolAH22^{BPA} in the solubilised inner membrane fraction (Figure 5.7.B).

Using Coomassie staining to observe the proteins expressed is necessary for the LC-MS/MS protocol. Therefore, the expression of TolAH22^{BPA} was again scaled up (0.8 L to 4.8 L) to achieve a higher yield of TolAH22^{BPA}. As before *ΔtolA::kan* cells expressing TolA or

TolAH22^{BPA} were grown in the presence of BPA and exposed to UV at 365 nm for 90 mins. A control expressing TolAHis22^{BPA} without UV exposure was also generated. Extraction and solubilisation of proteins in 2% β-OG was followed by nickel affinity pull-downs and SEC to separate proteins present by size and compare the elution profiles with the aim of identifying a peak corresponding to a crosslinked species. The SEC elution profiles of TolAH22^{BPA}, UV-TolAH22^{BPA} and UV-TolA were all consistent with one another (Figure 5.7.C). A western blot with anti-TolA of the SEC peaks again demonstrated the presence of TolA/TolAH22^{BPA}, however all bands found in UV-TolAH22^{BPA} were also present in in the non-UV TolAH22^{BPA} and UV-TolA controls (Figure 5.7.D). Therefore, no species consistent with a TolAH22^{BPA} crosslink could be observed.

The process of protein expression, UV exposure and purification was repeated for all three conditions, this time using Cymal-6 as the detergent. The rationale for this being that it was in Cymal-6 that the TolQ-TolR complex could be purified, and therefore using this detergent could allow for the pulldown of TolA crosslinked to TolQ. The SEC elution profiles of the UV-TolA, TolAH22^{BPA} and UV-TolAH22^{BPA} solubilised in 0.5% Cymal-6 were again consistent (Figure 5.7.E). The UV-TolA, TolAH22^{BPA} and UV-TolAH22^{BPA} samples were then concentrated to a low volume to allow for visualisation by Coomassie staining. When a longer exposure time was used during western blotting with anti-TolA, a weak band at a higher molecular weight than TolA could be observed in only UV-TolAH22^{BPA} (Figure 5.7.F). Sections of the corresponding Coomassie stained SDS-PAGE gel in TolAH22^{BPA} and UV-TolAH22^{BPA} were sent for peptide mass fingerprinting to determine if a species crosslinked to TolA was present. Mass spectrometry confirmed the presence of TolA at 44 kDa, however TolA was not one of the most abundant species present in any band of a higher molecular weight. Altogether, the data generated suggested under the current experimental conditions, no crosslinking between TolA and TolQ could be observed at the His22 site.

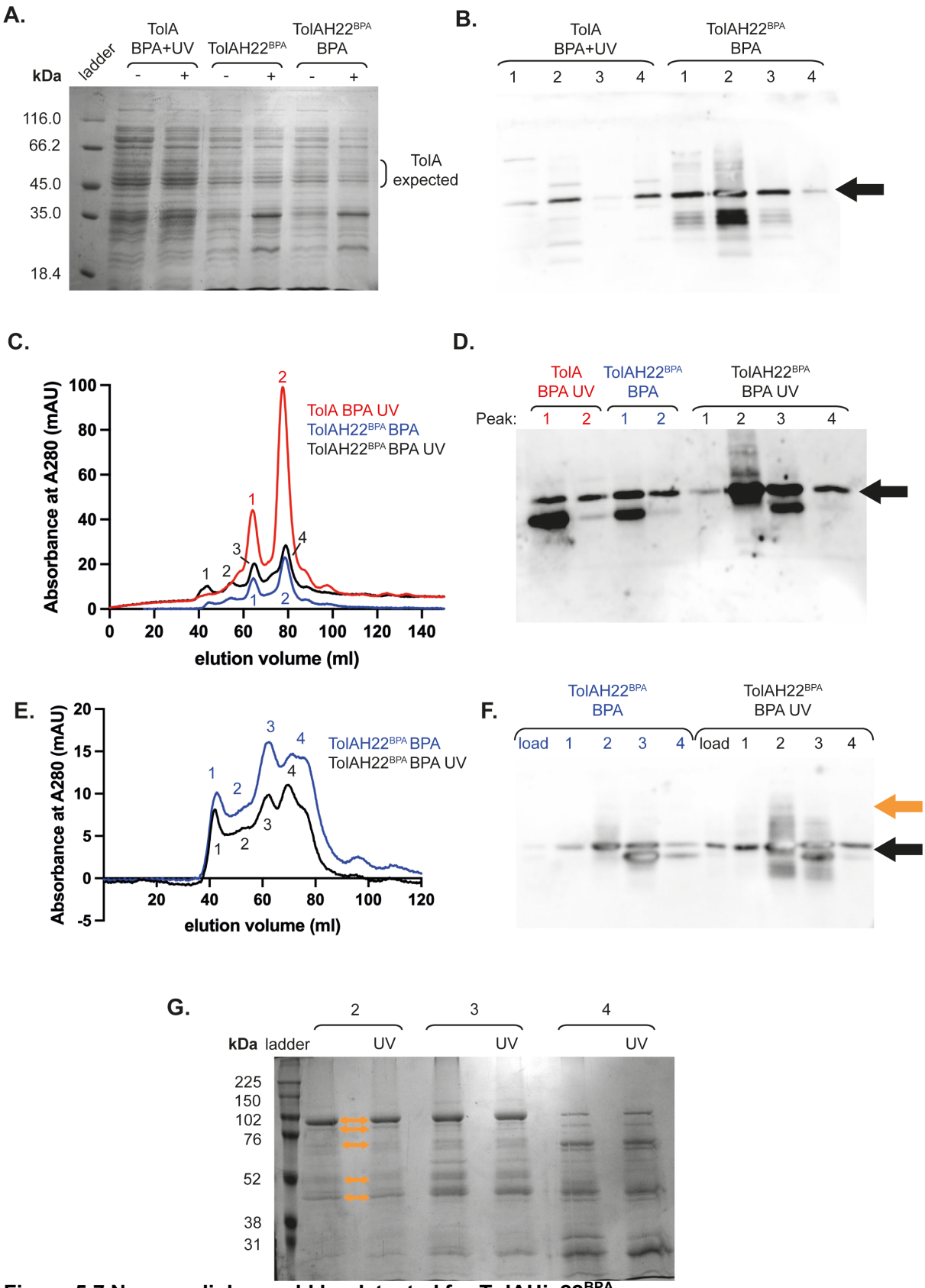


Figure 5.7 No crosslinks could be detected for TolAHis22^{BPA}.
A. Coomassie stained 12% SDS-PAGE gels were unable to identify expression of TolA or TolAHis22^{BPA} when comparing before (-) and after (+) induction overnight at 20°C with 0.125% arabinose and 1 mM IPTG. The region where the expected band for the 45 kDa TolA/TolAHis22^{BPA} is labelled and includes above 45 kDa as TolA constructs typically migrate slowly.
Figure legend continued p130.

Figure 5.7 continued. B. Western blot using anti-TolA with TolA and TolAHis22^{BPA} run on 12% SDS-PAGE gels. 1. Cell lysate soluble fraction. 2. Solubilised inner membrane fraction. 3. Nickel affinity resin sample after application of cell lysate soluble fraction and washing with lysis buffer. 4. Nickel affinity resin sample after application of solubilised inner membrane fraction and washing with lysis buffer. Bands for TolA could be observed in the solubilised inner membrane fraction and resin for TolA, and in both the soluble and solubilised inner membrane fractions for TolAHis22^{BPA}. The arrow indicated the 45 kDa band for TolA/TolAHis22^{BPA}. **C.** SEC profiles of TolA-UV (red), TolAHis22^{BPA} (blue) and TolAHis22^{BPA}-UV (black) eluted in 20 mM Tris/HCl pH 7.5, 100 mM NaCl, 1% β-OG on HiLoad 16/600 Superdex 200 column. **D.** Western blot using anti-TolA of 12% SDS-PAGE gel of peaks corresponding to SEC elution profile in C. TolA/TolAHis22^{BPA} was present in all peaks at 45 kDa, indicated by the black arrow. E. SEC profiles of TolAHis22^{BPA} (blue) and TolAHis22^{BPA}-UV (black) eluted in 20 mM Tris/HCl pH 7.5, 100 mM NaCl, 0.1% Cymal-6 on HiLoad 16/600 Superdex 200 column. **F.** Western blot using anti-TolA of 12% SDS-PAGE gel of peaks corresponding to SEC elution profile in E. TolA/TolAHis22^{BPA} was present in all peaks at 45 kDa, indicated by the black arrow. Faint bands at a higher molecular weight could be observed only in the UV exposed TolAHis22^{BPA} in peak 2, shown by orange arrow. **G.** Coomassie stained 10% SDS-PAGE gel of the concentrated peaks following SEC. The bands highlighted by the orange arrows were excised for both the UV sample and no UV control and sent for peptide mass fingerprinting analysis.

5.2.3 Modelling the TolQ-TolR-TolA assembly shows TolA TMH docks into groove composed of TolQ TMHs

As the BPA crosslinking approach was in progress, the models of TolA and TolQ-TolR described in Chapters 3 and 4 were used to investigate the interaction between TolQ-TolR and TolA through computational protein-protein docking. Residues 1-100 of the *E. coli* AlphaFold2 model were placed within 10 Å of the TolQ-TolR complex. Rosetta global and local refinement was used for docking to generate the TolQ-TolR-TolA$_{1\text{-}100}$ model, and the best scoring model was chosen for analysis (Webby *et al.*, 2022b) (Figure 5.8.A).

The TMH of TolA was found to dock within a groove created by the TMHs of two copies of TolQ, on the exterior of the pore structure (Figure 5.8.A). Assessment of the hydrophobicity showed the position of TolA's TMH aligns with the hydrophobic TMHs of TolQ, creating a belt of hydrophobicity consistent with insertion in IM (Figure 5.8.B). In addition, calculation of the electrostatic surfaces shows the extreme N-termini of TolA is electropositive, complementing the electronegative band on TolQ's exterior within the cytoplasmic domain (Figure 5.8.C). The SHLS motif of TolA-I sits on the face of the helix in contact with TolQ (Figure 5.8.D). Residues identified through suppressor mutagenesis in TolQ TMH 1 are found adjacent to the SHLS motif, consistent with previous work (Germon *et al.*, 1998) (Figure 5.8.D).

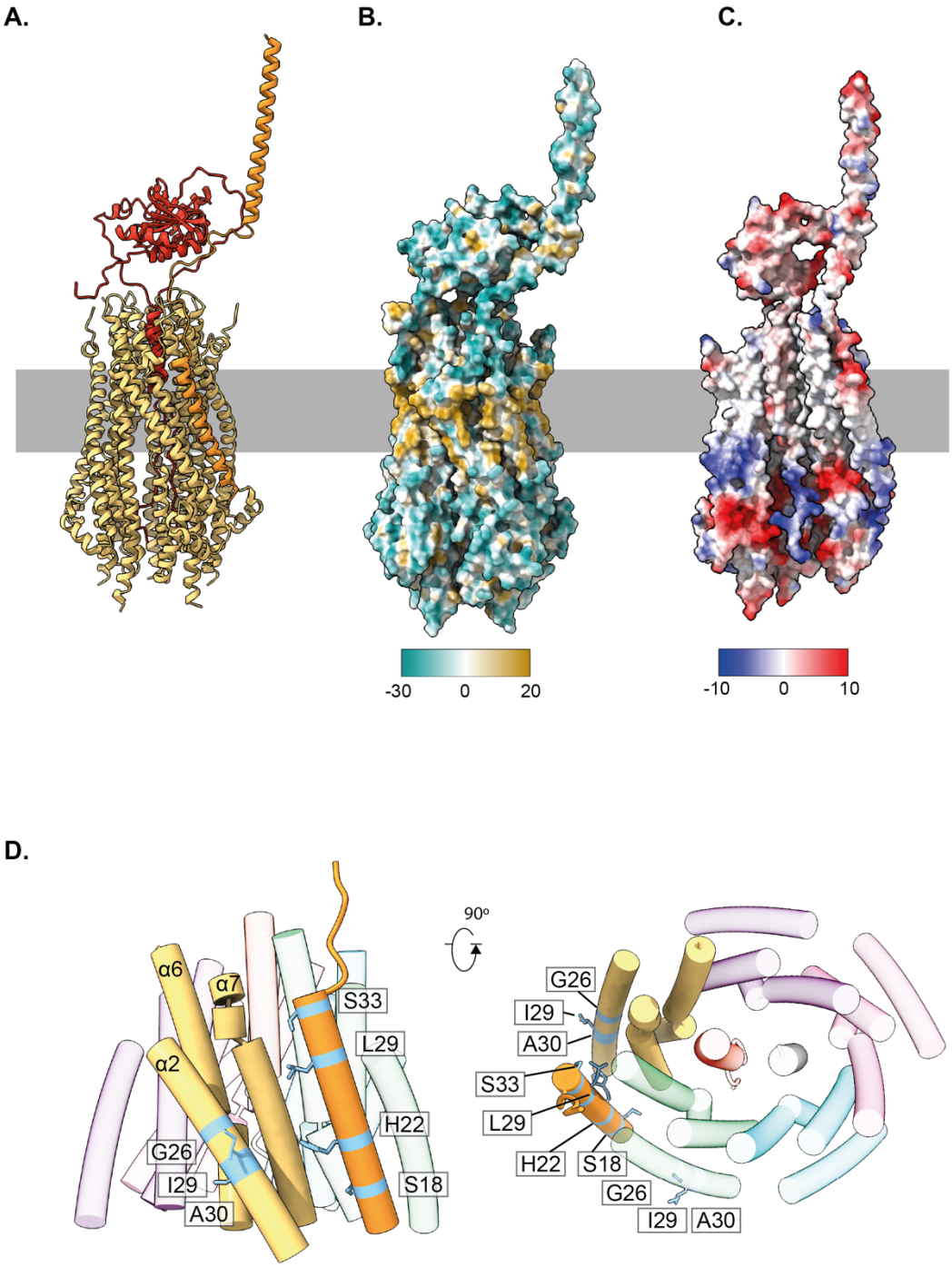


Figure 5.8 Docking TolQ-TolR and TolA models shows interaction of TolA TMH within TolQ's TMH groove

A. Docking of TolA (orange) to the TolQ-TolR (yellow/red) model shows TolA's TMH fits into a groove created by TMH1 of two adjacent copies of TolQ. **B.** Mapping the hydrophobic surfaces shows the TMH of TolA is placed in line with the TMHs of TolQ creating a consistent hydrophobic band (gold) in line with the IM. **C.** Calculation of the electrostatic surfaces of the TolQ-TolR-TolA model, with two chains removed to allow observation of the central pore. TolR's N-terminal residues are electropositive, complementing the electronegative TolQ pore in the cytoplasm. The N-terminal end of TolA prior to the TMH is also electropositive, complementing the negative band of TolQ at the IM-cytoplasm interface. **D.** The TolA (orange) SHLS motif (blue) is found adjacent to one copy of TolQ (yellow) residues (blue) identified by suppressor mutagenesis. Side (left) and top down (right) view shown. Panels A-C adapted from (Webby *et al.*, 2022b).

5.2.4 Discussion

The BPA crosslinking approach described in this work to investigate the interaction between TolA, and TolQ-TolR in the IM was unable to identify any crosslinked species. Assessment of the function of the TolAHis22^{BPA} mutant found it to be a functional mutant, contrary to expectation based upon previous mutational analysis (Germon *et al.*, 1998; Rassam *et al.*, 2018). Further analysis of the interaction using the computational models of TolQ-TolR and TolA described previously identified that TolA-I binds within a groove between TolQ's TMHs.

Interaction between TolQ and TolA within the TMHs has been previously investigated through formaldehyde crosslinking and mutagenesis approaches (Derouiche *et al.*, 1995; Germon *et al.*, 1998). This work set out to better understand the molecular detail of this interaction using a BPA crosslinking approach. The choice of His22^{BPA} as one of the mutation sites was first intended as a control as it has previously been shown that mutation of His22 to Ala, Pro or Arg infers a *tol* phenotype (Germon *et al.*, 1998; Vankemmelbeke *et al.*, 2009). It was therefore initially a surprise that expression of His22^{BPA} was able to complement the *ΔtolA* phenotype. However, the substitution of His22 with Tyr by Germon *et al*. did not confer a phenotype as severe as the Arg, Ala or Pro substitutions, inferring that some mutations of His22 can be tolerated (Germon *et al.*, 1998). Further to this, overexpression of TonB-His20Tyr recovered some sensitivity to bacteriophage Φ80 (Larsen, Thomas and Postle, 1999). BPA is a larger molecule than His or Tyr, with its two aromatic rings extending ~10 Å from the peptide backbone compared with ~4.4 Å and ~6.4 Å for His and Tyr, respectively. Inspection of the TolQ-TolR-TolA model around TolA His22 reveals the His22 fits within a pocket of residues created by TolQ TMH 1 and TMH2 (Figure 5.9). The pocket is large enough to allow for insertion of the ~10 Å long BPA in the place of His22, and placement of BPA above TolQ Trp32 would be close enough to allow for a π-π stacking interaction. Replacement of His22 with BPA does not, therefore, introduce steric clashes that may prevent complex formation. In addition, the role of the functional group of His22 as a positively charged side

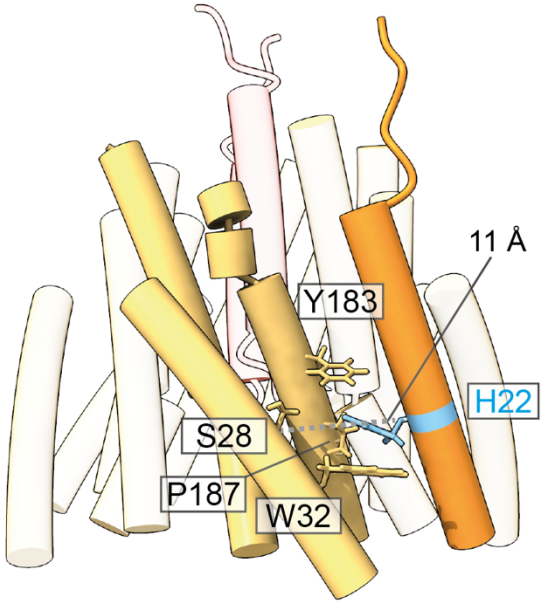


Figure 5.9 TolA His22 fits into a pocket formed by TolQ TMH1 and TMH2
The pocket TolA (orange) His22 (blue) fits into is 11 Å (grey dashed line), and therefore there is space to accommodate the His22^{BPA} mutation without a steric clash. In addition, His22 lies almost parallel to Trp32 of TolQ (yellow), in close enough proximity that π-π stacking could stabilise the interaction, both for His22 and His22^{BPA}.

chain in the conveyance of the PMF to TolA is unclear, especially as replacement with Arg yielded a *tol* phenotype. Why the aromatic, ketone containing BPA was functional is difficult to understand without more extensive mutational studies. Replacement of His22 with a negatively charged residue of similar size, Glu, or with alternative aromatic residues could give some insight into the attributes of His22 that make it important for Tol-Pal function.

Once the functionality of TolAHis22^{BPA} was established, the project moved to identify species crosslinked to TolA. Unfortunately, no such crosslinks could be identified, however there are several reasons that could explain the difficulties encountered. Firstly, the efficiency of BPA crosslinking is variable (Williams, Ackerman and Coleman, 1986; Chin *et al.*, 2002), the environment surrounding the BPA impacts the crosslinking efficiency (Lancia *et al.*, 2014) and examples of BPA crosslinking of TMH proteins are rare (Rubino *et al.*, 2020). Therefore, it is unknown of the TolAH22^{BPA} present in a cell, how many will successfully capture molecules within the vicinity of TolA. Secondly, at present how many TolA molecules are interacting with TolQ at any given time in an elongating cell has not been characterised.

Combined, these factors meant that the resulting crosslinked adduct would be a very low percentage of the total amount of TolA in the sample, and without good detection of un-crosslinked TolA, it suggests that searching for crosslinked species will remain challenging.

The expression system described here walks a fine line between under and overexpression of TolA/TolAHis22^{BPA}. Overexpression of TolA in *ΔtolA::kan* has been shown to result in a phenotype consistent with *ΔtolA* (Connolley *et al.*, 2021). To observe native interactions and ensure the TolA^{BPA} mutants were functional, it was important not to overexpress to the point of impairing system function. However, the use of an LC MS/MS approach requires the excision of bands of crosslinked species visualised by Coomassie stained SDS-PAGE. Under-expression of TolAHis22^{BPA} risks being unable to visualise TolA-crosslinked interactions. Further work on the system must continue to walk this fine line, using a systematic approach to increase expression whilst maintaining the complemented functional *tol* phenotype should allow for better visualisation of TolAHis22^{BPA}. One future option could be to utilise the colicin trapping system described in Rassam *et al.* 2018 to stall the Tol-Pal IM machinery whilst assembled (Rassam *et al.*, 2018). Any interaction observed would still be representative of native Tol-Pal interactions but would push the equilibrium of bound to unbound TolA:TolQ-TolR-TolA in the direction of the interacting complex.

The computationally generated model of TolQ-TolR-TolA interaction described in this chapter fits with previously published data as it shows the face of TolA-I, containing the SHLS motif, interacting with TolQ. Particularly, His22 is found in proximity to Glu26, Ile29 and Ala30, residues identified to suppress the effects of His22Pro and Ser18Leu (Germon *et al.*, 1998). Further substitution mutation analysis could reveal greater detail in the interaction between the helices of TolQ-TolA, especially if the crosslinking approach continues to prove challenging. In addition, a more directed cysteine crosslinking approach or non-specific LC MS/MS crosslinking approach could allow for investigation into how important the dynamicism of the TolQ-TolA interaction is.

The project described in this chapter displays great promise in shedding light on the details of the protein-protein interactions between TolA and TolQ-TolR. However, any further work will require considerable finessing of the expression conditions; balancing the expression levels to allow for the observation of native interactions that can be identified through the desired methods.

6 General Discussion

Identifying of the role of Tol-Pal in maintaining OM integrity provided an answer to the question of how a non-energised membrane invaginates at the septum during cell division. However, questions remained over the mechanism by which Tol-Pal achieves this function. Structural information on the IM components of the system remained elusive, which made understanding the mechanism of action and the role of the PMF challenging. The work described here aimed to further our knowledge of the IM proteins of the Tol-Pal system using a combination of biochemical, biophysical, structural, and computational approaches. The new data generated is here put into context within the system and a new 'helical hinge' mechanism proposed.

6.1 Updating the TolQ-TolR-TolA interaction mechanism

6.1.1 The TolQ-TolR rotary mechanism

For both MotA-MotB and ExbB-ExbD rotary mechanisms have been suggested, whereby MotA/ExbB rotates around the MotB/ExbD dimer that is anchored by binding the PG (Deme *et al.*, 2020; Santiveri *et al.*, 2020). By nature of the asymmetry introduced through the 5:2 complex stoichiometry, the key aspartate residue in each MotB/ExbD unit resides in a different environment. In the $F_O/V_O/A_O$ ATPases it is the neutralisation of a key residue that allows the C-ring to rotate through the hydrophobic lipid membrane (Kühlbrandt and Davies, 2016). Similarly, for the MotA-MotB/ExbB-ExbD stators it is in the state in which both Asp residues are neutralised where rotation of MotA/ExbB 36° around MotB/ExbD can occur.

The model of TolQ-TolR generated in this study allows for the extrapolation of the mechanism described for MotA-MotB to Tol-Pal (Figure 6.1). In sequence, the unplugging of the TolR dimer opens the proton channel through movement of the TolQ Leu175 residue. Asp23 in one TolR monomer is protonated and Asp23 in the neighbouring monomer is already protonated from the previous cycle. Now both Asp23 residues are neutral, allowing for the rotation of TolQ, driving a 'power stroke'. Through this movement Asp23 of monomer 2 becomes accessible to the cytoplasm and releases its proton. It then returns to its original

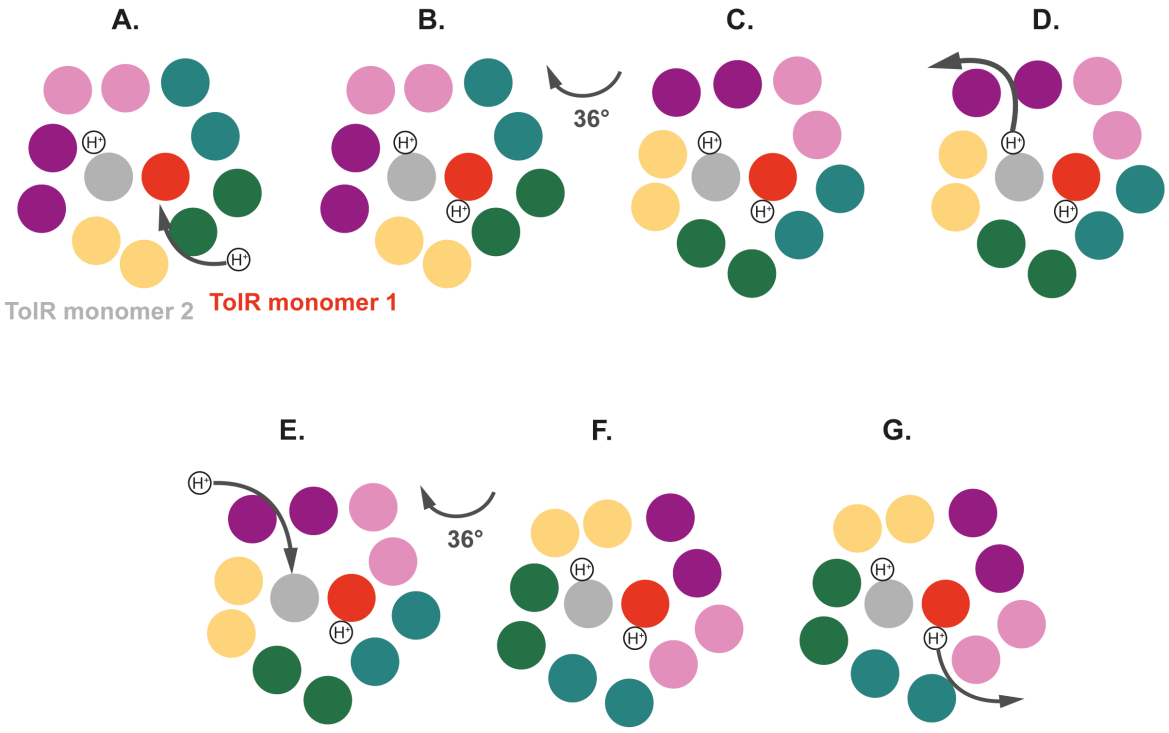


Figure 6.1 A model for TolQ-TolR proton transduction
A. View of TolQ-TolR from the periplasm. TolR chain 1 accepts a proton from the periplasm. The coloured circles represent TMH 2 and TMH 3 of each copy of TolQ in the pentamer. **B.** Both Asp23 residues of TolR monomers 1 and 2 are protonated, allowing for the rotation of TolQ 36° around TolR clockwise. **C.** The complex following clockwise rotation of TolQ's. **D.** Asp23 of the monomer 2 TolR is now accessible to the cytoplasm and releases its proton. **E.** TolR monomer 2 Asp23 can receive a proton from the periplasm. **F.** Now both Asp23 residues are neutralised, the TolQ pentamer rotates 36°. **G.** The proton on monomer 1 TolR Asp22 is now released, and the cycle can begin again. Based upon model presented in Santiveri et al 2020 Figure 7 (Santiveri *et al.*, 2020).

conformation to accept a new proton. The neutralisation of Asp23 causes TolQ to rotate 36° and now Asp23 from monomer 1 can release its proton, and so on (Figure 6.1).

The observation of a pentamer of TolQ in preliminary EM data supports the assumption that TolQ-TolR forms a complex with 5:2 stoichiometry like that of MotA-MotB and ExbB-ExbD. Major questions however remain when comparing the three systems and elucidating the mechanism of each. Santiveri *et al.* identified the role of Phe186 in acting as a gate to the proton channel through solving structures of MotA-MotB both with and without the plug helices of MotB. Observing that when the plug helices were removed, Phe186 moved to allow access to the proton channel (Santiveri *et al.*, 2020). Neither TolR nor ExbD have equivalent helices, suggesting the mechanism by which they are plugged must differ from MotA-MotB.

When building the computational model of TolQ-TolR, placement of the 'closed' strand-swapped dimer of TolR into the complex with the TPH residues as predicted proved sterically impossible. The first residue of the strand-swapped dimer Pro37 (Wojdyla *et al.*, 2015) is in the TolQ-TolR model the final residue of the TPH (Figure 6.2). The duality of this residue suggests the oscillation of TolR's periplasmic domain between its PG binding and non-PG binding states could be coupled to the piston-like movement of the TPHs in and out of the TolQ pore. This movement would allow for the formation of the strand-swapped dimer and act as the 'unplugging' event responsible for opening the proton channel through TolQ-TolR. What is still unclear is the role of the PMF in this proposed mechanism for TolR. Is the oscillation between states a constant occurrence or is it as a result of binding to TolA? In addition, does the formation of either state require energy? Answering these questions will take further experimentation across the three systems.

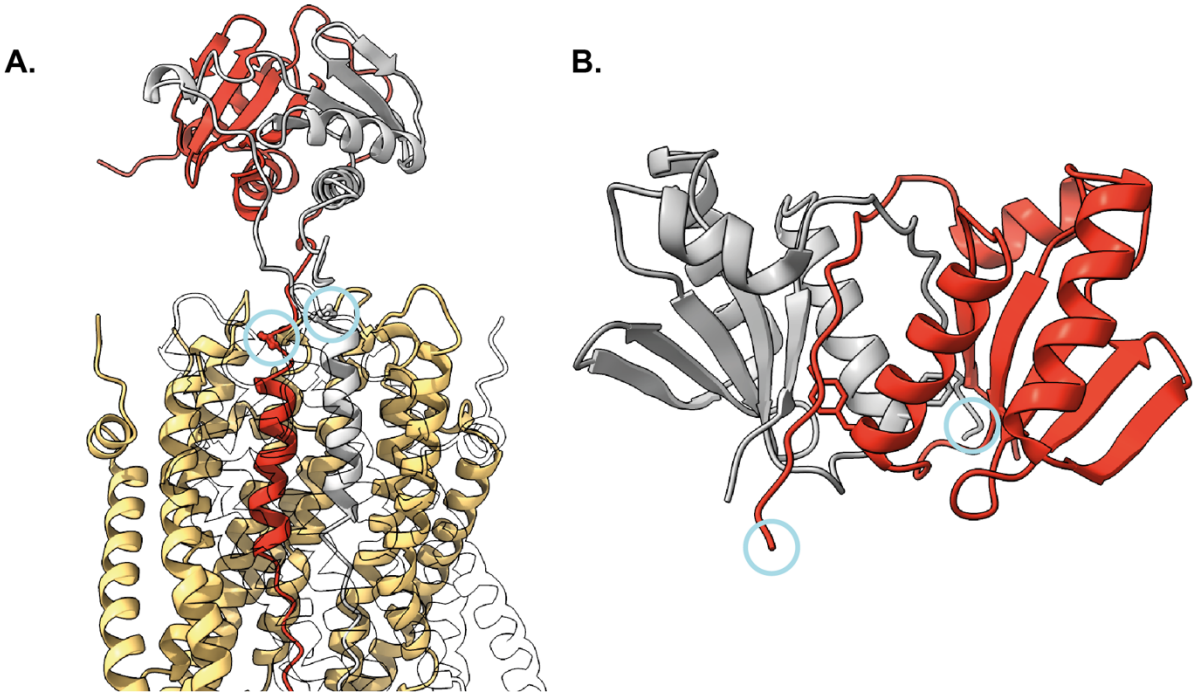


Figure 6.2 Proline 37 of *E. coli* TolR's position in different structural models implies piston action of TolR TPHs.
A. The Pro37 reside is circled in blue and shown in ball and stick form in the TolQ-TolR model generated in Chapter 4. This residue is found at the end of the TPHs. Two copies of TolQ are transparent to allow Pro37 residues to be viewed. **B.** The Pro37 residue is circled in blue in the *E. coli* strand-swapped dimer of TolR (PBD: 2JWK). Pro37 is found at the N-termini following the strand-swapped β-strand. For this structure to form, the TPHs must move 10 Å out of the pore, the implication being the TPHs undergo a piston-like movement within the pore of TolQ.

6.1.2 The TolA length paradox

The N-terminal GFP tagging of TolA gave rise to the observation that TolA can freely diffuse within the IM (Rassam *et al.*, 2018). However, TolA also has the capacity to reach the OM to bind to TolB-Pal or TolB in complex with a bacteriocin (Rassam *et al.*, 2018; Szczepaniak *et al.*, 2020). TolA cannot perform both functions simultaneously as when extended to the OM it becomes trapped within the holes within the PG. The mechanistic hypothesis that followed from this evidence was that TolA must exist in two forms: a low energy state folded in such a way that it can freely diffuse within the IM, and a high energy state that extends to the OM in response to the PMF.

In the central domain of the *E. coli* AlphaFold2 model the longest α-helix, α2, is the length of the entire periplasm, and the α1 helix spans the distance between the OM and PG. This could be taken as evidence to refute the hypothesis above. However, molecular dynamics simulations of the *E. coli* AlphaFold2 model inserted into the IM demonstrated that TolA can 'fold down' to interact with the IM. This reduces its extension into the periplasm from 300 Å to 130 Å, a distance that will keep it below the PG layer (Figure 3.8.A). It can therefore by suggested that when non-energised, TolA adopts this state where it stays close to the IM through interactions between domain II and the IM polar head groups. It is therefore under the PG layer, allowing for free diffusion in the IM.

The molecular dynamic simulations of TolA threaded through the PG layer provide a glimpse into what could be driving the pulling of TolB from Pal. The zippering of the two helices of TolA-II to form stable a hydrogen bonding network pulls the TolA-III domain along the PG surface (Figure 3.8.B). This demonstrates that if the two helices are separated, returning to the more energetically favourable state of interacting with one another could be what provides the force required to pull TolB from Pal. It therefore follows, that what the PMF is required to do is to separate these two helices into a less energetically favourable extended conformation. One limitation, however, of the experiment conducted in the presence of PG was the use of highly crosslinked PG that TolA could not pass through. Further experiments using PG with

holes TolA domains II and III could fit through would allow for investigation into whether TolA can relax from its energetically unfavourable extended conformation to a resting state via a hole in the PG.

6.1.3 The 'helical hinge' mechanism underpinning the role of TolQ-TolR-TolA in the Tol-Pal system

Here the new evidence gathered in this study together with evidence from previous work is amalgamated into an updated 'helical hinge' mechanism. The focus is on the interactions within the IM components. Therefore, this forms a more detailed look at steps within the 'mobilisation and capture' mechanism that explains Tol-Pal's role in accumulating Pal at the division site.

To start, TolQ-TolR and TolA are freely diffusing within the IM. At the OM, TolB-Pal is freely diffusing within the plane of the OM, or Pal is static and anchored to the PG (Figure 6.3.A). Interaction between the TMH of TolQ and TolA triggers a conformational change in the periplasmic domain of TolR (Figure 6.3.B-C). The TPHs move down, and the periplasmic dimer switches to the 'open' conformation and binds PG to anchor the stator (Figure 6.3.C). The proton channel through TolQ-TolR is now open, and a proton is transduced through the stator (Figure 6.3.D). The protonation of TolR Asp23 triggers the rotation of TolQ by 36°. This power stroke provides the energy required for TolA to change to its high energy state, hinging both helices up to bind TolB-Pal at the OM. The second consequence of TolQ's rotational movement is to break the contact between TolQ and TolA. The loss of this interaction triggers the return of TolR to its ground state where the proton channel is plugged (Figure 6.3.E-F). TolA will also, therefore, begin to relax back down to its ground state, pulling TolB from Pal in the process and releasing it beneath the PG layer (Figure 6.3.F-G). This then frees Pal to bind to PG and spatially separates TolB from Pal through the PG layer.

The description of the helices as 'hinging' is somewhat of an oversimplification. The TolA conformation observed as the endpoint of the CGMD simulations in the IM has a length of 230 Å. If this remained entirely rigid as it hinges 90° to an upright position, the hole in the

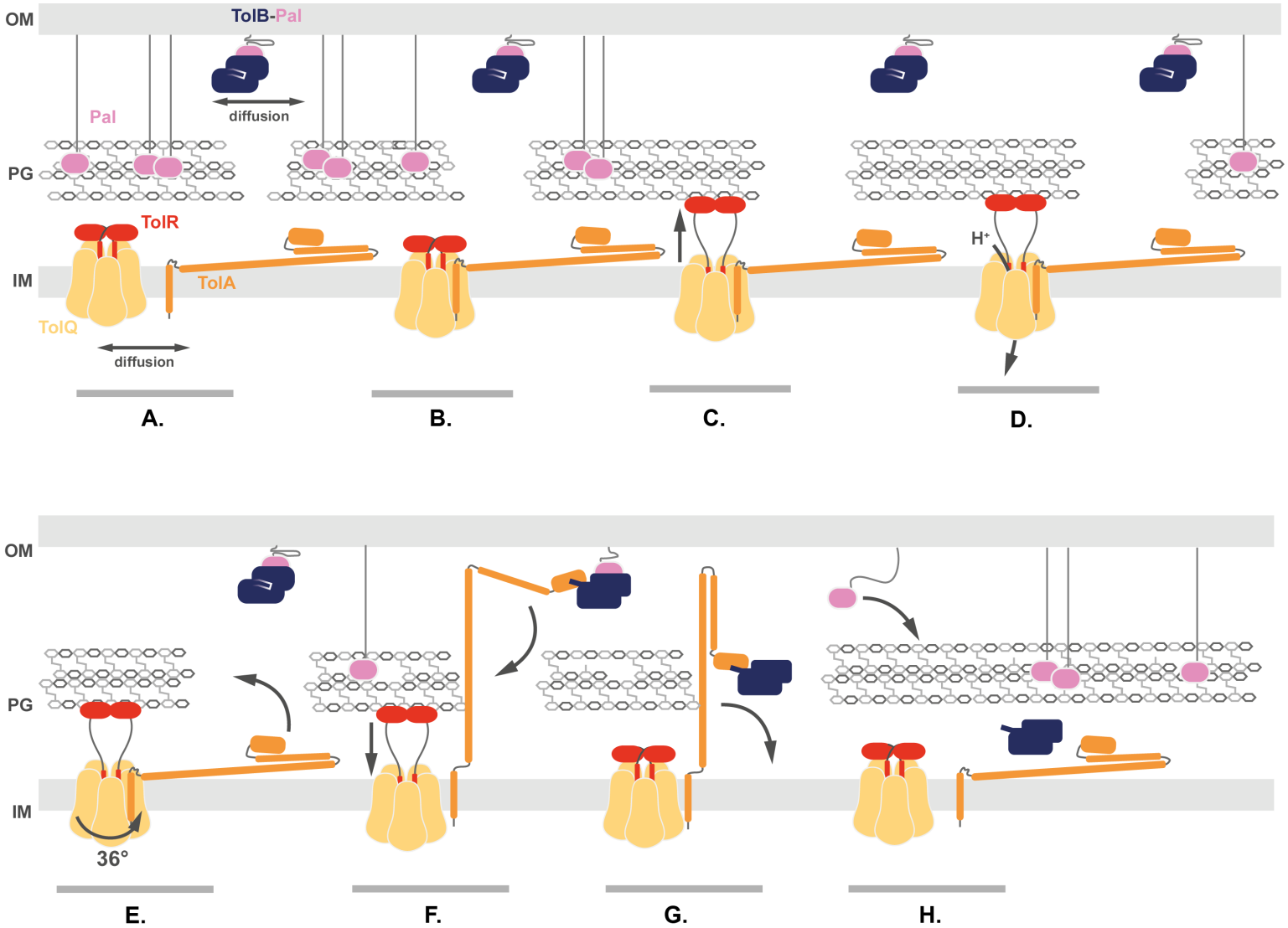


Figure 6.3 The 'helical hinge' mechanism.
Figure legend continued p 143.

Figure 6.3 continued A. TolQ-TolR and TolA diffuse within the IM, TolB-Pal diffuses in the plane of the OM, or Pal is bound to the PG. **B.** TolQ-TolR and TolA interact through their TMHs. **C.** The interaction triggers the movement of TolR's TPHs into the pore, opening the proton channel, and rearrangement of the TolR dimer to allow for PG binding. **D.** A proton is transduced through the open TolQ-TolR stator. **E.** TolQ rotates 36° around TolQ, providing the power stroke that causes TolA to hinge up to reach TolB-Pal at the OM. **F.** The rotation of TolQ has now broken the contact between TolQ and TolA. TolR returns to its resting state, with the two TPHs moving up to block the channel and the periplasmic domain forming the closed strand-swapped dimer. TolA begins to relax, driven by the re-association of the α helices in domain II. This relaxation event is what pulls TolB from its interaction with Pal. **G.** TolA continues to relax back down to ground state, pulling TolB under the PG layer. **H.** TolB is released and TolA returns to its interaction with the IM.

PG must be 113 Å in diameter to allow TolA's α2 to pass through it, assuming the PG layer to be 200 Å from the IM. Holes in the PG have been observed between 50 and 100 Å, so close to being large enough to accommodate this movement (Turner *et al.*, 2013). In reality, the helices of domain II responsible for this hinging motion are flexible and have the capacity to bend, as demonstrated by the CGMD simulations. It is more likely that a combination of hinging and 'unfurling', perhaps with guidance from interaction with the PG, is what feeds TolA through the PG mesh up to meet TolB-Pal at the OM.

When it comes to the mechanism of action in a dividing cell, it is plausible that the process differs from the elongating cell model. With TolQ-TolR present at a high concentration at the septum it is possible that another TolQ-TolR-TolA interaction could occur before the TolA has fully relaxed back to ground state. This would allow for a more rapid mechanism, depositing Pal and recycling TolB at a rate fast enough to keep up with the synthesis of PG and cell invagination. In addition, the role of TolA in lipid and PG biosynthesis suggests it might act as a scaffold. If always extended when at the septum this could explain its ability to interact with multiple systems at once. One further complexity that arises is the hyper-crosslinking of PG as induced by TolA binding of CpoB, in turn negating its effect on PBP1B, could prevent full retraction of TolA to the IM.

6.2 Future work

The work presented in this study furthers our mechanistic understanding of the Tol-Pal system, however many steps within the proposed hypothesis require validation. The most fundamental proposal to prove is the existence and importance of the helical hinge step.

Ideally a crystal structure of TolA-II followed by *in vivo* mutational investigations would answer whether the interaction between these two helices is important. Obtaining a crystal structure is, however, likely to continue to prove challenging. Therefore, using the AlphaFold2 structure as the basis for functional work provides an alternative avenue of research. For example, *in vivo* testing of the crosslinks detailed in Supplementary Figure 4 or domain II truncations based upon the AlphaFold2 secondary structure to remove interactions between α2 and α3 could validate the role of the helices in the above hypothesis.

In the helical hinge mechanism, in its resting state TolA interacts with the polar head groups of the IM to remain below the PG layer. This hypothesis was based upon the observation of the interaction in CGMD simulations. NMR experiments titrating mimetics of the polar head groups of TolA, or representative lipids, into solution with TolA-II could provide evidence for this interaction. In addition, in the CGMD simulations with PG present, TolA-II also interacted with the PG. The same NMR approach could show whether the extension of TolA up through the PG is guided by interaction with the cell wall.

Key to supporting the mechanism discussed above is showing that the relaxation process of TolA is separate to the PMF-induced extension step and is what drives TolB-Pal dissociation. The utilisation of engineered bacteriocins could provide a useful tool to investigate this. A disulfide-locked bacteriocin could be used to assemble the Tol-Pal machinery. Observation of killing after subsequent removal the PMF and reduction of the disulfide crosslink would provide evidence that it is the relaxation of TolA, not the PMF, that drives the pulling of the bacteriocin into the cell.

One of the major gaps in discussions on the Tol-Pal system is the lack of an experimental structure for TolQ-TolR. Solving this structure will allow for the confirmation of the complex stoichiometry and observation of the positions of the TPHs. Comparison of an experimentally determined structure to the model generated in this work will allow for validation of the mechanism discussed in 6.1.1. Little evidence exists for how the TolQ-TolR stator becomes plugged or unplugged. It is known that a TolR Tyr117Cys crosslink causes loss of system function (Goemaere *et al.*, 2007) and this crosslink is only feasible in the strand-

swapped dimer of TolR (Wojdyla *et al.*, 2015). Introducing and purifying the complex with this crosslink could therefore ensure the TPHs are in the 'plugged' conformation when solving a structure, even if the periplasmic domain is not resolvable. Comparison between a plugged and unplugged state of the TolQ-TolR stator would allow for further analysis of the TolQ Leu175 residue, to determine if it is responsible for access to the proton channel as suggested above, or if another residue plays this role.

The existence of an interaction between TolA and TolR is one aspect of the mechanism that has not well explored. Formaldehyde crosslinking identified a putative interaction between TolR and $TolA_{1-68}$ (Derouiche *et al.*, 1995). In the Ton system a BPA crosslinking approach mutating residues in the linker between ExbD's TPHs and periplasmic domain identified both PMF dependent and independent interactions between ExbD and TonB (Kopp and Postle, 2020). Furthermore, co-conservation analysis of TolR and TolA points to interaction sites between the two's periplasmic domains (unpublished data Daniel Williams-Jones). However, preliminary *in vitro* work through NMR and pull-down assays has been unable to validate the interaction (unpublished data Cara Press, Dr Melissa Webby). An *in vivo* approach may provide answers to establish the existence of, and molecular details of the proposed interaction. In addition, the work described in Chapter 5 provides a foundation to build the crosslinking approach upon.

6.3 Conclusions

Overall, the work detailed in this study progressed the understanding of the mechanisms underpinning the IM component of the Tol-Pal system. Biophysical and structural studies confirmed the helical and elongated nature of the TolA central domain, and these findings were supported by the generation of an AI model of TolA. Further examination of the model through CGMD simulations gave rise to a new hypothesis, that the interaction between the two helices of TolA-II are what drives pulling of TolB from Pal, and that these helices hinge from the IM to reach the OM.

The purification of the IM complex of TolQ-TolR was made possible through methodical optimisation including many iterations of construct design, purification steps and detergent choice. Although initial biochemical approaches to validate the complex were unsuccessful, native MS demonstrated the existence of oligomers of TolQ. Finally, further optimisation and preliminary EM data was able to identify the presence of TolQ as a pentamer, and density within the pore of TolQ consistent with TolR. Building upon the work in this study will make obtaining a high-resolution cryo-EM structure of the TolQ-TolR complex a real possibility in the near future.

Finally, the groundwork for a photoactivatable crosslinking approach to investigating the interaction between TolA and TolQ in the IM was laid in this study. The TolAHis22^{BPA} mutant was demonstrated to be functional, which in itself added to the picture of what role His22 plays in the interaction. Unfortunately, no crosslinks were obtained, however this result opens the door to further optimisation.

There are many unanswered questions on how the Tol-Pal system works to stabilise the OM and the role it plays in orchestrating cell division. The work described here has made inroads into understanding the mechanistic aspects of the system with respect to the IM components. It is hoped that this work can provide a foundation for further study. Ultimately helping us to understand the bacterial cell envelope, and the antibiotic targets within them, a little better.

7 Appendices

7.1 Appendix 1. Protein sequences

***Escherichia coli* TolA(48-310)-His$_6$**
MSSIDAVMVDSGAVVEQYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQ
LEKERLAAQEQKKQAEEAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAA
DAKKKAEAEAAKAAAEAQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKK
AAAEKAAADKKAAAEKAAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAA
AKAAAEADDIFGELSSGKNAPKTLEHHHHHH

***Escherichia coli* TolA(48-421)-His$_6$**
MSSIDAVMVDSGAVVEQYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQ
LEKERLAAQEQKKQAEEAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAA
DAKKKAEAEAAKAAAEAQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKK
AAAEKAAADKKAAAEKAAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAA
AKAAAEADDIFGELSSGKNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAGQIKSAI
ESKFYDASSYAGKTCTLRIKLAPDGMLLDIKPEGGDPALCQAALAAAKLAKIPKPPSQAVYE
VFKNAPLDFKPLEHHHHHH

***Proteus mirabilis* TolA(48-260)-His$_6$**
MSDGQVIDAIMVDPNAVVQEYNQQKAQQANVQKAEQERKARAQQQEEELRQQQIKEQER
LKALEVERLQAKEAAEAQRREAATAAAKAKEEQKQAEEAAAQAKAERDRILKEQADAKAKA
EAEAKKQAELAAKQKAEEAKAKAEAEAKAKAEADAKAKAEADAKAKAAAEAKAKAAAEAKA
KAAAQQQSKAVDSLLDGLVAQGDKQSGASAAGQLEHHHHHH

***Proteus mirabilis* TolA(48-357)-His$_6$**
MSDGQVIDAIMVDPNAVVQEYNQQKAQQANVQKAEQERKARAQQQEEELRQQQIKEQER
LKALEVERLQAKEAAEAQRREAATAAAKAKEEQKQAEEAAAQAKAERDRILKEQADAKAKA
EAEAKKQAELAAKQKAEEAKAKAEAEAKAKAEADAKAKAEADAKAKAAAEAKAKAAAEAKA
KAAAQQQSKAVDSLLDGLVAQGDKQSGASAAGQGGGNRSGANAAGIDRYTGQVKVAIEQ
KFFDSDLYKGRKCDLKISIAPDGMLVSAKAVGGDPALCQVALQAAKMATLPKPPKDIYEQVK
SPTIEFKPRLEHHHHHH

***Yersinia enterocolitica* TolA(45-291)-His$_6$**
MAGGEVIDAVMVDPGAVTEQYNRQQQQQTDAKRAEQQRQKKAEQQAEELQQKQAAEQQ
RLKELEKERLQAQEDAKQAAEEQKKQAAEQQKEAAEQQKAAAAAAAKAKEEQKQAEAAA
AQAKAEADKLVKAQAEAQKKAEAEAKKEAAVAAAAKKQADADAKKAADAAEKAATAAAEK
KAAADAEKKAAADAEKKAAADAKKAAAAEAKKKAAAEASAATDVDDLFGGLANSKNAPKS
GSGAGAAAAGKLEHHHHHH

***Yersinia enterocolitica* TolA(45-362)-His$_6$**
MAGGEVIDAVMVDPGAVTEQYNRQQQQQTDAKRAEQQRQKKAEQQAEELQQKQAAEQQ
RLKELEKERLQAQEDAKQAAEEQKKQAAEQQKEAAEQQKAAAAAAAKAKEEQKQAEAAA
AQAKAEADKLVKAQAEAQKKAEAEAKKEAAVAAAAKKQADADAKKAADAAEKAATAAAEK

KAAADAEKKAAADAEKKAAADAKKAAAAEAKKKAAAEASAATDVDDLFGGLANSKNAPKS
GSGAGAAAAGKGGGKKSGASAGDISGYLGQVTAAIQSKFYDADLYKGRTCNLRIKLAPDGL
LIDVTQEGGDPALCQAAIAAAKLAKIPKPPSQDVYEVFKNAPLVFKPQRLEHHHHHH

***Salmonella enterica* Typhimurium TolA(49-296)-His$_6$**

MSAIDAVMVDPGAVVQQYNRQQDQQASARRAEEERKKLQQQQAEELQQKQAAEQERLK
QLEKERLAAQEQQKQAEEAAKLAQQQQQQAEEAAKAAADAKKKAEAEAAKAAADAKKKA
EAEAVKAAADAKKKAEAEAAKAAADAKKKAEAEAAKAAAEAKKKAEAEAAKAAAEAKKKAD
AEAAKAAAEAKKKADAAAAKAAAEAKKKADAAAAKAAADAKKKAAAEKAAAAEGVDDLLG
DLSSGKNAPKTLEHHHHHH

***Salmonella enterica* Typhimurium TolA(49-407)-His$_6$**

MSAIDAVMVDPGAVVQQYNRQQDQQASARRAEEERKKLQQQQAEELQQKQAAEQERLK
QLEKERLAAQEQQKQAEEAAKLAQQQQQQAEEAAKAAADAKKKAEAEAAKAAADAKKKA
EAEAVKAAADAKKKAEAEAAKAAADAKKKAEAEAAKAAAEAKKKAEAEAAKAAAEAKKKAD
AEAAKAAAEAKKKADAAAAKAAAEAKKKADAAAAKAAADAKKKAAAEKAAAAEGVDDLLG
DLSSGKNAPKTGGGAKGNGQPSKDSGTSGANGGATGADISAYAKQIQVAIQSRLYDASLY
QGKQCVLHISLAPDGSLKSITSEGGDPALCQAALMAAKTAKIPKPPSQAVYEKIKDAKLDFKL
LEHHHHHH

***Escherichia coli* TolA(62-298)-TEV-His$_6$**

MEQYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKK
QAEEAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKA
AAEAQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAA
AEKAAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFEN
LYFQGLEHHHHHH

***Escherichia coli* TolA(72-298)-TEV-His$_6$**

MSSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQAEEAAKQAEL
KQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAAAEAQKKAEAA
AAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAAEKAAADKKAA
EKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFENLYFQGLEHHH
HHH

***Escherichia coli* TolA(82-298)-TEV-His$_6$**

MKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQAEEAAKQAELKQKQAEEAAA
KAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAAAEAQKKAEAAAAALKKKAEAA
EAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAAEKAAADKKAAEKAAAEKAAAD
KKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFENLYFQGLEHHHHHH

***Escherichia coli* TolA(92-298)-TEV-His$_6$**

MLREKQAAEQERLKQLEKERLAAQEQKKQAEEAAKQAELKQKQAEEAAAKAAADAKAKAE
ADAKAAEEAAKKAAADAKKKAEAEAAKAAAEAQKKAEAAAAALKKKAEAAEAAAAEARKKA
ATEAAEKAKAEAEKKAAAEKAAADKKAAAEKAAADKKAAEKAAAEKAAADKKAAAEKAAAD
KKAAAAKAAAEKAAAAKAAAEADDIFENLYFQGLEHHHHHH

***Escherichia coli* TolQ**

MVTDMNILDLFLKASLLVKLIMLILIGFSIASWAIIIQRTRILNAAAREAEAFEDKFWSGIELSRL
YQESQGKRDNLTGSEQIFYSGFKEFVRLHRANSHAPEAVVEGASRAMRISMNRELENLET
HIPFLGTVGSISPYIGLFGTVWGIMHAFIALGAVKQATLQMVAPGIAEALIATAIGLFAAIPAVM
AYNRLNQRVNKLELNYDNFMEEFTAILHRQAFTVSESNKG

***Escherichia coli* TolR**

MARARGRGRRDLKSEINIVPLLDVLLVLLLIFMATAPIITQSVEVDLPDATESQAVSSNDNPP
VIVEVSGIGQYTVVVEKDRLERLPPEQVVAEVSSRFKANPKTVFLIGGAKDVPYDEIIKALNLL
HSAGVKSVGLMTQPI

***Escherichia coli* TolA-strep-strep**
VSKATEQNDKLKRAIIISAVLHVILFAALIWSSFDENIEASAGGGGGSSIDAVMVDSGAV
VEQYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQ
AEEAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAA
AEAQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAA
EKAAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFGE
LSSGKNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAGQIKSAIESKFYDASSYAG
KTCTLRIKLAPDGMLLDIKPEGGDPALCQAALAAAKLAKIPKPPSQAVYEVFKNAPLDFK
PLELEENLYFQGQFGSWSHPQFEKGGGSGGGSGGGSWSHPQFEK

***Escherichia coli* TolR-strep-strep-His$_6$**
MARARGRGRRDLKSEINIVPLLDVLLVLLLIFMATAPIITQSVEVDLPDATESQAVSSNDNPP
VIVEVSGIGQYTVVVEKDRLERLPPEQVVAEVSSRFKANPKTVFLIGGAKDVPYDEIIKALNLL
HSAGVKSVGLMTQPI
LEENLYFQGQFGSWSHPQFEKGGGSGGGSGGGSWSHPQFEKHHHHHH

***Proteus mirabilis* TolQ**
VADMNVIDLFLKAGLLVKLIMLLLIGFSIASWAIIIQRTKVLNAADREAADFEDKFWSGTDLAR
LYKESQARRDELSGAEQIFHSGFKEFARLHQANVHAPDAVVNGASRAMRVSFNRELESLE
THIPFLGTVGSISPYIGLFGTVWGIMHAFIALGSVKQATLQMVAPGIAEALIATAIGLFAAIPAV
MAYNRFTQRVNKLEQSYDNFMEEFLTILHRQAFVSAEKK

***Proteus mirabilis* TolR**
MARGRSSRRGVKSEINIVPLLDVLLVLLLIFMATAPIISQSVEVDLPQATETQTVSTSDNPPVI
VEVAGVGQYNIRVEQEVLELLPQEQIMAEAKRQLEKNPKAVFLIGGAKDVPYDEIIKALNML
RQAGVKSVGLMTQPL

***Proteus mirabilis* TolA-strep-strep**
VGKKTKPIRSKLSGSVILSTVLHLLIIALLIWGSLTQKWNLGGGGGSDGQVIDAIMVDPNAVV
QEYNQQKAQQANVQKAEQERKARAQQQEEELRQQQIKEQERLKALEVERLQAKEAAEAQ
RREAATAAAKAKEEQKQAEEAAAQAKAERDRILKEQADAKAKAEAEAKKQAELAAKQKAE
EAKAKAEAEAKAKAEADAKAKAEADAKAKAAAEAKAKAAAEAKAKAAAQQQSKAVDSLLD
GLVAQGDKQSGASAAGQGGGNRSGANAAGIDRYTGQVKVAIEQKFFDSDLYKGRKCDLKI
SIAPDGMLVSAKAVGGDPALCQVALQAAKMATLPKPPKDIYEQVKSPTIEFKPLELEENLYF
QGQFGSWSHPQFEKGGGSGGGSGGGSWSHPQFEK

***Yersinia enterocolitica* TolQ**
VADMNILDLFLQASLFVKLIMLVLMGFSIASWAIIIQRTRILNAATRDAEAFEDKFWSGIELSRL
YQESQARRDTLGGSEQIFHSGFKEFARLHRANSHAPEAVIDGASRAMRISMNRELEALETH
IPFLGTVGSISPYIGLFGTVWGIMHAFISLGAVKQATLQMVAPGIAEALIATAIGLFAAIPAVMA
YNRLNQRVNKLEQNYDNFMEEFIAILHRQAFASAESK

***Yersinia enterocolitica* TolR**
MARVRGRKRRELKSEINIVPLLDVLLVLLLIFMATAPIITQSVEVDLPDATDSKTVSSDDNPPV
IVEVSGVGQYSLVIDHQRMEQLPSEQVVAEAQARLKSNPKTVFLIGGAKEVPYDEIIKALNM
LHQAGVTSVGLMTQPI

***Yersinia enterocolitica* TolA-strep-strep**
VGKATEQNDKLNRAVIVSVVLHIILIALLIWGSLTQTTEMGGGGAGGEVIDAVMVDPGAVTE
QYNRQQQQQTDAKRAEQQRQKKAEQQAEELQQKQAAEQQRLKELEKERLQAQEDAKQA
AEEQKKQAAEQQKEAAEQQKAAAAAAAKAKEEQKQAEAAAAQAKAEADKLVKAQAEAQK
KAEAEAKKEAAVAAAAKKQADADAKKAADAAEKAATAAAEKKAAADAEKKAAADAEKKAA

ADAKKAAAAEAKKKAAAEASAATDVDDLFGGLANSKNAPKSGSGAGAAAAGKGGGKKSG
ASAGDISGYLGQVTAAIQSKFYDADLYKGRTCNLRIKLAPDGLLIDVTQEGGDPALCQAAIA
AAKLAKIPKPPSQDVYEVFKNAPLVFKPQLEENLYFQGQFGSWSHPQFEKGGGSGGGSG
GGSWSHPQFEK

***Yersinia enterocolitica* TolR-strep-strep-His$_6$**
MARVRGRKRRELKSEINIVPLLDVLLVLLLIFMATAPIITQSVEVDLPDATDSKTVSSDDNPPV
IVEVSGVGQYSLVIDHQRMEQLPSEQVVAEAQARLKSNPKTVFLIGGAKEVPYDEIIKALNM
LHQAGVTSVGLMTQPILEENLYFQGQFGSWSHPQFEKGGGSGGGSGGGSWSHPQFEK

***Salmonella enterica* Typhimurium TolQ**
VTDMNILDLFLKASLLVKLIMLILIGFSIASWAIIIQRTRILNAAAREAEAFEDKFWSGIELSRLY
QESQGRRDSLSGSEQIFYSGFKEFVRLHRANSHAPEAVVEGASRAMRISMNRELETLETHI
PFLGTVGSISPYIGLFGTVWGIMHAFIALGAVKQATLQMVAPGIAEALIATAIGLFAAIPAVMA
YNRLNQRVNKLELNYDNFMEEFTAILHRQAFTVSESNKG

***Salmonella enterica* Typhimurium TolR**
MARTRGRGRRELKSEINIVPLLDVLLVLLLIFMATAPIITQSVEVDLPEANQSQAVSSNDDPP
VIIEVSGVGQYSVVVDKDRMDQLPSEQVIAEVKRHLQANPKTVFLIGGAKEVPYDEIIKALNL
LHSAGVKSVGLMTQPI

***Salmonella enterica* Typhimurium TolA-strep-strep**
MVDPGAVVQQYNRQQDQQASARRAEEERKKLQQQQAEELQQKQAAEQERLKQLEKERL
AAQEQQKQAEEAAKLAQQQQQQAEEAAKAAADAKKKAEAEAAKAAADAKKKAEAEAVKA
AADAKKKAEAEAAKAAADAKKKAEAEAAKAAAEAKKKAEAEAAKAAAEAKKKADAEAAKAA
AEAKKKADAAAAKAAAEAKKKADAAAAKAAADAKKKAAAEKAAAAEGVDDLLGDLSSGKN
APKTGGGAKGNGQPSKDSGTSGANGGATGADISAYAKQIQVAIQSRLYDASLYQGKQCVL
HISLAPDGSLKSITSEGGDPALCQAALMAAKTAKIPKPPSQAVYEKIKDAKLDFKLLEENLYF
QGQFGSWSHPQFEKGGGSGGGSGGGSWSHPQFEK

***Escherichia coli* TolA-His$_6$**
MVSKATEQNDKLKRAIIISAVLHVILFAALIWSSFDENIEASAGGGGGSSIDAVMVDSGAVVE
QYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQAE
EAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAAAE
AQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAAEK
AAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFGELSS
GKNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAGQIKSAIESKFYDASSYAGKTC
TLRIKLAPDGMLLDIKPEGGDPALCQAALAAAKLAKIPKPPSQAVYEVFKNAPLDFKPLEHHH
HHH

***Escherichia coli* TolAI15^{BPA}-His$_6$**
MVSKATEQNDKLKRA*IISAVLHVILFAALIWSSFDENIEASAGGGGGSSIDAVMVDSGAVVE
QYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQAE
EAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAAAE
AQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAAEK
AAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFGELSS
GKNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAGQIKSAIESKFYDASSYAGKTC
TLRIKLAPDGMLLDIKPEGGDPALCQAALAAAKLAKIPKPPSQAVYEVFKNAPLDFKPLEHHH
HHH
* Indicates the position of BPA.

***Escherichia coli* TolAH22^{BPA}-His$_6$**
MVSKATEQNDKLKRAIIISAVL*VILFAALIWSSFDENIEASAGGGGGSSIDAVMVDSGAVVEQ
YKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQAEE
AAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAAAEA

QKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAAEKA
AADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFGELSSG
KNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAGQIKSAIESKFYDASSYAGKTCTL
RIKLAPDGMLLDIKPEGGDPALCQAALAAAKLAKIPKPPSQAVYEVFKNAPLDFKPLEHHHH
HH

***Escherichia coli* TolAS32^{BPA}-His$_6$**
MVSKATEQNDKLKRAIIISAVLHVILFAALIW*SFDENIEASAGGGGGSSIDAVMVDSGAVVE
QYKRMQSQESSAKRSDEQRKMKEQQAAEELREKQAAEQERLKQLEKERLAAQEQKKQAE
EAAKQAELKQKQAEEAAAKAAADAKAKAEADAKAAEEAAKKAAADAKKKAEAEAAKAAAE
AQKKAEAAAAALKKKAEAAEAAAAEARKKAATEAAEKAKAEAEKKAAAEKAAADKKAAAEK
AAADKKAAEKAAAEKAAADKKAAAEKAAADKKAAAAKAAAEKAAAAKAAAEADDIFGELSS
GKNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAGQIKSAIESKFYDASSYAGKTC
TLRIKLAPDGMLLDIKPEGGDPALCQAALAAAKLAKIPKPPSQAVYEVFKNAPLDFKPLEHHH
HHH

7.2 Appendix 2. Supplementary Figures

Supplementary Table 7.1 Commercial crystallographic screens

Protein	Concentration (mg/ml)	Screen	Temperature
*Ec*TolA-II	4	JCSG+	4
		Morpheus	
		Proplex	
		MIDASplus	
		JCSG+	18
		Morpheus	
		Proplex	
		MIDASplus	
*Ec*TolA-II+III	16	JCSG+	4
		Morpheus	
		Proplex	
		MIDASplus	
		MemGold2	
		JCSG+	18
		Morpheus	
		Proplex	
		MIDASplus	
		MemGold2	
*Pm*TolA-II+III	18	JCSG+	18
		Morpheus2	
*Ye*TolA-II	9	JCSG+	18
		Morpheus2	
*ST*TolA-II	15	JCSG+	18
		Morpheus2	
*ST*TolA-II+III	15	JCSG+	18
		Morpheus2	
TolA 62-298	5.1	MIDASplus	18
		Morpheus	
		JCSG+	
		PACT	
TolA 92-298	27.8	MIDASplus	18
		Morpheus	
		JCSG+	
		PACT	
TolA 82-298	12	MIDASplus	18
		Morpheus	
		JCSG+	
		PACT	
TolA 72-298	145	Morpheus2	18
		JCSG+	
		PACT	
		STURA	
TolA-II	150	STURA	18
TolA 62-298	194	STURA	18
TolA 72-298 methylated	137	STURA	18
		JCSG+	18
		Morpheus2	18

1. Supplementary Figure 1. Manual crystallisation screens

A. Optimisation of STURA condition H8 (0.1 M sodium citrate pH 5.5, 20% w/v PEG 4000, 20% w/v 2-propanol)

	1	2	3	4	5	6	7	8	9	10	11	12	
A	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 4.8 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.2 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.6 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.8 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 24% 2-propanol	
B	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 4.8 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.2 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.6 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.8 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 22% 2-propanol	isopropanol change
C	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.8 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.2 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.6 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.8 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 20% 2-propanol	
D	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 4.8 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.2 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.6 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.8 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 18% 2-propanol	
E	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.8 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.2 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.6 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.8 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 24% PEG 4000 20% 2-propanol	
F	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.8 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.2 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.6 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.8 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 22% PEG 4000 20% 2-propanol	
G	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.8 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.2 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.6 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.8 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 20% 2-propanol	PEG 4000 change
H	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.8 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.2 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.6 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.8 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 18% PEG 4000 20% 2-propanol	
	original condition				pH change								

Plate with $TolA_{62-298}$ (194 mg/ml) in drop 1 and $TolA_{72-298}$ (145 mg/ml) in drop 2.

B. Further optimisation of STURA condition H8 (0.1 M sodium citrate pH 5.5, 20% w/v PEG 4000, 20% w/v 2-propanol)

	1	2	3	4	5	6	7	8	9	10	11	12	
A	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 4.75 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.25 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 5.75 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 24% 2-propanol	0.1 M sodium citrate pH 6.25 20% PEG 4000 24% 2-propanol	
B	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 4.75 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.25 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 5.75 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 22% 2-propanol	0.1 M sodium citrate pH 6.25 20% PEG 4000 22% 2-propanol	isopropanol change
C	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.75 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.25 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.75 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.25 20% PEG 4000 20% 2-propanol	
D	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 4.75 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.25 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 5.75 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 18% 2-propanol	0.1 M sodium citrate pH 6.25 20% PEG 4000 18% 2-propanol	
E	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.75 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.25 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.75 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 24% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.25 24% PEG 4000 20% 2-propanol	
F	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.75 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.25 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.75 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 22% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.25 22% PEG 4000 20% 2-propanol	PEG 4000 change
G	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.75 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.25 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.75 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.25 20% PEG 4000 20% 2-propanol	
H	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 20% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.5 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 4.75 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.0 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.25 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.5 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 5.75 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.0 18% PEG 4000 20% 2-propanol	0.1 M sodium citrate pH 6.25 18% PEG 4000 20% 2-propanol	
	original condition				pH change								

Plate set up with $TolA_{72-298}$ (145 mg/ml) with seed stock (made from original H8 condition in STURA screen) in drop 1 and $TolA_{72-298}$ (145 mg/ml) drop 2.

Second plate set up with $TolA_{72-298}$ (145 mg/ml) with Hampton silver bullet additive screen.

C. Buffer screening STURA condition H8 (0.1 M sodium citrate pH 5.5, 20% w/v PEG 4000, 20% w/v 2-propanol)

Plate with $TolA_{72-298}$ (145 mg/ml) with seed stock from optimisation screen in Supplementary Figure 1.B condition D12 (0.1 M sodium citrate pH 6.25, 20% PEG 4000, 18% 2-propanol) in drop 1, $TolA_{72-298}$ (145 mg/ml) drop 2

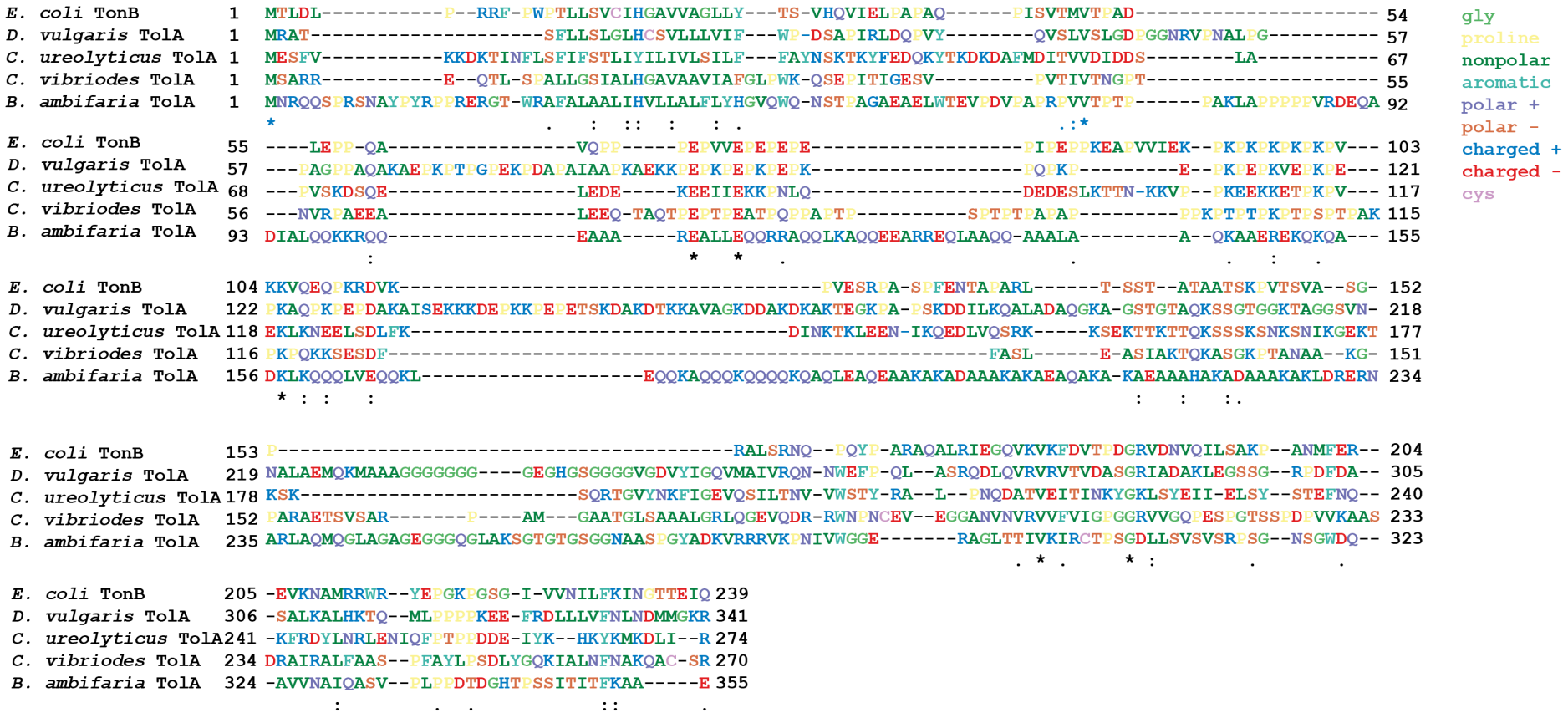
Whole plate reservoir condition STURA H8.

A-B: Resuspended in MQ water

C-D: Resuspended in 20 mM HEPES

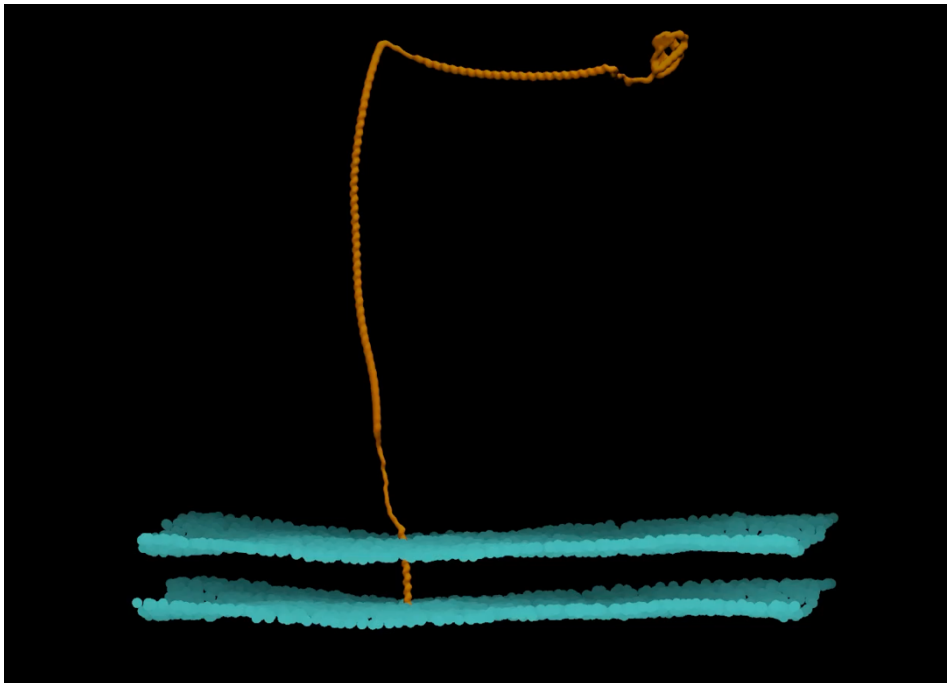
E-F: Resuspended in 20 mM HEPES, 100 mM NaCl

G-H: Resuspended in MQ water, crosslinked in 0.1% glutaraldehyde

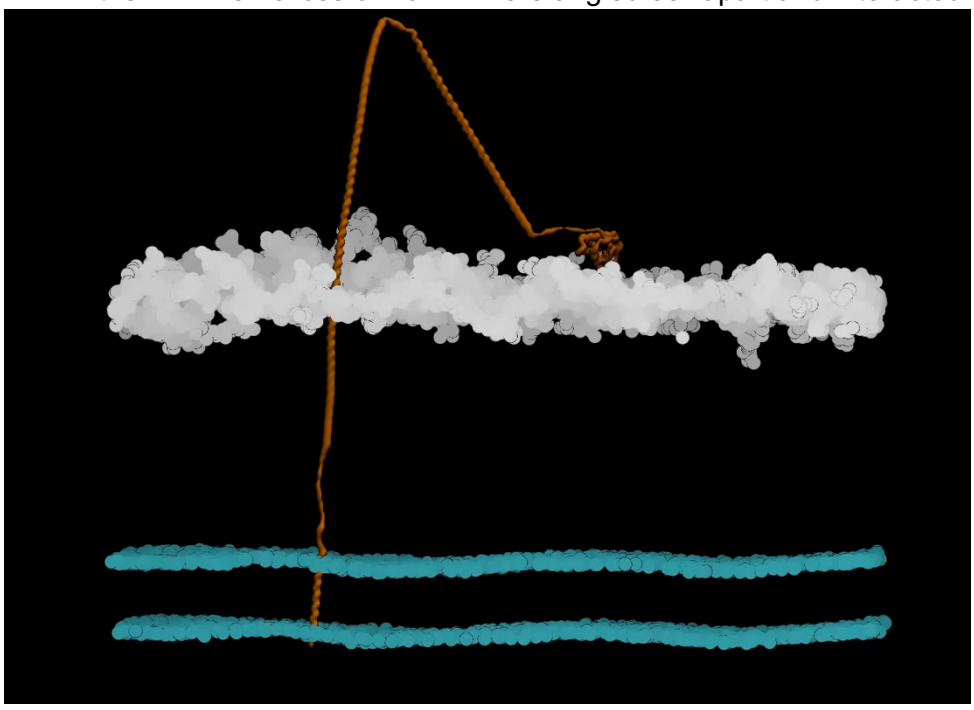


2. **Supplementary Figure 2.** Alignment of TolA of example species from different classes of proteobacteria demonstrate differences in primary sequence. *B. ambifaria* has a propensity for Ala, Lys, Gln, Asp and Glu in its central region. *C. ureolyticus, C. vibriodes* and *D. vulgaris* all display high Pro content more like *E. coli*'s TonB.

3. Supplementary Figure 3. mp4 videos of a representative of the coarse-grained molecular dynamics simulations conducted in Chapter 3.



A. A clip of the CGMD simulation of the *E. coli* TolA AlphaFold2 model embedded in the IM. The helices of TolA-II were angled 85° apart and interacted after 60 ns.



B. A clip of the CGMD simulation of the *E. coli* TolA AlphaFold2 model embedded in the IM and threaded through the PG. The helices of TolA-II were angled 40° apart and interacted after 250 ns.

CGMD simulation mp4 files also attached as additional files for submission.

4. Supplementary Figure 4. Predicted disulfide pair locations between *E. coli* TolA α2 and α3 in the AlphaFold2 model following CGMD simulations.

Residue Number 1	Amino Acid	Residue Number 2	Amino Acid
159	ALA	272	ASP
159	ALA	275	ALA
159	ALA	276	ALA
171	ALA	264	ALA
175	ALA	260	ALA
178	ALA	256	GLU
178	ALA	260	ALA
182	ALA	252	LYS
183	GLN	253	ALA
186	ALA	248	LYS
186	ALA	252	LYS
190	ALA	244	ALA
190	ALA	248	LYS
194	LYS	240	ALA
194	LYS	244	ALA
198	GLU	240	ALA
199	ALA	237	LYS
202	ALA	232	ALA
202	ALA	233	ALA
206	GLU	228	ALA
206	GLU	229	ALA

Atomistic version of AlphaFold2 *E. coli* TolA following CGMD was run through Disulfide by Design (http://cptweb.cpt.wayne.edu/DbD2/index.php) (Craig and Dombkowski, 2013) to predict residues that could be used to make novel disulfide bonds between α2 and α3 for functional assessment of TolA *in vivo*.

Bibliography

Abdelhamed, H., Lu, J., Lawrence, M. L. and Karsi, A. (2016) 'Involvement of tolQ and tolR genes in Edwardsiella ictaluri virulence', *Microbial Pathogenesis,* 100, pp. 90-94.

Abergel, C., Bouveret, E., Claverie, J. M., Brown, K., Rigal, A., Lazdunski, C. and Bénédetti, H. (1999) 'Structure of the Escherichia coli TolB protein determined by MAD methods at 1.95 Å resolution', *Structure,* 7(10), pp. 1291-1300.

Abergel, C., Walburger, A., Chenivesse, S. and Lazdunski, C. (2001) 'Crystallization and preliminary crystallographic study of the peptidoglycan-associated lipoprotein from Escherichia coli', *Acta Crystallographica Section D: Biological Crystallography,* 57(2), pp. 317-319.

Angelini, A., Cendron, L., Goncalves, S., Zanotti, G. and Terradot, L. (2008) 'Structural and enzymatic characterization of HP0496, a YbgC thioesterase from Helicobacter pylori', *Proteins: Structure, Function, and Bioinformatics,* 72(4), pp. 1212-1221.

Atanaskovic, I. and Kleanthous, C. (2019) 'Tools and approaches for dissecting protein bacteriocin import in Gram-negative bacteria', *Frontiers in Microbiology,* 10(MAR), pp. 646-646.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. and Mori, H. (2006) 'Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection', *Molecular Systems Biology,* 2(1), pp. 2006.0008-2006.0008.

Behrens, H. M., Six, A., Walker, D. and Kleanthous, C. (2017) 'The therapeutic potential of bacteriocins as protein antibiotics', *Emerging Topics in Life Sciences,* 1(1), pp. 65-65.

Bernstein, A., Rolfe, B. and Onodera, K. (1972) 'Pleiotropic Properties and Genetic Organization of the tolA, B Locus of Escherichia coli K-12', *Journal of Bacteriology,* 112(1), pp. 74-74.

Biou, V., Adaixo, R. J. D., Chami, M., Coureux, P.-D., Laurent, B., Enguéné, V. Y. N., De Amorim, G. C., Izadi-Pruneyre, N., Malosse, C., Chamot-Rooke, J., Stahlberg, H. and Delepelaire, P. (2022) 'Structural and molecular determinants for the interaction of ExbB from Serratia marcescens and HasB, a TonB paralog', *Communications Biology,* 5(1), pp. 1-15.

Boags, A. T., Samsudin, F. and Khalid, S. (2019) 'Binding from Both Sides: TolR and Full-Length OmpA Bind and Maintain the Local Structure of the E. coli Cell Wall', *Structure,* 27(4), pp. 713-724.

Bonsor, D. A., Grishkovskaya, I., Dodson, E. J. and Kleanthous, C. (2007) 'Molecular Mimicry Enables Competitive Recruitment by a Natively Disordered Protein', *Journal of the American Chemical Society,* 129(15), pp. 4800-4807.

Bonsor, D. A., Hecht, O., Vankemmelbeke, M., Sharma, A., Krachler, A. M., Housden, N. G., Lilly, K. J., James, R., Moore, G. R. and Kleanthous, C. (2009) 'Allosteric Beta-Propeller Signalling in TolB and its Manipulation by Translocating Colicins', *EMBO Journal,* 28(18), pp. 2846-2857.

Bouveret, E., Derouiche, R., Rigal, A., Lloubès, R., Lazdunski, C. and Bénédetti, H. (1995) 'Peptidoglycan-associated lipoprotein-TolB interaction: A possible key to explaining the formation of contact sites between the inner and outer membranes of Escherichia coli', *Journal of Biological Chemistry,* 270(19), pp. 11071-11077.

Bowe, F., Lipps, C. J., Tsolis, R. M., Groisman, E., Heffron, F. and Kusters, J. G. (1998) 'At Least Four Percent of the Salmonella typhimurium Genome Is Required for Fatal Infection of Mice', *Infection and Immunity,* 66(7), pp. 3372-3372.

Braun, V. (1975) 'Covalent lipoprotein from the outer membrane of Escherichia coli', *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes,* 415(3), pp. 335-377.

Braun, V. and Herrmann, C. (2004) 'Point mutations in transmembrane helices 2 and 3 of ExbB and TolQ affect their activities in Escherichia coli K-12', *Journal of Bacteriology,* 186(13), pp. 4402-4406.

Caesar, J., Reboul, C. F., Machello, C., Kiesewetter, S., Tang, M. L., Deme, J. C., Johnson, S., Elmlund, D., Lea, S. M. and Elmlund, H. (2020) 'SIMPLE 3.0. Stream single-particle cryo-EM analysis in real time', *Journal of Structural Biology: X,* 4, pp. 100040.

Carr, S., Penfold, C. N., Bamford, V., James, R. and Hemmings, A. M. (2000) 'The structure of TolB, an essential component of the tol-dependent translocation system, and its protein-protein interaction with the translocation domain of colicin E9', *Structure,* 8(1), pp. 57-66.

Cascales, E., Bernadac, A., Gavioli, M., Lazzaroni, J. C. and Lloubes, R. (2002) 'Pal lipoprotein of Escherichia coli plays a major role in outer membrane integrity', *Journal of Bacteriology,* 184(3), pp. 754-9.

Cascales, E., Gavioli, M., Sturgis, J. N. and Lloubes, R. (2000) 'Proton motive force drives the interaction of the inner membrane TolA and outer membrane Pal proteins in Escherichia coli', *Molecular Microbiology,* 38(4), pp. 904-915.

Cascales, E. and Lloubès, R. (2004) 'Deletion analyses of the peptidoglycan-associated lipoprotein Pal reveals three independent binding sequences including a TolA box', *Molecular Microbiology,* 51(3), pp. 873-885.

Cascales, E., Lloubès, R. and Sturgis, J. N. (2001) 'The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB', *Molecular Microbiology,* 42(3), pp. 795-807.

Celia, H., Botos, I., Ni, X., Fox, T., De Val, N., Lloubes, R., Jiang, J. and Buchanan, S. K. (2019) 'Cryo-EM structure of the bacterial Ton motor subcomplex ExbB–ExbD provides information on structure and stoichiometry', *Communications Biology,* 2(1), pp. 1-6.

Celia, H., Noinaj, N., Zakharov, S. D., Bordignon, E., Botos, I., Santamaria, M., Barnard, T. J., Cramer, W. A., Lloubes, R. and Buchanan, S. K. (2016) 'Structural insight into the role of the Ton complex in energy transduction', *Nature,* 538(7623), pp. 60-65.

Chang, C., Mooser, A., Plückthun, A. and Wlodawer, A. (2001) 'Crystal Structure of the Dimeric C-terminal Domain of TonB Reveals a Novel Fold', *Journal of Biological Chemistry,* 276(29), pp. 27535-27540.

Chin, J. W., Martin, A. B., King, D. S., Wang, L. and Schultz, P. G. (2002) 'Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America,* 99(17), pp. 11020-11024.

Chorev, D. S., Baker, L. A., Wu, D., Beilsten-Edmands, V., Rouse, S. L., Zeev-Ben-Mordehai, T., Jiko, C., Samsudin, F., Gerle, C., Khalid, S., Stewart, A. G., Matthews, S. J., Grünewald, K. and Robinson, C. V. (2018) 'Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry', *Science,* 362(6416), pp. 829-834.

Clarke, T., Tari, L. and Vogel, H. (2005) 'Structural Biology of Bacterial Iron Uptake Systems', *Current Topics in Medicinal Chemistry,* 1(1), pp. 7-30.

Clavel, T., Germon, P., Vianney, A., Portalier, R. and Lazzaroni, J. C. (1998) 'TolB protein of Escherichia coli K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA', *Molecular Microbiology,* 29(1), pp. 359-367.

Clavel, T., Lazzaroni, J. C., Vianney, A. and Portalier, R. (1996) 'Expression of the tolQRA genes of Escherichia coli K-12 is controlled by the RcsC sensor protein involved in capsule synthesis', *Molecular Microbiology,* 19(1), pp. 19-25.

Clifton, L. A., Skoda, M. W. A., Le Brun, A. P., Ciesielski, F., Kuzmenko, I., Holt, S. A. and Lakey, J. H. (2015) 'Effect of Divalent Cation Removal on the Structure of Gram-Negative Bacterial Outer Membrane Models', *Langmuir,* 31(1), pp. 404-412.

Connolley, L., Szczepaniak, J., Kleanthous, C. and Murray, S. M. (2021) 'The quantitative basis for the redistribution of immobile bacterial lipoproteins to division septa', *PLoS Computational Biology,* 17(12), pp. e1009756.

Cox, J. and Mann, M. (2008) 'MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification', *Nature Biotechnology,* 26(12), pp. 1367-1372.

Craig, D. B. and Dombkowski, A. A. (2013) 'Disulfide by Design 2.0: A web-based tool for disulfide engineering in proteins', *BMC Bioinformatics,* 14(1), pp. 1-7.

Cramer, W. A., Sharma, O. and Zakharov, S. D. (2018) 'On mechanisms of colicin import: the outer membrane quandary', *Biochemical Journal,* 475(23), pp. 3903-3915.

Cukierman, S. (2003) 'The transfer of protons in water wires inside proteins', *Frontiers in Bioscience,* 8(6), pp. s1118-1139.

Datsenko, K. A. and Wanner, B. L. (2000) 'One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products', *Proceedings of the National Academy of Sciences of the United States of America,* 97(12), pp. 6640-6645.

De Mot, R. and Vanderleyden, J. (1994) 'The C-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both Gram-positive and Gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan', *Molecular Microbiology,* 12(2), pp. 333-334.

DeCoursey, T. E. and Hosler, J. (2014) 'Philosophy of voltage-gated proton channels', *Journal of The Royal Society Interface,* 11(92), pp. 20130799.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) 'NMRPipe: a multidimensional spectral processing system based on UNIX pipes', *Journal of Biomolecular NMR,* 6(3), pp. 277-93.

Delano, W. L. (2010) *The PyMOL Molecular Graphics System*. Available at: https://pymol.org/2 (Accessed: 2018).

Deme, J. C., Johnson, S., Vickery, O., Aron, A., Monkhouse, H., Griffiths, T., James, R. H., Berks, B. C., Coulton, J. W., Stansfeld, P. J. and Lea, S. M. (2020) 'Structures of the stator complex that drives rotation of the bacterial flagellum', *Nature Microbiology,* 5(12), pp. 1553-1564.

Dennis, J. J., Lafontaine, E. R. and Sokol, P. A. (1996) 'Identification and characterization of the tolQRA genes of Pseudomonas aeruginosa', *Journal of Bacteriology,* 178(24), pp. 7059-7059.

Deprez, C., Lloubès, R., Gavioli, M., Marion, D., Guerlesquin, F. and Blanchard, L. (2005) 'Solution structure of the E. coli TolA C-terminal domain reveals conformational changes upon binding to the phage g3p N-terminal domain', *Journal of Molecular Biology,* 346(4), pp. 1047-1057.

Derouiche, R., Bénédetti, H., Lazzaroni, J.-C., Lazdunski, C. and Lloubès, R. (1995) 'Protein Complex within Escherichia coli Inner Membrane: TolA N-terminal domain interacts with TolQ and TolR proteins', *Journal of Biological Chemistry,* 270(19), pp. 11078-11084.

Derouiche, R., Lloubès, R., Sasso, S., Bouteille, H., Oughideni, R., Lazdunski, C. and Loret, E. (1999) 'Circular dichroism and molecular modeling of the E. coli TolA periplasmic domains', *Biospectroscopy,* 5(3), pp. 189-198.

Dietsche, T., Tesfazgi Mebrhatu, M., Brunner, M. J., Abrusci, P., Yan, J., Franz-Wachtel, M., Schärfe, C., Zilkenat, S., Grin, I., Galán, J. E., Kohlbacher, O., Lea, S., Macek, B., Marlovits, T. C., Robinson, C. V. and Wagner, S. (2016) 'Structural and Functional Characterization of the Bacterial Type III Secretion Export Apparatus', *PLOS Pathogens,* 12(12), pp. e1006071.

Dormán, G., Nakamura, H., Pulsipher, A. and Prestwich, G. D. (2016) 'The Life of Pi Star: Exploring the Exciting and Forbidden Worlds of the Benzophenone Photophore', *Chemical Reviews,* 116(24), pp. 15284-15398.

Dubuisson, J. F., Vianney, A., Hugouvieux-Cotte-Pattat, N. and Lazzaroni, J. C. (2005) 'Tol-Pal proteins are critical cell envelope components of Erwinia chrysanthemi affecting cell morphology and virulence', *Microbiology,* 151(10), pp. 3337-3347.

Dubuisson, J. F., Vianney, A. and Lazzaroni, J. C. (2002) 'Mutational Analysis of the TolA C-Terminal Domain of Escherichia coli and Genetic Evidence for an Interaction between TolA and TolB', *Journal of Bacteriology,* 184(16), pp. 4620-4620.

Ducret, A., Quardokus, E. M. and Brun, Y. V. (2016) 'MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis', *Nat Microbiol,* 1(7), pp. 1-7.

Egan, A. J. F., Maya-Martinez, R., Ayala, I., Bougault, C. M., Banzhaf, M., Breukink, E., Vollmer, W. and Simorre, J. P. (2018) 'Induced conformational changes activate the peptidoglycan synthase PBP1B', *Molecular Microbiology,* 110(3), pp. 335-356.

Evans, J. S., Levine, B. A., Trayer, I. P., Dorman, C. J. and Higgins, C. F. (1986) 'Sequence-imposed structural constraints in the TonB protein of E. coli', *FEBS Letters,* 208(2), pp. 211-216.

Fortney, K. R., Young, R. S., Bauer, M. E., Katz, B. P., Hood, A. F., Munson, Jr. and Spinola, S. M. (2000) 'Expression of Peptidoglycan-Associated Lipoprotein Is Required for Virulence in the Human Model of Haemophilus ducreyi Infection', *Infection and Immunity,* 68(11), pp. 6441-6441.

Garcia-Herrero, A., Peacock, R. S., Howard, S. P. and Vogel, H. J. (2007) 'The solution structure of the periplasmic domain of the TonB system ExbD protein reveals an unexpected structural homology with siderophore-binding proteins', *Molecular Microbiology,* 66(4), pp. 872-889.

Gerding, M. A., Ogata, Y., Pecora, N. D., Niki, H. and De Boer, P. A. J. (2007) 'The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli', *Molecular Microbiology,* 63(4), pp. 1008-1025.

Germon, P., Clavel, T., Vianney, A., Portalier, R. and Lazzaroni, J. C. (1998) 'Mutational Analysis of the Escherichia coli K-12 TolA N-Terminal Region and Characterization of Its TolQ-Interacting Domain by Genetic Suppression', *Journal of Bacteriology,* 180(24), pp. 6433-6433.

Germon, P., Ray, M. C., Vianney, A. and Lazzaroni, J. C. (2001) 'Energy-dependent conformational change in the TolA protein of Escherichia coli involves its N-terminal domain, TolQ, and TolR', *Journal of Bacteriology,* 183(14), pp. 4110-4114.

Goemaere, E. L., Cascales, E. and Lloubès, R. (2007) 'Mutational Analyses Define Helix Organization and Key Residues of a Bacterial Membrane Energy-transducing Complex', *Journal of Molecular Biology,* 366(5), pp. 1424-1436.

Goemaere, E. L., Devert, A., Lloubès, R. and Cascales, E. (2007) 'Movements of the TolR C-terminal domain depend on TolQR ionizable key residues and regulate activity of the Tol complex', *Journal of Biological Chemistry,* 282(24), pp. 17749-17757.

Gratia, J. P. (1964) 'Resistance to colicin B in Escherichia coli specificity relations among colicins B, I and V and phage T-4. genetic study', *Annales de l'Institut Pasteur,* 107, pp. 132-51.

Gray, A. N., Egan, A. J. F., van't Veer, I. L., Verheul, J., Colavin, A., Koumoutsi, A., Biboy, J., Altelaar, M. A. F., Damen, M. J., Huang, K. C., Simorre, J. P., Breukink, E., den Blaauwen, T., Typas, A., Gross, C. A. and Vollmer, W. (2015) 'Coordination of peptidoglycan synthesis and outer membrane constriction during Escherichia coli cell division', *eLife,* 4.

Grizot, S. and Buchanan, S. K. (2004) 'Structure of the OmpA-like domain of RmpM from Neisseria meningitidis', *Molecular Microbiology,* 51(4), pp. 1027-1037.

Gully, D. and Bouveret, E. (2006) 'A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in Escherichia coli', *Proteomics,* 6(1), pp. 282-293.

Gumbart, J., Wiener, M. C. and Tajkhorshid, E. (2007) 'Mechanics of Force Propagation in TonB-Dependent Outer Membrane Transport', *Biophysical Journal,* 93(2), pp. 496-496.

Hannavy, K., Barr, G. C., Dorman, C. J., Adamson, J., Mazengera, L. R., Gallagher, M. P., Evans, J. S., Levine, B. A., Trayer, I. P. and Higgins, C. F. (1990) 'TonB protein of Salmonella typhimurium: A model for signal transduction between membranes', *Journal of Molecular Biology,* 216(4), pp. 897-910.

Hellman, J., Roberts, J. D., Tehan, M. M., Allaire, J. E. and Shaw Warren, H. (2002) 'Bacterial Peptidoglycan-associated Lipoprotein Is Released into the Bloodstream in Gram-negative Sepsis and Causes Inflammation and Death in Mice', *Journal of Biological Chemistry,* 277(16), pp. 14274-14280.

Hickman, S. J., Cooper, R. E. M., Bellucci, L., Paci, E. and Brockwell, D. J. (2017) 'Gating of TonB-dependent transporters by substrate-specific forced remodelling', *Nature Communications 2017 8:1,* 8(1), pp. 1-12.

Housden, N. G., Hopper, J. T. S., Lukoyanova, N., Rodriguez-Larrea, D., Wojdyla, J. A., Klein, A., Kaminska, R., Bayley, H., Saibil, H. R., Robinson, C. V. and Kleanthous, C. (2013) 'Intrinsically disordered protein threads through the bacterial outer-membrane porin OmpF', *Science,* 340(6140), pp. 1570-1574.

Housden, N. G. and Kleanthous, C. (2012) 'Colicin translocation across the Escherichia coli outer membrane', *Biochemical Society Transactions,* 40, pp. 1475-1479.

Housden, N. G., Loftus, S. R., Moore, G. R., James, R. and Kleanthous, C. (2005) 'Cell entry mechanism of enzymatic bacterial colicins: Porin recruitment and the thermodynamics of receptor binding', *Proceedings of the National Academy of Sciences,* 102(39), pp. 13849-13854.

Housden, N. G., Webby, M. N., Lowe, E. D., El-Baba, T. J., Kaminska, R., Redfield, C., Robinson, C. V. and Kleanthous, C. (2021) 'Toxin import through the antibiotic efflux channel TolC', *Nat Commun,* 12(1), pp. 1-12.

Jordan, L. D., Zhou, Y., Smallwood, C. R., Lill, Y., Ritchie, K., Yip, W. T., Newton, S. M. and Klebba, P. E. (2013) 'Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane', *Proceedings of the National Academy of Sciences of the United States of America,* 110(28), pp. 11553-11558.

Journet, L., Rigal, A., Lazdunski, C. and Bénédetti, H. (1999) 'Role of TolR N-terminal, central, and C-terminal domains in dimerization and interaction with TolA and TolQ', *Journal of Bacteriology,* 181(15), pp. 4476-4484.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. and Hassabis, D. (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature,* 596(7873), pp. 583-589.

Kampfenkel, K. and Braun, V. (1993) 'Membrane topologies of the TolQ and TolR proteins of Escherichia coli: inactivation of TolQ by a missense mutation in the proposed first transmembrane segment', *Journal of Bacteriology,* 175(14), pp. 4485-4491.

Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popović, Z., Baker, D. and Players, F. (2011) 'Algorithm discovery by protein folding game players', *Proceedings of the National Academy of Sciences,* 108(47), pp. 18949-18953.

Kim, Y., Joachimiak, A., Skarina, T., Savchenko, A. and Edwards, A. 2004. Crystal Structure of Hypothetical Protein EC709 from Escherichia coli. Protein Data Bank.

Kleanthous, C. (2010) 'Swimming against the tide: progress and challenges in our understanding of colicin translocation', *Nature Reviews Microbiology,* 8(12), pp. 843-848.

Ködding, J., Killig, F., Polzer, P., Howard, S. P., Diederichs, K. and Welte, W. (2005) 'Crystal Structure of a 92-Residue C-terminal Fragment of TonB from Escherichia coli Reveals Significant Conformational Changes Compared to Structures of Smaller TonB Fragments', *Journal of Biological Chemistry,* 280(4), pp. 3022-3028.

Koebnik, R. (1993) 'The molecular interaction between components of the TonB–ExbBD–dependent and of the TolQRA-dependent bacterial uptake systems', *Molecular Microbiology,* 9(1), pp. 219-219.

Koebnik, R., Bäumler, A. J., Heesemann, J., Braun, V. and Hantke, K. (1993) 'The TonB protein of Yersinia enterocolitica and its interactions with TonB-box proteins', *Molecular and General Genetics MGG,* 237(1), pp. 152-160.

Köhler, S. D., Weber, A., Howard, P. S., Welte, W. and Drescher, M. (2010) 'The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria', *Protein Science,* 19(4), pp. 625-630.

Kojima, S., Imada, K., Sakuma, M., Sudo, Y., Kojima, C., Minamino, T., Homma, M. and Namba, K. (2009) 'Stator assembly and activation mechanism of the flagellar motor by the periplasmic region of MotB', *Molecular Microbiology,* 73(4), pp. 710-718.

Kopp, D. R. and Postle, K. (2020) 'The intrinsically disordered region of ExbD is required for signal transduction', *Journal of Bacteriology,* 202(7), pp. e00687-19.

Krachler, A. M., Sharma, A., Cauldwell, A., Papadakos, G. and Kleanthous, C. (2010) 'TolA Modulates the Oligomeric Status of YbgF in the Bacterial Periplasm', *Journal of Molecular Biology,* 403(2), pp. 270-285.

Kühlbrandt, W. and Davies, K. M. (2016) 'Rotary ATPases: A New Twist to an Ancient Machine', *Trends in Biochemical Sciences,* 41(1), pp. 106-116.

Lakey, B. D., Myers, K. S., Alberge, F., Mettert, E. L., Kiley, P. J., Noguera, D. R. and Donohue, T. J. (2022) 'The essential Rhodobacter sphaeroides CenKR two-component system regulates cell division and envelope biosynthesis', *PLOS genetics,* 18(6), pp. e1010270-e1010270.

Lancia, J. K., Nwokoye, A., Dugan, A., Joiner, C., Pricer, R. and Mapp, A. K. (2014) 'Sequence context and crosslinking mechanism affect the efficiency of in vivo capture of a protein–protein interaction', *Biopolymers,* 101(4), pp. 391-397.

Larsen, R. A., Deckert, G. E., Kastead, K. A., Devanathan, S., Keller, K. L. and Postle, K. (2007) 'His20 provides the sole functionally significant side chain in the essential TonB transmembrane domain', *Journal of Bacteriology,* 189(7), pp. 2825-2833.

Larsen, R. A. and Postle, K. (2001) 'Conserved Residues Ser16 and His20 and Their Relative Positioning Are Essential for TonB Activity, Cross-linking of TonB with ExbB, and the Ability of TonB to Respond to Proton Motive Force', *Journal of Biological Chemistry,* 276(11), pp. 8111-8117.

Larsen, R. A., Thomas, M. G. and Postle, K. (1999) 'Protonmotive force, ExbB and ligand-bound FepA drive conformational changes in TonB', *Molecular Microbiology,* 31(6), pp. 1809-1824.

Larsen, R. A. and Wood, G. E. (1993) 'The conserved proline-rich motif is not essential for energy transduction by Escherichia coli TonB protein', 10(5), pp. 943-953.

Lazdunski, C. J., Bouveret, E., Rigal, A., Journet, L., Lloubès, R. and Bénédetti, H. (1998) 'Colicin import into Escherichia coli cells', *Journal of Bacteriology,* 180(19), pp. 4993-5002.

Lazzaroni, J.-C. and Portalier, R. (1992) 'The excC gene of Escherichia coli K-12 required for cell envelope integrity encodes the peptidoglycan-associated lipoprotein (PAL)', *Molecular Microbiology,* 6(6), pp. 735-742.

Lazzaroni, J. C., Vianney, A., Popot, J. L., Bénédetti, H., Samatey, F., Lazdunski, C., Portalier, R. and Géli, V. (1995) 'Transmembrane α-Helix Interactions are Required for the Functional Assembly of the Escherichia coli Tol Complex', *Journal of Molecular Biology,* 246(1), pp. 1-7.

Levengood, S. K., Beyer, W. F. and Webster, R. E. (1991) 'TolA: A membrane protein involved in colicin uptake contains an extended helical region', *Proceedings of the National Academy of Sciences of the United States of America,* 88(14), pp. 5939-5943.

Li, C., Zhang, Y., Vankemmelbeke, M., Hecht, O., Aleanizy, F. S., Macdonald, C., Moore, G. R., James, R. and Penfold, C. N. (2012) 'Structural evidence that colicin A protein binds to a novel binding site of TolA protein in Escherichia coli periplasm', *The Journal of Biological Chemistry,* 287(23), pp. 19048-19057.

Lim, B. L. (2010) 'TonB-Dependent Receptors in Nitrogen-Fixing Nodulating Bacteria', *Microbes and Environments,* 25(2), pp. 67-74.

López-Sánchez, A., Leal-Morales, A., Jiménez-Díaz, L., Platero, A. I., Bardallo-Pérez, J., Díaz-Romero, A., Acemel, R. D., Ilán, J. M., Jiménez-López, J. and Govantes, F. (2016) 'Biofilm formation-defective mutants in Pseudomonas putida', *FEMS Microbiology Letters,* 363(13), pp. 1-10.

Lubkowski, J., Hennecke, F., Plückthun, A. and Wlodawer, A. (1999) 'Filamentous phage infection: Crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA', *Structure,* 7(6), pp. 711-722.

Maguire, J. B., Haddox, H. K., Strickland, D., Halabiya, S. F., Coventry, B., Griffin, J. R., Pulavarti, S. V. S. R. K., Cummins, M., Thieker, D. F., Klavins, E., Szyperski, T., Dimaio, F., Baker, D. and Kuhlman, B. (2021) 'Perturbing the energy landscape for improved packing during computational protein design', *Proteins: Structure, Function, and Bioinformatics,* 89(4), pp. 436-449.

Maki-Yonekura, S., Matsuoka, R., Yamashita, Y., Shimizu, H., Tanaka, M., Iwabuki, F. and Yonekura, K. (2018) 'Hexameric and pentameric complexes of the ExbBD energizer in the Ton system', *eLife,* 7, pp. e35419.

Mandela, E., Stubenrauch, C. J., Ryoo, D., Hwang, H., Cohen, E. J., Torres, V. L., Deo, P., Webb, C. T., Huang, C., Schittenhelm, R. B., Beeby, M., Gumbart, J. C., Lithgow, T. and Hay,

I. D. (2022) 'Adaptation of the periplasm to maintain spatial constraints essential for cell envelope processes and cell viability', *eLife,* 11, pp. 11-11.

Marmon, L. (2013) 'Elucidating the origin of the ExbBD components of the TonB system through Bayesian inference and maximum-likelihood phylogenies', *Molecular Phylogenetics and Evolution,* 69(3), pp. 674-686.

Marze, N. A., Roy Burman, S. S., Sheffler, W. and Gray, J. J. (2018) 'Efficient flexible backbone protein–protein docking for challenging targets', *Bioinformatics,* 34(20), pp. 3461-3469.

Masilamani, R., Cian, M. B. and Dalebroux, Z. D. (2018) 'Salmonella Tol-Pal reduces outer membrane glycerophospholipid levels for envelope homeostasis and survival during bacteremia', *Infection and Immunity,* 86(7), pp. e00173-18.

Matias, V. R. F., Al-Amoudi, A., Dubochet, J. and Beveridge, T. J. (2003) 'Cryo-transmission electron microscopy of frozen-hydrated sections of Escherichia coli and Pseudomonas aeruginosa', *Journal of Bacteriology,* 185(20), pp. 6112-6118.

Meury, J. and Devilliers, G. (1999) 'Impairment of cell division in tolA mutants of Escherichia coli at low and high medium osmolarities', *Biology of the Cell,* 91(1), pp. 67-75.

Mitchell, P. (1966) 'Chemiosmotic coupling in oxidative and photosynthetic phosphorylation', *Biological Reviews,* 41(3), pp. 445-501.

Mizuno, T. (1979) 'A Novel Peptidoglycan-Associated Lipoprotein Found in the Cell Envelope of Pseudomonas aeruginosa and Escherichia coli', *The Journal of Biochemistry,* 86(4), pp. 991-1000.

Nagel de Zwaig, R. and Luria, S. E. (1967) 'Genetics and physiology of colicin-tolerant mutants of Escherichia coli', *Journal of Bacteriology,* 94(4), pp. 1112-1123.

Nagle, J. F. and Tristram-Nagle, S. (1983) 'Hydrogen bonded chain mechanisms for proton conduction and proton pumping', *The Journal of Membrane Biology,* 74(1), pp. 1-14.

Nikaido, H. (2003) 'Molecular Basis of Bacterial Outer Membrane Permeability Revisited', *Microbiology and Molecular Biology Reviews,* 67(4), pp. 593-593.

Nivón, L. G., Moretti, R. and Baker, D. (2013) 'A Pareto-Optimal Refinement Method for Protein Design Scaffolds', *PLoS ONE,* 8(4), pp. e59004.

Notredame, C., Higgins, D. G. and Heringa, J. (2000) 'T-Coffee: A novel method for fast and accurate multiple sequence alignment', *Journal of Molecular Biolology,* 302(1), pp. 205-17.

O'Neill, J., Xie, M., Hijnen, M. and Roujeinikova, A. (2011) 'Role of the MotB linker in the assembly and activation of the bacterial flagellar motor', *Acta Crystallographica Section D: Biological Crystallography,* 67, pp. 1009-1016.

Ollis, A. A., Manning, M., Held, K. G. and Postle, K. (2009) 'Cytoplasmic membrane protonmotive force energizes periplasmic interactions between ExbD and TonB', *Molecular Microbiology,* 73(3), pp. 466-481.

Parsons, L. M., Grishaev, A. and Bax, A. (2008) 'The periplasmic domain of TolR from Haemophilus influenzae forms a dimer with a large hydrophobic groove: NMR solution structure and comparison to SAXS Data', *Biochemistry,* 47(10), pp. 3131-3142.

Parsons, L. M., Lin, F. and Orban, J. (2006) 'Peptidoglycan recognition by Pal, an outer membrane lipoprotein', *Biochemistry,* 45(7), pp. 2122-2128.

Paterson, G. K., Northern, H., Cone, D. B., Willers, C., Peters, S. E. and Maskell, D. J. (2009) 'Deletion of tolA in Salmonella Typhimurium generates an attenuated strain with vaccine potential', *Microbiology,* 155(1), pp. 220-228.

Pawelek, P. D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C. M., Moiseeva, N., Allaire, M. and Coulton, J. W. (2006) 'Structure of TonB in complex with FhuA, E. coli outer membrane receptor', *Science,* 312(5778), pp. 1399-1402.

Peacock, S. R., Weljie, A. M., Peter Howard, S., Price, F. D. and Vogel, H. J. (2005) 'The Solution Structure of the C-terminal Domain of TonB and Interaction Studies with TonB Box Peptides', *Journal of Molecular Biology,* 345(5), pp. 1185-1197.

Petiti, M., Serrano, B., Faure, L., Lloubes, R., Mignot, T. and Duché, D. (2019) 'Tol Energy-Driven Localization of Pal and Anchoring to the Peptidoglycan Promote Outer-Membrane Constriction', *Journal of Molecular Biology,* 431(17), pp. 3275-3288.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) 'UCSF Chimera: A visualization system for exploratory research and analysis', *Journal of Computational Chemistry,* 25(13), pp. 1605-1612.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H. and Ferrin, T. E. (2021) 'UCSF ChimeraX: Structure visualization for researchers, educators, and developers', *Protein Science,* 30(1), pp. 70-82.

Postle, K. and Good, R. F. (1983) 'DNA sequence of the Escherichia coli tonB gene', *Proceedings of the National Academy of Sciences of the United States of America,* 80(17), pp. 5235-5239.

Postle, K., Kastead, K. A., Gresock, M. G., Ghosh, J. and Swayne, C. D. (2010) 'The TonB dimeric crystal structures do not exist in vivo', *mBio,* 1(5), pp. 307-317.

Postle, K. and Skare, J. T. (1988) 'Escherichia coli TonB protein is exported from the cytoplasm without proteolytic cleavage of its amino terminus', *Journal of Biological Chemistry,* 263(22), pp. 11000-11007.

Prakaash, D., Cook, G. P., Acuto, O. and Kalli, A. C. (2021) 'Multi-scale simulations of the T cell receptor reveal its lipid interactions, dynamics and the arrangement of its cytoplasmic region', *PLOS Computational Biology,* 17(7), pp. e1009232.

Pravda, L., Sehnal, D., Toušek, D., Navrátilová, V., Bazgier, V., Berka, K., Svobodová Vařeková, R., Koča, J. and Otyepka, M. (2018) 'MOLEonline: a web-based tool

for analyzing channels, tunnels and pores (2018 update)', *Nucleic Acids Research,* 46(W1), pp. W368-W373.

Punjani, A., Rubinstein, J. L., Fleet, D. J. and Brubaker, M. A. (2017) 'cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination', *Nature Methods,* 14(3), pp. 290-296.

Punjani, A., Zhang, H. and Fleet, D. J. (2020) 'Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction', *Nature Methods,* 17(12), pp. 1214-1221.

Raetz, C. R. and Dowhan, W. (1990) 'Biosynthesis and function of phospholipids in Escherichia coli', *Journal of Biological Chemistry,* 265(3), pp. 1235-1238.

Raetz, C. R. H. (1990) 'Biochemistry of Endotoxins', *Annual Review of Biochemistry,* 59(1), pp. 129-170.

Raetz, C. R. H. and Whitfield, C. (2002) 'Lipopolysaccharide Endotoxins', *Annual Review of Biochemistry,* 71(1), pp. 635-700.

Ranjith, K., Ramchiary, J., Prakash, J. S. S., Arunasri, K., Sharma, S. and Shivaji, S. (2019) 'Gene targets in ocular pathogenic Escherichia coli for mitigation of biofilm formation to overcome antibiotic resistance', *Frontiers in Microbiology,* 10, pp. 1308-1308.

Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M., Garrod, T. J., Sansom, M. S. P., Piehler, J., Baumann, C. G. and Kleanthous, C. (2015) 'Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria', *Nature,* 523(7560), pp. 333-336.

Rassam, P., Long, K. R., Kaminska, R., Williams, D. J., Papadakos, G., Baumann, C. G. and Kleanthous, C. (2018) 'Intermembrane crosstalk drives inner-membrane protein organization in Escherichia coli', *Nature Communications,* 9(1), pp. 1082-1082.

Ratliff, A. C., Buchanan, S. K. and Celia, H. (2022) 'The Ton Motor', *Frontiers in Microbiology,* 0, pp. 1240-1240.

Reeves, P. (1966) 'Mutants resistant to colicin CA42-E: cross resistance and genetic mapping of a special class of mutants', *Australian Journal of Experimental Biology and Medical Science,* 44, pp. 301-316.

Reid, S. W., Leake, M. C., Chandler, J. H., Lo, C.-J., Armitage, J. P. and Berry, R. M. (2006) 'The maximum number of torque-generating units in the flagellar motor of Escherichia coli is at least 11', *Proceedings of the National Academy of Sciences,* 103(21), pp. 8066-8071.

Ridley, H. and Lakey, J. H. (2015) 'Antibacterial toxin colicin N and phage protein G3p compete with TolB for a binding site on TolA', *Microbiology,* 161(3), pp. 503-515.

Rojas, E. R., Billings, G., Odermatt, P. D., Auer, G. K., Zhu, L., Miguel, A., Chang, F., Weibel, D. B., Theriot, J. A. and Huang, K. C. (2018) 'The outer membrane is an essential load-bearing element in Gram-negative bacteria', *Nature,* 559(7715), pp. 617-621.

Roujeinikova, A. (2008) 'Crystal structure of the cell wall anchor domain of MotB, a stator component of the bacterial flagellar motor: Implications for peptidoglycan recognition',

Proceedings of the National Academy of Sciences of the United States of America, 105(30), pp. 10348-10353.

Rubino, F. A., Mollo, A., Kumar, S., Butler, E. K., Ruiz, N., Walker, S. and Kahne, D. E. (2020) 'Detection of Transport Intermediates in the Peptidoglycan Flippase MurJ Identifies Residues Essential for Conformational Cycling', *Journal of the American Chemical Society,* 142(12), pp. 5482-5486.

Sankaran, K. and Wu, H. C. (1994) 'Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol', *Journal of Biological Chemistry,* 269(31), pp. 19701-19706.

Santiveri, M., Roa-Eguiara, A., Kühne, C., Wadhwa, N., Hu, H., Berg, H. C., Erhardt, M. and Taylor, N. M. I. (2020) 'Structure and Function of Stator Units of the Bacterial Flagellar Motor', *Cell,* 183(1), pp. 244-257.e16.

Sauter, A., Howard, S. P. and Braun, V. (2003) 'In vivo evidence for TonB dimerization', *Journal of Bacteriology,* 185(19), pp. 5747-5754.

Schendel, S. L., Click, E. M., Webster, R. E. and Cramer, W. A. (1997) 'The TolA protein interacts with colicin E1 differently than with other group A colicins', *Journal of Bacteriology,* 179(11), pp. 3683-3690.

Shrivastava, R., Jiang, X. E. and Chng, S. S. (2017) 'Outer membrane lipid homeostasis via retrograde phospholipid transport in Escherichia coli', *Molecular Microbiology,* 106(3), pp. 395-408.

Shultis, D. D., Purdy, M. D., Banchs, C. N. and Wiener, M. C. (2006) 'Outer membrane active transport: Structure of the BtuB:TonB complex', *Science,* 312(5778), pp. 1396-1399.

Silhavy, T. J., Kahne, D. and Walker, S. (2010) 'The Bacterial Cell Envelope', *Cold Spring Harbor Perspectives in Biology,* 2(5).

Steimle, A., Autenrieth, I. B. and Frick, J. S. (2016) 'Structure and function: Lipid A modifications in commensals and pathogens', *International Journal of Medical Microbiology,* 306(5), pp. 290-301.

Sturgis, J. N. (2001) 'Organisation and evolution of the tol-pal gene cluster', *Journal of Molecular Microbiology and Biotechnology,* 3(1), pp. 113-22.

Sun, T. P. and Webster, R. E. (1987) 'Nucleotide sequence of a gene cluster involved in entry of E colicins and single-stranded DNA of infecting filamentous bacteriophages into Escherichia coli', *Journal of Bacteriology,* 169(6), pp. 2667-2674.

Sverzhinsky, A., Chung, J. W., Deme, J. C., Fabre, L., Levey, K. T., Plesa, M., Carter, D. M., Lypaczewski, P. and Coulton, J. W. (2015) 'Membrane Protein Complex ExbB4-ExbD1-TonB1 from Escherichia coli Demonstrates Conformational Plasticity', *Journal of Bacteriology,* 197(11), pp. 1873-1873.

Szczepaniak, J., Holmes, P., Rajasekar, K., Kaminska, R., Samsudin, F., Inns, P. G., Rassam, P., Khalid, S., Murray, S. M., Redfield, C. and Kleanthous, C. (2020) 'The lipoprotein Pal stabilises the bacterial outer membrane during constriction by a mobilisation-and-capture mechanism', *Nature Communications,* 11(1), pp. 1-14.

Szczepaniak, J., Press, C. and Kleanthous, C. (2020) 'The multifarious roles of Tol-Pal in Gram-negative bacteria', *FEMS Microbiology Reviews,* 44, pp. 490-506.

Traub, I., Gaisser, S. and Braun, V. (1993) 'Activity domains of the TonB protein', *Molecular Microbiology,* 8(2), pp. 409-423.

Turner, R. D., Hurd, A. F., Cadby, A., Hobbs, J. K. and Foster, S. J. (2013) 'Cell wall elongation mode in Gram-negative bacteria is determined by peptidoglycan architecture', *Nature Communications,* 4(1), pp. 1-8.

Urner, L. H., Liko, I., Yen, H. Y., Hoi, K. K., Bolla, J. R., Gault, J., Almeida, F. G., Schweder, M. P., Shutin, D., Ehrmann, S., Haag, R., Robinson, C. V. and Pagel, K. (2020) 'Modular detergents tailor the purification and structural analysis of membrane proteins including G-protein coupled receptors', *Nature Communications,* 11(1), pp. 1-10.

Vankemmelbeke, M., Zhang, Y., Moore, G. R., Kleanthous, C., Penfold, C. N. and James, R. (2009) 'Energy-dependent Immunity Protein Release during tol-dependent Nuclease Colicin Translocation', *Journal of Biological Chemistry,* 284(28), pp. 18932-18941.

Vianney, A., Lewin, T. M., Beyer, W. F., Lazzaroni, J. C., Portalier, R. and Webster, R. E. (1994) 'Membrane topology and mutational analysis of the TolQ protein of Escherichia coli required for the uptake of macromolecules and cell envelope integrity', *Journal of Bacteriology,* 176(3), pp. 822-829.

Vianney, A., Muller, M. M., Clavel, T., Lazzaroni, J. C., Portalier, R. and Webster, R. E. (1996) 'Characterization of the tol-pal region of Escherichia coli K-12: translational control of tolR expression by TolQ and identification of a new open reading frame downstream of pal encoding a periplasmic protein', *Journal of Bacteriology,* 178(14), pp. 4031-4038.

Vollmer, W., Blanot, D. and De Pedro, M. A. (2008) 'Peptidoglycan structure and architecture', *FEMS Microbiology Reviews,* 32(2), pp. 149-167.

Vollmer, W. and Seligman, S. J. (2010) 'Architecture of peptidoglycan: more data and more models', *Trends in Microbiology,* 18(2), pp. 59-66.

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J. and Laue, E. D. (2005) 'The CCPN data model for NMR spectroscopy: development of a software pipeline', *Proteins,* 59(4), pp. 687-96.

Walburger, A., Lazdunski, C. and Corda, Y. (2002) 'The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB', *Molecular Microbiology,* 44(3), pp. 695-708.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R. and Schwede, T. (2018) 'SWISS-MODEL: homology modelling of protein structures and complexes', *Nucleic Acids Research,* 46(W1), pp. W296-W303.

Webby, M. N., Oluwole, A. O., Pedebos, C., Inns, P. G., Olerinyova, A., Prakaash, D., Housden, N. G., Benn, G., Sun, D., Hoogenboom, B. W., Kukura, P., Mohammed, S., Robinson, C. V., Khalid, S. and Kleanthous, C. (2022a) 'Lipids mediate supramolecular outer membrane protein assembly in bacteria', *Scientific Advances,* 8(44), pp. eadc9566.

Webby, M. N., Williams-Jones, D. P., Press, C. and Kleanthous, C. (2022b) 'Force-Generation by the Trans-Envelope Tol-Pal System', *Frontiers in Microbiology,* 13, pp. 347-347.

Weigand, R. A. and Rothfield, L. I. (1976) 'Genetic and physiological classification of periplasmic-leaky mutants of Salmonella typhimurium', *Journal of Bacteriology,* 125(1), pp. 340-340.

White, P. (2017) *Bacterial protein import mediated by an iron transporter.* DPhil thesis, University of Oxford, Oxford.

White, P., Joshi, A., Rassam, P., Housden, N. G., Kaminska, R., Goult, J. D., Redfield, C., McCaughey, L. C., Walker, D., Mohammed, S. and Kleanthous, C. (2017) 'Exploitation of an iron transporter for bacterial protein antibiotic import', *Proceedings of the National Academy of Sciences of the United States of America,* 114(45), pp. 12051-12056.

Williams, N., Ackerman, S. H. and Coleman, P. S. (1986) 'Benzophenone-ATP: a photoaffinity label for the active site of ATPases', *Methods in Enzymology,* 126, pp. 667-82.

Witty, M., Sanz, C., Shah, A., Grossmann, J. G., Mizuquchi, K., Perham, R. N. and Luisi, B. (2002) 'Structure of the periplasmic domain of Pseudomonas aeruginosa TolA: Evidence for an evolutionary relationship with the TonB transporter protein', *EMBO Journal,* 21(16), pp. 4207-4218.

Wojdyla, J. A., Cutts, E., Kaminska, R., Papadakos, G., Hopper, J. T. S., Stansfeld, P. J., Staunton, D., Robinson, C. V. and Kleanthous, C. (2015) 'Structure and function of the Escherichia coli Tol-Pal stator protein TolR', *The Journal of Biological Chemistry,* 290(44), pp. 26675-26687.

Yakhnina, A. A. and Bernhardt, T. G. (2020) 'The Tol-Pal system is required for peptidoglycan-cleaving enzymes to complete bacterial cell division', *Proceedings of the National Academy of Sciences of the United States of America,* 117(12), pp. 6777-6783.

Yeh, Y. C., Comolli, L. R., Downing, K. H., Shapiro, L. and McAdams, H. H. (2010) 'The Caulobacter Tol-Pal Complex Is Essential for Outer Membrane Integrity and the Positioning of a Polar Localization Factor', *Journal of Bacteriology,* 192(19), pp. 4847-4847.

Yonekura, K., Maki-Yonekura, S. and Homma, M. (2011) 'Structure of the flagellar motor protein complex pomAB: Implications for the torque-generating conformation', *Journal of Bacteriology,* 193(15), pp. 3863-3870.

Zhang, X. Y. Z., Goemaere, E. L., Seddiki, N., Célia, H., Gavioli, M., Cascales, E. and Lloubes, R. (2011) 'Mapping the interactions between Escherichia coli TolQ transmembrane segments', *Journal of Biological Chemistry,* 286(13), pp. 11756-11764.

Zhang, X. Y. Z., Goemaere, E. L., Thomé, R., Gavioli, M., Cascales, E. and Lloubés, R. (2009) 'Mapping the Interactions between Escherichia coli Tol Subunits', *Journal of Biological Chemistry,* 284(7), pp. 4275-4282.

Zhou, J., Lloyd, S. A. and Blair, D. F. (1998) 'Electrostatic interactions between rotor and stator in the bacterial flagellar motor', *Proceedings of the National Academy of Sciences of the United States of America,* 95(11), pp. 6436-6441.

Zhuang, Z., Song, F., Martin, B. M. and Dunaway-Mariano, D. (2002) 'The YbgC protein encoded by the ybgC gene of the tol-pal gene cluster of Haemophilus influenzae catalyzes acyl-coenzyme A thioester hydrolysis', *FEBS Letters,* 516(1-3), pp. 161-163.