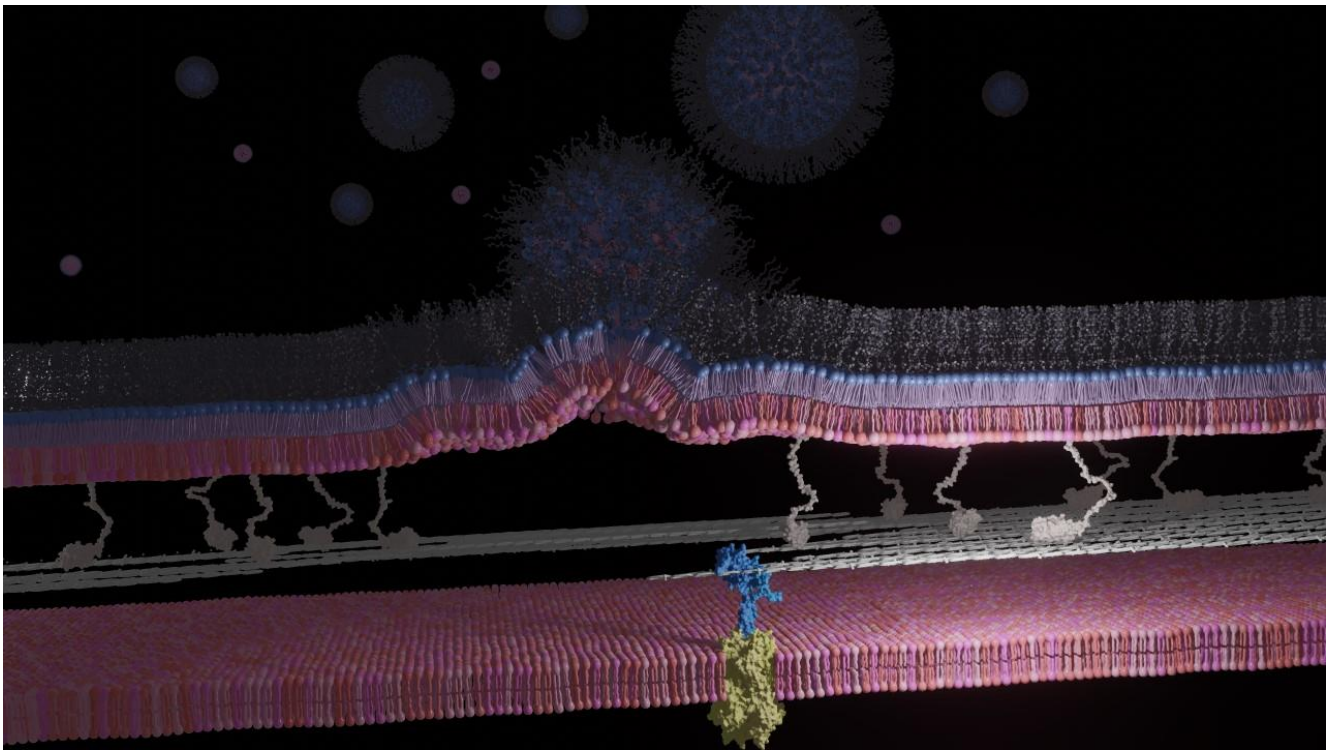
Force Transduction Through The Bacterial Cell Envelope

Daniel Williams-Jones

University of Oxford

2024



Contents

Force Transduction Through The Bacterial Cell Envelope 1
Figures 6
Acknowledgements 9
Abbreviations 10
Abstract 12
Declaration 13
1.0 General Introduction 14
1.1 The Cell Envelope of *Escherichia coli* 15
1.1.1 The outer membrane 17
1.1.2 Peptidoglycan 18
1.1.3 The inner membrane 20
1.1.4 PMF-energised IM complexes 21
1.1.5 ExbBD-TonB drives nutrient import 25
1.2 The Tol-Pal system 26
1.2.1 Tol-Pal proteins 26
1.2.2 $TolQ_5R_2$-TolA stabilises the OM at the septum 30
1.3 Aims of This Work 32
2.0 Materials and Methods 33
2.1 Microbiology 33
2.1.1 Strains, media, and growth conditions 34
2.1.2 Plasmids 35
2.2 Molecular Biology 37
2.2.1 Molecular cloning 37
2.2.2 Chemically competent cell preparation 37
2.2.3 Bacterial heat shock transformation 37
2.2.4 DNA gel electrophoresis 38
2.2.5 Plasmid preparation and sequencing of DNA 38
2.3 Protein expression and purification 39
2.3.1 Purification of $ColN_{1-162}$-Cys and $AlexFluor_{647}$ conjugation 39
2.3.2 TolQR purification optimisation 39
2.3.3 Purifying and UV crosslinking TolQ-TolR BPA mutants 40
2.3.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 40
2.3.5 In-gel digestion of UV-crosslinked TolQR 41
2.3.6 Mapping TolQR BPA crosslinks 41
2.3.7 Protein quantification 41
2.4 Electron microscopy 42

2.4.1 Sample preparation 42
2.4.2 Grid preparation 42
2.4.3 Imaging conditions 42
2.4.4 Data processing 42
2.5 Fluorescence microscopy 43
2.5.1 Sample preparation 43
2.5.2 Imaging conditions 43
2.5.3 Image analysis 43
2.5.4 Colicin-fusion protein (FP) labelling 44
2.5.5 ColN-AlexaFluor (AF) import assay 44
2.6 Tol-Pal functional assays 45
2.6.1 Colicin toxicity assay 45
2.6.2 Time resolved colicin CFU assay 45
2.6.3 Detergent sensitivity assays 46
2.6.4 Western blot 46
2.7 Computational analyses 47
2.7.1 Structural predictions 47
2.7.2 Tol-Pal/TonB locus search (manual) 47
2.7.3 Tol-Pal/TonB locus search (algorithm) 47
2.7.4 Co-conservation analyses 48
2.7.5 α-Helix:Proline ratio calculations 48
3.0 Cellular assays of TolA-TonB chimeras 49
3.1 Introduction 49
3.1.1 The Tol and Ton systems are structurally homologous 50
3.2 Results 53
3.2.1 TolA and TonB chimeras transduce force 53
3.2.2 TolA chimeras demonstrate OM stability is a quantitative trait 56
3.2.3 TolA chimeras reveal a spectrum of septal Pal localisation 58
3.2.4 TolAII secondary structure influences colicin uptake rate 61
3.2.5 Group B constructs complement *ΔtonB* growth rates 62
3.3 Discussion 63
3.3.1 Domain II structures influence force transduction efficiency 63
3.3.2 The Tol-Pal system stabilises the septal OM 65
3.3.3 The TolA-TolB interaction may comprise a catch-bond 67
3.3.4 A “twist-to-writhe” model of tensile force generation 68
4.0 The origin, organisation and distribution of the Tol-Pal system 72
4.1 Introduction 72

4.1.1 Annotation of *tolA* requires proximal *tolB* and/or *pal genes* 72
4.2 Results **74**
4.2.1 Tol-Pal systems are conserved throughout Proteobacteria 74
4.2.2 Tol-Pal systems are conserved in Gracilicutes 79
4.2.3 Distribution of Tol/Ton systems suggest TolA origin 81
4.3 Discussion **84**
4.3.1 The inferred origin of TolA from a TonB paralogue 84
5.0 PMF-dependent rearrangements of TolQ-TolR **89**
5.1 Introduction **89**
5.1.1 Identifying TolQR contacts using chemical crosslinking 89
5.2 Results **91**
5.2.1 Co-conservation analyses of TolQ-TolR interface 91
5.2.2 The TolQR cryo-EM complex 93
5.2.3 Designing TolQR mutants for *in vivo* crosslinking 94
5.2.4 TolQR BPA mutants confer colicin sensitivity 95
5.2.5 Strep-Tactin purification yields pure TolQR 97
5.2.6 UV crosslinking solubilised TolQR BPA mutants *in vitro* 99
5.2.7 UV crosslinking solubilised TolQR mutants *in vivo* 100
5.2.8 Optimising proteolytic TolQR digests for LC-MS-MS 102
5.2.9 TolQR I39X forms intermolecular crosslinks 103
5.2.10 Chemical de-energisation alters UV-crosslink formation 104
5.3 Discussion **105**
5.3.1 Higher-order protein species suggest trimer formation 105
6.0 Interactions of TolQR with TolA **107**
6.1 Introduction **107**
6.1.1 PMF-dependent interactions of TolA with TolQR 107
6.2 Results **108**
6.2.1 Modelling the TolQ-TolA interaction 108
6.2.2 Conserved TolA TMH His_{22} drives TolA-TolQ interaction 109
6.2.3 A novel motif for transducer-stator interaction 110
6.2.4 The motor-box is sequence-agnostic but sensitive to mutation 115
6.2.5 Modelling suggests the motor-box drives TolR rearrangement 118
6.3 Discussion **120**
7.0 Cryo-ET of the energised TolQRA assembly **122**
7.1 Introduction **122**
7.1.1 TolQRA energisation may be visualised *in situ* by cryo-ET 122
7.2 Results **123**

7.2.1 *E. coli* strain PB114 is sensitive to Colicins N & E9 123
7.2.2 ColN-GFP labels minicells 125
7.2.3 ColN-mCherry traps TolA-GFP *in situ* 126
7.2.4 Fluorescently labelled ColN-AF_{647} is translocation competent 127
7.2.5 $ColN_{1-162}$ uptake is Δψ dependent 130
7.2.6 Large scale minicell ColN-GFP labelling 131
7.2.7 Gold nanoparticles are used as a secondary label for cryo-ET 132
7.2.8 AuNPs label cells heterogeneously 134
7.2.9 SPOTs are a superior secondary label to AuNPs 135
7.2.10 SPOTs improve cryo-EM resolution over AuNPs 137
7.2.11 Labelling SPOTs in two steps improves efficiency 138
7.2.11 Optimised ColN-GFP-SPOT labelled samples 140
7.2.12 SPOT-labelled minicell tilt series collection 141
7.3 Discussion 142
8.0 Concluding remarks 143
References 146

Figures

1.1 | The *E. coli* cell envelope is a crowded and diverse environment.

1.2 | The *E. coli* cell surface.

1.3 | The PG network is organised circumferentially.

1.4 | $MotA_5B_2$ is the archetypal 5:2 motor that drives the BFM complex.

1.5 | Essential threonine ring and protonatable aspartate residues are proposed to drive rotation of 5:2 motors.

1.6 | Rotation of 5:2 motors likely proceeds through proton half-channels.

1.7 | The Ton system drives uptake of extracellular substrates.

1.8 | TolB binds and mobilises Pal.

1.9 | TolA dissociates the TolB-Pal complex by applying tensile force.

1.10 | $TolQ_5R_2$ forms a 165 kDa heptamer.

1.11 | Overview of *E. coli* OM-PG tethering systems

1.12 | The Tol-Pal system uses a mobilisation-and-capture mechanism to accumulate Pal at the division site.

3.1 | Homologous aspects of the Tol and Ton systems suggest similar force transduction mechanisms.

3.2 | Structural divergence of *E. coli* TonB and TolA periplasm-spanning domains

3.3 | Comparison of α-helix and PPII-helix properties.

3.4 | TolA-TonB chimeric mutants categorised by CTD.

3.5 | Colicin killing assays demonstrate force transduction ability of TolA/TonB chimeras.

3.6 | Group A constructs show intermediate SDS tolerance.

3.7 | Relative accumulation of Pal-mCherry at division septa varies between group A mutants.

3.8 | Quantification of group A mutant localisation in *ΔtolA* cells.

3.9 | TolAII secondary structure influences colicin uptake rates.

3.10 | Growth curves of *E. coli ΔtonB* cells expressing group B mutants suggest secondary structures within domain II restore efficient TBDT-driven substrate uptake.

3.11 | Structural models of TolA, ABA and AIA.

3.12 | TolA efficiency drives colicin import, septal Pal accumulation, and OM stability.

3.13 | A catch-bond model of TolB-Pal dissociation.

3.14 | Energised extension model of Tol-Pal force transduction.

3.15 | A twist-to-writhe model of tensile force generation.

3.16 | Possible tensioned domain II conformations may explain differences in mutant efficiency.

4.1 | The *E. coli tol-pal* operon.

4.2 | *tol-pal* gene organisations and α-helix: proline ratios from representative proteobacterial species.

4.3 | Synteny of proteobacterial *tol-pal* flanking genes.

4.4 | TolA and TonB domain II length and secondary structure compositions both vary amongst Proteobacteria.

4.5 | Quantification of TolAII/TonBII structural diversity.

4.6 | Validated *tol-pal* loci from Gracilicute taxa.

4.7 | TolAII structural composition within Gracilicute species reveals resemblances to proteobacterial TonBII.

4.8 | Distribution of Tol and Ton systems suggest the Ton system emerged first.

4.9 | Unrooted FoldTree phylogenetic tree of TolA/TonB AF prediction models reveals TolA proteins comprise a separate clade from most TonB proteins.

4.10 | A hypothetical route for TonB-to-TolA speciation.

5.1 | The “Open” and “Strand-swapped” conformations of TolR PGB dimers.

5.2 | Mapping the TolQR interaction interface.

5.3 | The TolQR complex modelled from cryo-EM data.

5.4 | TolR mutants designed to incorporate BPA residues within the TolQ lumen.

5.5 | Expression of TolQR BPA mutants elicits ColE9 sensitivity.

5.6 | Nickel-affinity purification of $TolQR_{L29X}$ yields insufficient purity for crosslink detection.

5.7 | UV-treating the $TolQR_{L12X}$ mutant in solubilised membranes enriches high-weight protein species.

5.8 | $TolQR_{L12X}$ exhibits *in vivo* UV crosslinking.

5.9 | TolQR BPA UV-treated *in vivo* exhibit crosslinking.

5.10 | Chymotrypsin-Trypsin double digest fails to resolve the TolR N-terminus.

5.11 | $TolR_{I39X}$ subunits forms crosslinks within the TolQR complex between adjacent $TolR_{I39X}$ sidechains.

5.12 | Chemical depolarisation alters crosslink profiles of TolQR BPA mutants.

5.13 | The PG-binding cleft of Pal, MotB, and PorE may be present in TolR.

6.1 | Co-conservation analysis of the putative TolQ-TolA binding site.

6.2 | GFP-TolAI is recruited to the septum by interaction with TolQR, driven by $TolA_{H22}$.

6.3 | TolR-TolA interactions predicted in periplasmic region between TolR and TolA residues 45-60.

6.4 | Co-conservation maxima correlate with PsiPred 2.0 β-strand prediction sites.

6.5 | Co-conservation scores of *H. influenzae* TolR-TolA mapped to respective Robetta prediction models.

6.6 | The *E. coli* TolA IDR is predicted by AF2 to interact with TolR via β-addition.

6.7 | Motor-box predictions suggest formation of a continuous, extended β-sheet

6.8 | AF3 prediction of Ton and Tol motor-transducer complexes predicts TolA/TonB motor-box threading between single-strand-swapped TolR/ExbD dimers.

6.9 | Motor-box mutants designed to perturb TolA function.

6.10 | Motor-box mutants vary in OM stabilisation abilities.

6.11 | TolA motor-box mutants retain colicin sensitivity.

7.1 | ColN uses OmpF as both receptor and translocon, whereas ColE9 uses BtuB as a receptor and OmpF as a translocon.

7.2 | Minicells can be labelled with ColN-GFP.

7.3 | GFP-TolA is immobilised by ColN-mCherry.

7.4 | $ColN_{1-186}$-cys purification and AF_{647} conjugation.

7.5 | Purified PB114 minicells labelled with ColN-AF.

7.6 | Trypsin assays suggest that ColN-AF is internalised into minicells while ColN-GFP constructs are not.

7.7 | Valinomycin reduces ColN-AF import into PB114 cells.

7.8 | ColN-GFP labelling minicells after upscaling growth in LB.

7.9 | AuNP labelling does not remove ColN-GFP from minicell surface.

7.10 | AuNP labels on minicells exhibit heterogeneity.

7.11 | Cryo-TEM shows excess BSA slightly improves AuNP specificity, but contrast artefacts remain.

7.12 | Expected compliment of proteins labelled by SPOTs.

7.13 | AuNPs impart resolution-limiting contrast artefacts in cryo-TEM.

7.14 | Validation of SPOT-labelling in minicells by cryo-TEM.

7.15 | ColN-GFP-labelling is inefficient at low temperatures.

7.16 | Overnight labelling with ColN-GFP in TAE followed by SPOTs yields well-labelled but plasmolysed cells.

7.17 | Optimised two-step labelling yields a more ideal SPOT concentration.

7.18 | Tomogram of *E. coli* minicell labelled with colicin-GFP and SPOTs.

Acknowledgements

I owe a great deal of thanks to many individuals. First and foremost, I am deeply grateful to Professor Colin Kleanthous for granting me this incredible opportunity and for his invaluable guidance and unwavering support throughout. I am also profoundly appreciative of the Wellcome Trust for their generous funding and continued support, as well as the Cellular Structural Biology programme for presenting me with this life-changing experience.

To the Kleanthous lab, my heartfelt thanks. In particular, I would like to acknowledge Dr. Melissa Webby for supervising my project, for her infectious enthusiasm, mentorship, advice, scientific discussion and contribution, and for always helping me up during the tough times. I am also extremely thankful to Dr. Joanna Szczepaniak, whose guidance was instrumental to my project, and whose passion for science motivated me to pursue curiousity and love my work. I extend my sincere thanks to Dr. Nick Housden and Dr. Renata Kaminska for their practical advice and for running a tight ship. Special thanks go to past and present lab members—Gideon, Vivian, Shreya, Abbie, Haowei, Arc, and Veronika—for making the lab a delightful place to work. Special shout out to my deskmate Angela, for her deadpan humour and wry wit, and for always bringing the best snacks. I am also indebted for the deep discussions over one too many beers (*or ciders, oops!*) with Dr Jonathon Goult, though it is probably best that we have wrapped those up! I am deeply grateful to Dr. Patrick Inns for helping me find my feet in the lab, for entertaining fun hypotheticals, and helping me get to grips with coding. Similarly, thanks to Dr. Cara Press for her advice throughout my PhD, and for making me feel at home in the lab. Additionally, I would like to express my gratitude to those outside the lab who have offered me scientific support, particularly Prof. Lindsay Baker and Dr. Emma Silvester, for their tutelage in cryo-ET and for providing SPOTs that were essential to this project. Many thanks also to Basile Beaud for his guidance in all things phylogenetics.

I am incredibly lucky for my friends, who made my time at Oxford unforgettable. Thank you, Jack Whitehead, for your solidarity, advice, encouragement, fitness motivation, and for always being an excellent listener. I would also like to thank my friend Luis “Chuy” Orenday Tapia for always making being there to celebrate and commiserate at a moment’s notice. My housemates—Dr. Charlie Cook, William Rochira, and Bjarne Feddersen—deserve special mention for always making home a welcoming, joyful place, and for pushing me out of my comfort zone, both socially and with our fun adventures. I look forward to catching up over a beer or two for many years to come.

I cannot forget my family, starting with my brother, Dr. Robert Smyth, whose unwavering support has been a constant throughout my life, and I trust will continue for many years to come (*despite the Gangbeasts incident*!). Thank you, Sara, for welcoming me into your family, and to Jack, Will, Jess, and Sam for being such wonderful siblings. Lastly, to my parents, Susan and Rhys, I am grateful for the strength and resilience you gave me, and I hope you are both resting in peace.

Finally, I would like to thank my partner, Julia, for everything. I could not have done this without you.

Abbreviations

AF	AlexaFluor
ATP	Adenosine triphosphate
AU	Arbitrary units
CCCP	Carbonyl cyanide m-chlorophenyl hydrazine
ColB	Colicin B
ColE9	Colicin E9
ColIA	Colicin Ia
ColN	Colicin N
Cryo-EM	Electron cryo-microscopy
Cryo-ET	Electron cryo-tomography
DNA	Deoxyribonucleic acid
FA	Formaldehyde
FP	Fluorescent protein
FSC	Fourier shell correlation
GFP	Green fluorescent protein
GlcNAc	N-Acetyl-Glucosamine
GNB	Gram negative bacteria
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
IM	Inner membrane
IMP	Inner membrane protein
IMAC	Immobilised metal affinity chromatography
kDa	KiloDalton
Kdo	2-keto-3-deoxy-*d*-octanoate
LB	Lysogeny broth
LPS	Lipopolysaccharide
Lpp	Braun's lipoprotein
mDa	MegaDalton
mDAP	*meso*-diaminopimelate
MS	Mass spectrometry
MurNAc	N-Acetylmuramic acid
NS-EM	Negative stain electron microscopy
OM	Outer membrane
OMP	Outer membrane protein
ONI	Oxford nanoimager
ORF	Open reading frame

Pal	Peptidoglycan associated lipoprotein
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PG	Peptidoglycan
PGB	PG-binding (domain)
PMF	Proton motive force
PSF	Point spread function
SDS	Sodium dodecyl sulfate
SE	Standard error
STA	Subtomogram averaging
sPG	Septal peptidoglycan
SPT	Single particle tracking
TBDT	TonB-dependent transporter
TEM	Transmission electron microscopy
TM	Transmembrane
TMH	Transmembrane helix
TPH	Transpore helix
TPR	Tetratricopeptide repeat
WT	Wild type

Abstract

The Tol-Pal system is a conserved machine in double-membraned bacteria that stabilises the Gram-negative bacterial outer membrane. In this thesis, I aim to elucidate the mechanism by which the Tol-Pal system dynamically stabilises this important protective layer. I examine how assembly of the *trans*-periplasmic motor-transducer complex recruits a membrane stabilising lipoprotein, Pal, to the division site. To contextualise the experimental aspects of this thesis, I first examine the physiological role of the Tol-Pal system, and its cellular function in *E. coli*. Through use of mutagenesis, fluorescence microscopy, and live cell assays, I demonstrate that system function can be modulated by altering the secondary structure of the periplasm spanning transducer protein, TolA. I also examine the cell-wide implications of reducing TolA function in such a manner, with respect to membrane stability and cell division (chapter three). Subsequent bioinformatic investigations into the phylogeny of the Tol-Pal system reveal its unexpectedly broad distribution throughout the kingdom of bacteria and hint at its evolutionary origin (chapter four). Further molecular insights were gleaned through *in vivo* crosslinking, where I probe the conformational rearrangements that occur within the TolQR motor in response to proton flux (chapter five). Additional modelling and bioinformatic data reveal the putative TolA binding site on the TolQR motor, and I demonstrate the biological relevance of a previously unknown interaction motif involved in TolA-TolR interactions and force transduction (chapter six). Finally, I describe the optimisation of electron cryotomography (cryo-ET) to structurally visualise the TolQR-TolA assembly in an energetically stalled state (chapter seven). Overall, this work reveals the function, distribution, conservation and organisation of bacterial Tol-Pal systems and their likely evolutionary origin. I also propose a model for the mechanism by which these protein machines use proton flux to stabilise the outer membrane in *E. coli.*

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.

Daniel Pryse Williams-Jones,

October 2024

Word count excluding references: 30,972

1.0 General Introduction

Life fundamentally comprises cells, the boundaries of which delineate and separate the “self” from the environment. This compartmentalisation requires synthesis and maintenance of lipid membranes, a ubiquitous feature of all living organisms. Cell membranes invariably fulfil two roles: i) sequestration of biochemical activity, and ii) protection/separation from the extracellular environment. Bacteria are routinely categorised by their envelope structure into *monoderms* (“single skin”) and *diderms* (“double-skinned”), denoting the presence of either one or two cell membranes. In diderms, these are termed the inner and outer membrane (IM and OM, respectively). Of the 30 identified bacterial phylae, 27 are didermal, reflecting the extraordinary resilience of the double-membraned cell envelope [1, 2]. The first identified diderms were described as “Gram-negative” bacteria (GNB) by the microbiologist Hans Christian Gram, who found that while they can be stained with Crystal Violet, GNB are quickly destained in alcohol, unlike Gram-positive bacteria which remain stained due to their thicker cell wall [3]. However, as is often true of biological classifications, not all diderms are GNB, and vice versa owing to species-specific idiosyncrasies. A small number of GNB constitute a large portion of AMR-associated infections that become particularly dangerous during prolonged hospitalisation. In this work, I predominantly refer to the cell envelope architecture of Proteobacteria, with a specific focus on the model organism *Escherichia coli* [4, 5].

1.1 The Cell Envelope of *Escherichia coli*

In *E. coli*, the tripartite superstructure of the cell envelope comprises two membranes and the intervening periplasm, which contains a mesh of peptidoglycan (PG) (Fig. 1.1). The OM of such species typically comprises an asymmetric proteolipid bilayer, where the inner leaflet is primarily composed of simple phospholipids (PLs), while the outer leaflet is composed of clusters of lipopolysaccharide (LPS) molecules, intercalated between outer membrane proteins (OMPs) which are immersed in both leaflets. This cell envelope is typically around 30-50 nm thick, 20-30 nm of which comprises the periplasm [6-8]. The OM is typically tethered to the PG by OMPs and lipoproteins (*lipo*ylated proteins), an additional class of soluble periplasmic proteins with lipid anchors embedded in the inner leaflet of the OM [9].

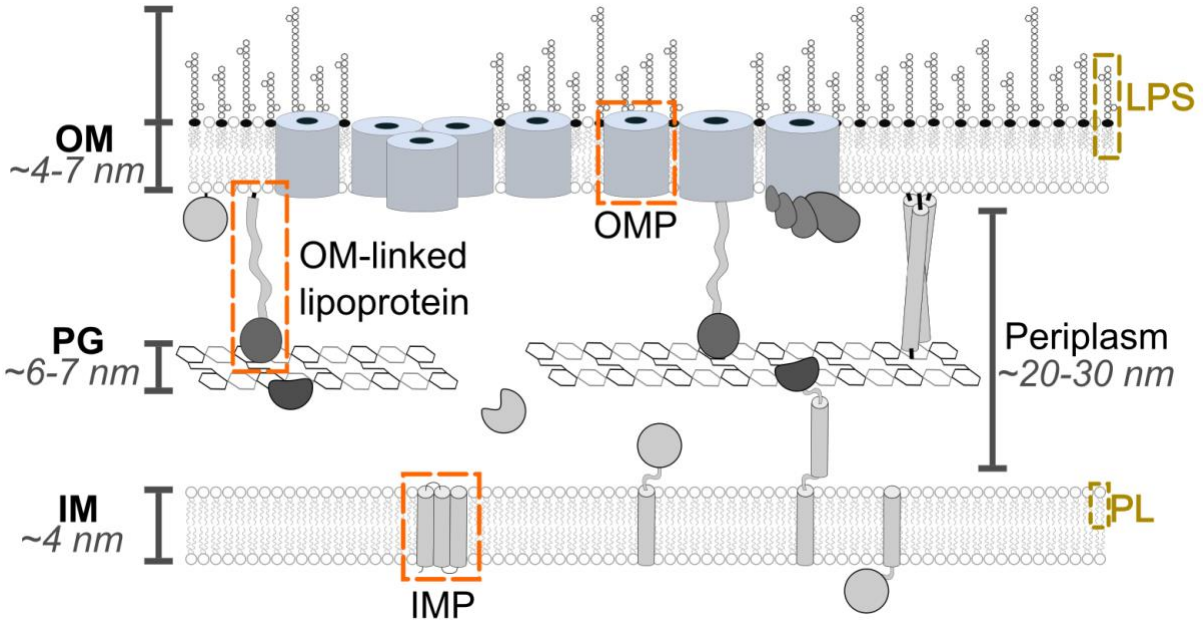


Figure 1.1 | The *E. coli* cell envelope is a crowded and diverse environment. The OM is an asymmetric proteolipid bilayer, where β-barrelled OMPs are intercalated by LPS molecules, surrounded by lipid patches. Beneath the OM is the periplasm, a crowded gel-like compartment containing many solutes. Within the periplasm are 1-3 layers of PG, which form a fibrous mesh encasing the entire cell. The IM contains predominantly PLs and α-helical integral membrane proteins (IMPs).

The Gram-negative cell envelope is a significant challenge for antimicrobial compounds, occluding a plethora of otherwise-potent antibiotics [10, 11]. Molecules which efficiently kill GNB must bypass the OM, typically by diffusion, translocation, piercing, or dissolving the OM directly. Due to its recalcitrance to antimicrobial compounds, the biogenesis and maintenance of the cell envelope is of increasing

medical relevance. The Tol-Pal system is important for maintaining the integrity of the cell envelope, particularly in Proteobacteria where the system is conserved [12]. The Tol-Pal system plays a pivotal role in the fitness of many GNB, and loss of Tol-Pal function is associated with debilitating fitness costs, including reduced growth rates, growth defects, and antibiotic sensitisation [13]. Illustrating this is the use of vancomycin in assaying *tol* mutants, where loss of function mutants exhibit increased sensitivity [14]. Fortuitously, a class of protein toxins referred to as bacteriocins, specifically utilise the Tol-Pal system to enter and kill GNB [15, 16]. This creates an opportunity: these toxins are species-selective and extremely potent, and resistance to these toxins through loss of function mutants often incurs fitness costs. These mutants often involve loss of the OM receptor for the toxin, or the OM-stabilising and/or nutrient uptake systems [16, 17]. Given this potential lifeline, further characterisation of the Tol-Pal system is prudent. Bacteriocins are also ideal tools for probing Tol-Pal functionality, as illustrated throughout this thesis.

1.1.1 The outer membrane

In most GNB, the OM is an asymmetric proteolipid bilayer, and the cell's erstwhile defence. Recent atomic force microscopy (AFM) and mass spectrometry (MS) data reveal that in *E. coli* the OM comprises an imperfect lattice of OMPs noncovalently bridged by adjacent LPS molecules, with interspersed lipid patches (Fig. 1.2) [18, 19]. The OM is estimated to be around 75% protein by surface area and 50% by dry mass, with a ratio of three LPS molecules per OMP [20, 21]. The OM is predominantly protective, acting as a physical barrier to toxic substances in the environment. OMPs typically exist as monomers, trimers and occasionally dimers in *E. coli* [22].

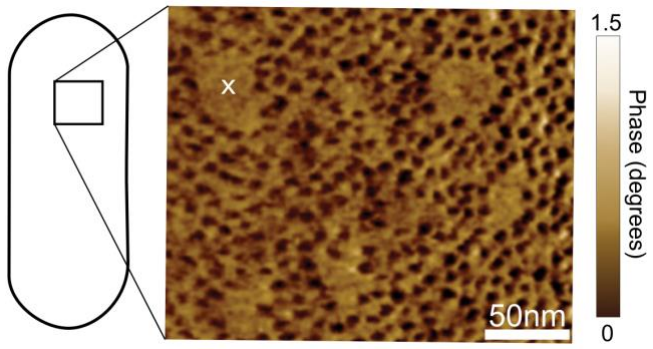


Figure 1.2 | The *E. coli* cell surface. A typical phase map of *E. coli* MG1655 cells as observed by AFM features an imperfect lattice of OMPs (seen as dark holes) and interspersed LPS patches (white "x"). "Phase" refers to measured differences in mechanical properties such as adhesion, which differ between proteins and LPS. Cartoon of a typical *E. coli* cell is shown for approximate scale. The raised lipid patch is approximately 1 nm in height. Data adapted from Benn *et al*., [18].

In *E. coli*, many OMPs are trimeric, and tessellate to form an imperfect lattice, mediated through noncovalent OMP-LPS-OMP associations [18, 19]. The intervening LPS is composed of diverse chains of covalently linked sugars, often containing charged groups such as phosphates, linked to the lipid A core. The length of LPS differs by number of repeat units at the exposed surface of the molecule (Fig. 1.1). The sugars within each repeat unit also differ, producing serotypes with varied immunogenicity to facilitate immune system evasion [23, 24].

The OM is a physical but selectively permeable barrier, through which uptake of small molecules is orchestrated and regulated by OMPs [21, 25]. Porins confer semi-permeability to the OM by allowing passive diffusion of small molecules across the OM but restricting diffusion of molecules of 550 Da or larger, such as oligosaccharides, hydrophilic sugars, hydrophobic antimicrobials and dyes [26]. High-affinity energy dependent transporters are predominantly involved in the selective uptake of substrates, in a process coupled to the IM through *trans*-envelope protein assemblies, such as the Ton system [27]. The OM also plays an important mechanical role in

resisting turgor pressure and shear forces through its tight association with the PG layer [28-30].

1.1.2 Peptidoglycan

Between the IM and OM lies the PG layer, a heteropolymeric network of polysaccharide chains. PG is a highly conserved feature of the periplasm, found even in some chloroplast lineages [31]. The PG layer is tethered to the OM, generating a robust internal structure which protects the cell from both extrinsic shear stress and intrinsic osmotic pressure [28]. PG comprises repeating chains of aminated sugars-N-Acetyl glucosamine (GlcNAc) and N-Acetyl muramic acid (MurNAc). MurNAc residues typically feature polypeptide moieties, which crosslink adjacent chains to form a non-crystalline mesh [32]. In *E. coli*, the PG is typically 1-3 layers thick [33, 34]. In Proteobacteria these chains of crosslinked sugars are generally arranged circumferentially (Fig. 1.3) [31, 35-39].

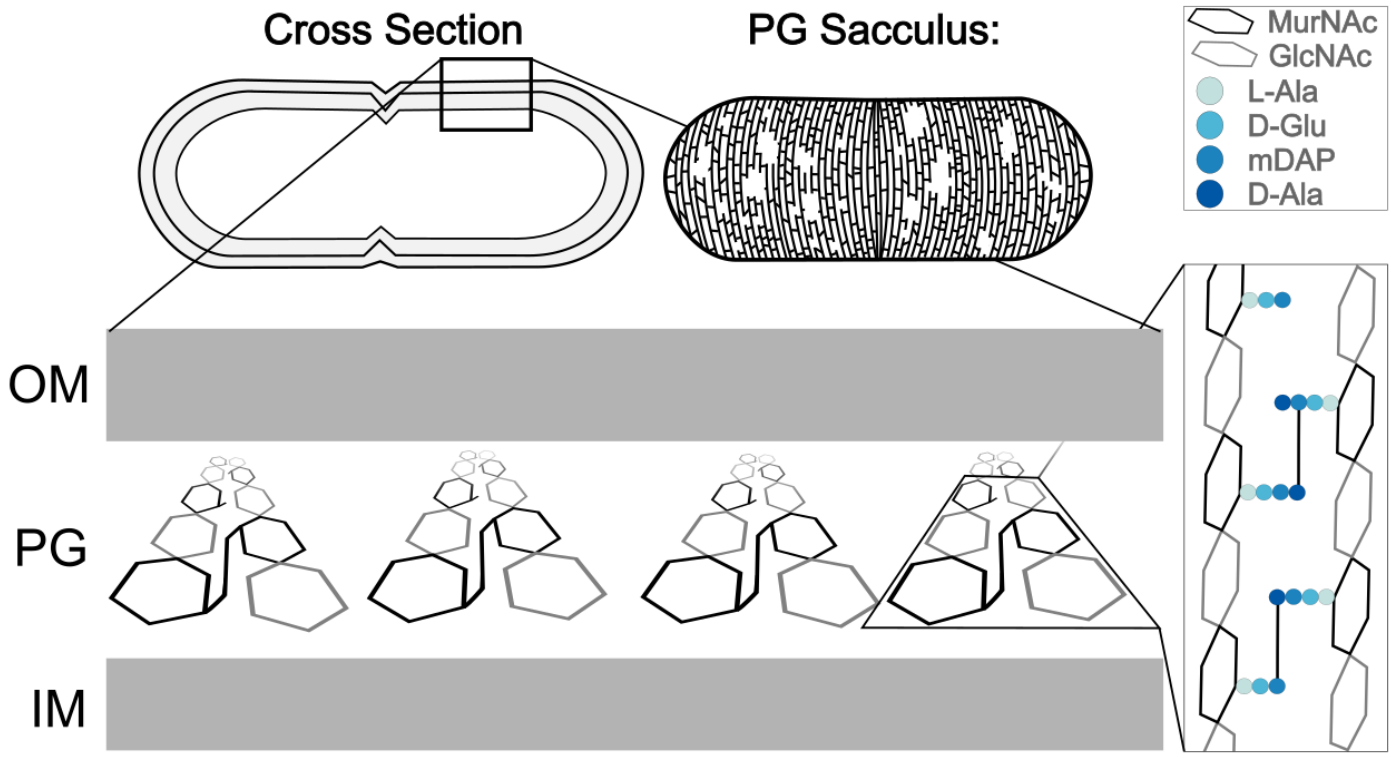


Figure 1.3 | The PG network is organised circumferentially. PG comprises linear chains of alternating GlcNAc and MurNAc residues, crosslinked through adjacent mDAP moieties. The cell sacculus has a defined grain, generated by the ring-like circumferential synthesis of PG chains at the cell septum. This cartoon attempts to illustrate the circumferential orientation of PG within the cell wall cross section, as indicated by the PG sacculus. PG model adapted from Garde *et al.*, along with the presence of PG pores [31, 40]. In all other figures a less accurate, laterally oriented PG configuration is shown for simplicity.

The ring-like orientation of rigid glycans around the cell, with polypeptide sidechains oriented along the long axis, is predicted to account for the anisotropic elasticity observed in purified sacculi [41]. This circumferential organisation has profound implications for proteins which utilise PG-binding in their localisation. It was further revealed through AFM of purified sacculi that there is a distribution of pores in the PG network, ranging from 4-20 nm in diameter [40]. It has further been suggested that the PG plays an organisational role within the periplasm by acting as a semi-permeable barrier to compartmentalise the inner and outer periplasm, in addition to its canonical mechano-structural role [42]. During cell division nascent PG is incorporated at the septum in parallel with the localised insertion of nascent OMPs [43, 44] and synchronised with OM constriction and invagination. This process is orchestrated by the “divisome”, an assembly of cell division proteins found chiefly within the IM [45].

1.1.3 The inner membrane

The IM provides a second level of cellular protection, as a functionalised barrier around the cytoplasm. In *E. coli,* the phospholipid composition of the IM bilayer is similar to that of the inner leaflet of the OM, comprising phosphatidylethanolamine (~75%), phosphatidylglycerol (~20%) and cardiolipin-rich regions at the cell poles [46, 47]. This composition makes the IM more fluid than the OM. Antimicrobial compounds that bypass the OM are typically removed from the periplasm by efflux pumps: *trans*-envelope protein assemblies comprising an IMP and an OMP component, e.g. those involved in chloramphenicol resistance [48]. This process is endergonic, and like most endergonic functions in the periplasm, is energetically coupled to the IM due to the absence of ATP and other energetic substrates in the periplasm. The IM contains integral proteins with cytoplasmic ATPase domains that utilise cytoplasmic ATP to energise their periplasmic functions. Other integral IMPs utilise chemiosmosis to energise their functions. This involves harvesting homeostatic proton and electronic charge gradients, which together comprise the proton motive force (PMF). This describes the potential (Δp) for protons to pass from the acidic / positively charged periplasm to the less acidic (and more negatively charged) cytoplasm:

$$\Delta p = \Delta\Psi - (0.059)\Delta\mathrm{pH}$$

The PMF is generated by the electron transport chain, and serves as a communal energy pool for a diverse range of proteins, primarily utilised through IMPs [49]. Δp is established and maintained through the continuous functions of the electron transport chain [49]. How this energy potential is converted to function varies between proteins.

1.1.4 PMF-energised IM complexes

The best characterised family of PMF-transducing proteins are the F-type ATPases, conserved between bacteria, mitochondria and chloroplasts [50-52]. This complex rotates in a process driven by proton flux, which can be likened by analogy to the water/air-driven rotation of a turbine. Proton translocation causes conformational rearrangements within the protein that in turn energises endergonic functions. Other PMF-energised complexes include periplasmic efflux pumps [53], the Twin Arginine Translocon (Tat) system [54], and the Sec ancillary factors SecDF [55].

A number of PMF-harvesting pentamer-dimer complexes also exist, henceforth described as 5:2 motors [56]. The archetypal 5:2 motor complex is the $MotA_5B_2$ stator unit, which drives motility via the bacterial flagellar motor (BFM) complex [57-60] (Fig. 1.4). Homologous motility system motors include the lateral flagellar stator unit $LafT_5U_2$ [61], the sodium ion-driven $PomA_5B_2$ [62], the smaller, more distantly related gliding motility complexes predominantly found in Myxobacteria ($AglR_5QS$) [63], and the type IX secretion systems of Bacteroidetes, $GldL_5M_2$ and $PorL_5M_2$ [64, 65]. Multiple $MotA_5B_2$ stator units assemble within the BFM, where they power rotation of the C-ring (Fig. 1.4). In the BFM each $MotA_5B_2$ complex synchronously rotates clockwise, driving rotation of the C-ring and flagellum (Fig. 1.4B) [57-59, 66].

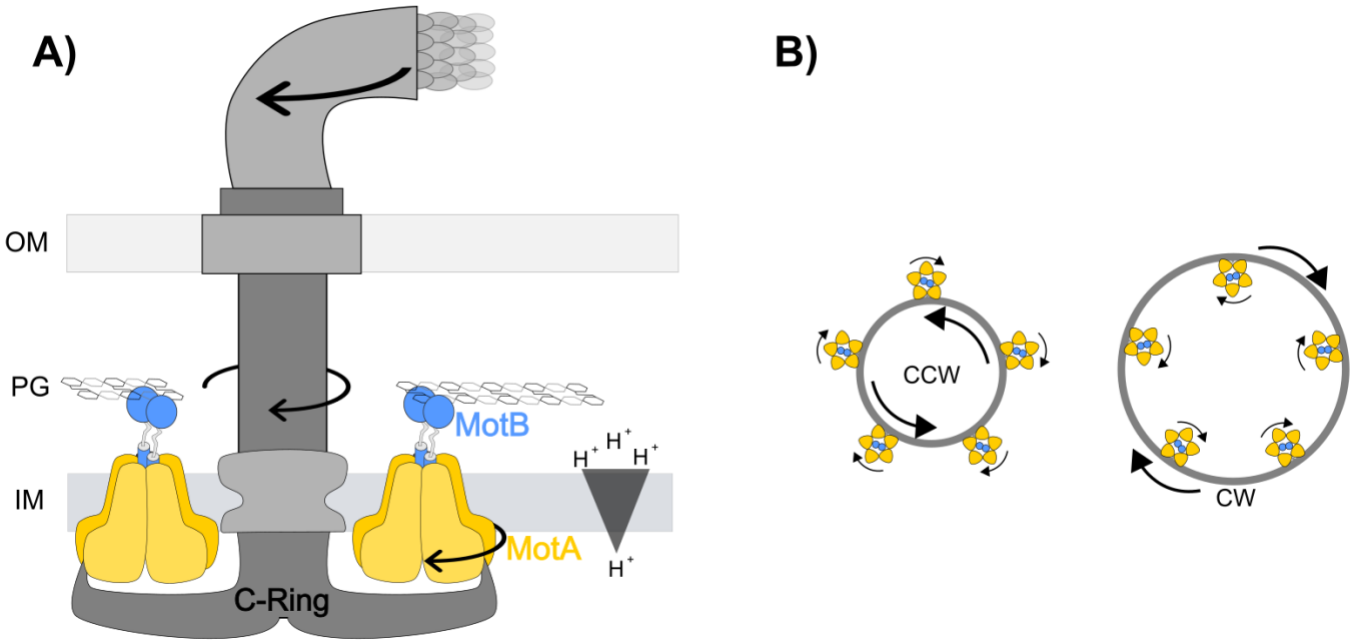


Figure 1.4 | MotA$_5$B$_2$ is the archetypal 5:2 motor that drives the BFM complex. A) Cartoon of BFM *trans*-envelope assembly. MotA$_5$B$_2$ utilises the PMF to drive cell motility, where rotation of MotA$_5$ about the PG-bound MotB$_2$ drives rotation of the flagellar C-ring and the entire flagellum. The flagellar C-ring binds MotA$_5$ in the cytoplasmic region through non-covalent interactions. **B)** The C-ring determines direction of flagellar rotation by alternating the bound face of MotA$_5$B$_2$. The BFM C-ring (*grey*) travels counter-clockwise (CCW) when bound at the inner face of the MotAB ring, or clockwise (CW) if bound at the outside of the MotAB ring. Some species form heterogeneous BFMs incorporating PomA$_5$B$_2$ (68).

This 5:2 stoichiometry complements the observation that rotary complexes favour mismatched symmetries, which yields energetic minima of low stability to reduce the activation energy per rotation step [67]. Although rotation is assumed to occur in all 5:2 motors, it has only been indirectly observed for MotA$_5$B$_2$ (through rotation of the flagellum) and evidence for rotation in the other 5:2 motor complexes is circumstantial [68]. Rotation is inferred, in that key residues essential in MotA$_5$B$_2$ for flagellar rotation are also essential to the various functions of these homologous motors.

Despite the wealth of motility-related 5:2 motors, some drive completely orthogonal processes. For example, the ZorA$_5$B$_2$ complex was recently structurally characterised as part of the Zorya phage detection system [69]. This complex features a unique cytoplasmic tail, and is suggested to detect dsDNA phage attack through cell envelope deformations. Another homologue, SiiB$_{(5)}$A$_{(2)}$, influences *Salmonella* virulence by regulating the SPI-4 Type 1 secretion system [70]. The structure of this complex remains unknown, but prediction models indicate a cytoplasmic tail highly similar to that of ZorA [71]. The combination of ever-growing structural prediction databases with

fold-based search algorithms is currently revolutionising the identification of novel 5:2 motors [71, 72].

A $MotA_5B_2$ protonation mimetic from *Campylobacter jejuni*, was generated by mutating the conserved Asp residue to Asn [58]. MotB D22N is poised downwards compared to D22, suggesting that charge neutralisation of D22 by protonation subtly reorients the side chain towards the cytoplasm during movement (Fig. 1.6). Recent models suggest that these motors form two half-channels, one of which is exposed to the periplasm (where Asp accepts a proton), and one which is exposed to the cytoplasm, where Asp-H^+ deposits a proton [68]. Physical separation of proton half-channels is essential to prevent PMF leakage, and is also observed in F-type ATPases [50]. The mechanism of rotation derived from Asp protonation and deprotonation remain to be elucidated, as well as the extent of vertical movement exhibited by the transpore helix (TPH). The TPH is a conserved α-helix found at the N-terminus of the dimeric component of the 5:2 motor. For each motor, two TPHs reside within the lumen of the pentamer, and are assumed to rotate with respect to the pentamer.

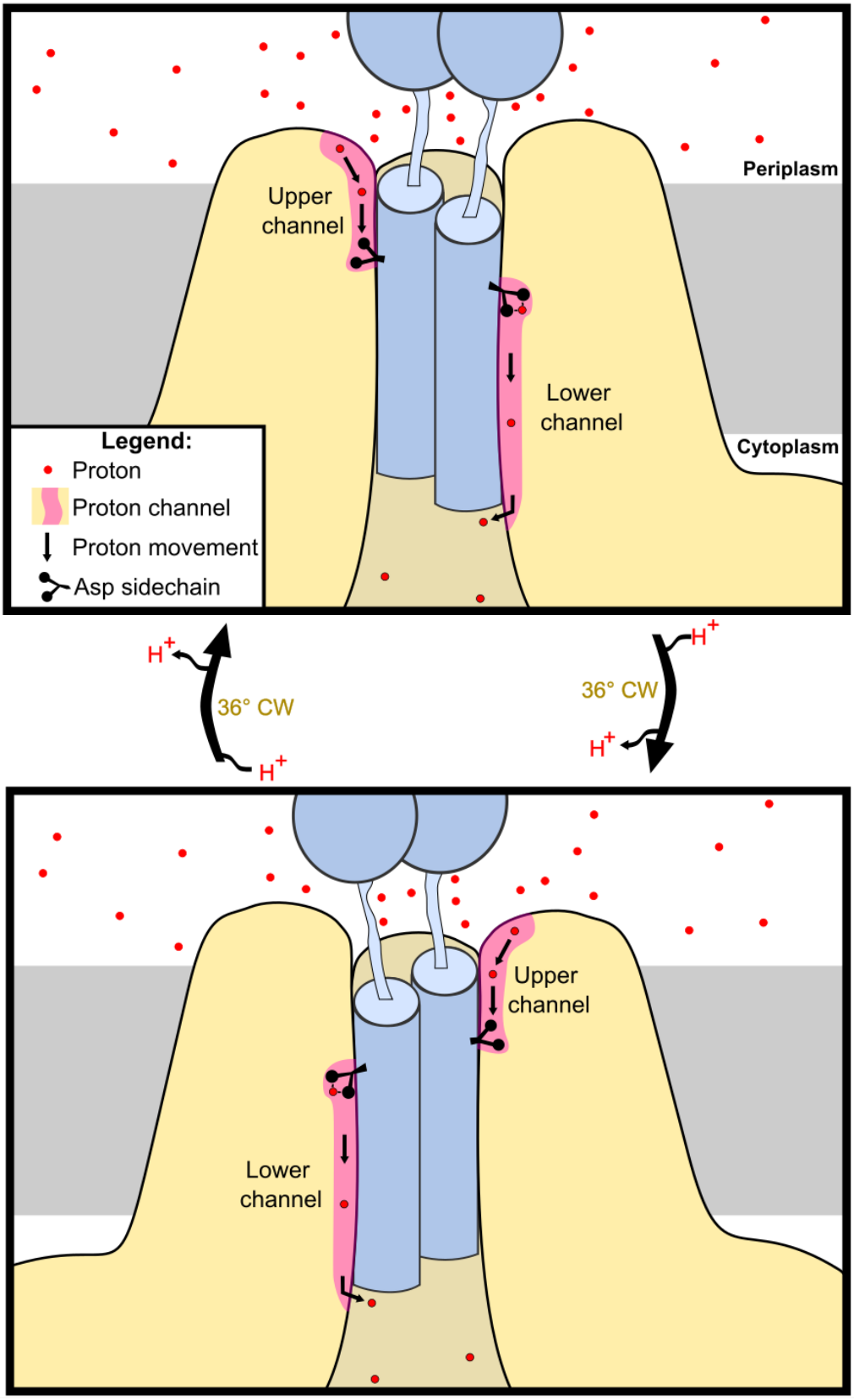


Figure 1.6 | Rotation of 5:2 motors likely proceeds through proton half-channels. As periplasmic protons enter the upper channel, coordination with the negatively charged essential Asp residue allows a conformational flip upon charge neutralisation. This may allow rotation by altering the stereochemistry of the pentamer-dimer interface: e.g. the protonated Asp enters a groove where lateral rotation can occur, akin to a mechanical spline [68]. Direction of rotation is assumed from observations of MotAB but may differ between motors.

1.1.5 ExbBD-TonB drives nutrient import

Of the various 5:2 motor systems, the Ton system is the most comparable homologue to the Tol system. The Ton system is a promiscuous energisation system which drives import of a diverse range of extracellular substrates, such as metals, sugars and organic compounds [27, 73]. Like the Tol system, the Ton system is parasitised by colicins and bacteriophages, and was thus largely characterised through resistance mutants [17, 74]. In the Ton system, the $ExbB_5D_2$ motor associates with periplasm spanning TonB to allow for the import of substrates across the OM through TonB-dependent transporters (TBDTs). The N-terminal transmembrane helix (TMH) of TonB associates with the $ExbB_5D_2$ motor, whilst the C-terminal domain binds an exposed β-strand of the TBDT (The TonB-box; Fig.1.7)) [75-78]. The conformational rearrangements which occur within TBDTs during substrate uptake involve channel opening, and are understood to proceed through the application of tensile force to the TonB-box via TonB [79], but how $ExbB_5D_2$-TonB generates tensile force at the OM remains poorly understood.

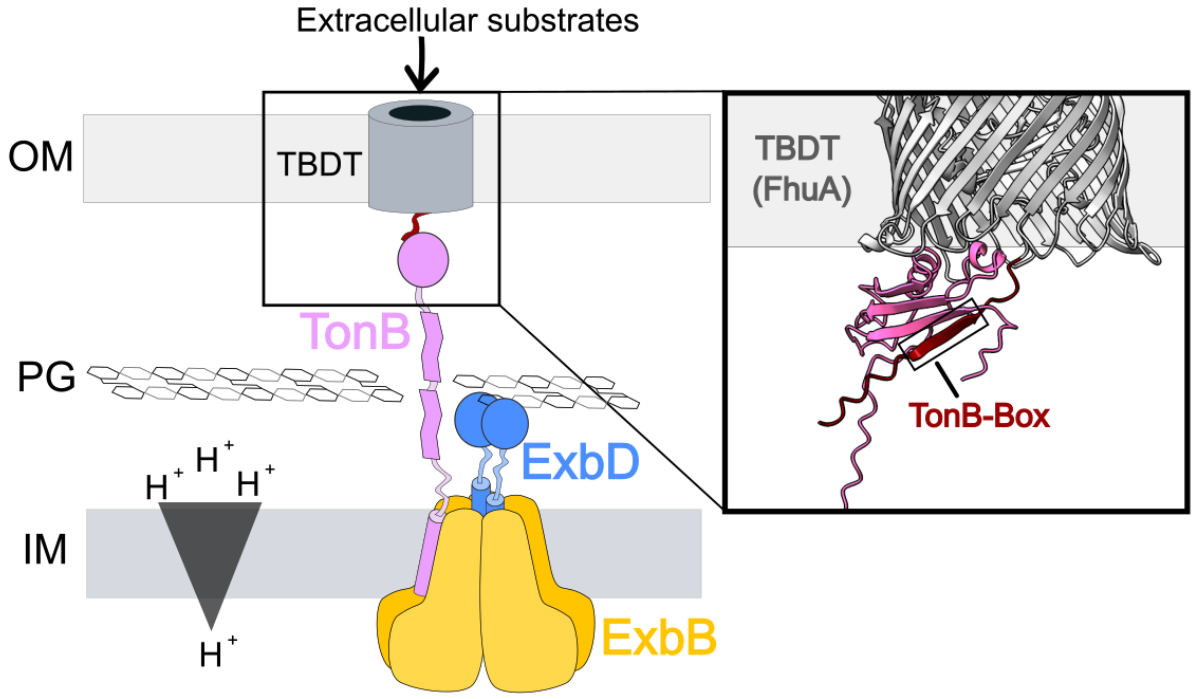


Figure 1.7 | The Ton system drives uptake of extracellular substrates. ExbBD-TonB applies tensile force to TBDTs, through conformational rearrangements within the monomeric force transducer TonB (*left*). TonB binds its target motif, the TonB-box, by β-addition (AF3 model; *right*). How the $ExbB_5D_2$-TonB complex uses the PMF to drive substrate import is unclear, but is assumed to require rotation of the $ExbB_5D_2$ motor. This may involve TonB wrapping around the motor unit while simultaneously bound to the TBDT [80].

1.2 The Tol-Pal system

1.2.1 Tol-Pal proteins

The *tol* genes were first discovered through mutagenesis studies in *E. coli* that caused gene defects that in turn conferred *tolerance* to a family of bacterial toxins [81-83]. The toxins in question, colicins, are soluble protein toxins typically many kDa in weight- far too large to diffuse through the OM. Colicins are imported into the cell by parasitising the PMF-conducting activity of the Ton or Tol systems [15], which applies tension to the colicin, driving its unfolding and translocation across the OM. Once inside the cell, the colicin refolds and kills the cell via its cytotoxic domain. Typical examples include cytotoxicity through pore-formation and cell depolarisation (colicin N), or through nuclease activity following transport across the inner membrane (ColE9) [84, 85]. Research into the Tol-Pal system is inexorably linked to colicins, which are indispensable tools. The Tol-Pal system itself is a dynamic envelope-tethering system driven by 5:2 motors, which actively redistributes the OM lipoprotein Pal (*peptidoglycan associated lipoprotein*) [42]. Pal links the OM to the PG, via its C-terminal PG-binding domain (PGB). These proteins not only recruit Pal to the mid-cell, but also fulfil ancillary functions at the interface of PG remodelling and OM invagination [86, 87]. The core proteins vary in copy number, with known estimates at ~600 TolA [88], ~2500 TolR [89], ~6,000 TolB, and ~35,000 Pal copies per cell [90].

TolB and Pal

TolB is a soluble periplasmic protein, comprising an N-terminal Rossmann fold domain and a C-terminal β-propeller domain. Pal is an OM lipoprotein, comprising a lipoylated N-terminus, disordered central linker, and a globular PGB. While Pal is responsible for dynamic OM-PG tethering, it is immobile while bound to the PG. TolB plays a more regulatory role; by binding Pal via its β-propeller domain, TolB disrupts the PG binding activity of Pal to form a mobilised TolB-Pal complex (Fig. 1.8) [91]. The TolB β-propeller domain demonstrates some homology to β-lactamase enzymes, but no enzymatic functions have thus far been ascribed to TolB [92].

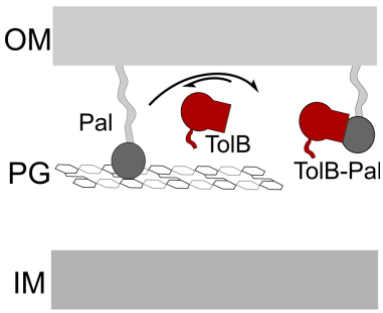


Figure 1.8 | TolB binds and mobilises Pal. The PG-binding activity of Pal is modulated by the soluble periplasmic protein, TolB. The TolB-Pal interaction precludes the OM-PG tethering activity of Pal. TolB is later dissociated from Pal through the activity of the $TolQ_5R_2A$ motor-transducer assembly.

TolA

The mobilised TolB-Pal complex is dissociated by the action of TolA, which uses the PMF via an unknown mechanism [42, 56, 93]. TolA is a predominantly α-helical protein that spans the periplasm and is embedded in the IM by a 20-residue N-terminal TMH. TolA has traditionally been divided into three domains, bordered by disordered regions (Fig. 1.9) [88, 94]. Domain I contains the TMH, believed to bind the $TolQ_5R_2$ motor [95]. Domain II is predominantly α-helical [88, 96], and is predicted by AF2 to form a periplasm-spanning hairpin [71, 97]. Rosetta modelling predicts no hinge, but rather a single 35 nm helix, that is longer than the width of the periplasm (Fig. 1.9) [98]. Finally, domain III has a globular fold which binds an exposed β-strand of TolB, to dissociate the TolB-Pal complex (Fig. 1.9) [42, 91, 99]. This is a structurally homologous interaction to that of TonB with TBDTs, and like TonB, requires the PMF for completion (Fig. 1.7) [42]. How TolA can diffuse around the cell but also span the periplasm to contact the OM without becoming trapped by the PG layer is also unclear.

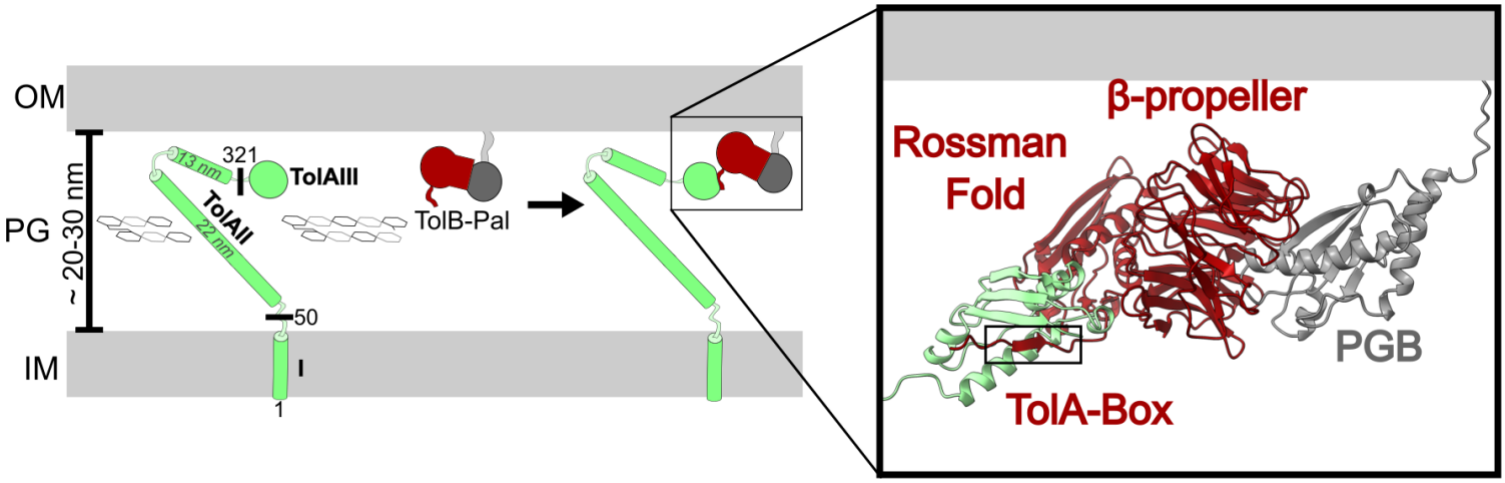


Figure 1.9 | TolA dissociates the TolB-Pal complex by applying tensile force. **A)** Black bands indicate approximate domain boundaries at respective positions in the *E. coli* TolA amino acid sequence that occur within disordered regions. Domain boundaries are poorly defined in the literature. In *E. coli*, TolAII is extremely α-helical, with a predicted length of TolAII-III at around 40 nm, if the helical hinge were to open. This is significantly longer than the width of the periplasm. TolAIII interacts directly with the exposed N-terminus of TolB from the TolB-Pal complex, by β-addition (*right*). Structural model produced by AF3 [100], with individual TolB (*red*) and Pal (*grey*) domains labelled. "TolA-Box" here refers to the TolA-bound β-strand formed by TolB residues 23-28.

TolQ$_5$R$_2$

TolQ$_5$R$_2$ comprises the integral 5:2 motor complex of the Tol system, residing in the IM [89, 101-103]. This complex is thought to rotate in a PMF-dependent manner generating mechanical force from ion flux [104]. TolQ forms a pentameric outer ring, each subunit of which contains three TMHs (Fig. 1.5) [101]. At the N-terminus of TolR is a positively charged, disordered tail followed by a short α-helical TPH that resides in the centre of the TolQ$_5$ rotor (Fig. 1.5). Connected to the TPH by a disordered linker is the TolR C-terminal domain (CTD), which constitutes a putative PGB domain [105]. This is postulated to form the stator unit of the complex, by binding the PG and forming a pivot point around which TolQ$_5$ rotates (Fig. 1.10).

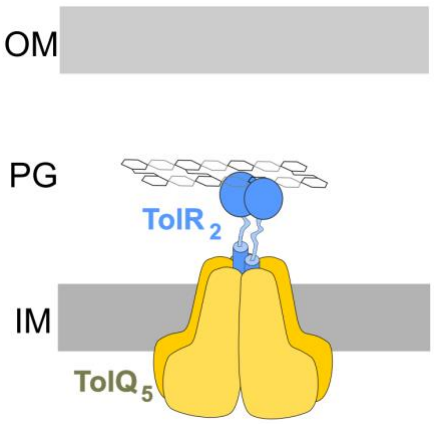


Figure 1.10 | TolQ$_5$R$_2$ forms a 165 kDa heptamer. TolQ (*yellow*) forms a pentameric ring around the TolR dimer (*blue*). This is an integral IM complex. The C-termini of TolR are predicted to bind the PG. Multiple confirmations of TolR have been identified, suggesting conformational rearrangements occur during PG-binding and/or force transduction. Each TolQ monomer comprises TMHs, and a bulky soluble region in the cytoplasm. The structure of TolQ$_5$R$_2$ was determined during this work and is discussed in further detail in chapter five.

It has also been proposed that TolR$_2$ transitions between a “strand-swapped” and an “open” conformation of TolR to influence its affinity for PG [97]. It is unclear if this cycle occurs once (i.e. upon initiation of the TolQ$_5$R$_2$-TolA assembly) or multiple times during the TolA energisation cycle, or whether one of these conformations is a crystallisation/truncation artefact [97]. It is possible that TolR exhibits a preference for denuded PG, indicative of septal PG (sPG), because the PG-binding specificity of TolR has not yet been determined [106]. Of the 5:2 motor family, TolQ$_5$R$_2$ is most closely related to ExbB$_5$D$_2$ [107]. Unlike ExbB$_5$D$_2$ however, TolQ$_5$R$_2$ is trafficked to the septum during cell division [108, 109]. How this is achieved is also unknown.

CpoB

Of the ancillary genes, only one possesses a demonstrable role in cell division. CpoB (Coordinator of PG synthesis and OM constriction), previously known as YbgF, is a soluble periplasmic protein with two domains; a coiled-coil domain involved in trimerisation, and a tetratricopeptide repeat (TPR) domain involved in protein-protein interactions [110]. *ΔcpoB* mutants do not exhibit the classical *Δtol-pal* OM instability phenotype, but they do exhibit increased sensitivity to cefsulodin [111]. SPR and NMR data suggest that CpoB plays a regulatory role in cell wall synthesis, involved in the modulation of the transpeptidase activity of PBP1B-LpoB [111]. At present it appears that a low affinity interaction with TolA near the CTD modulates the trimeric state and activity of CpoB, negatively regulating its influence on PG synthesis [110]. This effect may be conferred in a concentration-dependent manner, with septal localisation of TolA eliciting CpoB dissociation.

YbgC

The remaining ancillary gene in γ-proteobacteria, YbgC, is an acyl-thioesterase, thought to be involved in lipid metabolism [112]. The gene is not syntenic with *tol-pal* genes in many GNB, and deletion of the gene gives no traditional *Δtol* phenotypes, but results in PL accumulation in the OM [12, 113]. It was thus proposed that the Tol-Pal system plays a role in retrograde phospholipid transport, but PL accumulation at the OM could equally be a pleiotropic or regulatory consequence of reduced OM stability, to alleviate the difficulty of invaginating the OM in those mutants.

1.2.2 $TolQ_5R_2$-TolA stabilises the OM at the septum

GNB require coordinated synthesis and maintenance of the entire cell envelope [43, 44]. There is typically moderate turgor pressure within both the cytoplasm and the periplasm, both saturated with soluble molecules to a gel-like phase [114]. To avoid detachment of the OM, three distinct OM-PG tethering systems exist in *E. coli*: Braun's lipoprotein (Lpp), OmpA, and Pal (Fig. 1.11) [115-118]. All three specifically bind non-crosslinked *meso*-diaminopimelic acid (mDAP) moieties of the PG [119-121]. Unlike the non-covalent associations of OmpA and Pal, Lpp is covalently linked to mDAP in an enzymatically reversible reaction [121].

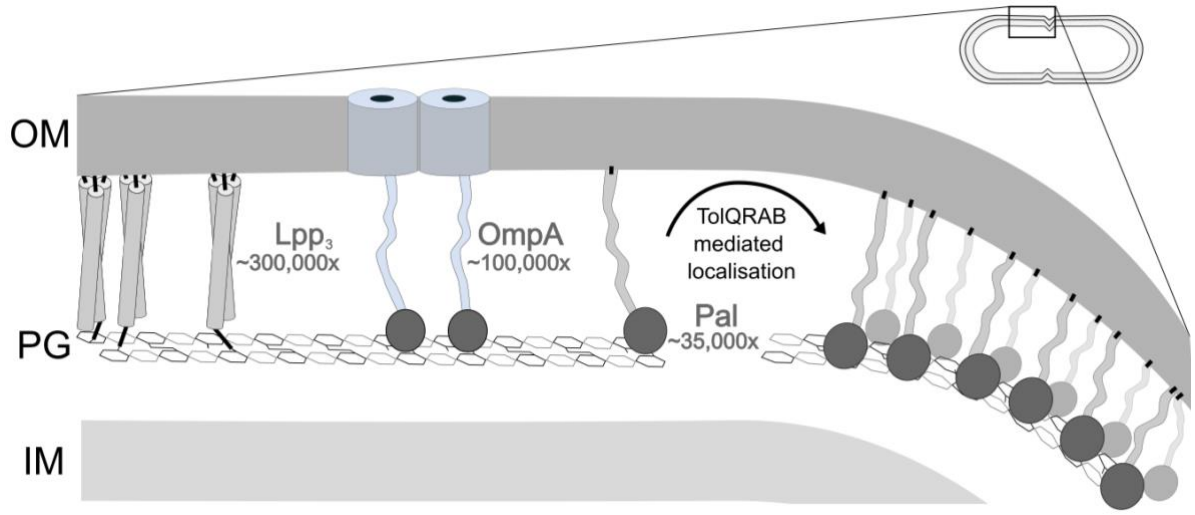


Figure 1.11 | Overview of *E. coli* OM-PG tethering systems. In Proteobacteria three proteins tether the OM to the PG. All three proteins recognise non-crosslinked mDAP moieties of the PG, presumably competing to occupy mDAP residues. Lpp is covalently attached to mDAP [115], and OmpA/Pal bind mDAP non-covalently [119, 120]. They vary significantly in copy number (Lpp_3:~300,000 [122]; OmpA: ~100,000 [123]; Pal: ~35,000 [90]). Estimates of copy number likely vary depending on cell volume, growth conditions, and quantification techniques.

OmpA and Lpp are typically distributed along the barrel of the cell, while Pal uniquely displays dynamic localisation along the barrel during elongation and septa during cell division. Lpp has been found to occupy the septum only in Pal deficient strains [124]. Depletion of any of these OM-PG linkages causes OM instability, periplasmic leakage, OM blebbing, and increased detergent/antibiotic sensitivity. OM-PG linkage deficient mutants, such as *Δlpp*, *ΔompA* or *Δpal* mutants, demonstrate increased vancomycin sensitivity [125, 126]. Typically, *Δpal* mutants exhibit cell chaining, while *Δlpp Δpal* double mutants exhibit compounded defects that lead to total loss of OM invagination [124]. Overexpression of Lpp or OmpA does not complement *Δpal,* but Pal overexpression can complement both *Δlpp* and *ΔompA*, supporting Pal's dual role in general OM-PG tethering and in sPG tethering. As a lipoprotein, Pal can diffuse

around the periplasm because it is linked to the fluid inner leaflet of the OM. In contrast, OmpA, which is an integral OMP, displays minimal lateral diffusion due to the immobile nature of the OMP-LPS lattice [127-129]. Pal likely evolved from OmpA after the emergence of the Lol lipoprotein trafficking pathway [2].

The Tol-Pal system primarily redistributes Pal to the division septum for localised OM-PG tethering. During cell division the OM requires invagination, followed by PG remodelling, septation, and scission of the two daughter cells. Critically, the Tol-Pal system uses the PMF to drive OM stabilisation via the redistribution of Pal (Fig. 1.9) [90, 109, 130]. During elongation, both $TolQ_5R_2$-TolA and Pal are distributed throughout the periplasm. TolB binds and mobilises Pal, and the mobile Pal-TolB complexes are dissociated by $TolQ_5R_2$-TolA. This continues during cell division, except the $TolQ_5R_2$-TolA proteins are each trafficked to the septum, thus their localisation is mirrored by the septal recruitment of Pal (Fig. 1.12) [42, 93, 108].

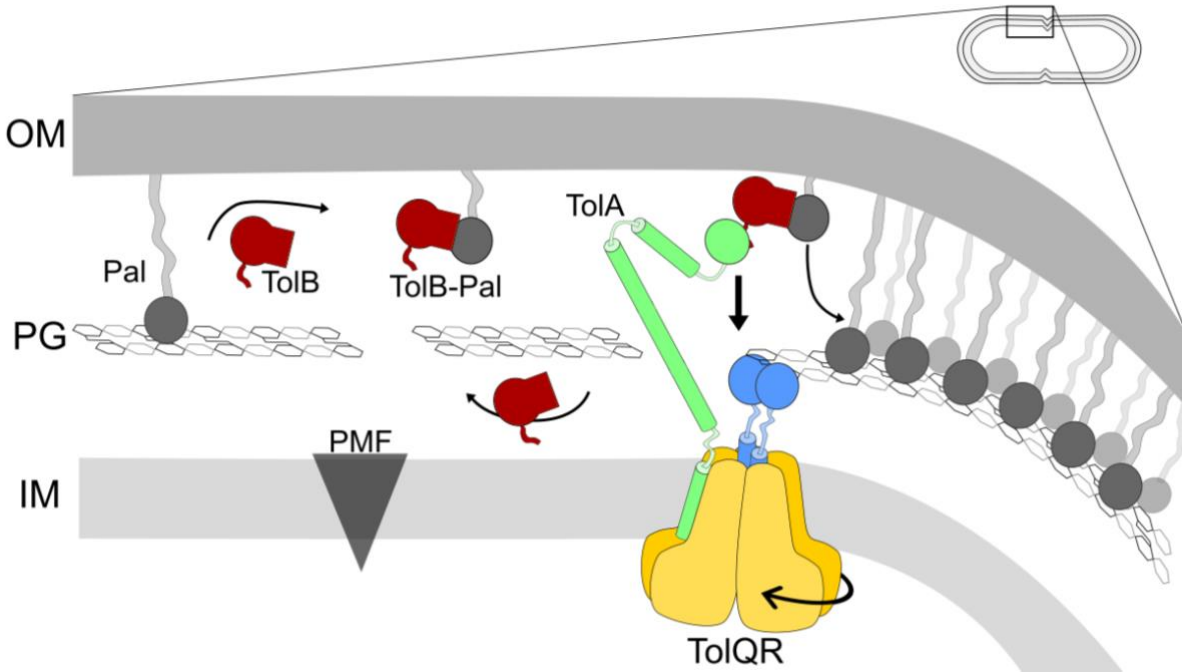


Figure 1.12 | The Tol-Pal system uses a mobilisation-and-capture mechanism to accumulate Pal at the division site. Pal tethers the inner leaflet of the OM to the PG layer. The soluble periplasmic protein TolB displaces Pal from the PG, effectively mobilising Pal, which can then diffuse around the cell. During cell division, $TolQ_5R_2$ and the force transducer TolA are independently recruited to the mid-cell, where the $TolQ_5R_2$-TolA complex strips TolB from Pal. This results in local deposition of Pal, the accumulation of OM-PG tethers and invagination of the OM. Free TolB is recycled, which diffuses through the periplasm to mobilise additional Pal molecules for mid-cell localisation. Figure adapted from Szczepaniak *et al*. [42], Webby *et al*. [97] and Williams-Jones *et al.*, [131].

Many questions remain; how are each of the various proteins recruited to the mid-cell? What orchestrates $TolQ_5R_2$-TolA association and energisation? How does $TolQ_5R_2$-TolA apply force to the TolB-Pal complex?

1.3 Aims of This Work

Despite the importance of the Tol-Pal system, studies have been hindered by the need for synergised presence of all requisite components. These include an intact cell envelope/PG layer (modified appropriately to signal cell division), an intact PMF, and the as-yet unknown trafficking factors that recruit $TolQ_5R_2$-TolA to the septum. These are difficult to reconstitute *in vitro*, hence the importance of parallel work in live cells. Structural characterisation of purified Tol-Pal proteins has been informative, but many molecular details remain unresolved and their functions *in vivo* remain enigmatic. Here we aim to better understand how the Tol-Pal system contributes to cell envelope stability in *E. coli* and other GNB. First, I probe the localisation of Pal in response to chimeric TolA mutants with reduced force transduction efficiency. This work reveals that force transduction efficiency is influenced by TolA secondary structure, and correlates with both septal recruitment of Pal and relative OM stability (chapter three). Based on those findings, I perform a comprehensive analysis of *tol-pal* gene clusters throughout the kingdom of bacteria to sample the structural diversity of TolA sequences, and propose that TolA emerged from a TonB paralogue (chapter four). I also focus on the molecular interactions of the *E. coli* Tol-Pal system, beginning with a combination of computational analyses and site-specific crosslink mapping to investigate the rotary motion within $TolQ_5R_2$ in response to proton flux (chapter five), followed by a combination of bioinformatic and mutagenesis experiments to study the interaction of TolA with $TolQ_5R_2$ (chapter six). I then describe our progress towards 3D reconstruction of these assemblies *in vivo* and identify the conformational rearrangements occurring within $TolQ_5R_2$-TolA during energisation, by performing cryo-ET on stalled complexes within live bacterial cells (chapter seven). These data are then evaluated in light of recently published work, highlighting promising new avenues for future lines of study.

2.0 Materials and Methods

2.1 Microbiology

Inoculants were thawed, precultured in high salt LB (lysogeny broth) media at pH 7.2, transformed by 45 second heat shock and plated onto LB agar selection plates with the appropriate antibiotics, before storage at 4 °C [132]. Chemically competent *E. coli* PB114 cells were prepared from a glycerol stock, streaked onto agarose plates (with 50 µg/ml kanamycin), by following the $CaCl_2$ chemical competence protocol [133]. For antibiotic selection, media were typically supplemented with ampicillin (100 µg/mL (w/v)), chloramphenicol (35 µg/mL (w/v)), or kanamycin (50 µg/mL (w/v)), as determined by selection markers on bacterial strains and plasmids.

2.1.1 Strains, media, and growth conditions

Media used	**Purpose**
High Salt Lysogeny Broth (LB)	General Cell Growth
M9 Minimal Media	Microscopy
TB Rich Media	Protein Purification

Strains used:

A *ΔtolA::kan* mutant in the BW25113 background was used for TolA investigations, obtained from the Keio collection [134]. This cell line was further chromosomally modified to create a Pal-mCherry fusion reporter strain by Dr. Joanna Szczepaniak.

Strain	Abbreviation	Usage	Reference
BL21 (DE3)		Protein purification	[135]
DH5α		Cloning	[136]
BW25113		WT controls	[137]
JW0729-3	*ΔtolA*	Group A expression	[134]
RKCK10	*ΔtolA::Pal-mCherry*	Group A expression	[42]
JW5195-1	*ΔtonB*	Group B expression	[134]
JS10	*ΔtolQR*	TolQR *tag* expression tests	This work
PB114 / λDB173	*ΔminCDE*	Minicell purification	[138]
MG1655		WT controls	[139]

Table 1 | Strains used in this work.

2.1.2 Plasmids

Plasmids published in Williams-Jones *et al*., 2023, are available to order from Addgene [131]. TolA plasmids were uninduced, with leaky expression levels approximately equivalent to chromosomal expression [140].

Title	Genotype	Abbreviation	Reference
pCEP13	pT12-*tolQ*-*tolR*	TolQR	[57]
pJS61	pBAD24-*GFP*-*tolA*	TolA	[140]
pJS70	pBAD24-*GFP*-*tolA* (H22A)	$TolA_{H22A}$	[131]
pJS65	pBAD24-*GFP*-GGGGS-*tonB*(2-239)	TonB	[141]
pJS96	pBAD24- *GFP*-GGGGS-*tonB* (H20A)	$TonB_{H20A}$	[131]
pJS104	pBAD24-*GFP*-*tolA*(1-47)-*tolA*(311-421)	TolAΔII	[131]
pJS109	pBAD24-*GFP*-*tolA*(1-45)-*tonB*(33-168)-*tonB*(33-168)-*tolA*(311-421)	ABBA	[131]
pJS110	pDWJ26 (H22A)	AIA_{H22A}	[131]
pJS119	pBAD24-GFP-*tonB*(2-40)-*tonB*(168-239)	TonBΔII	[131]
pJS124	PDWJ27 (H20A)	BIB_{H20A}	[131]
pJMG1	pBAD24-*GFP*-GGGGS-*tonB*(2-62)-*tonB*(105-239)	TonBΔPII	[131]
pJMG2	pBAD24-*GFP*-GGGGS-*tonB*(2-62)-*tolA*(48-310)-*tonB*(105-239)	BAB	[131]
pDWJ1	ColN(1-162)-Cys	$ColN_{1-162}$	This Work
pDWJ2	pBAD24-*GFP*-*tolA* AAAAAAAAAAAA	TolA 12-Ala	This Work
pDWJ3	pBAD24-*GFP*-*tolA* AAAAAA	TolA 6-Ala	This Work
pDWJ4	pBAD24-*GFP*-*tolA*(1-34)-*tonB*(34-50)-*tolA*(51-421)	$TolA_{TonB\text{-}IDR}$	This Work
pDWJ5	pBAD24-*GFP*-*tolA* GSGSGS	TolA 3GS	This Work
pDWJ6	pBAD24-*GFP*-*tolA* GAGGAG	TolA 2GAG	This Work
pDWJ7	pBAD24-*GFP*-*tolA* GAGSGSA	TolA 2GSA	This Work
pDWJ8	pBAD24-*GFP*-*tolA* GKGSGSK	TolA 2GSK	This Work
pDWJ9	pBAD24-*GFP*-*tolA*(1-44)	TolAI	This Work
pDWJ11	pBAD24-*GFP*-*tolA*(12A-deleted)	TolAΔII	This Work
pDWJ12	pBAD24-*GFP*-*tolA*(1-47)-*tolA* (188-421)	TolAΔII	This Work
pDWJ13	pBAD24-*GFP*-*tolA*(1-320)-*tonB*(139-239)	AAB	[131]
pDWJ14	pBAD24-*GFP*-*tolA*(1-220)-*tonB*(39-239)	AABB	[131]
pDWJ15	pBAD24-*GFP*-*tonB*(2-139)-*tolA*(321-421)	BBA	[131]
pDWJ16	pBAD24-*GFP*-*tonB*(2-45)-*tolA*(46-421)	BAA	[131]
pDWJ17	pBAD24-*GFP*-*tolA*(1-45)-*tonB*(46-239)	ABB	[131]
pDWJ18	pBAD24-*GFP*-*tolA*(1-45)-*tonB*(33-168)-*tolA*(311-421)	ABA	[131]
pDWJ21	pBAD24-*GFP*-*tolA* $(KLEALEG)_5$	TolA $KLEALEG_5$	This Work
pDWJ22	pBAD24-*GFP*-*tolA*(1-220)-*tonB*(39-139)-*tolA*(311-421)	AABA	[131]

pDWJ25	pBAD24-*GFP-tolA*(1-50)-*tolA*(319-421)	TolAΔII	This Work
pDWJ26	pBAD24-*GFP-tolA*(1-45)-IDP129-*tolA*(311-421)	AIA	[131]
pDWJ27	pBAD24-*GFP-tonB*(2-40)-IDP129-*tonB*(139-239)	BIB	[131]
pDWJ29	pBAD24-*GFP-tolA* Δ280-313	TolA ΔCBS	This Work
pDWJ31	pBAD24-*GFP-tolA* D51P	TolA D51P	This Work
pDWJ32	pBAD24-*GFP-tolA* A52P	TolA A52P	This Work
pDWJ33	pBAD24-*GFP-tolA* V53P	TolA V53P	This Work
pDWJ34	pBAD24-*GFP-tolA* M54P	TolA M54P	This Work
pDWJ35	pBAD24-*GFP-tolA* V55P	TolA V55P	This Work
pDWJ36	pBAD24-*GFP-tolA* D56P	TolA D56P	This Work
pDWJ37	pBAD24-*GFP-tolA* D51P A52P	TolA 51-52P	This Work
pDWJ38	pBAD24-*GFP-tolA* V53P M54P	TolA 53-54P	This Work
pDWJ39	pBAD24-*GFP-tolA* V55P D56P	TolA 55-56P	This Work
pDWJ40	pBAD24-*GFP-tolA* V53P M54P V55P	TolA 53-55P	This Work
pDWJ41	pT12–*tolQ-tolR 4-tag*	TolR A4X	This Work
pDWJ43	pT12–*tolQ-tolR 12-tag*	TolR L12X	This Work
pDWJ44	pT12–*tolQ-tolR 19-tag*	TolR V19X	This Work
pDWJ45	pT12–*tolQ-tolR 29-tag*	TolR L29X	This Work
pDWJ46	pT12–*tolQ-tolR 39-tag*	TolR I39X	This Work
pDWJ56	pT12–*tolQ-tolR 7-tag*	TolR R7X	This Work
pDWJ57	pT12-*TolQ* (no TolR)	TolQ	This Work
pDWJ76	pBAD24-*GFP-tolA(Δ41-62)*	$TolA_{\Delta IDR}$	This work
pDWJ77	pBAD24-*GFP-tolA(Δ41-62)*-	$TolA_{IDR\text{-}helix}$	This work

Table 2 | Plasmids used in this work. Amino acid residues of protein products are indicated (*brackets*), as well as deletions (*Δ*) of different residues and/or domains of TolA and TonB (*I/II/III*, *PII = polyproline region*, "*X"* indicates *tag* mutation for BPA incorporation, *IDR* indicates intrinsically disordered region).

2.2 Molecular Biology

2.2.1 Molecular cloning

A range of cloning techniques were used, including polymerase chain reaction (PCR), restriction digestion and ligation, ligation-independent cloning, and gene synthesis. Synthesised genes were codon optimised for *E. coli*, and purchased from ThermoFisher Geneart.

2.2.2 Chemically competent cell preparation

To generate competent cells, cells were first grown on LB agar selection plates. A single colony was picked, and cultured overnight in 5 ml LB with antibiotics. This was used to inoculate a 50 ml LB day culture the following day (1:100). The culture was grown at 37 °C with shaking, until OD_{600}: 0.4. Cells were pelleted at 4 °C, and resuspended on ice, in 20 mM Tris HCl, pH 8, with 200 mM $CaCl_2$. Cells were incubated on ice for 1 hour, then resuspended in 2 ml of the same buffer supplemented with 20% glycerol (v/v). Cells were aliquoted, snap frozen and stored at -80 °C. Competent BL21 and DH5α cells were purchased from NEB.

2.2.3 Bacterial heat shock transformation

Cells were transformed by 45 second heat shock at 42 °C [132]. Competent cells were first thawed on ice, then 1-2 µl plasmid was added depending on plasmid/cell concentration. The tube was mixed gently by flicking, then allowed to incubate on ice for 30-60 minutes. Tubes were then heated for approximately 45 seconds at 42 °C, followed by incubation on ice for 2-5 minutes. 800 µl SOC outgrowth media or LB media was then added, and cells incubated at 37 °C with shaking for 60 minutes. Cells were then pelleted by centrifuging for 1 minute at 5,000 *g*, then resuspended and spread onto LB agar selection plates and allowed to dry, before overnight incubation at 37 °C. Plates were stored at 4 °C for up to one month.

2.2.4 DNA gel electrophoresis

DNA was resolved by agarose electrophoresis, typically for 40 minutes at 80 V [142]. Agarose gels were typically cast between 0.8-1.5% (w/v) depending on expected product size. Gels were cast at 50 ml total volume and dyed with 5 µl Syber Safe (Thermo Fisher).

2.2.5 Plasmid preparation and sequencing of DNA

Cloning was performed in *E. coli* DH5 α-Select (NEB). Plasmids were typically purified from 10 ml overnight culture using a DNA miniprep kit (NEB). Purified plasmids were posted under ambient conditions for either Sanger sequencing (at 100 ng/µl DNA) or Nanopore sequencing (30 ml at 30 ng/µl) [143, 144]. Upon sequence confirmation, plasmids were midi-prepped from 50ml overnight culture (Promega).

2.3 Protein expression and purification

2.3.1 Purification of ColN$_{1\text{-}162}$-Cys and AlexFluor$_{647}$ conjugation

After confirmation of DNA sequence, *E. coli* BL21 (DE3) cells were transformed with the ColN$_{1\text{-}186}$-cys plasmid pDWJ1. Six litres of LB (with ampicillin) were inoculated from precultures. Cells were grown at 37°C with shaking. At OD_{600}: 0.6, expression was induced with 1 mM IPTG, then cells grown for a further three hours at 37°C, at 140 RPM before harvesting. Cells were pelleted at 4,200 *g* for ten minutes, resuspended in 35 ml 20 mM Tris HCl buffer pH 8.5 with 5 mM imidazole, 500 mM NaCl, 1mM PMSF, and stored at -20°C. The following day, cells were thawed and sonicated. Lysate was clarified by 30 minutes of centrifugation at 16,000 *g*, before application to a His-Trap column. Protein was step eluted with 500 mM imidazole, and ColN fractions were spiked with 0.2 mM EDTA and dialysed overnight at 4°C into gel filtration buffer to remove imidazole. Dialysed protein was centrifuged for 10 minutes at 12,000 *g* to pellet aggregated protein, then filtered through a 0.2 µm filter, before gel filtration with 25 mM Tris HCl buffer pH 8.5 with 100 mM NaCl and 2mM DTT. DTT was removed from 1 ml of purified ColN-cys (150 µM), via a desalting column. Desalted ColN-cys was immediately spiked with 167 µM AF_{647} and incubated for 30 minutes, to conjugate AF_{647} to the C-terminal cysteine in absence of DTT, desalted to remove excess AF_{647}, then concentrated to 500µl for gel filtration. Protein fractions were identified through SDS-PAGE and in-gel fluorescence. Protein was concentrated to 33.6 µM, snap frozen and stored at -20 C.

2.3.2 TolQR purification optimisation

TolQR was expressed in *E. coli* BL21 (DE3) cells transformed with pCEP13, grown in TB media with kanamycin for selection, as described in Williams-Jones *et al.,* [131]. The *tolQ-tolR* plasmid sequence includes a Strep-II tag at the C-terminal end of TolR, followed by a TEV protease cleavage site and a terminal His_6 sequence. Overnight cultures were grown in LB, then 50 ml day cultures were grown in TB at 37°C with shaking. Upon reaching OD_{600} : 2, one litre of TB was inoculated to an OD_{600} of 0.02, and induced with rhamnose (0.1% w/v). Cells were grown overnight at 20°C, then pellets harvested and optionally frozen. Pellets were resuspended in ice cold lysis buffer (50 mM Tris HCl pH 7.5 300 mM NaCl supplemented with 1 mM PMSF), then lysed by sonication. The resulting lysate was centrifuged at 10,000 *g* for 10 mins at 4

°C. The membrane fraction was pelleted from the soluble lysate by ultracentrifugation at 45,000 *g* for 45 minutes at 4 °C. The membrane pellet was then homogenised in 10 ml solubilisation buffer (50 mM Tris HCl pH 7.5 300 mM NaCl supplemented with 1% LMNG (w/v)). Membranes were solubilised overnight on a roller at 4 °C, then ultracentrifuged as previously done, to separate solubilised and insoluble membrane fractions. The solubilised membranes (supernatant) were applied to Strep Tactin II resin equilibrated with solubilisation buffer, on a gravity flow column. The sample was incubated on the resin for 2 hours at 4 °C, with gentle rolling. The sample was then washed with buffer containing 0.02% LMNG, and eluted with 3 mM desthiobiotin. Samples were stored overnight at 4 °C, then concentrated and snap frozen.

2.3.3 Purifying and UV crosslinking TolQ-TolR BPA mutants

TolQR amber stop codon mutants were expressed similarly to the protocol described above, for benzoyl-phenylalanine (BPA) incorporation. The BPA plasmids were co-transformed into BL21 (DE3) with the pEVol (pBpF) plasmid generated by Chin *et al*., [145], with a chloramphenicol selection marker. Three litres of cultures of BL21 (DE3) containing the TolQR-*tag* plasmid and the pEvol plasmid were induced overnight with 0.1% rhamnose (w/v) and 0.15% arabinose (w/v) respectively, in the presence of 0.5 mM BPA, at an OD_{600} of 0.02 at 20 °C. BPA was first dissolved in 1.2M NaOH, 500 µl of which was added to the TB media, and shaken until dissolved, before neutralisation with 500 µl of 1.2M HCl solution. Cells were pelleted after overnight growth at 20° C in TB media, and then each 1 L cell pellet was resuspended in 10 ml cell lysis buffer (Tris-HCl 150 mM NaCl pH8 with 1 mM PMSF). These cells were crosslinked on ice for 2 hours in plastic Petri dishes (10 ml per dish), at maximum UV power. Cells were mixed every 20 minutes, and the ice pack was replaced at 60 minutes. After UV-crosslinking, proteins were purified by Strep-Tactin affinity as described above.

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

Proteins were typically diluted into 1x SDS buffer, boiled for 10 minutes, then pelleted by spinning for 1 minute at 10,000 *g*. Solubilised proteins were then resolved by SDS PAGE through 10-16% polyacrylamide gels at 35 mA for 40 minutes, as first described by Laemmli [146].

2.3.5 In-gel digestion of UV-crosslinked TolQR

After UV crosslinking and resolution by SDS-PAGE, gel bands were excised and the denatured protein samples digested *in situ*. This was achieved by destaining the band in 50% acetonitrile (v/v), reducing with TCEP (10 mM), carbamoylating the protein with 2-chloroacetamide (50 mM), and then digesting with trypsin (100 ng) or chymotrypsin (250 ng) + GluC (250 ng) or elastase (300 ng) overnight in either 25 mM ammonium bicarbonate + 1 mM $CaCl_2$ (trypsin) or 50mM Tris HCl pH 8 + 10 mM $CaCl_2$ (chymotrypsin/elastase). Trypsin and elastase digests were performed at 37 °C and 25 °C respectively with shaking at 1000 rpm. The digestion was subsequently quenched 5% formic acid (v/v), and peptides were desiccated in a rotary evaporator. All steps utilised MS-grade reagents including ^{18}C filtered deionised water and LowBind tubes.

2.3.6 Mapping TolQR BPA crosslinks

BPA incorporation utilises the amber stop codon *tag*, which ordinarily ends protein translation, but with the addition of the BPA-specific aminoacyl-tRNA synthetase on a helper plasmid (pEvol), is read as a codon specific to BPA, which is taken up from the growth media [145]. Digested peptides were solvated and resolved by a 60 minute size exclusion gradient LC-MS/MS. Spectra were analysed in MaxQuant to detect peptides containing BPA at expected positions, using an *E. coli* FASTA reference database. Following confirmation of BPA incorporation, pLINK was used to identify and map the covalently crosslinked peptides.

2.3.7 Protein quantification

Proteins were predominantly quantified by biophotometry. A multi-wavelength scan was performed and the absorbance at 280 nm was used to estimate protein concentration. Absorbance at 320 nm was used to estimate aggregation where appropriate. For fluorescent proteins, the absorbance maxima of the fluorophore was also used. AF_{647}-labelled protein concentrations were estimated from absorption spectra using a formula that features a correction factor for A_{280} absorption of AF_{647}:

$$\frac{(A_{280} - 0.03(A_{650}))x\ dilution\ factor}{A_{280}\ of protein\ at\ 1mg/ml}$$

2.4 Electron microscopy

2.4.1 Sample preparation

For negative stain EM, 20 µl of sample was applied to a carbon coated copper TEM grid, which was previously glow-discharge treated for 30 seconds. After 1 minute the grid was washed with deionised water, then negatively stained with 2% (w/v) uranyl acetate 20 seconds. Samples were imaged on a 120 kV FEI Tecnai12. For cryo-EM, PB114 cells were grown to OD_{600}: 1, then cells were pelleted by centrifugation at 10,000 *g* and discarded. The supernatant was filtered with a 1.2 µm filter and centrifuged at 21,000 *g* to pellet minicells. Minicells were resuspended in 200 µl LB and ColN-GFP was added to a final concentration of 1 µM. Minicells were labelled in the dark for 1 hour, then washed three times with PBS. The final minicell pellet was labelled with SPOTs in TAE buffer containing 10 mM $MgCl_2$ and incubated on ice for 30 minutes. Gold fiducials were then added in TAE buffer including bovine serum albumin (BSA; 1% w/v). For cryo-EM of whole cells, the same protocol was applied to PB114 cell pellets.

2.4.2 Grid preparation

Cell samples were spotted onto Quantifoil gold grids plasma cleaned for two minutes. Cells were blotted onto grids using a Mark VI Vitrobot, for 3 seconds with a blot force of 5, 100% humidity at 8 ^{o}C, before plunge freezing in liquid ethane.

2.4.3 Imaging conditions

Grids were screened for ice thickness and cell density on a Talos Arctica (200 kV). Tilt series were collected on a Titan Krios (300 kV) at 10° tilt angles, using serial EM software.

2.4.4 Data processing

Tilt series were stacked and aligned in IMOD, using fiducial markers for frame alignments. Fiducials were corrected out from the final tomograms.

2.5 Fluorescence microscopy

2.5.1 Sample preparation

In general, cells grown for microscopy were cultured overnight in M9 minimal media, then a day culture of 5 ml M9 media was inoculated 1:100 (v/v). Typically, the day culture was grown to mid-log phase (OD_{600} : ~0.5), then cells were pelleted for 1min at 5,000 *g*, optionally labelled and washed with M9 (see below), and finally resuspended in a small volume of M9 media (typically 10-50 µl). Of the resuspended cells, 5 µl was typically spotted on a 1.5% agarose pad (w/v) prepared with M9 media, and a coverslip placed on top. These agarose pads were set on glass slides, and a glass coverslip pressed on top, to adhere to the Geneframe adhesive. Cells were imaged with the Oxford Nanoimager-S with a 100x/1.49 oil immersion objective lens, and a pixel size of 117 nm.

2.5.2 Imaging conditions

Typically, 200 frames were taken at a frequency of 10 Hz, with the first 100 frames under transillumination and the second 100 frames under fluorescence. For GFP-labelled samples, the 473 nm laser was used, with transillumination at 520 nm. For mCherry-labelled samples, the 560 nm laser was used. For ColN-AF-labelled samples, the 640 nm laser was used.

2.5.3 Image analysis

During image analysis, the first 50 frames were averaged as a representative transillumination image, and the first 50 frames from the point of fluorescence activation were averaged as a representative fluorescence image. During image processing, the “find peak maxima” function of ImageJ was used to save minicell coordinates under transillumination, and a 3x3 pixel square was measured for fluorescence intensities. MicrobeJ was used for whole cell segmentation and fluorescence quantification. Single molecule tracking was performed using the Trackmate plugin for ImageJ [147]. Sub-pixel localisation and 2 frame gap-filling (max 300 nm) was applied, and tracks of ≤15 frames were discarded. Tracks were analysed manually by producing MSD curves in Microsoft Excel and GraphPad.

2.5.4 Colicin-fusion protein (FP) labelling

ColN-GFP and ColN-mCherry (colicin-FPs) were characterised and provided by Dr Patrick Inns [148]. Primary labelling with ColN-GFP/ColN-mCherry was performed by addition of ColN-GFP to a final concentration of 1 µM, and incubating for 20-60 minutes. Cells were washed and optionally fixed with 4% formaldehyde (v/v) in PBS buffer pH 7.24, followed by three pelleting/wash steps, before resuspension in 5-30 µL PBS. Colicin-FPs were generally used to label cells at room temperature at 1 µM final concentration. Cells were labelled in darkness and washed 2-3 times after labelling steps, before fluorescence microscopy. Labelling was found to work in all media/buffers tested (M9, LB, PBS, TAE).

2.5.5 ColN-AlexaFluor (AF) import assay

For valinomycin and nigericin assays (and their controls), cells were pre-incubated in PBS pH 7.24, with 1 mM EDTA and 150 mM KCl to equilibrate. EDTA permeabilises cells to ionophores, while KCl salts are required for depolarisation in the case of valinomycin (absence of salts cause hyperpolarisation). Cells were then resuspended in the same buffer with the relevant ionophore, equilibrated for minutes, then incubated for 20 hours with 1µM ColN-AF at room temperature, in darkness. Labelling was followed by sample splitting, optional trypsin treatment, and fixation. For trypsin treatment, cells were incubated for 1 hour at 30 °C in PBS + 1 µM trypsin with shaking, then fixed in PBS + 4% formaldehyde (v/v) and washed thrice in PBS. Cells were immediately imaged in PBS, as described above.

2.6 Tol-Pal functional assays

2.6.1 Colicin toxicity assay

LB agar selection plates were cast, then 100 µl mid-log cell culture (OD_{600} normalised to 0.6) grown in LB was resuspended in 6 ml soft-LB agar (0.7% agar (w/v)), supplemented with antibiotics and inducers where appropriate. A soft-agar lawn was cast onto the surface of LB agar selection plates. After setting, 1 µl spots of bacteriocin dilution series were applied onto the lawn and allowed to dry, then plates were incubated overnight at 37 °C. For Colicin E9 and Colicin N, a fourfold dilution series was typically used, starting from 8 µM. For Colicin Ia, a fivefold dilution series was used starting from 10 µM. Native colicins were previously purified by colleagues (Dr. Melissa Webby, Dr. Nick Housden).

2.6.2 Time resolved colicin CFU assay

Time resolved killing assays were performed by growing cells in LB to mid-log phase (normalised to 1 ml LB at OD_{600}: 0.4), at which point growth was inhibited by addition of 20 µg/ml chloramphenicol (w/v). Cells were incubated in a benchtop shaker at 37 °C for 2 minutes with shaking, and then a 150 µl aliquot was taken. The aliquot was then spiked with ColE9 (100 nM final concentration) and the sample further incubated in the benchtop shaker. At each time point, 1 µl trypsin (1 mg/ml final concentration) was added to a 9 µl aliquot of cells and incubated for a further 30 minutes in the shaker, before storing the aliquot at 4 °C. The trypsin stock was freshly suspended in 50 mM ammonium bicarbonate at a concentration of 10 mg/ml. After all time points were acquired and trypsin-treated, a tenfold dilution series of each aliquot was prepared in a 96-well plate, diluted into fresh LB lacking chloramphenicol. For each dilution in series, a 2 µl aliquot was spotted onto LB agar, supplemented with kanamycin and ampicillin where appropriate. Three technical repeats were performed per dilution, and the experiment was biologically repeated three times per sample. The number of colony forming units (CFU) were estimated by multiplying the number of colonies counted by the dilution factor. The natural log of the CFU count was then plotted against time the sample had been incubated with colicin (prior to trypsin treatment). The translocation half-life was derived from the slope (-λ) of the data on this semi-log plot, using the formula: $t_{1/2} = \frac{0.693}{\lambda}$

2.6.3 Detergent sensitivity assays

Cells were first cultured to stationary phase in LB, with antibiotic selection where applicable. Cells were then day cultured (1:100) and grown to mid-log phase (OD_{600}: ~0.6). Cell density was normalised to OD_{600}: 0.2, and a tenfold dilution series prepared in LB, with selection markers. 1 µl of each dilution in series was spotted onto LB agar plates, supplemented with variable concentrations of SDS. The SDS was added during preparation of the LB agar in pellet form, and dissolved by autoclaving. No antibiotic selection was included on these plates, to accommodate the BW25113 control. Variable concentrations of arabinose were applied to the plates, to facilitate tuneable plasmid expression of different constructs of *tolA* in the PBAD24 plasmid. For BPA mutants, variable plates were made including BPA, rhamnose, arabinose and IPTG.

2.6.4 Western blot

During SDS assays cell aliquots were taken, normalised to a final volume of 500 µl at OD_{600}: 0.5. Cells were pelleted, the supernatant discarded, and the pellets stored at -20 °C. Cells were later lysed with BugBuster and boiled for 20 minutes in 1x SDS buffer. Boiled cell lysates were resolved by SDS PAGE (12% polyacrylamide gel), alongside ECL rainbow ladder (Amersham). The resolved proteins were then transferred onto a PVDF membrane using a Powerblot, for 45 minutes at 45 mA, at room temperature. The membranes were then blocked overnight (α-TolA) or 1 hour (α-GFP) with gentle rolling, at room temperature. The blocking buffer was Tris-buffered saline containing 0.1% Tween® 20 Detergent (TBST) and 8% milk powder (w/v). Membranes were then briefly washed with TBST and labelled with either α-GFP (1:1,000 v/v, overnight) or α-TolA (1:1,000 v/v, 1 hour). The labelling buffer contained 4% milk powder in TBST. After another round of washing with TBST, the membranes were labelled for 1 hour with a secondary anti-rabbit antibody, conjugated with horseradish peroxidase (1:1,000 v/v). This was again in the presence of 4% milk powder in TBST. The membranes were washed once more, then developed using an ECL chemiluminescence kit. Membranes were imaged at several time points post-development.

2.7 Computational analyses

2.7.1 Structural predictions

Secondary structural predictions were performed using PsiPred 2.0 & 3.2 [149]. PPII-helix propensity was predicted using PPIIPred [150]. Tertiary structural predictions of proteins were performed using AF2 on the Google Colab server embedded within ChimeraX [151], or later using the DeepMind AF3 server [100]. Typically, PDB templating and relaxation steps were enabled for AF2. Additional homologues of motor proteins were detected using FoldSeek, and identities were verified by proximity within Genbank [72, 152]. TMHs were predicted with TMHMM 2.0 where unclear [153, 154]. Cellular localisation and lipoylation were predicted using SignalP 6.0 [155].

2.7.2 Tol-Pal/TonB locus search (manual)

To detect and validate *tol-pal* loci in genomes either poorly represented in the pubMLST database or presenting as split systems, a manual search was also implemented. This search was performed by searching for "*species* TolB" in UniProt, then investigating the AF2-predicted structure (if available) for both TolB domains (33, 35). If both the N-terminal domain and the β-propeller were present, the corresponding genome was then accessed via Genbank [152], and inspected for nearby presence of *motA/ExbB/tolQ*, *motB/exbD/tolR*, *tonB/tolA*, *pal/ompA*, and *ybgF/cpoB/TPR*. If nearby proteins were labelled "hypothetical protein", the gene was opened in Uniprot to investigate the AF2 prediction. If no prediction was available in UniProt, the structure was predicted using AF3 [100]. For non-contiguous or poorly annotated genomes, the genome was searched using tBLASTn or BlastP using protein sequences from the nearest identified homologue (Fig. S2). The same genome accessions were probed for *tonB* genes at other genomic loci, to prevent mislabelling a *tolA* gene as *tonB*. UniProt was used to identify TonB/TBDT genes, then located within the same genome.

2.7.3 Tol-Pal/TonB locus search (algorithm)

To detect *tol-pal* loci within the pubMLST genome database, a search algorithm was designed to identify and flag genes encoding the MotA-ExbB-TolQ proton channels (PF01618), the PGB subunit of the motor MotB-ExbB-TolR (PF02472), the TolA-TonB force transducer (PF06519/PF03544/PF13103), β-propeller TolB (PF07676/PF04052), a PG-binding lipoprotein Pal (PF00691), a TPR repeat

(PF16331/PF13432/PF13174), and an acyl-thioesterase YbgC (PF03061). The output is available (Supplementary Spreadsheet). Genomes flagged for the minimal complement of genes (*tolQRAB* and *pal*) were inspected to confirm the presence of the expected gene products. In cases of poor annotation and/or low sequence homology, protein structures were screened using AF3 to check they were full length products with appropriate folds. Upon confirmation of a minimal *tol-pal* complement, the TolA protein sequence was analysed using PsiPred 3.2. Genomes of taxa which were not flagged by the search algorithm were searched using genome BLAST to detect the presence of TolB homologues, which were then manually checked for the remaining *tol-pal* complement within genomic vicinity. These loci were also screened by AF3 structure prediction. Manual curation of *tol-pal* loci is outlined in SI. 1. A similar algorithm was also written to scan the pubMLST database for PFAM domains corresponding to TonB (PF16031 (the TonB polyproline region) and PF03544 (the TonB-CTD), and TBDTs (PF00593 (the TBDT β-barrel) and PF07715 (the TBDT Plug domain). Outputs available in Supplementary Spreadsheet (SS. 1). Flagged species were then manually inspected (SI. 2).

2.7.4 Co-conservation analyses

Co-conservation analyses were performed using the RaptorX submission server and data were analysed in Microsoft Excel [156]. For each protein the pairwise inter-residue scores were ranked from highest to lowest to determine regions of highest interaction probability. The highest score observed per residue of each protein was coloured using a custom attribute script, which was imported into ChimeraX [157].

2.7.5 α-Helix:Proline ratio calculations

FASTA sequences for TolA and TonB proteins were input into the PsiPred 3.2 server with default parameters [149]. The output was saved as an SVG file (Supplementary Zip), and domain II was analysed for α-helix: proline ratio (SI. 1). The end of the predicted TMH was used to identify the beginning of domain II, and the final 100 residues were designated as domain III. Within domain II residues predicted to be α-helical vs the number of proline residues, were expressed as a percentage of the total number of residues in domain II. Scores were conditionally formatted to colour by relative α-helix: proline, by the maximum and minimum observed α-helix: proline scores (SI. 1-2; SS. 1).

3.0 Cellular assays of TolA-TonB chimeras

Much of the work in this chapter was published in Williams-Jones *et al.,* 2023 [131]. Work presented here is my own, with full permission from fellow co-authors. The stoichiometry of $TolQ_5R_2$ and $ExbB_5D_2$ is henceforth implied in “TolQR” and “ExbBD”.

3.1 Introduction

Since discovery of the *tol* genes in the 1960s the mechanism of the Tol-Pal system has been incrementally elucidated. The TolQR complex constitutes an IM motor that energises TolA, which recruits Pal to the septum by stripping TolB from the mobilised TolB-Pal complex (Fig. 1.9) [42, 91, 99]. TolQR likely drives this process through rotation, given its shared constellation of residues to those essential for flagellar rotation in MotAB (Fig. 1.4) [57-60, 68]. The same residues are conserved in ExbBD, and thus rotation is also suggested to underpin the Ton system [68, 80, 158]. Similarly, the force transduction proteins TolA and TonB share essential residues within their TMH that are essential for motor-binding [95, 159]. Moreover, TolA can promiscuously interact with ExbBD in a *ΔtolQR* mutant, and TonB can interact with TolQR in a *ΔexbBD* mutant via their conserved unilateral **Ser**-X_3-**His**-X_6-**Leu**-X_3-**Ser** (SHLS) motif (Fig. 3.1B) [95, 159-162]. It has thus been proposed that TolA and TonB utilise the same fundamental forces to drive their disparate functions.

The conserved three-domain architecture of TolA and TonB varies most apparently in the central domain (Fig. 3.1C) [13, 42, 91]. In *E. coli* and many related γ-proteobacterial species, the periplasm-spanning TolA domain II (TolAII) is predominantly α-helical [88, 163, 164] while TonB domain II (TonBII) is proline-rich, and is expected to form Type II polyproline (PPII) helices [165-168]. TolAII when fully extended is around 40 nm, whereas TonBII is 179 amino acid residues shorter than TolAII in *E. coli*, and yet is also predicted to span the periplasm (20-30 nm), with a total reach of around 30 nm. Given these differences in domain II length and composition, are there functional and mechanistic implications to this structural divergence?

3.1.1 The Tol and Ton systems are structurally homologous

Despite their divergent functions, TolA and TonB bind their targets at the OM by similar mechanisms through their conserved CTDs (domain III) (Fig. 3.1). The interaction of TolA with TolB at the OM is achieved through β-augmentation, as is that of TonB with the TonB-box of TBDTs [91, 169-171]. Both TolA and TonB then apply force to their respective partners perpendicular to the plane of the membrane; TolA dissociates TolB-Pal, while TonB dislodges the TBDT plug domain [13, 42, 75-79, 91].

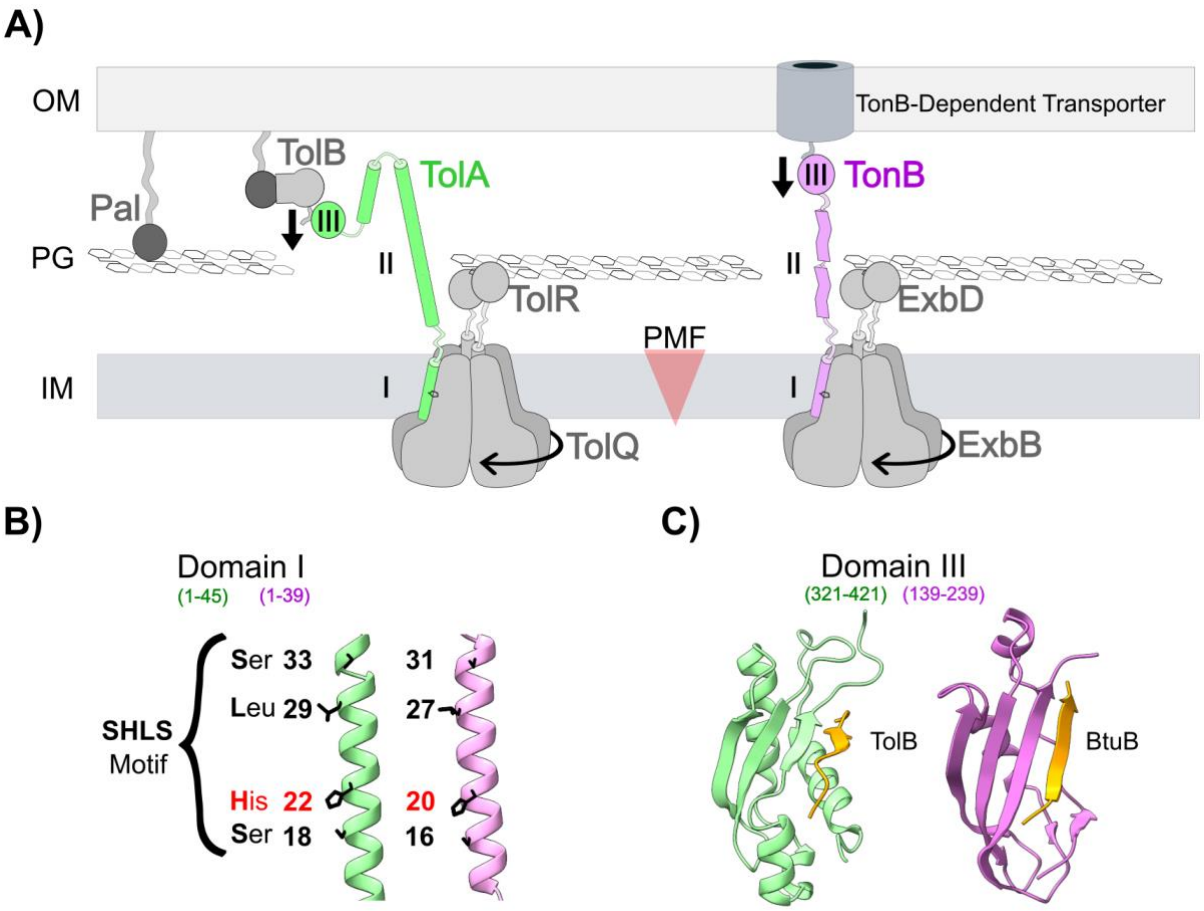


Figure 3.1 | Homologous aspects of the Tol and Ton systems suggest similar force transduction mechanisms. A) TolA and TonB each comprise three-domain structures (I-III), which transduce force from a motor complex to their respective OM-bound targets, to apply force (*downwards arrow*). **B)** TolA (*green*) and TonB (*pink*) domain I share a conserved SHLS motif on a single face of the TMH found in domain I. Of these residues, the conserved His residue is essential for force transduction, as mutation to alanine abrogates motor association. The TMHs depicted are taken from AF2 predictions. **C)** TolAIII and TonBIII both participate in β-addition interactions. Models depicted are *Pseudomonas aeruginosa* TolAIII in complex with the N-terminal peptide of TolB (*orange*) determined by solution NMR (PDB: 63SW), and *E. coli* TonBIII in complex with the TonB-box peptide from BtuB determined by X-ray crystallography (PDB: 2GSK) [77]. Figure adapted from Williams-Jones *et al.*, [131].

The majority of genes identified as *tolA* contain α-helical repeats within domain II, comprising repeating K, E, and A residues [172]. Unlike TolAII, TonBII is often predicted to be disordered by structural modelling such as AF, Rosetta and PsiPred

[71, 173]. However, these proline-rich regions of the protein are typical of sequences that form PPII-helices, comprising Pro-X_{1-3}-Pro motifs [174-176]. Despite this, certain structural prediction tools are more capable of modelling PII-helices, such as ESMFold and PPIIPred [150, 177]. Moreover, purified TonBII has been demonstrated to adopt PII-helices by double electron-electron resonance and circular dichroism measurements [165]. Notably, PPII-helices are also found in structural multimeric fibrous proteins such as collagen [178], elastin [179], and keratin [180]. In *E. coli* TolAII a helical hairpin is predicted by AF, whereas TonBII is predicted to form two oppositely charged PPII-helices by ESMFold (Fig. 3.2)[71, 177]. Charged residues are not required for PPII-helices, and may thus reflect a functional role within TonBII.

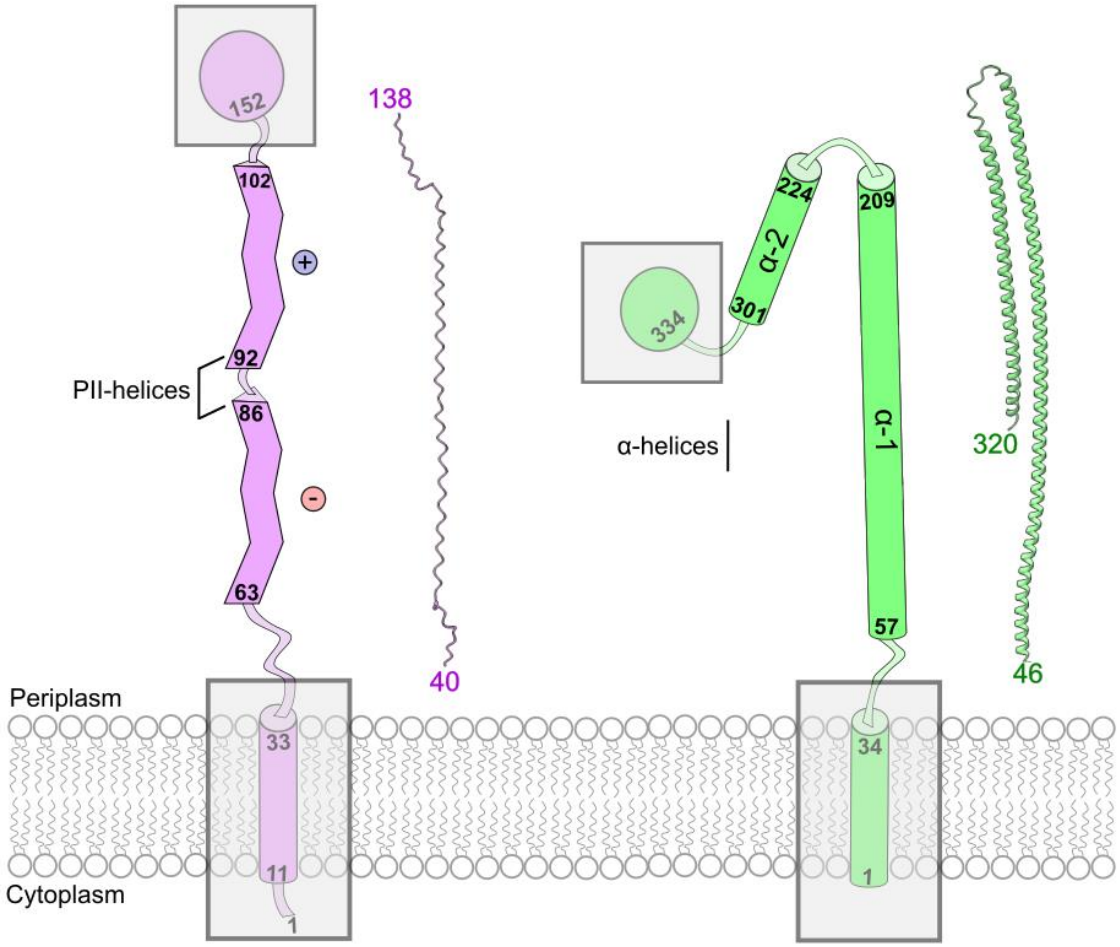


Figure 3.2 | Structural divergence of *E. coli* TonB and TolA periplasm-spanning domains. TonB (*pink*) features two oppositely-charged PPII-helices, separated by a short flexible linker. TolA (*green*) is predicted to form a helical hairpin in domain II. Domains I and III are greyed out, and approximate domain borders are numbered by *E. coli* sequences. The TonB PPII-helix was predicted by ESMFold [177], while the TolA helical hairpin was predicted with AF2 [71].

PPII-helices and α-helices exhibit markedly different structural geometries and properties, including opposite-handedness. Unlike the 3.6 amino acid residues per turn in α-helices, PPII-helices have 3 residues per turn [181, 182]. This produces a

narrower and more extended structure in PPII-helices, hence fewer residues are required to reach the OM than for α-helices (Fig. 3.3). This is reflected in the total number of domain II residues of *E. coli* TolA (*289*) vs TonB (*110*). Further, α-helices are significantly more rigid than PPII-helices because they have a greater cross-sectional area and a strong internal hydrogen bond network. This yields a higher elastic modulus than that of a PII-helix. PPII-helices lack such hydrogen bonds, with the exception of glutamine residues [175], hence they adopt a more flexible structure that has lower resistance to stretching, torsion, and bending. These differences in mechanical properties are likely to affect the conformational rearrangements that presumably occur during force transduction.

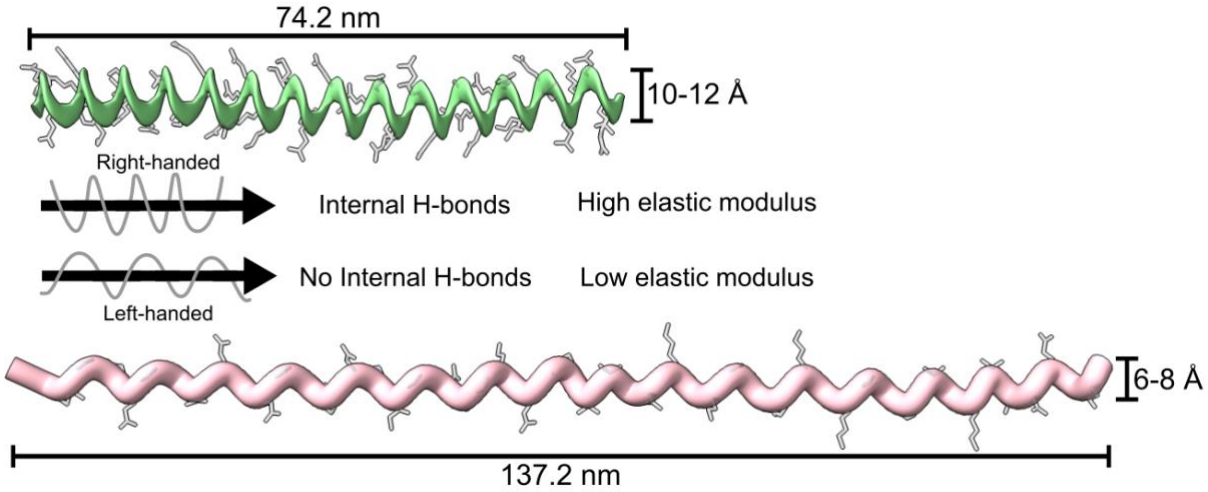


Figure 3.3 | Comparison of α-helix and PPII-helix properties. Fifty TolA α-helix (*green*) and TonB PPII-helix (*pink*) residues are shown (*residues 61-111*), predicted by AF2 and ESMFold respectively [71, 177]. The extended conformation of PPII-helices yields a greater end-to-end length than an α-helix with an equal number of residues. Side chains are indicated in grey.

3.2 Results

3.2.1 TolA and TonB chimeras transduce force

To assess the role that domain II secondary structures play in force transduction, I generated chimeric mutants of each. Because the function of TolA and TonB is determined by the CTD, mutants with TolAIII were deemed “group A”, and mutants with TonBIII were deemed “group B” (Fig. 3.4) [16, 17]. Group A chimeras were tested in a *ΔtolA* background, while group B chimeras were tested in *ΔtonB* background. To control for the differences in total residue numbers between TolAII and TonBII, mutants with additional TonBII repeats were also tested (“AABA/ABBA”). Additional mutants were designed to test how the absence of secondary structure affects force transduction, by replacing domain II with a 127-residue intrinsically disordered protein (IDP) sequence (“AIA/BIB”). Similarly, mutants with entirely deleted central domains were designed to test whether their decreased length would still permit force transduction (“ΔII”). The collection of mutants designed is indicated below (Fig. 3.4). The IDP sequence selected contains six repeats of a known disordered region of the bacteriocin ColE9, that is not involved in OM binding or translocation, and lacks cytotoxic activity [183].

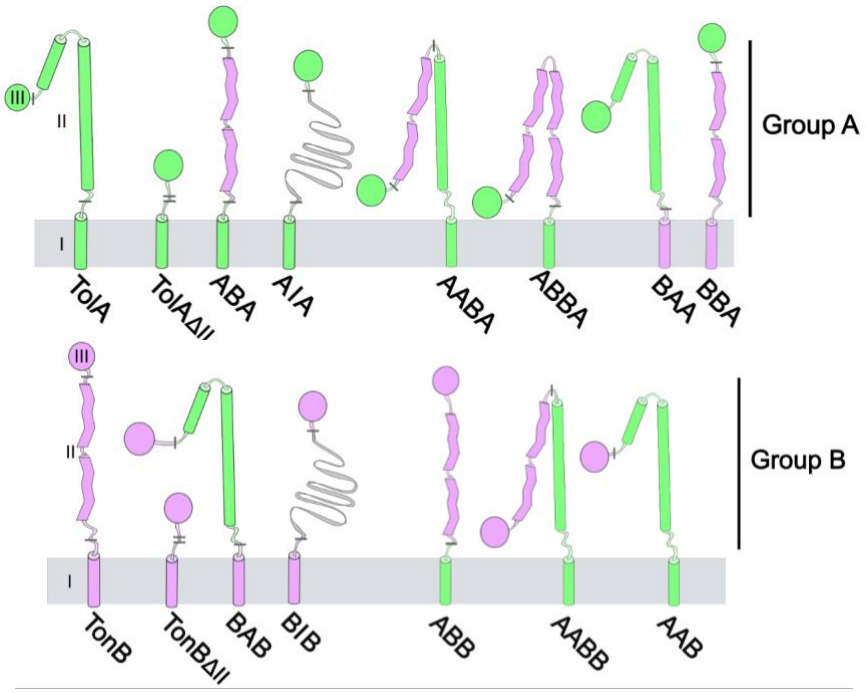


Figure 3.4 | TolA-TonB chimeric mutants categorised by CTD. Top row indicates mutants ending in TolAIII, which bind the TolB-Pal complex. Bottom row indicates mutants ending in TonBIII, which bind TBDTs. Domain boundaries indicated in Table 1.

I tested these mutants for their ability to transduce force, using colicins dependent on either TolA or TonB for import. Colicins are bacterial toxins that bind to the *E. coli* OM and interact with either TolA (Group A colicins) or TonB (Group B colicins) to energise their translocation into the cell [81-83]. Once inside the cell, which is presumed to be as unfolded polypeptides, they refold and perform their cytotoxic functions [15]. In this plate-based killing assay, a lawn of bacterial cells is spotted with a dilution series of colicins. After overnight incubation, the presence of clearance zones indicate that the cells are sensitive to the colicin, and therefore that the TolA/TonB chimera is capable of force transduction (Fig. 3.5). Clearance zones indicate that the WT controls were killed by both ColE9 and ColIA, whereas *ΔtolA* exhibits tolerance to ColE9 and *ΔtonB* exhibits tolerance to ColIA (Fig. 3.5A). This demonstrates that the deletion mutants cannot import colicins dependent on TolA and TonB, respectively. Upon transformation of those mutants with plasmids that express *tolA* and *tonB*, sensitivity to ColE9 and ColIA was restored respectively, confirming that colicin import is driven by TolA and TonB proteins (Fig. 3.5B-C; *left*). In contrast, expression of the same proteins featuring an additional His-Ala mutation in the TMH did not restore colicin sensitivity, because this essential histidine is required for interaction with their respective motors (Fig. 3.5B-C; *right*). Mutants featuring domain II deletions (TolA/TonB$_{\Delta II}$) were resistant to colicins, confirming that domain II is required to span the periplasm. In contrast, all full-length chimeras restored sensitivity to their respective colicins, indicating that each is capable of force transduction. This included the centrally disordered mutants AIA and BIB, which suggests that secondary structure is not essential for colicin import. To confirm that AIA and BIB import colicins by transducing force from ExbBD/TolQR to the OM, additional His-Ala mutants were generated (AIA$_{H22A}$/BIB$_{H20A}$). These were unable to import colicins, confirming that colicin import in AIA and BIB utilises their respective motor complexes. Of the colicin-sensitive mutants, AIA and BIB were noticeably less sensitive.

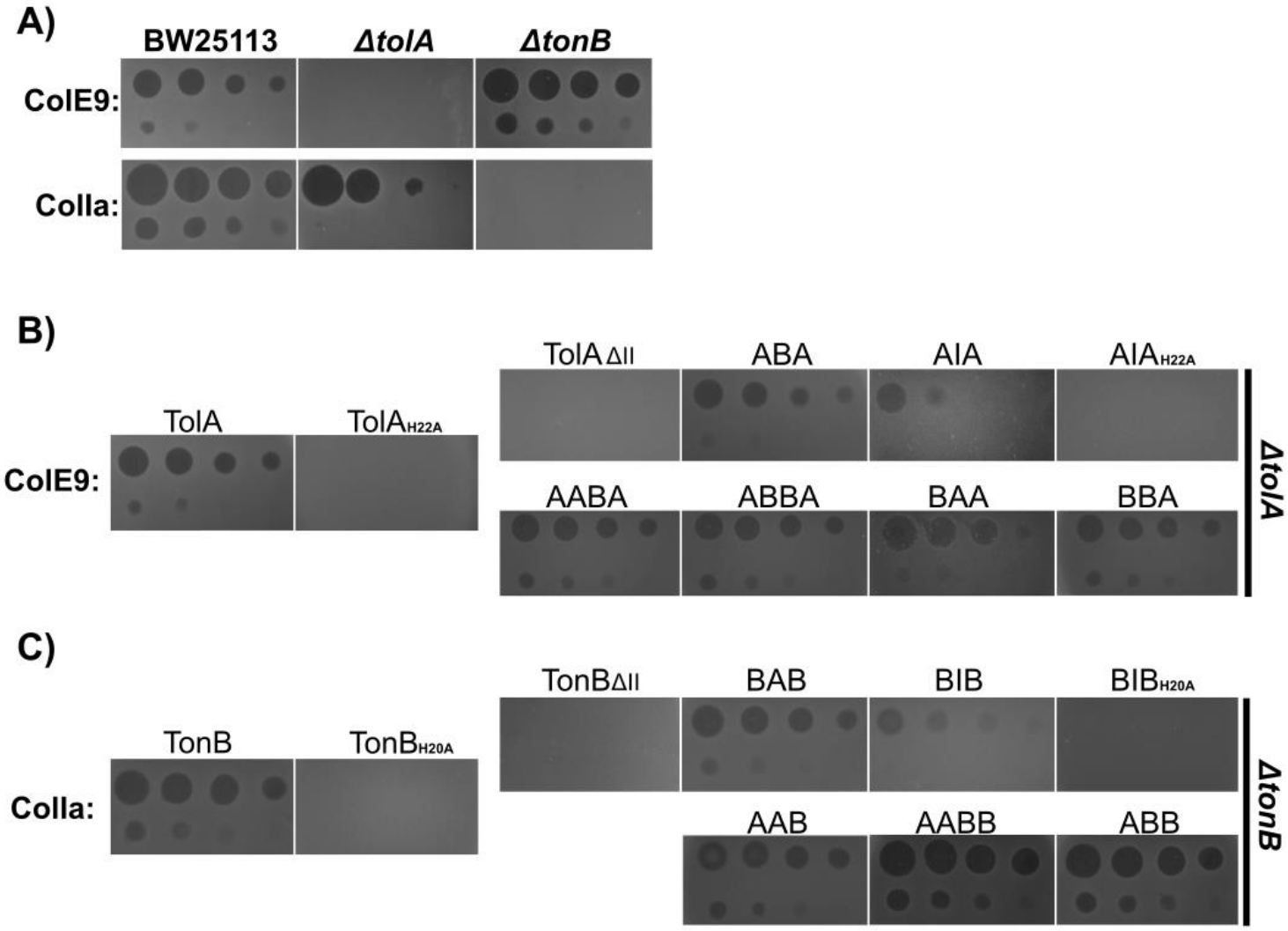


Figure 3.5 | Colicin killing assays demonstrate force transduction ability of TolA/TonB chimeras. **A)** *tolA* and *tonB* deletions render cells resistant to ColE9 and ColIA, respectively. Cells were treated with a fourfold dilution series of ColE9, starting from 8µM, or a fivefold dilution series of ColIA starting from 10 µM. For each plate, the highest colicin concentration was spotted at the top left, and the dilution series applied in two rows. **B)** Full length group A constructs restore sensitivity to ColE9. **C)** Full length group B constructs restore sensitivity to ColIA.

3.2.2 TolA chimeras demonstrate OM stability is a quantitative trait

The ability to import colicins is not equivalent to the native function of the Tol-Pal system, which is OM-stabilisation. In order to achieve this, TolA mutants must be both recruited to the septum and transduce sufficient force to recruit Pal [42, 93]. To test the ability of group A mutants to stabilise the OM, an SDS assay was performed (Fig. 3.6) [83]. In this assay, *ΔtolA* cells expressing these mutants were cultured, then spotted onto LB agar plates in a tenfold dilution series. The plates contained variable concentrations of SDS (0, 0.4, and 2% (w/v)). The ability to grow on plates containing detergent reflects the relative stability of the OM [83]. WT *E. coli* exhibited growth on 2% SDS, indicative of a stable OM (Fig. 3.6). In contrast, *ΔtolA* was unable to grow on SDS-impregnated plates, reflecting its destabilised OM. Complementation of *ΔtolA* with plasmid-expressed TolA restored SDS tolerance, confirming that it is sufficient to stabilise the OM. Complementation of *ΔtolA* with mutants conferred various levels of OM stability, as indicated by their different levels of growth on SDS-impregnated plates. BAA, which features a domain II nearly identical to WT TolA, yields growth on 2% SDS almost equivalent to WT TolA, indicating a stable OM. In contrast, most of the other mutants were unable to grow on 2% SDS but were able to tolerate 0.4% SDS, indicative of partial OM-stabilisation. This phenotype varied between constructs, with AABA, ABA and ABBA exhibiting moderate growth on 0.4% SDS, BBA exhibiting weak growth on 0.4% SDS, and AIA exhibiting no SDS-tolerance whatsoever. Taken together these data demonstrate that OM stability is a quantifiable phenotypic trait, and that the ability of TolA to fulfil its OM-stabilising role is heavily influenced by the secondary structure of domain II.

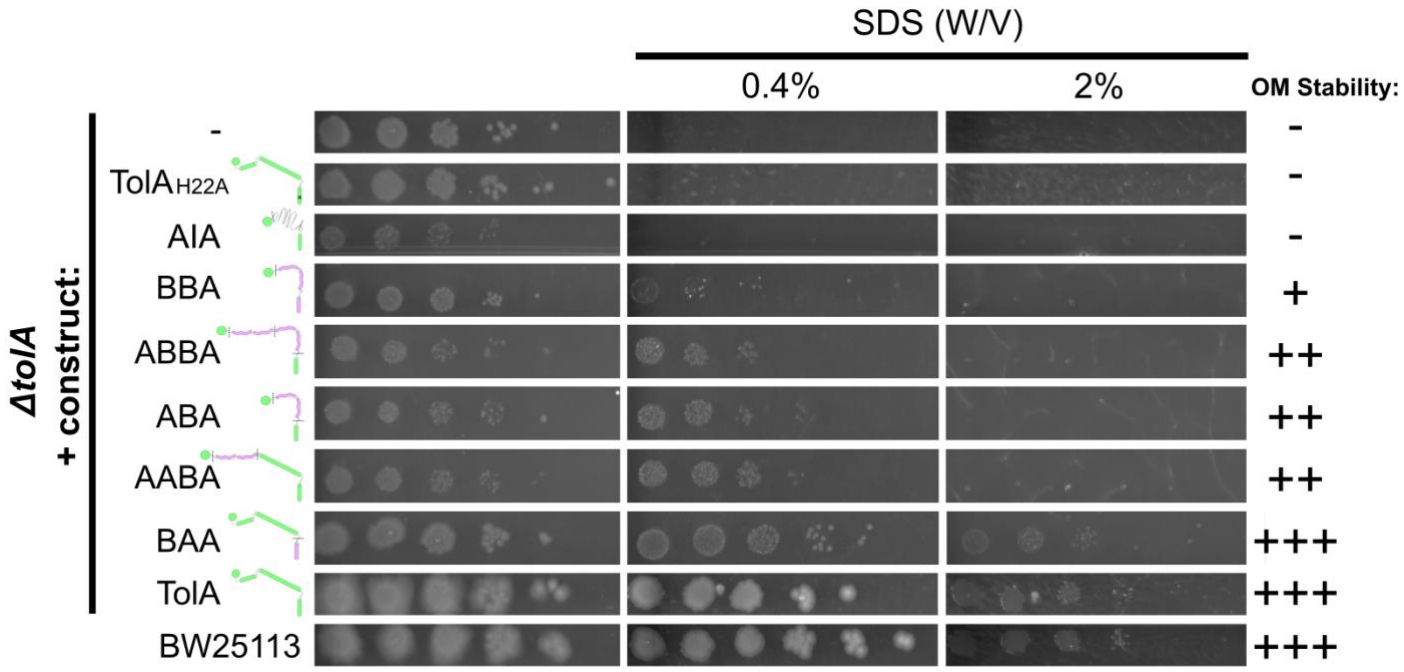


Figure 3.6 | Group A constructs show intermediate SDS tolerance. 1 µl of cells was spotted in a tenfold dilution series starting from the left of each plate, then incubated overnight at 37 °C. Equivalent aliquots of cells were spotted on each plate. Growth on 0% SDS indicates cell viability, while growth on SDS plates reflects both viability and OM stability.

3.2.3 TolA chimeras reveal a spectrum of septal Pal localisation

The current model of Tol-Pal-mediated OM-stability describes how Pal is recruited *en masse* to the septum to stabilise the OM and assist invagination [42, 93]. These data reveal a spectrum of OM stability, as indicated by the SDS assay (Fig. 3.6). To validate these findings, the ability of each mutant to recruit Pal to the mid-cell was investigated using fluorescence microscopy. Here, mutants were expressed in a *ΔtolA::pal-mCherry* mutant strain. These data reveal a spectrum of Pal-mCherry recruitment to the septum that reflects the relative ability of mutants to stabilise the OM (Fig 3.7). WT TolA was found to efficiently recruit Pal to the septum, as does BAA. Those mutants previously demonstrating intermediate OM stabilities also yielded intermediate levels of septal Pal recruitment (ABA, AABA, ABBA, BBA). Conversely, AIA and *ΔtolA* failed to localise Pal to the septum, reflective of their unstable OMs. To determine whether differences in Pal recruitment arise from poor recruitment of the TolA-like constructs to the septum, the expression and localisation of each construct was examined by fluorescence microscopy, via its N-terminal GFP fusion protein (Fig 3.8). These data suggest that in the *ΔtolA* background, all group A mutants are recruited to the septum. However, the extent of recruitment varies; TolA and AABA were the most septally-localised, followed by AIA, BBA, ABA, BAA then ABBA. Surprisingly, this does not correlate with their septal Pal recruitment ability particularly well. This is best illustrated by BAA, which yields the greatest septal Pal recruitment and OM-stabilisation of all mutants, despite being one of the least septally-recruited mutants. In contrast, AIA is relatively well recruited to the septum, yet confers no OM-stability or septal Pal recruitment. Meanwhile ABBA confers intermediate OM stability and septal Pal recruitment, despite the lowest expression and weak septal-trafficking. This suggests that although septal recruitment of mutants is required for septal Pal localisation and OM stabilisation, it is not their sole determinant.

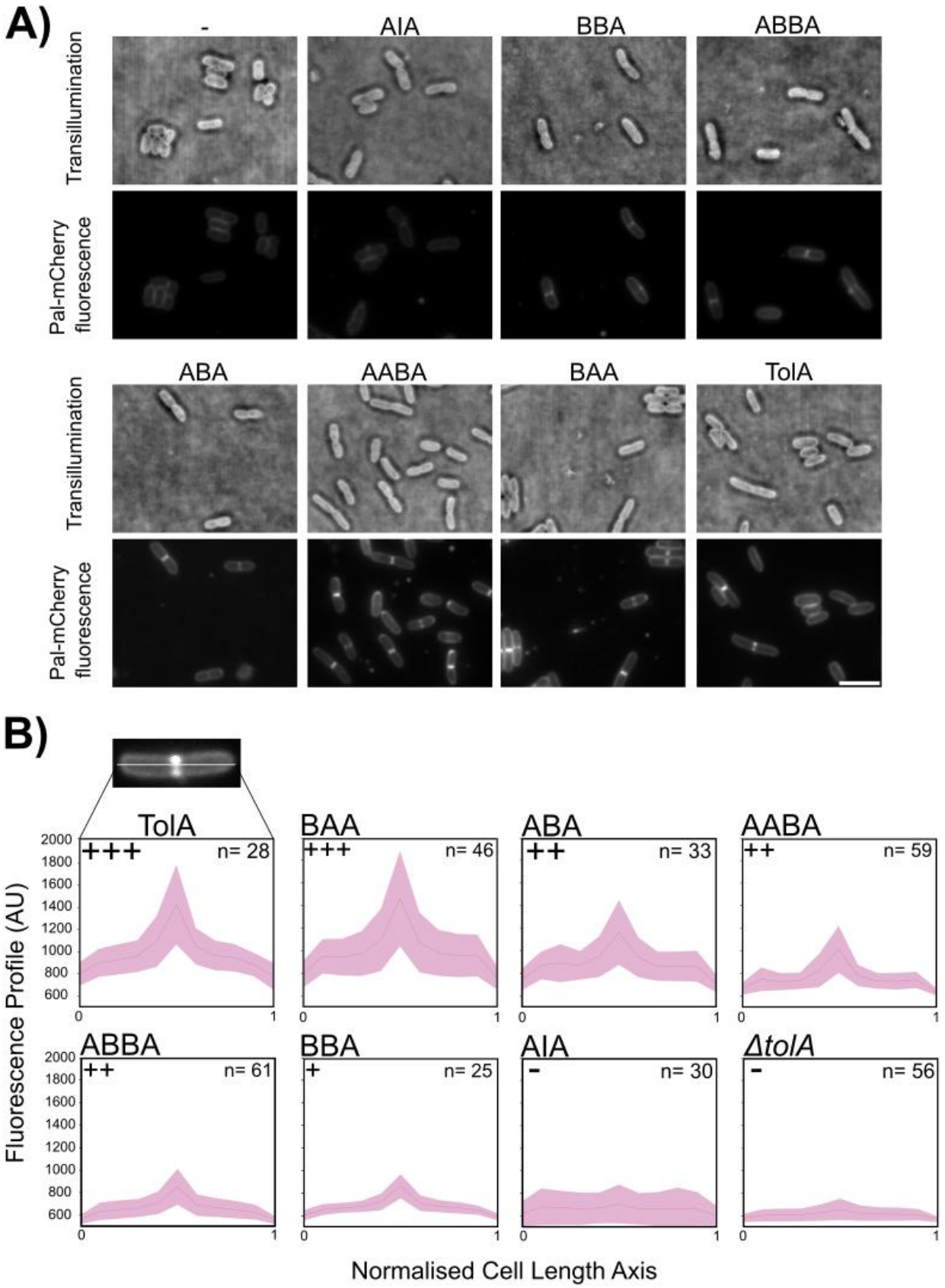


Figure 3.7 | Relative accumulation of Pal-mCherry at division septa varies between group A mutants. **A)** Cells were imaged by transillumination, and Pal-mCherry imaged at 560 nm. Micrographs are shown on an equivalent brightness and contrast scale. Scale bar indicates 5 µm. These data indicate that the ability to recruit Pal to the septum varies between TolA mutants. **B)** Cells in A) were segmented and labelled as either dividing or non-dividing cells. The dividing cells (featuring indentations indicative of septation) were quantified for fluorescence intensity profiles plotted along the medial cell length axis (*indicated above TolA*), to quantify the amount of Pal at the septum. Fluorescence profiles are on equivalent intensity axes. SDS tolerance is indicated as categorised previously (Fig. 3.5).

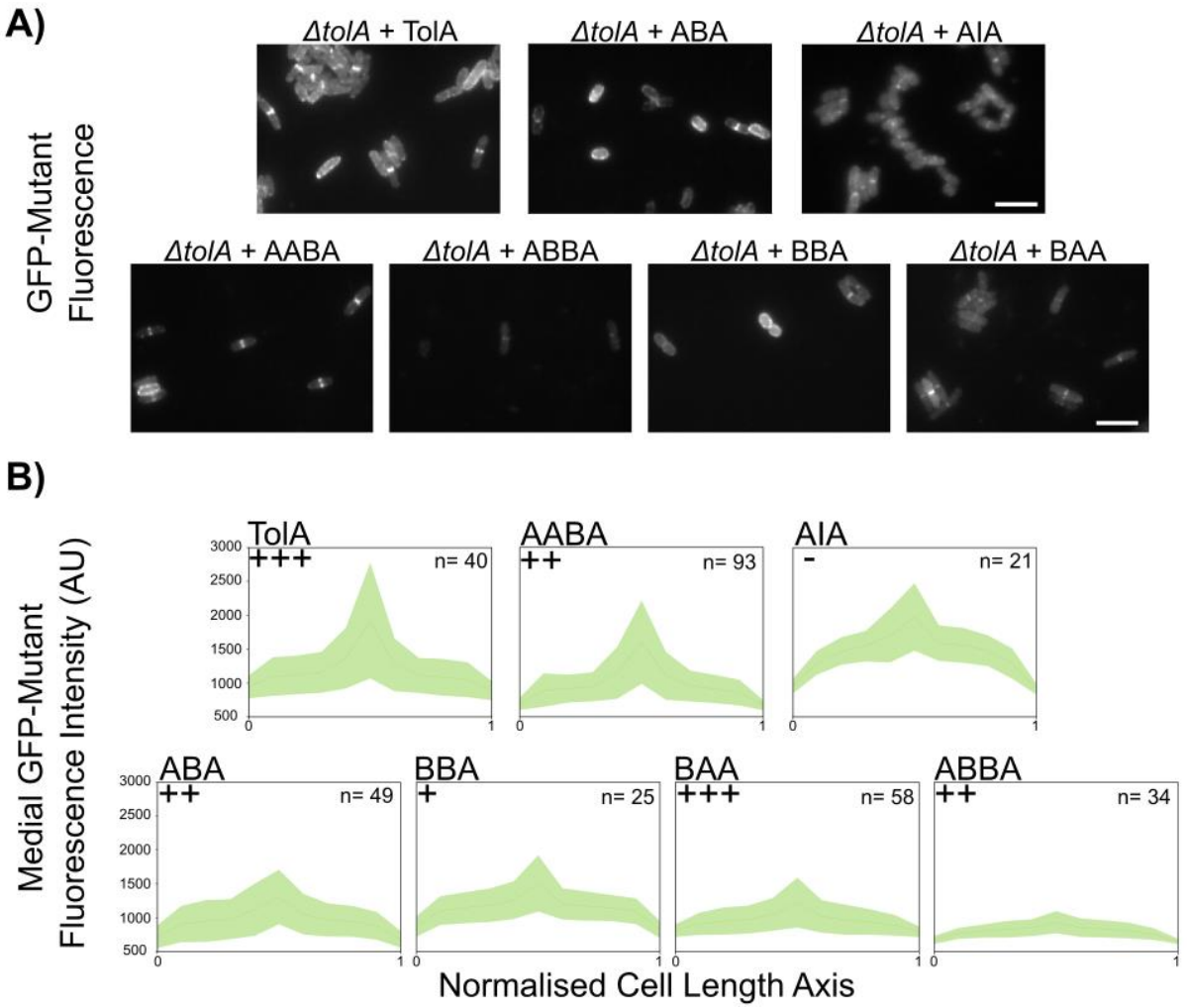
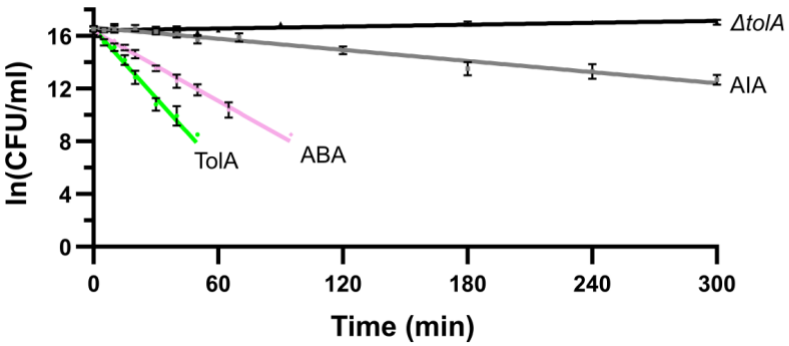


Figure 3.8 | Quantification of group A mutant localisation in *ΔtolA* cells. **A)** Cells expressing group A mutants, where a cytoplasmic GFP fusion is present at the NTD, were imaged with fluorescence at 473 nm, then segmented. Cells are shown on different brightness contrast settings for visibility. **B)** The dividing cells' GFP distribution profiles were plotted against the medial length axis as performed previously (Fig. 3.7B). GFP fluorescence profiles are shown on equivalent fluorescence intensity axes. Relative ability to stabilise the OM is indicated in top left of each panel.

3.2.4 TolAII secondary structure influences colicin uptake rate

These assays demonstrated that both OM stability and septal Pal recruitment are continuous traits. However, further work was required to explain how these intermediate phenotypes arise, given that each group A mutant demonstrated the ability to transduce force by importing colicins (Fig. 3.5). I proposed that these differences arose due to differences in efficiency of force transduction between mutants. Overnight plate-killing assays take place on a long timescale likely to mask inefficient import. To differentiate strong (efficient) force transduction from weak (inefficient) transduction, a reduced time frame for colicin-induced killing was measured. Time-resolved CFU assays were carried out in liquid culture where growth was stalled with chloramphenicol and cells expressing group A mutants were treated with 100 nm ColE9. At specific time points, an aliquot of cells was removed, treated with trypsin, to degrade colicins that had not fully imported, and treated cells were then diluted in fresh LB, and spotted on LB agar to grow overnight. Colonies were then counted to determine colony forming units (CFU) at each time point. Translocation is the rate-limiting step of colicin cytotoxicity, and as demonstrated here, follows pseudo-first-order kinetics [183]. The rates measured here are slightly lower than those previously estimated for WT TolA ($t_{1/2}$ ~ 2.5 mins), which may be explained by drag due to the N-terminal GFP fused to TolA here [183]. This assay revealed that WT TolA imports colicins twice as fast as ABA, and tenfold faster than AIA (Fig. 3.9), mirroring their respective OM-stability and septal Pal recruitment abilities.



Construct	$t_{1/2}$ (mins)	95% CI (mins)
TolA	4.0	3.6 – 4.5
ABA	7.8	7.2 – 8.5
AIA	49.1	43.3 – 56.8

Figure 3.9 | TolAII secondary structure influences ColE9 uptake rates. Translocation half-life was determined from the slope of ln(CFU)/time. The shorter the half-life, the faster each ColE9 molecule is imported and thus the more "efficient" the construct is deemed at force transduction. N= 3 biological repeats per sample.

3.2.5 Group B constructs complement *ΔtonB* growth rates

To assay whether group B constructs functionally complement *ΔtonB* mutants, a growth assay was performed. Here, *ΔtonB* cells were transformed with group B constructs, and the ability of those mutants to uptake substrates such as iron-bound siderophores and vitamin B12 was indirectly assayed through growth rates in minimal media (M9) [27, 73]. Here I observed that both α-helical and PPII-helical secondary structures are sufficient to drive physiological levels of substrate import, as indicated by complemented growth rates (Fig. 3.12). As with group A, the ability of group B mutants to fulfil native functions by transducing force is influenced by domain II secondary structures, which is likely to determine the import rate of group B colicins.

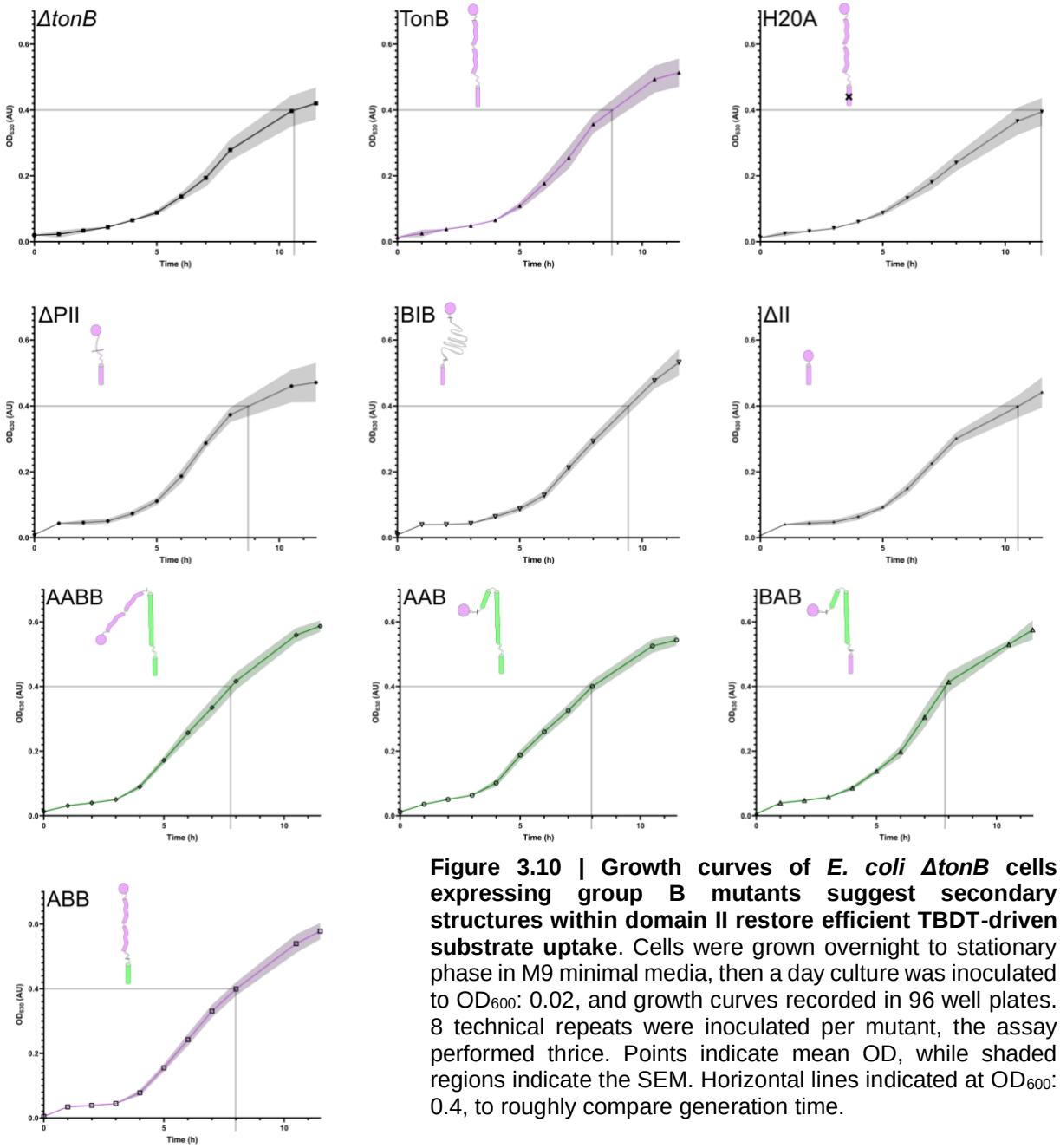


Figure 3.10 | Growth curves of *E. coli ΔtonB* cells expressing group B mutants suggest secondary structures within domain II restore efficient TBDT-driven substrate uptake. Cells were grown overnight to stationary phase in M9 minimal media, then a day culture was inoculated to OD_{600}: 0.02, and growth curves recorded in 96 well plates. 8 technical repeats were inoculated per mutant, the assay performed thrice. Points indicate mean OD, while shaded regions indicate the SEM. Horizontal lines indicated at OD_{600}: 0.4, to roughly compare generation time.

3.3 Discussion

3.3.1 Domain II structures influence force transduction efficiency

These data demonstrate that regardless of domain II secondary structure, force transducers can apply force at the OM if they have sufficient length to span the periplasm. However, the degree of force transduction varies greatly between mutants. Of those generated, ABA and AIA are key examples. In each mutant the conformational rearrangements that occur during force transduction likely vary due to the physical properties of their various secondary structures (Fig. 3.11). Biophysical measurements of TonBII confirm that this proline-rich sequence adopts an extended PPII-helical conformation (~15 nm), that likely reaches the OM more efficiently than a random coil [165]. While AIA features only 30 more residues than ABA, AIA is approximately sixfold less efficient at importing colicins (Fig. 3.9). This is likely due to differences in the intrinsic structural properties of the PPII-helix compared to the random coil. Similarly, TolA was found to import colicins twice as fast as ABA suggesting that the mechano-structural properties of α-helices confer even greater efficiency than PPII-helices.

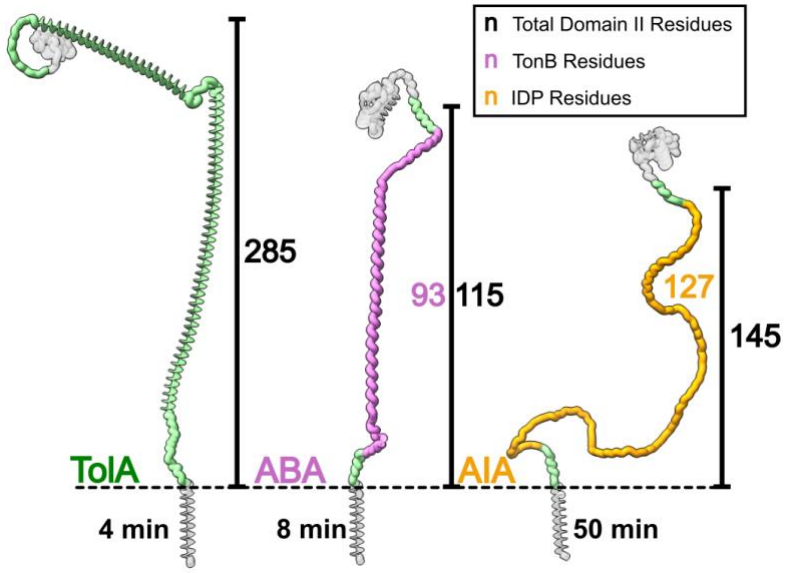
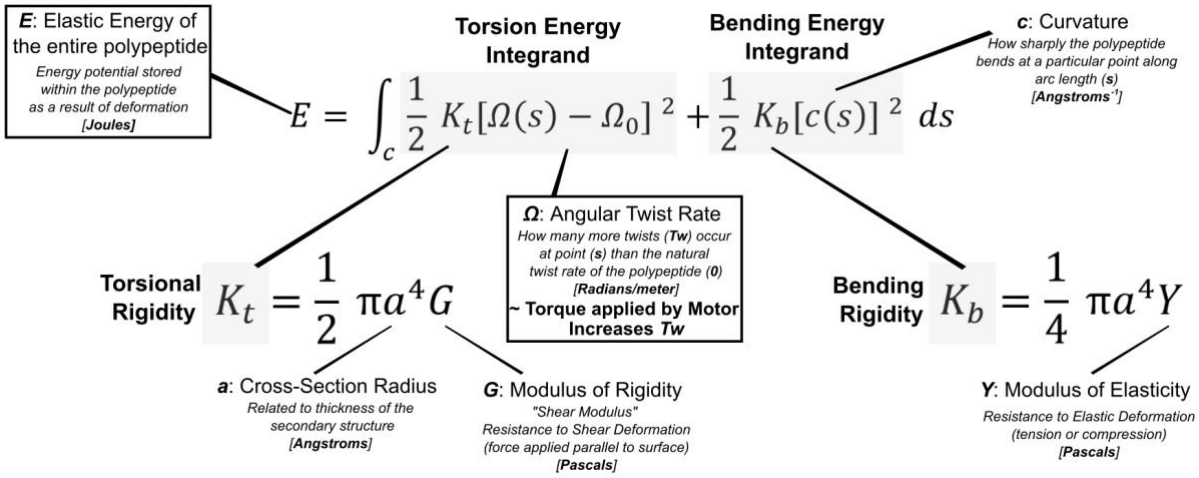


Figure 3.11 | Structural models of TolA, ABA and AIA. These are AF3 prediction models, except for ABA where the ESM-fold model for TonB residues 61-111 is integrated. Bonds were rotated in disordered regions to reorient each protein. These proteins vary in length and structure, with ABA occupying a more extended conformation than AIA. The total number of domain II residues is indicated (*black*), as well as those from TonBII (*pink*) and the ColE9 IDP (*orange*). Domain II regions derived from TolA are indicated (*green*) as well as domains I and III (*grey*). The dashed line indicates the IM border. These proteins vary in force transduction efficiency (*colicin translocation half-lives indicated*) (Fig. 3.9).

As the TolQR motor is likely rotary in nature, this suggests that these force transducers are torque-driven. Interaction with TolQR was indirectly confirmed for TolA and AIA through their H22A variants, which became incapable of force transduction. Assuming that force transducers utilise torque, we can better delineate the specific physical properties of domain II secondary structures likely to influence force transduction. Notably, α-helices have the greatest rigidity and cross-sectional area, conferring the greatest resistance to torsion and bending. In contrast, the random coil has the lowest rigidity and cross-sectional area, and the least resistance to torsion and bending. Because α-helices are the stiffest and rod-like, they provide the most resistance to rotation, whereas the flexibility of a random coil provides the least resistance to rotation. Thus, more twists are absorbed by a random coil than an α-helix. The elastic potential of a polypeptide can be described in terms of torsion and bending, which suggest that resistance to torsion and bending yields greater elastic potential (Eq. 1).



Equation 1 | Elastic Potential is the sum of Torsion and Bending energies, influenced by the physical properties of secondary structures. Constants and variables are annotated with their meanings, properties and units. *Torsion* and *Bending* are both quantifications of stored energy, measured in Joules. Rigid, structured polypeptides with thicker diameters tend to have greater *torsional* and *bending rigidity* (K_t and K_B), due to their increased cross-sectional area (*A*), and greater moduli of rigidity (*G*) and elasticity (*Y*) compared to a random coil of equivalent length. These structural parameters thus provide greater resistance to torque applied by the motor. As the motor rotates domain II, the number of Twists (*Tw*) increases, increasing the angular twist rate (*Ω).* As *Ω* increases, the total elastic energy of domain II (*E*) increases. Compared to α-helices, a random coil has lower values for K_t, K_b, *A*, *G* and *Y*, resulting in a lower total value for *E* for an equivalent number of twists. This suggests that the random coil "absorbs" twists due to its low resistance parameters. Additionally, for a longer polypeptide each twist yields a lower average *Ω* value, because each twist is divided by a greater arc length. This suggests inserting additional disordered residues introduces inefficiency. Equation adapted from Ricca [184].

3.3.2 The Tol-Pal system stabilises the septal OM

Through these different TolA mutants I also establish that OM stability is a quantitative trait, proportional to the accumulation of Pal at the septum. These data suggest that the rate of colicin uptake reflects the rate at which Pal is deposited at the septum, which in turn influences the relative stability of the OM (Fig. 3.14A,C). This model also suggests that in *ΔtolA,* the septum is the site of OM instability. This is supported by observations that outer membrane vesicles (OMVs) often form at the septum in *ΔtolA* mutants. This is well-documented, with electron micrographs showing that large OM detachments primarily occur at the septum in both *ΔtolA* and *Δpal* mutants, compared to the small blebs that occur ubiquitously in *Δlpp* and *ΔompA* [185-188]. Together these data suggest that the proportion of Pal at the septum is inversely correlated with septal OM detachment and vesicle formation. Moreover, the blebbing phenotype concomitant with an unstable OM explains the observation that mutants conferring low OM-stability tend to have lower total Pal-mCherry fluorescence, as Pal-mCherry is lost from the cell in OMVs.

I also demonstrate that recruitment of Pal to the septum is achieved by TolA, although the exact mechanism for this remains unknown. The ability of the centrally disordered construct AIA to import colicins suggests that the secondary structure of TolAII is dispensable for force transduction, but contributes dramatically to its efficiency. The loss of killing observed with AIA_{H22A} confirms that AIA functions through interaction with the TolQR motor. This therefore led me to propose a model where TolA transduces force by directly connecting OM targets to TolQR rotation (Fig. 3.12), as has been suggested for the Ton system [80]. The same trend was observed for BIB/BIB_{H20A}, supporting the hypothesis that the Ton and Tol-Pal systems utilise fundamentally similar mechanisms for force transduction.

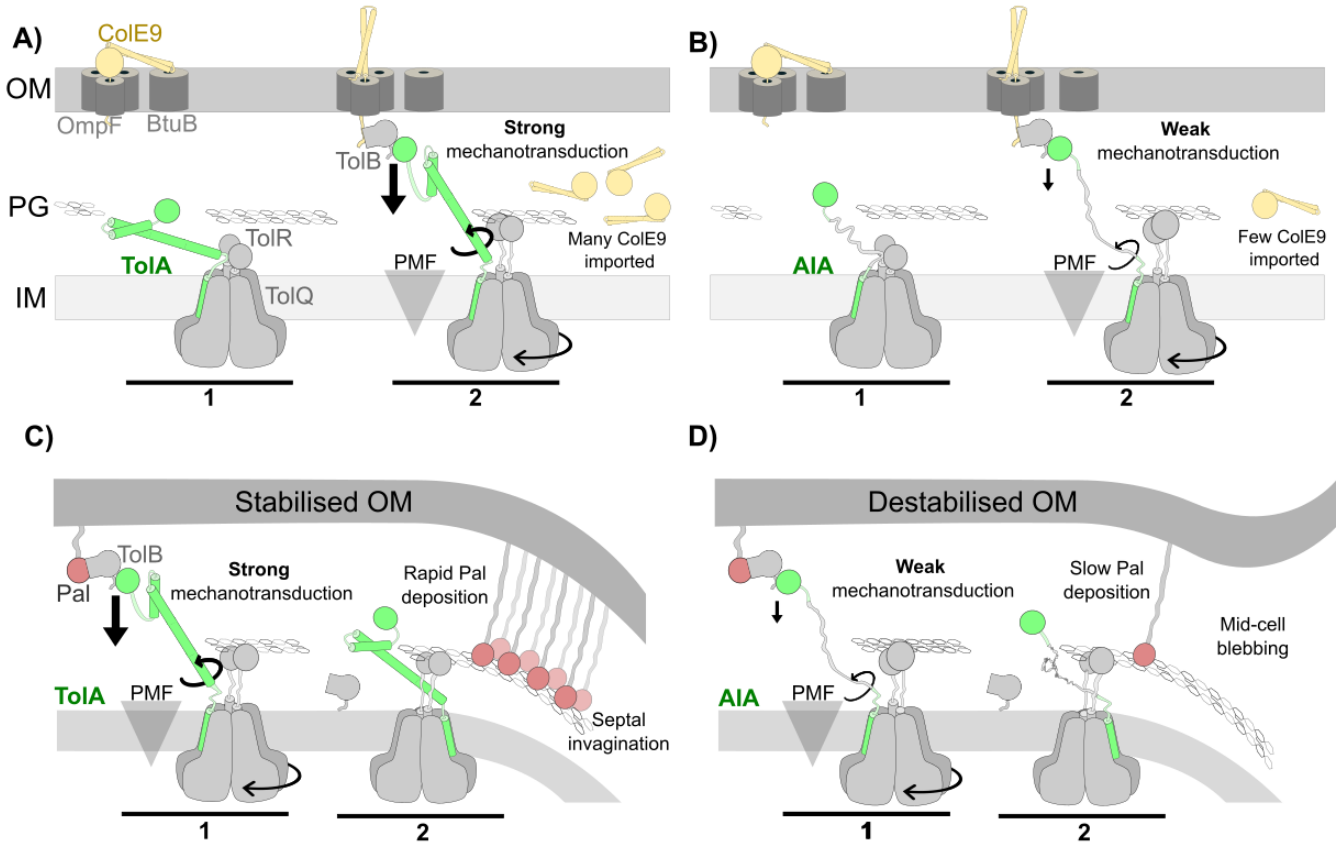


Figure 3.12 | TolA efficiency drives colicin import, septal Pal accumulation, and OM stability. **A)** TolA rapidly imports colicins. **I)** TolA engages TolQR below the PG layer. **2)** TolQR energises conformational rearrangements within TolAII to apply tensile force to the TolB-ColE9 complex (*large arrow*) and entry of unfolded ColE9 to the periplasm from where it refolds and causes cell death. The precise conformational changes that take place during force transduction are currently unknown but may, for example, involve a combination of TolA refolding and/or wrapping about the stem of TolR, together generating tensile force perpendicular to the OM. **B)** AIA imports colicins slowly. **1)** AIA interacts with TolQR. **2)** Force transduction via AIA occurs less efficiently than via TolA (*small arrow*). This is sufficient to import some cytotoxic colicins, but their passage across the OM is tenfold slower than for TolA. **C)** TolA rapidly recruits Pal to the division site, stabilising and invaginating the OM at the septum. **I)** TolQRA is trafficked to the division septum to capture mobilised TolB-Pal complexes. **2)** TolA efficiently dissociates TolB-Pal complexes, recruiting many Pal monomers that stabilise the OM. **D)** AIA exhibits poor Pal recruitment at the septum, which fails to stabilise the OM. **1)** AIA is a weak/slow force transducer. **2)** The inefficiency of AIA force transduction yields slow Pal accumulation at the septum, resulting in OM destabilisation and blebbing. The model assumes that the force requirements of colicin import are fundamentally similar to those involved in the dissociation of the TolB-Pal complex. Figure published in Williams-Jones *et al.,* [131].

3.3.3 The TolA-TolB interaction may comprise a catch-bond

The primary mechanistic role of TolA is dissociation of the TolB-Pal complex, which occurs via the TolA-TolB β-addition interaction. Due to the apparent essential role of PMF and tensile force that were shown in this study (Figure 3.12), it suggests that the TolA-TolB β-addition may form a catch bond [189]. These bonds are formed where a protein and a ligand form an interaction that tightens under tension, such as the bacterial surface adhesin FimH, which also utilises β-addition [190-192]. Catch-bonds are analogous to a finger trap toy, where pulling tightens the bond but relaxation permits dissociation. A TolA-TolB catch bond would mean that TolA only dissociates after TolB has been plucked away from Pal, and no further force can be applied to TolA-TolB due to mechanical resistance dissipating once the complex is no longer tethered to the OM (Fig. 3.13). Further biophysical work is required to test whether TolA-TolB forms a catch-bond, which are typically measured by AFM force curves and MD simulations [193, 194].

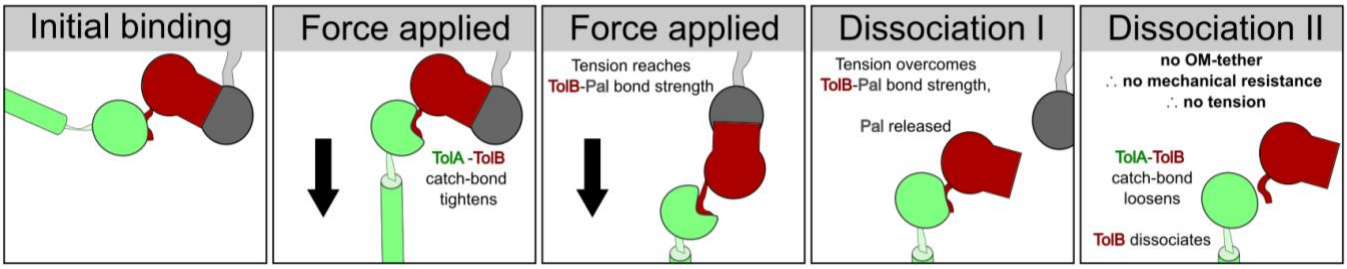


Figure 3.13 | A catch-bond model of TolB-Pal dissociation. TolAIII (*green*) interacts directly with an exposed β-strand of TolB (*red*) from the TolB-Pal complex by β-addition, to liberate Pal (*grey*). The TolB-Pal complex is dissociated via force applied through TolAIII, likely proceeding via a TolA-TolB catch-bond. The site of TolB that TolA binds is expected to exist in conformational exchange between sequestered and exposed states [43].

The interaction of TonB with TBDTs is equally likely to comprise a catch-bond, given the homologous β-addition interaction mode. Unlike the TolA-TolB catch-bond, the TonB-TBDT catch-bond opens the pore in the TBDT target, and likely dissociates after the tensile force exceeds the strength of the TonBIII-TBDT interaction, rather than due to loss of resistance. Here, the growth assay performed for group B mutants was an indirect assay used only to demonstrate whether chimeras are functional for ligand import, but is insufficient to measure the kinetics of TonB-dependent substrate import (Fig. 3.12). A radiolabelled ligand import assay or CFU assay with colicin B would better assess whether the kinetics rules determined in Group A mutants hold true.

3.3.4 A “twist-to-writhe” model of tensile force generation

The TolA/TonB chimeric mutants provide unique insights into the mechanism through which they apply tensile force at the OM. Firstly, the ability of the centrally disordered mutants, AIA/BIB, to import colicins at all suggests that in both cases the motor complex transduces force through synchronous tension. That is to say that TolA/TonB rotates at the same time as it is bound to its OM target, which drives contraction of domain II. If this were not true, then the disordered mutants AIA and BIB would not be capable of colicin import. Previous work suggested a hinge-opening mechanism where TolQR drives extension of TolAII, followed by collapsing back to a helical hairpin upon interaction with the target (Fig. 3.14) [97]. This was predicated on the observation that translocation-incompetent ColE9 immobilises GFP-TolA, suggesting that TolA extends to bind the bacteriocin. When PMF energisation was inhibited by chemical depolarisation TolA remained mobile, suggesting that it no longer extended to bind the bacteriocin [195]. However, if colicins also form a catch bond with TolA, this explains why TolA is not tethered to the colicin in the absence of PMF/tension.

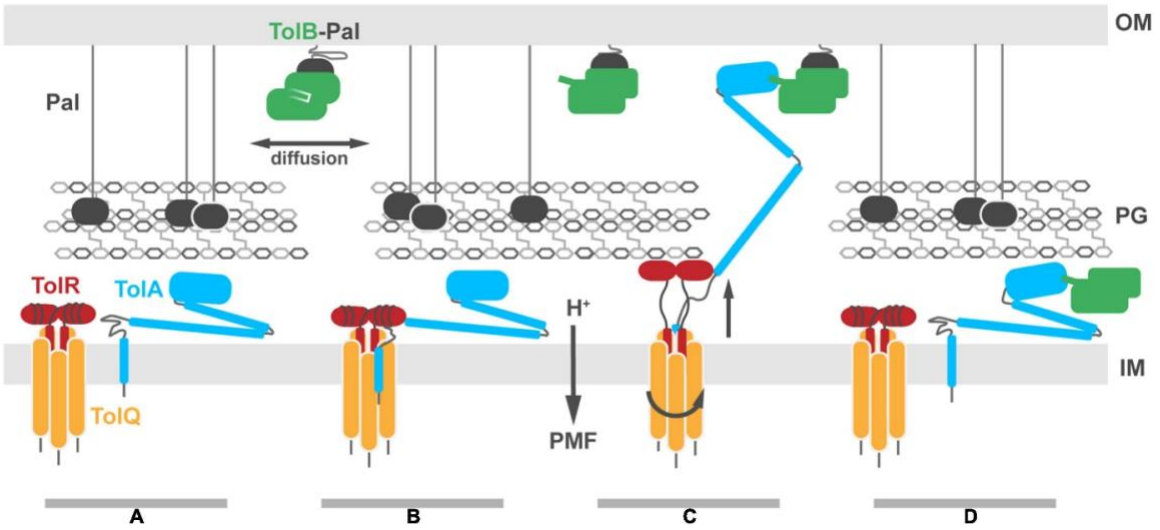


Figure 3.14 | Energised extension model of Tol-Pal force transduction. **A)** The TolQR motor and TolA diffuse freely within the IM. The abundant OM lipoprotein Pal is predominantly bound to the PG, stabilising its connection to the OM. A sub-population of Pal molecules are bound to TolB, blocking association of those Pal molecules to PG and enhancing their diffusion in the OM. **B)** TolA associates with the TolQR motor, inserting between monomers of TolQ. **C)** In response to proton flux, TolR extends and binds the PG layer, allowing rotation of the TolQ helices, which in turn drives extension of the TolA helical hairpin through pores in the PG layer. The TolB NTD exhibits a sequestered and an exposed conformation, which exist in dynamic equilibrium. TolAIII associates with the N-terminus of TolB in its complex with Pal, when the TolB NTD enters its exposed conformation. **D)** Proton flux drives retraction of the TolR periplasmic domain away from the PG. Collapse of the TolA helical hairpin toward the IM could be either through coupling to this structural change or refolding of the hairpin, following dissociation of TolA from the motor. We speculate that retraction of the TolA hairpin provides the driving force for dissociating TolB-Pal complexes at the OM and pulling TolB below the PG layer. The figure does not show recruitment of the TolQRA complex to the divisome which liberates Pal at division sites to stabilise and invaginate the OM. Figure published previously [97].

The observed ability of AIA/BIB to import colicins is incompatible with some aspects of this model, because the disordered and flexible central domain could not feasibly be extended through rotation in the way a structured domain might. The alternative hypothesis is that TolA/TonB randomly extends across the periplasm and when binding to the OM target occurs at the same time as torque is applied via the motor, torsion accumulates within TolAII/TonBII. This causes twisting that drives TolAII/TonBII contraction. Peptide backbones are routinely modelled as open ribbons, because peptide bonds are planar and cannot freely rotate [196-198]. When a ribbon is twisted from one end while pinching the other, this shortens the end-to-end length of the ribbon by forming *writhes* (*see Ricca,* [184]). The accumulated twists within TolA/TonB would rapidly unravel when domain III is no longer constrained, returning the protein to to its relaxed, starting state. Crucially, this mechanism is equally applicable to any domain II comprising any combination of α-helices, PPII helices or random coils. This is illustrated below with TonB, and assumes that simultaneous tethering of TonB at the IM and OM applies topological constraints, as suggested by colicin trapping (Fig. 3.15) [195]. It is not currently known whether the pentameric or the dimeric motor components apply torque to TonB, the latter of which was recently proposed (discussed in chapter six) [199].

The ability of both AIA/BIB to import colicins indirectly supports synchronous rotation and contraction, but this is difficult to test directly. One means to identify the conformational rearrangements that occur within force transducers is explored in chapter seven, where we trap TolA/TonB in their energised conformations to analyse their structures by cryo-ET.

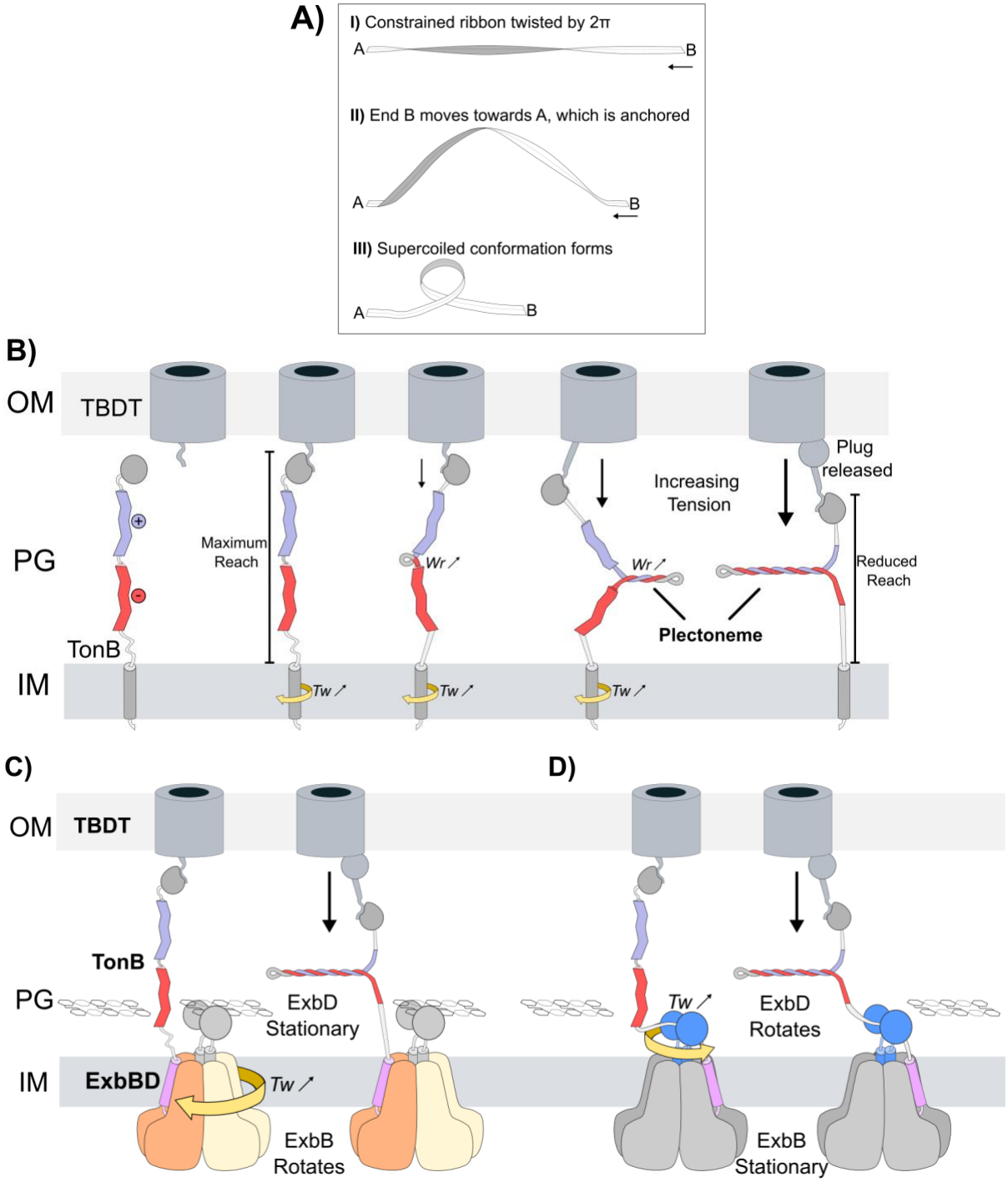


Figure 3.15 | A twist-to-writhe model of tensile force generation. A) Application of torsion to a ribbon with constrained ends causes writhing and contraction (*adapted from Au* [200]). **B)** TonBII features negatively (*red*) and positively (*blue*) charged PPII-helices connected by a short linker. In its non-constrained, non-twisted state, TonB exists in an extended conformation which spans the periplasm. The distance spanned (***Reach***; ~30 nm) is close to its theoretical maximum. When TonB binds to its OM target via diffusion, this spatially constrains TonB. As TonBI rotates via ExbBD, the number of twists within TonBII increases accordingly (***Tw↗***). As the number of twists in TonBII increases, torsional stress accumulates. Consequently, the centre of TonBII relieves torsional stress by writhing (***Wr***) to form a *plectoneme*. As the plectoneme grows with *Tw*, the total reach decreases, thus applying tensile force at the OM to dislodge the TBDT plug domain. The flexible properties of PPII-helices and random coils theoretically permit writhing with little resistance. TolAII possess α-helices, which are unlikely to writhe in exactly the same way. It is more likely that entire α-helices form a tightly-wound coiled-coil plectoneme as *Tw* increases (Fig. 3.16). TonBIII likely dissociates from the TBDT plug domain when the tensile force exceeds the strength of the TonBIII-TBDT catch-bond. **C)** Application of torque via the ExbB pentamer. ExbD binds PG and remains stationary. Axis of rotation is inferred from MotAB. **D)** Application of torque via the ExbD dimer. ExbD PG-binding activity is inhibited by its interaction with TonB, and clamps onto TonB. Rotation of the ExbD dimer applies torque to TonB, while ExbB remains stationary. The axis of rotation is inverted from B), because the stationary component is switched.

Applying this model to force transducers explains observations that short truncations are tolerable, because torsion can occur regardless of short deletions, as long as TolAIII/TonBIII can still reach the OM [78, 201]. If we apply these concepts to TolA a similar conformational rearrangement is suggested, but because α-helices are thicker and stiffer than PII-helices, their secondary structure is likely better maintained. This may result in a wider plectoneme comprising a tightly-wound coiled-coil (Fig. 3.16).

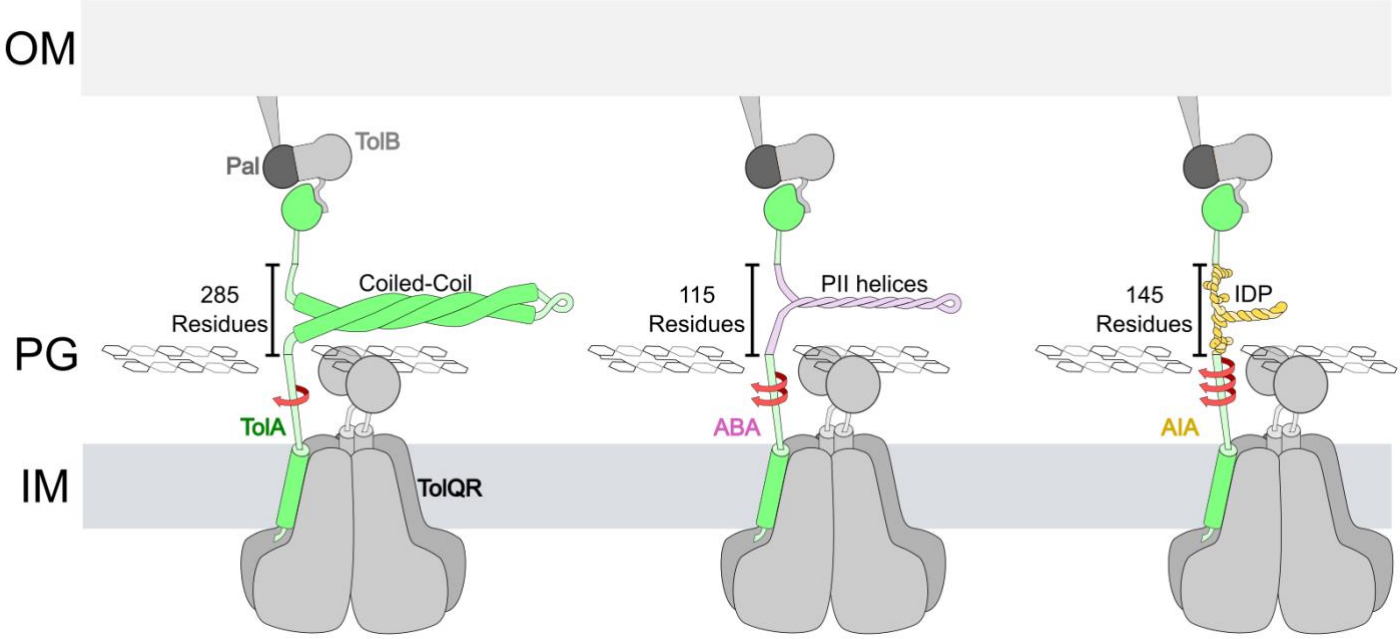


Figure 3.16 | Possible tensioned domain II conformations may explain differences in mutant efficiency. TolA, ABA and AIA all transduce force with varying efficiency, denoted by the number of twists required to generate equivalent tensile force (*red arrows*). TolAII is depicted as tightly-wound coiled-coil as a product of writhing, reducing its reach and applying tension. This is based on biophysical observations that α-helices largely resist unfolding as tensile force is applied, and reconciles the observation that TolA is more efficient at force transduction than ABA, despite having over twice the number of domain II residues. ABA is illustrated as previously for TonB (Fig. 3.15). The twisted conformation for the IDP sequence of AIA was physically modelled with string. Unlike the TonB-TBDT catch-bond the dissociation of the TolB-Pal complex dissipates mechanical tension at the OM, allowing TolAIII-TolB dissociation. Overall, this model assumes all three force transducers utilise a similar torque-driven mechanism to generate tensile force.

Despite the structural diversity of the chimeras explored here, force transduction is achieved in all mutants that span the periplasm. This begs the question, how diverse are domain II secondary structures throughout nature?

4.0 The origin, organisation and distribution of the Tol-Pal system

Much of the work in this chapter was submitted in Williams-Jones *et al.,* 2025. Work presented here is my own, with full permission from fellow co-authors.

4.1 Introduction

4.1.1 Annotation of *tolA* requires proximal *tolB* and/or *pal genes*

In the previous chapter, I discovered that domain II swaps of TolA and TonB are capable of force transduction, even when swapped with disordered sequences. Here, I sought to identify the limits of domain II structures between TolA and TonB sequences in nature, but to do so I first needed to gather a library of novel *tolA* genes. *tol-pal* genes were initially discovered through mutants tolerant to colicins, such as those used in chapter three [15-17]. The mutants were also found to exhibit membrane permeability defects, cell chaining, and detergent sensitivity [82, 83, 202]. It was subsequently determined that the Tol-Pal system is a specialised OM tethering system found primarily in Proteobacteria, but also in other GNB [12, 86]. The *tol-pal* operon in *E. coli* features five core genes (*coloured/grey*) and two flanking genes (*black/light blue;* Fig. 4.1), within two open reading frames (ORFs) [203].

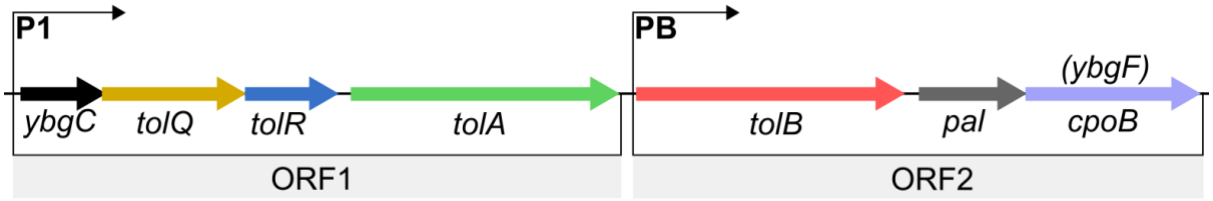


Figure 4.1 | The *E. coli tol-pal* operon. The five core *tol-pal* genes (*tolQRAB* & *pal*) are coloured as their protein products depicted throughout. Ancillary genes *ybgC* (*black*) and *cpoB* (*violet*) are also shown. Two promoters are present within the operon, indicated by P1 and PB. Figure adapted from Sun and Webster [204]; Vianney *et al.*, [205], and Sturgis [12].

This operon is conserved in most Proteobacteria. *tol-pal* genes are essential in *Caulobacter crescentus* and *Pseudomonas aeruginosa* [206, 207]. In *E. coli* K12 strains where the LPS is truncated these genes are not strictly essential. However, loss-of-function mutants of *tolQRA, tolB* or *pal* tolerate neither SDS nor bile salts, suggesting these core genes are conditionally essential to *E. coli* within the gut. The *tol-pal* genes are essential in *E. coli* O7 with full-length LPS, possibly because the OM

is more rigid which further impedes OM invagination [13, 208]. Loss of any of these key genes greatly impacts the virulence of many Proteobacteria [209-213].

The challenge in specifically identifying *tolA* lies in the fact that many species typically have multiple *tonB* genes but only a single *tolA* [12]. Misannotation is common in databases, because *tolA* and *tonB* share homologous CTDs and have particularly poor sequence conservation in domain II. To obtain TolA sequences, one must therefore verify *tolA* identity by detecting a proximal compliment of *tol-pal* genes, such as *tolB* and *pal*. While it is known that the *tol-pal* cluster is generally prevalent throughout Proteobacteria, the extent of that prevalence is poorly quantified. Here, I sought to readdress this question using a combination of modern bioinformatic tools and updated databases. After identifying *tol-pal* gene clusters from a diverse range of species, I used secondary structural prediction tools to quantify domain II structure, overcoming the problem of poor sequence conservation. This was crucial for analysing domain II because the secondary and tertiary structures tend to be more conserved than sequence alone. Here, I qualify the extent of *tol-pal* prevalence and organisation throughout the kingdom of bacteria, and analyse the secondary structures of TolAII from these data.

4.2 Results

4.2.1 Tol-Pal systems are conserved throughout Proteobacteria

In *E. coli* and many related γ-proteobacteria, TolAII is predominantly α-helical [88, 163, 164] while TonBII is predominantly expected to form PPII-helices [165-168] (Fig. 4.2A). The reason for divergence in domain II structure is unknown, as a result the identity of the last common TolA/TonB ancestor is also unknown. Of taxa beyond Proteobacteria, only Acidobacteria, Chlamydiae and Chlorobium have documented *tol-pal* clusters [12, 13]. Here I used a combination of GenBank [152], UniProt [172], BLAST [214], AlphaFold [100] and FoldSeek [72] to catalogue taxa with putative Tol-Pal systems, to assess the organisation of their *tol-pal* in 84 species selected in a recently published consensus phylogeny tree [215] (Fig. 4.2B). I was then able to structurally analyse all detected TolA proteins using PsiPred (147) (SI. 1). Within domain II, the proportion of residues predicted to form α-helices was quantified and compared with the proportion of proline residues. The latter is reflective of the minimum number of residues implicated in PPII-helices, which comprise either homopolymeric proline repeats or Pro-X_{1-3}-Pro motifs (Fig. 4.2A). These data reveal that as well as the α-helical rich TolA sequences known from *E. coli* and its close relatives (γ-proteobacteria), some feature PPII-helices (α-proteobacteria), a mixture of PPII-helices and α-helices (β-proteobacteria), or more disordered residues (ε-proteobacteria) (Fig. 4.1B). Moreover, both TolAII length and structural composition appear clustered by phylogeny, where closely related species feature similar TolAII structures.

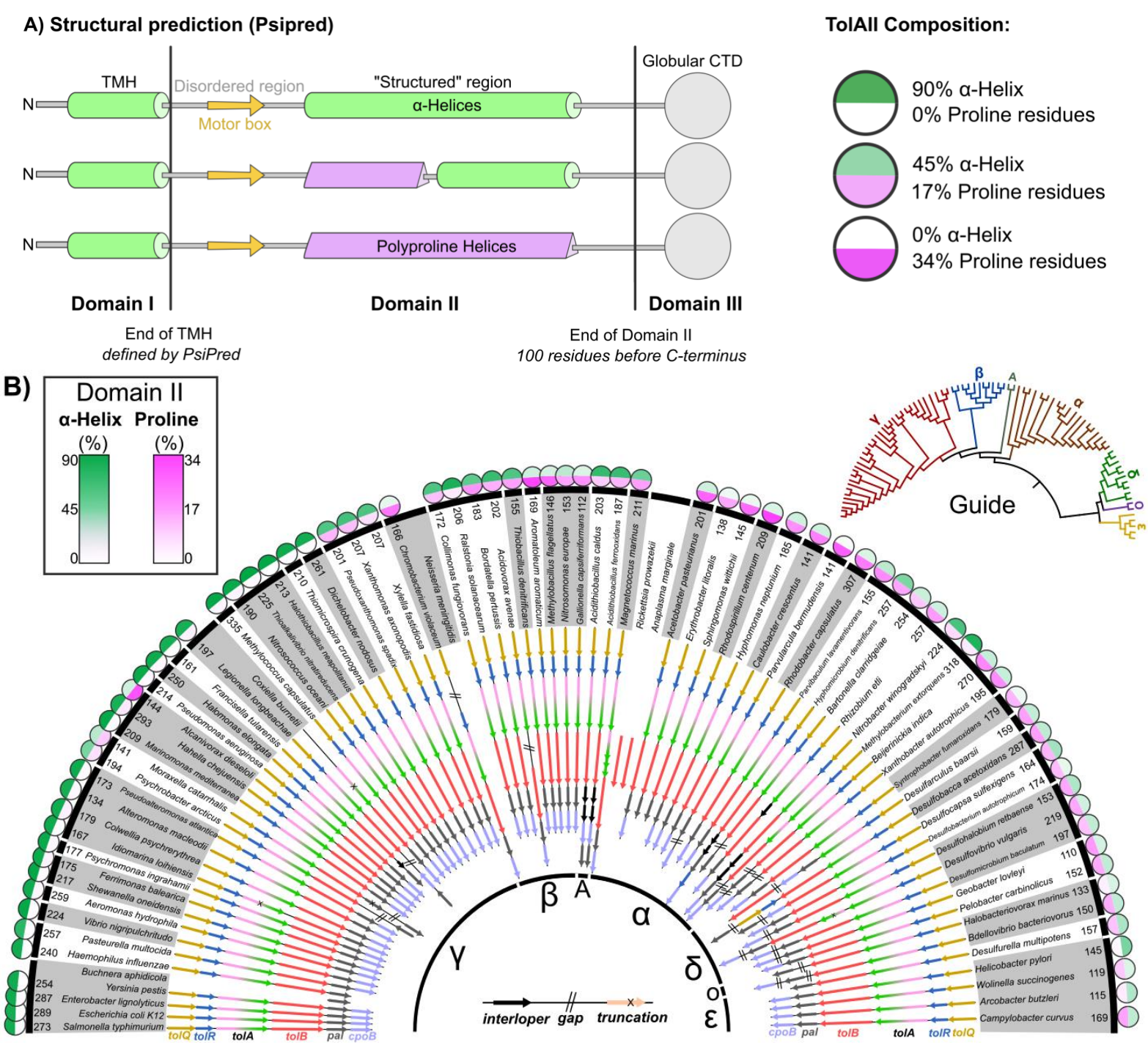


Figure 4.2 | *tol-pal* gene organisations and α-helix:proline ratios from representative proteobacterial species. **A)** TolA proteins have a three-domain structure (*left*), where the TMH is followed by a disordered linker, an α-helical/proline-rich region, followed by the globular C-terminus (final 100 residues). PsiPred predictions often show a mixture of α-helical and proline-rich regions, not defined by a specific ratio or relative position within domain II (SI. 1). The vast majority of TolA species feature a predicted β-strand (*"motor-box", yellow*) within a disordered region invariably found after the TMH, that is suggested to interact with the dimeric component of the motor (*discussed in chapter six*) [199, 216]. The ratio of α-helix:proline residues within domain II of each protein in B) is coloured according to the key (*top right*). **B)** The *tol-pal* loci from a range of species were curated by identifying a contiguous region containing a motor, force transducer, *tolB* and *pal* (*loci aligned in centre; gaps indicate absent genes*). Species were clustered relative to a circularised phylogenetic tree adapted from Sharma *et al*. [215] (*"Guide"*), with different proteobacterial orders indicated (*grey/white background*). Greek letters separate proteobacterial classes; "A" denotes Acidothiobacillia and "O" denotes Oligoflexia. TolAII secondary structures were predicted using PsiPred 3.2, and the α-helix: proline ratios of each protein are indicated by colour (*outer circles*)(SI. 1). Species with faint α-helix and proline scores are suggested to have a greater propensity for disorder within domain II. Circumferential numbers indicate respective length of domain II. Some species have putative "split" *tol-pal* systems (denoted by "//"), where different components are not proximal on the genome, while others lack some/all *tol-pal* genes. Evidence for loci and TolAII composition analysis indicated in SI. 1. FASTA sequences of detected *tol-pal* genes are also collated (SS. 1). Data submitted in Williams-Jones *et al.,* 2025.

These loci demonstrate considerably greater conservation of core gene organisation/synteny than expected, the majority featuring a similar organisation to that of *E. coli.* In some cases, "split" *tol-pal* loci were observed, where *tolQRAB* and *pal-cpoB* loci were separated on the genome, while a few appear to have lost their *tol-pal* clusters. Overall, these outliers suggest that while the "core" locus organisation is a suitable guide for *tol-pal* system detection, exceptions exist. The flanking genes of each *tol-pal* cluster were semi-syntenic but varied between classes, suggestive of positional changes on the chromosome over time (Fig. 4.3) (SI. 1).

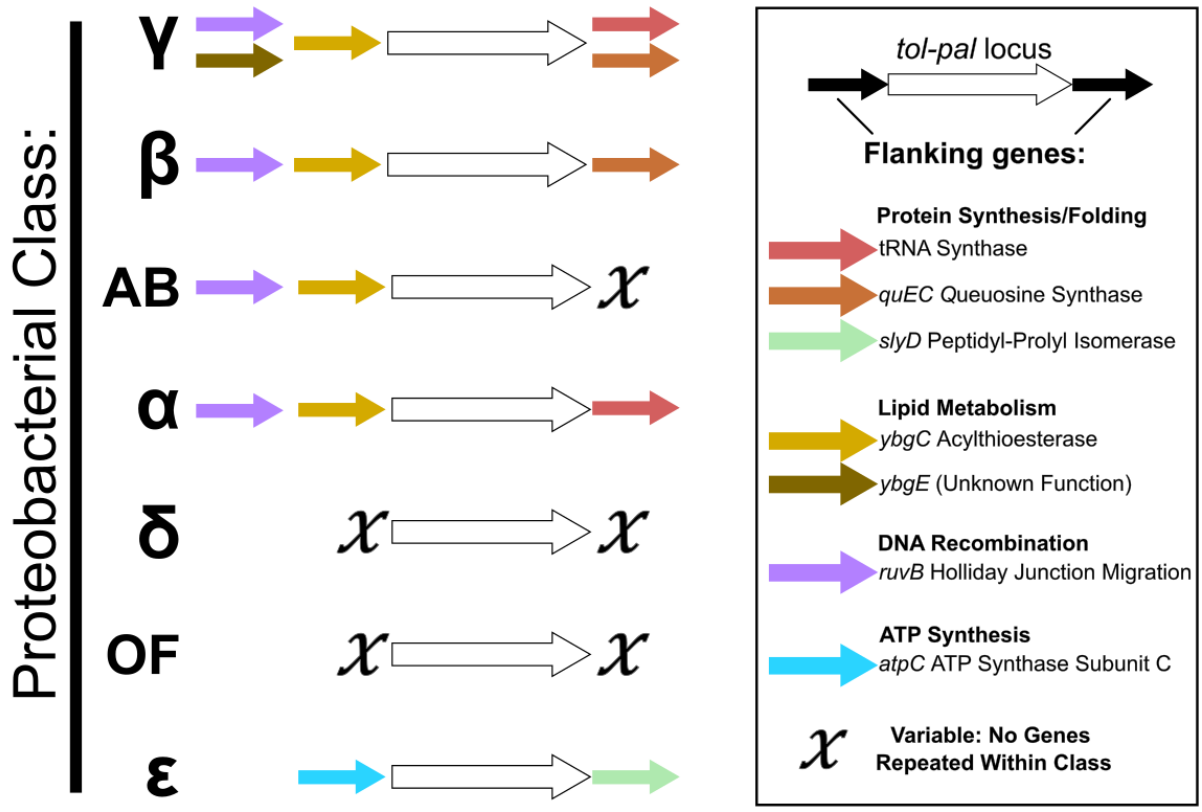


Figure 4.3 | Synteny of proteobacterial *tol-pal* flanking genes. The general functions of syntenic *tol-pal* flanking genes vary between proteobacterial classes including *Acidothiobacillia* (*AB*) and *Oligoflexia* (*OF*). Paired arrows in γ-proteobacteria indicate alternative syntenic genes within the class. "***x***" denotes the absence of any genes reoccurring more than once within the class. Overall, the genes adjacent to *tol-pal* are generally conserved within proteobacterial classes, but vary between them.

Some species, most notably of α-proteobacteria, were found to feature proline-rich TolAII reminiscent of *E. coli* TonBII. The same structural composition analysis was thus performed on TonB sequences detected within the same genome accessions, to compare the structural diversity of TonBII. To verify that these were *tonB* genes, neighbouring genes were inspected to ensure that no *tolB/pal* genes were present. Many species feature multiple *tonB* genes, but only one was analysed per species,

thus those selected may structurally differ from alternative TonB paralogues. The calculated TonBII α-helix:proline ratios were then mapped to the same phylogenetic tree (Fig. 4.4) [215]. These data demonstrate that in Proteobacteria, the majority of TonBII sequences are proline-rich, but α-helical and disordered TonBII sequences were also detected.

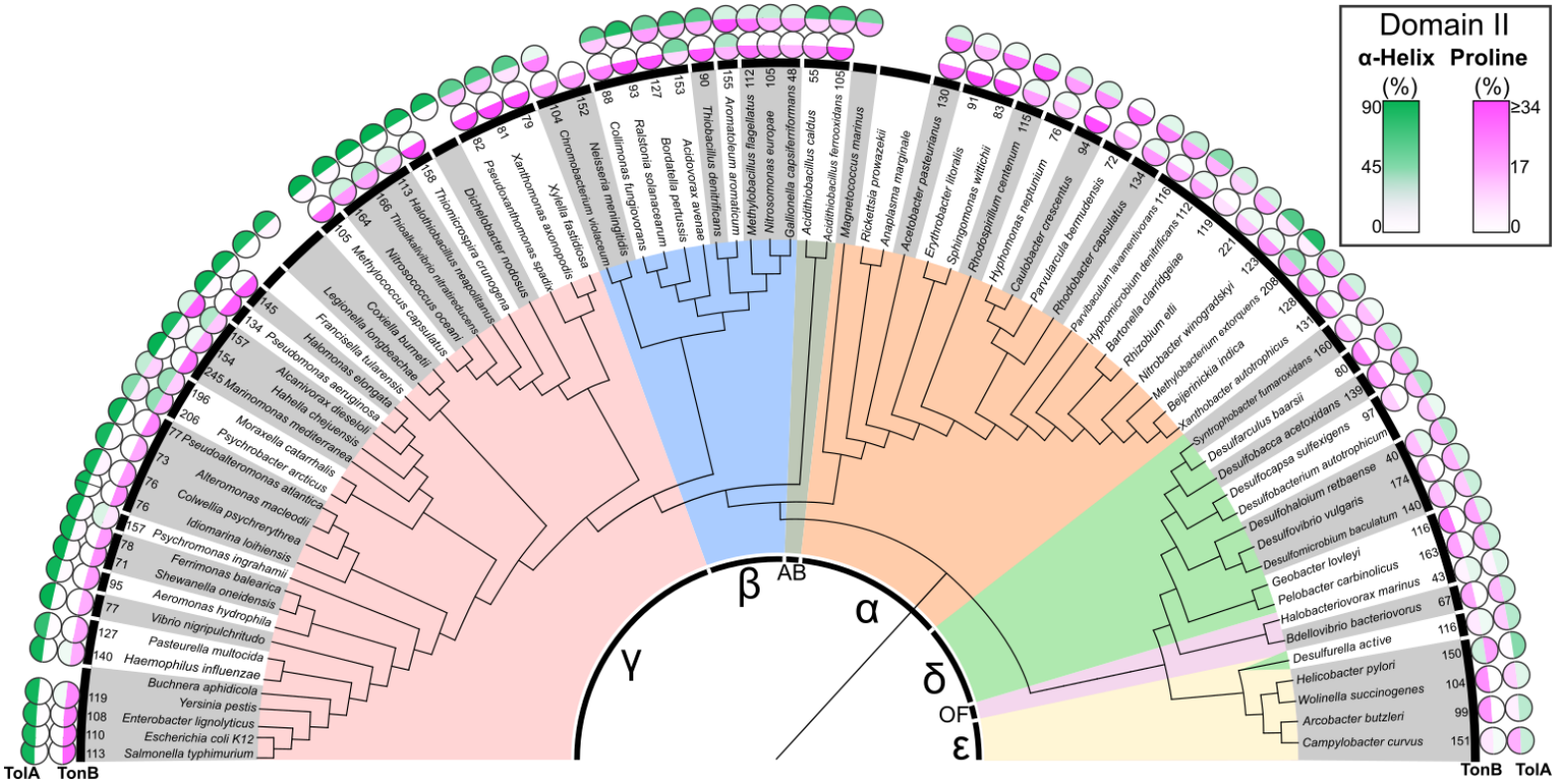


Figure 4.4 | TolA and TonB domain II length and secondary structure compositions vary amongst Proteobacteria. Genome accessions from Fig. 4.1B were screened for both *tonB* and TBDT genes (SI. 2; SS. 1). Each representative TonB was structurally analysed by the same method as for TolA (SI. 2). Proline content of some TonBII sequences exceeded 34%, but they are coloured the same for consistency with Fig. 4.1. Phylogenetic tree adapted from Sharma *et al*., [215]. Circumferential numbers indicate respective length of TonBII per sequence analysed.

This analysis implies that the synthetic *E. coli* domain II chimeras generated earlier are perhaps not as unusual as suspected, because many proline-rich and disordered TolA sequences exist naturally in nature, as do examples of TonBII featuring α-helices (SI. 1-2). The diversity of these TolAII and TonBII sequences suggest that there are many possible functional secondary structures for force transducers, including some that may not have been detected here such a 3_{10}-, π-, or ω-helices [217-219]. Further investigation of TonBII sequences reveals that the two oppositely charged PPII-helix segments observed in *E. coli* are only found in TonBII sequences tested from *Enterobacterales*, *Pasteuralleles*, and *B. pertussis*. Many Proteobacteria feature more neutral PPII-helix motifs with a preponderance for alanine and valine residues (e.g. *N. oceani* or *C. violaceum*), or entirely negative PII-helices (glutamate-rich, e.g. *M.*

mediterranea). TonBII PII-helices also varied in organisation, with most featuring Pro-X_{1-4}-Pro motifs, but some with long homopolymeric stretches of proline residues alone (e.g. *X. axonopodis*, *T. denitrificans*, and *C. crescentus*). The length of TonBII was also variable; some species feature very short TonBII sequences while the longest observed was 245 residues. In contrast, TolAII ranges between 110-335 residues. The median TonBII and TolAII lengths were 113 and 192 residues, respectively. The respective lengths of TolAII and TonBII per species are weakly positively correlated (Fig. 4.5). This may be linked to the width of the periplasm in each species, but very few measurements of bacterial periplasms are available to test this.

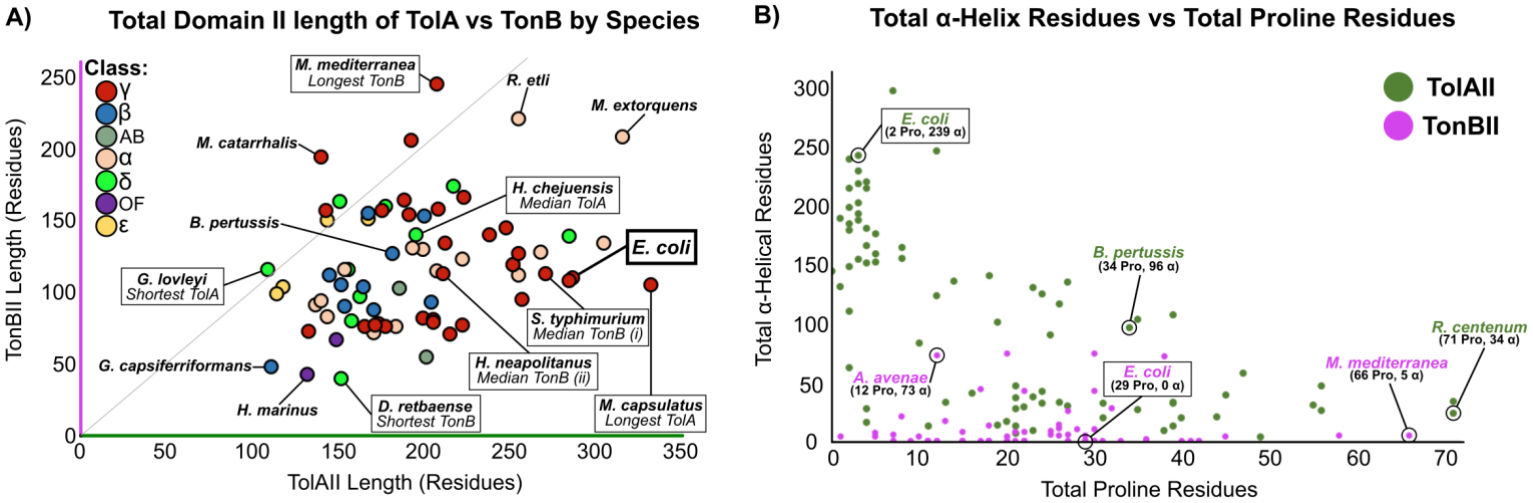


Figure 4.5 | Quantification of TolAII/TonBII structural diversity. A) The total length of domain II for selected TolA and TonB sequences is plotted here by species (SI. 1-2). Each point is coloured by proteobacterial class as in Fig. 4.3 ("*AB*"= *Acidothiobacillia*, "*OF*"= *Oligoflexia*). A Pearson's correlation test suggests that the length of TolAII and TonBII is weakly positively correlated in a given species (Score = 0.28). These data show that lengths of both TolAII and TonBII also vary greatly within proteobacterial classes, but the majority of TolAII sequences are longer than TonBII (*grey line*). Species of note are indicated, to illustrate the variation in TolAII/TonBII length. Two TonB medians were observed (*i/ii*). **B)** For each selected TolAII (*green*) and TonBII (*pink*) sequence, the total number of α-helical residues is plotted against the total number of proline residues. Many TolA sequences are α-helix dominant (e.g. *E. coli*), but many mixed (e.g. *B. pertussis*) and proline-dominant sequences (*R. centenum*) are also observed. For TonB sequences, the majority are proline-dominant or mixed, but some α-helix dominant species are present (e.g. *A. avenae*). Note that proline residues are a minimal proxy for the number of PII-helical residues present.

These data highlight that the function of TolA and TonB are practically indistinguishable from sequence and structural predictions alone; TolAII can be short and PPII-helix dominant, TonBII can be long and α-helix dominant, and both proteins often feature mixtures of both helix types. Protein function must therefore be inferred from genomic context. The close resemblance of some of these proteins reiterates the question of their evolutionary relationship. To contextualise these data and address these questions, a broader sample of bacterial *tol-pal* systems must first be validated.

4.2.2 Tol-Pal systems are conserved in Gracilicutes

To structurally analyse TolA sequences outside the Proteobacteria phylum, I used a combination of manual locus curation and a bespoke search algorithm to identify *tol-pal* loci in the pubMLST database [220]. The algorithm preliminarily detects *tol-pal* constituent genes by PFAM annotations, then measures their relative genomic proximity. Species predicted to feature *tol-pal* loci were manually validated through a combination of UniProt and GenBank to curate the correct complement of genes for each species [152, 172](SI. 1). Recent work by Witwinowski *et al*., traced the evolutionary emergence of multiple membrane peptidoglycan tethering proteins, including Pal [2]. I thus adapted the *pal*-containing branch of their phylogenetic tree to map *tol-pal* locus organisation and TolAII composition (Fig. 4.6).

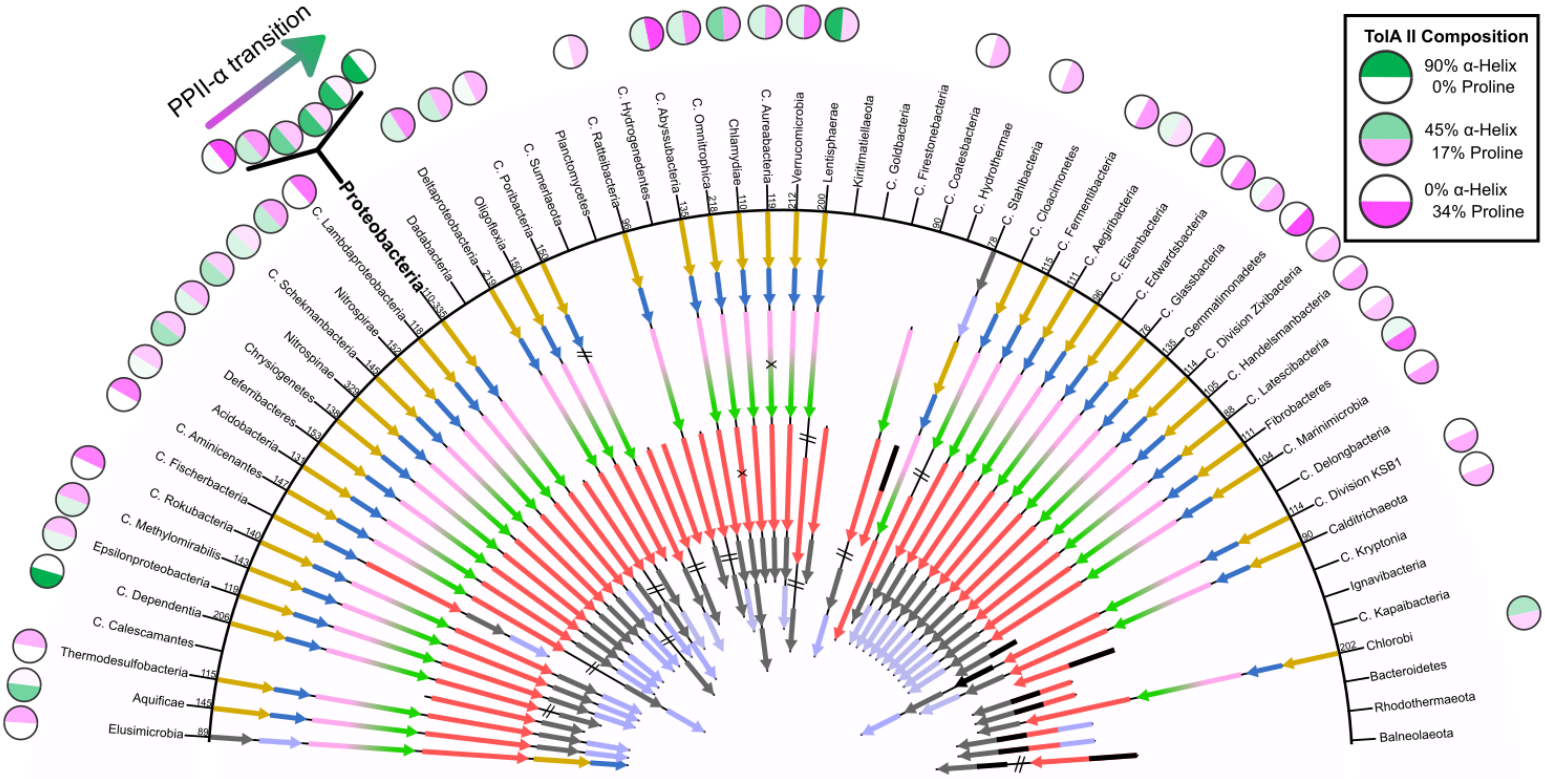


Figure 4.6 | Validated *tol-pal* loci from Gracilicute taxa. Gracilicutes with identified *tol-pal* loci generally feature similar gene organisations to those of Proteobacteria (*arrows, coloured as in Fig. 4.1*) (*species indicated in SI. 1*). Blank regions indicate absent/undetected *tol-pal* genes. Some taxa lacking *tol-pal* loci feature genes with homology to multiple *tol-pal* genes (*multi-coloured arrows*). TolAII length (*outer numbers*) and α-helix:proline structural analysis (*circles*) reveals that many Gracilicutes possess proline-rich TolAII sequences. "PPII-α transition" refers to the enriched representation of α-helices within proteobacterial TolAII compared to other Gracilicutes. This map is adapted from a portion of the phylogenetic tree featuring *pal*, published by Witwinowski *et al.,* [2]. Some fusion proteins are predicted to be lipoproteins with homology to the Pal PGB, suggesting they fulfil similar OM-PG tethering roles in *pal*-deficient taxa. One such taxon is Bacteroidetes, where *porE* was identified as an OM-PG tethering candidate [221](SI. 1). Crosses indicate genes that are likely non-functional due to frameshift mutations or truncations. Data submitted in Williams-Jones *et al.,* 2025. The *tol-pal* search algorithm was written and executed by Dr. James Bray (Department of Biology, University of Oxford).

These data suggest that the α-helical bias of γ-proteobacterial TolAII is anomalous amongst the Gracilicutes. These Gracilicute TolAII sequences are generally shorter than Proteobacterial TolAII sequences, more akin to Proteobacterial TonBII sequences (Fig. 4.7A), with further analysis revealing that the total number of α-helical residues vs proline residues also resembles the distribution of proteobacterial TonBII sequences more closely than proteobacterial TolAII sequences (Fig. 4.7B-C).

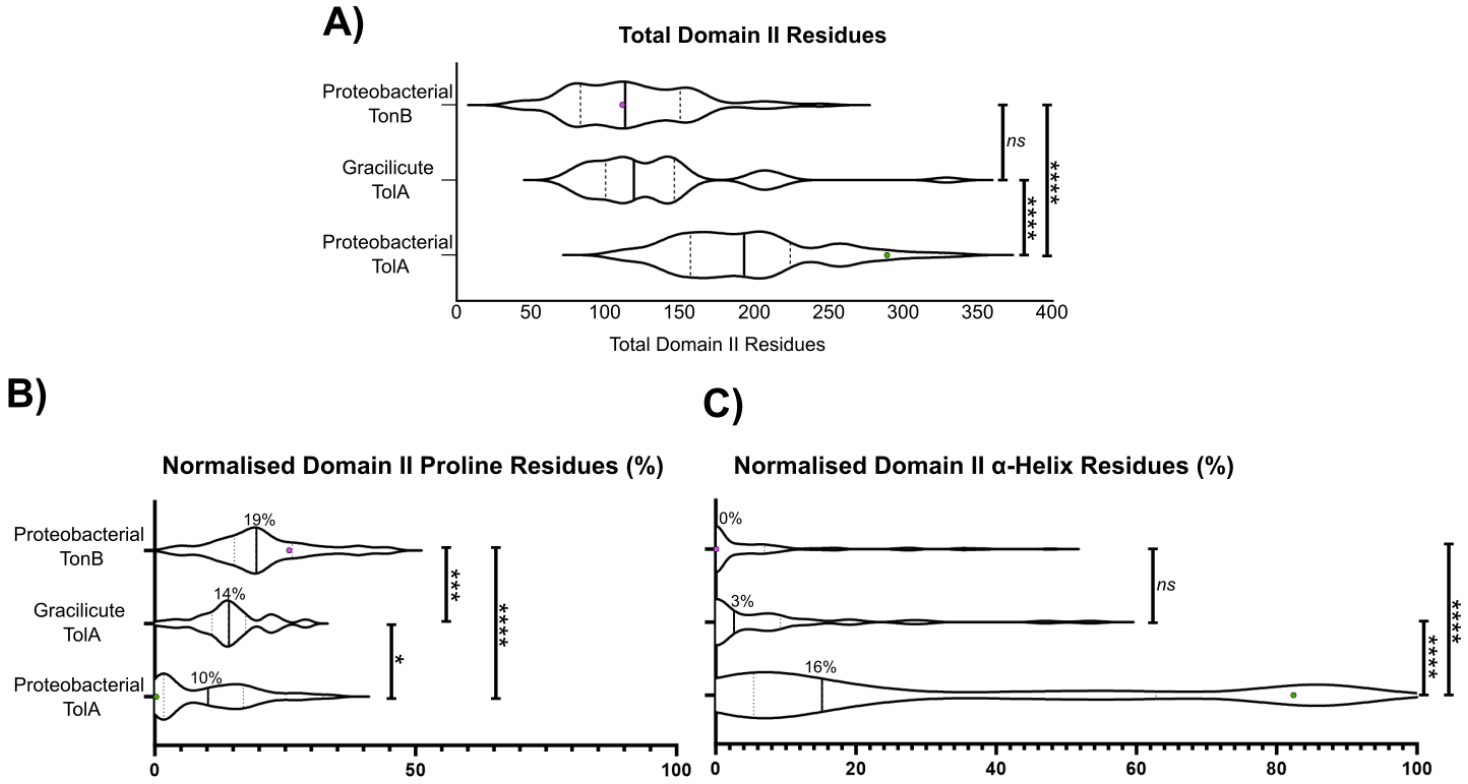


Figure 4.7 | TolAII structural composition within Gracilicute species reveals resemblances to proteobacterial TonBII. A) The distribution of Gracilicute TolAII lengths more closely resembles proteobacterial TonBII than TolAII. The median lengths (*solid bars*) and upper/lower quartiles (*dashed lines*) are indicated. A Mann-Whitney U Test suggested that proteobacterial TonB and Gracilicute TolA domain II length distributions do not differ significantly (P= 0.1487), while proteobacterial TolAII length distributions differ significantly from both (P< 0.0001). **B)** The distribution of Gracilicute TolAII sequences features a significantly lower proportion of proline residues than proteobacterial TonBII sequences (P= 0.0003), but greater than proteobacterial TolAII sequences (P= 0.013). **C)** The distribution of Gracilicute TolAII sequences indicates a similar proportion of α-helical residues to proteobacterial TonBII sequences (P= 0.2704), both of which are significantly lower in α-helical residues than proteobacterial TolAII sequences (P< 0.0001). For both B) and C), median values are labelled. For all three panels, specific values for *E. coli* TolA (*green*) and TonB (*pink*) are shown.

This apparent structural similarity between Gracilicute TolAII and proteobacterial TonBII sequence distributions reveals putative proline-rich TolA sequences that are virtually indistinguishable from TonB. Since most Gracilicute TolA sequences sampled are proline-rich, this suggests that the common ancestor of TolA was also proline-rich. This reiterates the question of which emerged first, TolA or TonB?

4.2.3 Distribution of Tol/Ton systems suggest TolA origin

To identify which of the Ton system or Tol-Pal system emerged first, an investigation into the phylogenetic distribution of these two systems was required. Previous analyses identified *tonB* in Proteobacteria, Cyanobacteria, Bacteroidetes, Chlorobi, Chlamydiae, Verrucomicrobia and Fusobacteria [168, 222]. Homologues of *exbBD* were more widely detected, but the full complement of Ton-system proteins were not assessed [168, 222]. To achieve this, a second algorithm was used to search for co-occurring *tonB* and TBDT genes within each accession, which are not always proximal. These were manually validated as previously (SI. 2; SS. 1). These data were then mapped onto the phylogenetic tree generated by Witwinowski *et al.*, demonstrating that Ton systems are far more widely distributed than Tol-Pal (Fig. 4.8) [2]. Overall these data suggest that the last common bacterial ancestor, which was likely a diderm, featured a Ton system [2]. The Tol-Pal system likely emerged significantly later.

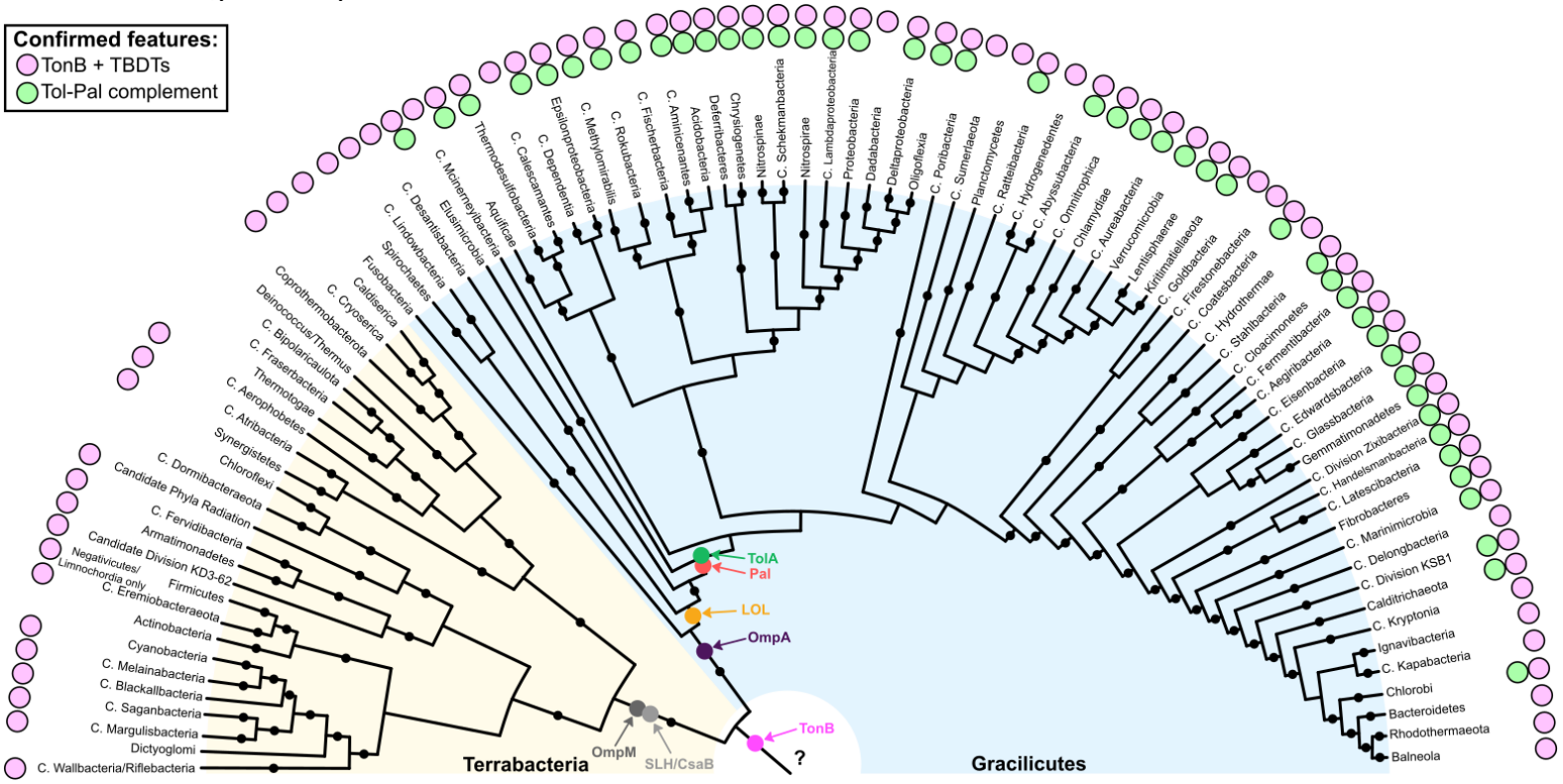


Figure 4.8 | Distribution of Tol and Ton systems suggest the Ton system emerged first. Taxa featuring curated Tol-Pal systems (*green circles*) are found within Gracilicutes only, whereas most didermal species across both Terrabacteria and Gracilicutes feature Ton systems (*pink circles*). This suggests that the earliest common didermal ancestor possessed a functional Ton system (*pink arrow*). Species flagged for Ton systems are noted (SS. 1). Of the Terrabacteria, only didermal species possessed both *tonB* and TBDT genes. The most likely point of TolA emergence (*green arrow*) occurs near the proposed origin of Pal [2]. Phylogenetic tree adapted from Witwinowski *et al*. [2]. Figure/data submitted in Williams-Jones *et al*., 2025. The *tol-pal* and *tonB*/TBDT search algorithm was written and executed by Dr. James Bray (Department of Biology, University of Oxford).

It is increasingly well-known that protein folds tend to remain structurally conserved over evolutionary timescales that greatly surpass the timescales of nucleotide/amino acid sequence conservation. This is particularly well-illustrated by TolA and TonB structural predictions, which exhibit little sequence conservation despite their conserved folds, particularly in domain III. To address the phylogenetic problem of TolA and TonB despite their low sequence conservation, I utilised FoldTree [223]. This software uses PDB structures (including AF prediction models) to model phylogenetic trees based on tertiary structural similarity, with significantly improved accuracy over sequence-based methods for fast-evolving or evolutionarily-distant proteins. I generated an unrooted phylogenetic tree using full-length AF predictions of a single representative TolA and TonB sequences from each taxon, from UniProt-embedded AF predictions [71]. TolA structures segregate into a single clade, with the exception of one example from Candidatus Aegiribacteria. Proteins labelled TolA were identified by neighbouring genes of *tolB* or *pal*, but the prevalence of split loci outlined previously suggests that some of the proteins labelled TonB (due to lack of adjacent *tolB/pal*) may represent chromosomally isolated *tolA* genes. Further still, the presence of TolA-like proteins in Terrabacteria where Tol-Pal systems are not found, suggests horizontal gene transfer of *tolA* genes, possibly followed by evolutionary reversion to reacquire the ancestral TonB activity. Far more taxa feature TonB than TolA, but these were not well-segregated by superphylum; Terrabacteria TonB sequences were found evenly distributed across the tree, suggesting that horizontal gene transfer of *tonB* genes occurs relatively frequently. Moreover, the neighbours of each taxon (for both TolA and TonB proteins) do not always match their close relatives on the consensus tree (Fig. 4.8), further suggesting horizontal mobility and rapid evolution of force transducers.

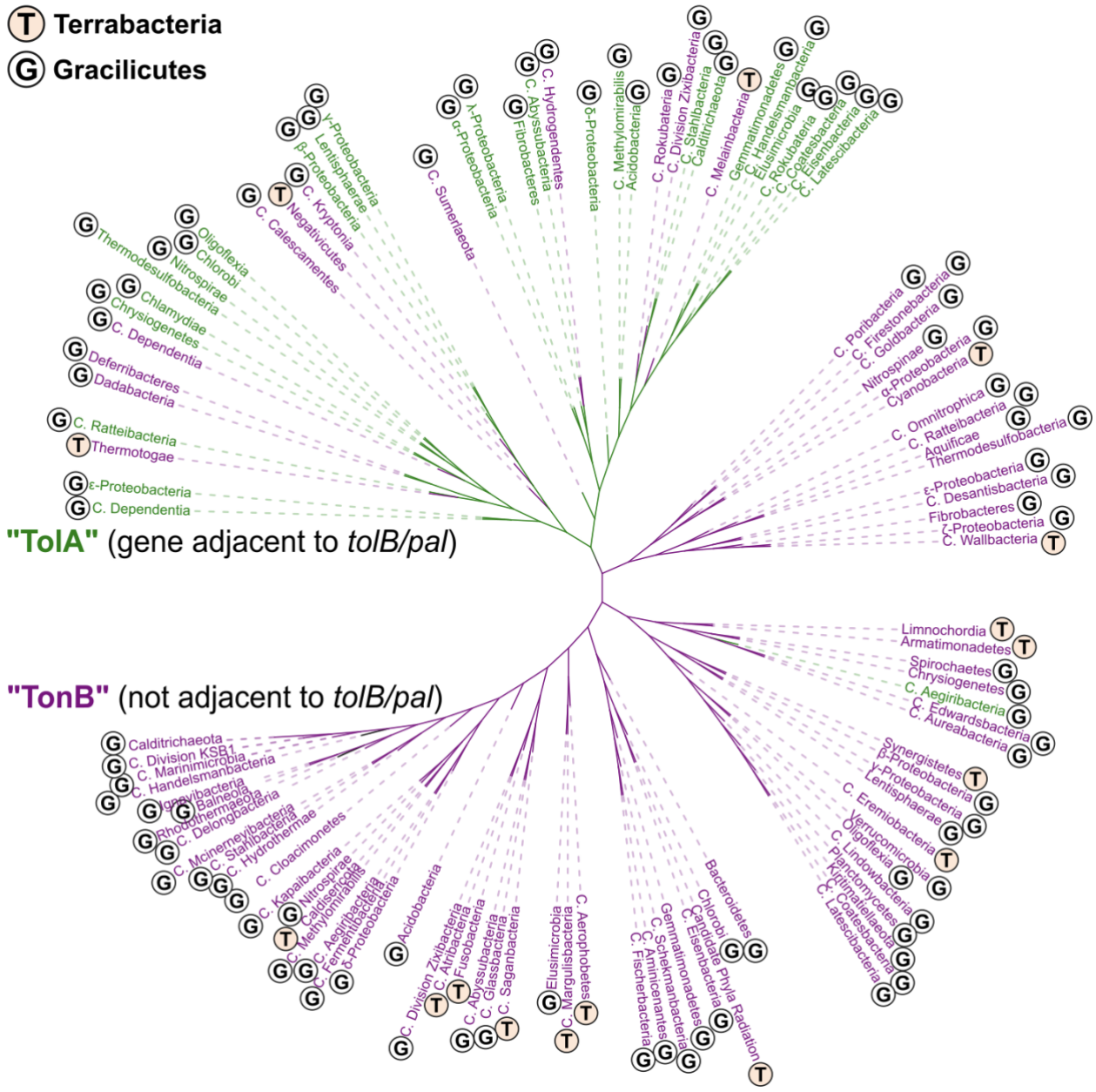


Figure 4.9 | Unrooted FoldTree phylogenetic tree of TolA/TonB AF prediction models reveals TolA proteins comprise a separate clade from most TonB proteins. TolA proteins are found in a single clade (*top left, green*), while TonB sequences are found to comprise multiple clades (*pink*), which may vary in TBDT specificity. TolA proteins were identified/labelled by presence of neighbouring *tolB*/*pal* genes, while TonB proteins were labelled by absence of neighbouring *tolB/pal* genes. While informative, some *tolA* genes may be misannotated if found distally from *tolB/pal*, and thus the protein is mislabelled as TonB despite fulfilling the role of TolA.

4.3 Discussion

4.3.1 The inferred origin of TolA from a TonB paralogue

The Tol-Pal and Ton systems have long been understood to be homologous, but the divergence of the two systems to fulfil their different functions has remained unresolved for decades [12, 19, 107]. Here, I reveal that the Ton system is found in the majority of bacterial taxa, indicating that the Ton system precedes the origin of the Tol-Pal system (Fig. 4.7). Moreover, the joint prevalence of both systems throughout Gracilicutes suggests the Tol-Pal system likely evolved from an ancestral Ton system via duplication and speciation. Previous phylogenetic analysis detected *exbBD* homologues in some species of archaea, further suggesting that the ExbBD-TonB complex may have existed in the last bacterial common ancestor [222]. In contrast, Tol-Pal systems can only be found in taxa that evolved after the emergence of Pal, a defining element of the system [2]. To avoid biasing the *tol-pal* distribution, the algorithm searched all taxa nonetheless, reaffirming those findings. Here I also demonstrate the use of secondary structure composition to compare TolA and TonB, circumventing their low-sequence conservation. The structural diversity of TolAII sequences range from a mixture of predominantly α-helical to predominantly proline, the latter of which are more widely conserved in Gracilicutes, which may reflect a TonB-like common ancestor (Fig. 4.6). In contrast to TolA, analysis of TonBII sequences revealed a general lack of α-helices in domain II, suggesting that prevalence of TolAII α-helices in Proteobacteria is a product of further TolA speciation (Fig. 4.5, 4.6; SI. 1-2). Many species examined also exhibit Ton system paralogues, where duplication and speciation of TonB leads to multiple TonB proteins with unique TBDT targets, such as HasB in *Serratia marcescens* [224]. Some species exhibit a great number of TonB paralogues, as illustrated in *Gemmatirosa kalamazoonensis* where there are twelve TonB paralogues that likely vary in TBDT-specificity (SI. 2). The sheer number of TonB paralogues observed and their tendency to speciate presents ample opportunity for emergence of novel functions (*neofunctionalisation*). In contrast, no more than one *tol-pal* system was identified within any genome. Taken together, these data suggest the Tol-Pal system evolved from a paralogue of the Ton system, emerging to stabilise the OM during cell division (Fig. 4.10).

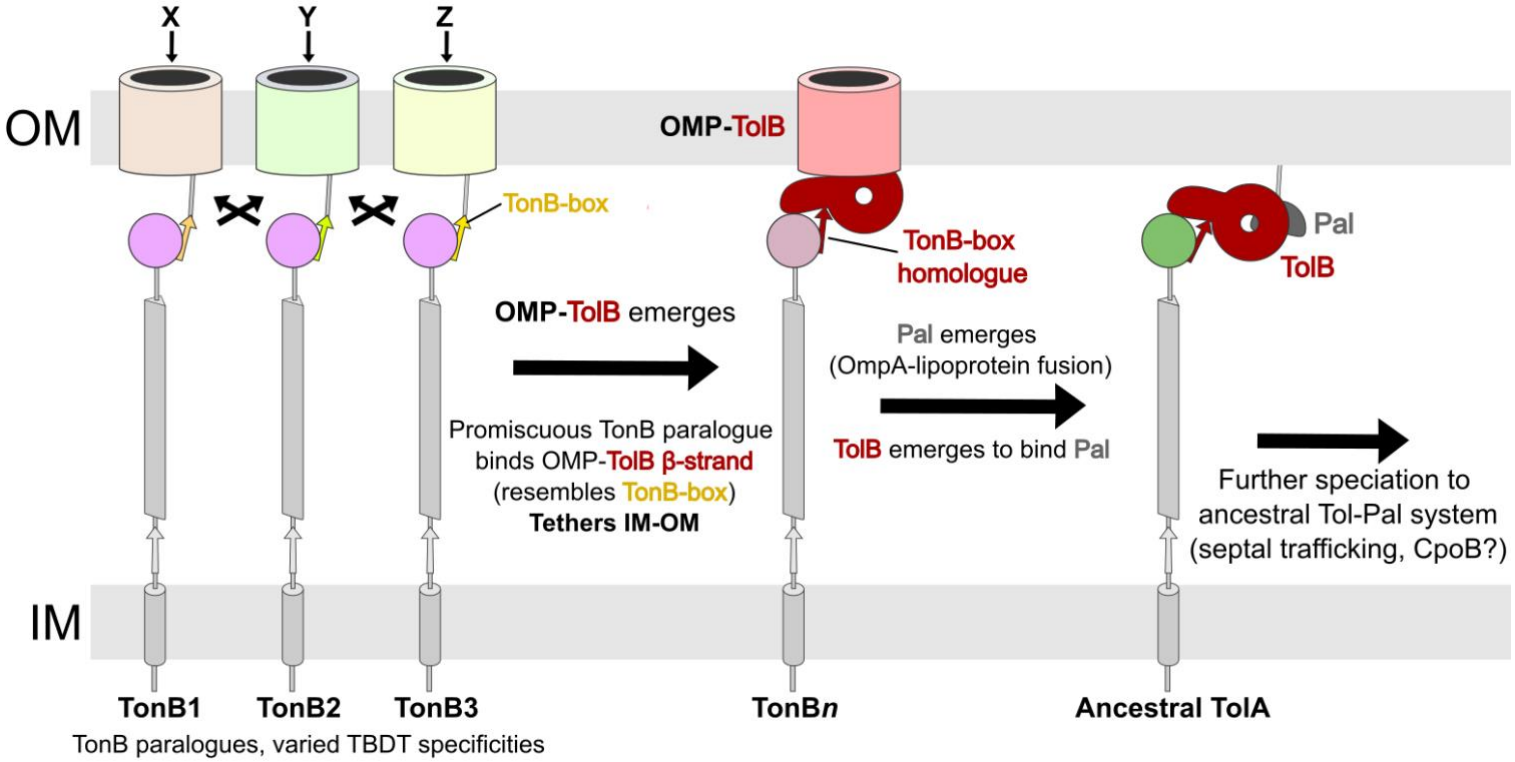


Figure 4.10 | A hypothetical route for TonB-to-TolA speciation. TonB duplication generates paralogues with varied affinities for different TBDT TonB-boxes (*crossed arrows*). OMP-TolB fusion proteins could reflect the ancestral progenitor from which TolB emerged, where the β-addition interaction might stabilise the OM by tethering it to the IM. Emergence of Pal and the TolB-Pal interaction leads to greater complexity of the system within the ancestral Gracilicute (*right*). The β-addition interaction of TonBIII remains conserved in TolAIII. Alternatively, a TonB-box from a TBDT gene may have been genetically fused to the ancestral TolB gene via recombination. It is unclear whether TolB or Pal was first-recruited to the division site, or how the Tol-system achieves septal trafficking.

If the Tol-Pal system did indeed evolve from the Ton system, further questions arise. The first, is how? The evolution of TonB from a TBDT-binding protein to a TolB-binding protein is plausible, since their β-addition interactions are similar. It is less clear how the Ton system transitioned from being uniformly distributed around the cell envelope as observed in *E. coli* [225], to being actively transported to the division site, because the mechanism by which TolQRA is trafficked to the septum remains unclear. It may involve interaction (possibly indirectly) with late-stage divisome components such as FtsN, and/or other as-yet unknown factors [109, 226, 227]. Since the divisome is conserved between bacteria and archaea, septal localisation apparatus clearly preceded the emergence of the bacterial kingdom [228]. Thus, neofunctionalisation of the ancestral TolQRA proteins to interact with the divisome may explain how TolQRA evolved septal localisation.

The emergence of *pal* has been described previously, by fusion of the *ompA* PGB domain with a lipoprotein, after the evolution of the LOL pathway [2]. In contrast, it is unclear how and when TolB emerged. TolB has a two-domain structure, where the

NTD adopts a Rossmann-fold (near the TolA-box), and the CTD forms a six-bladed β-propeller (Fig. 1.9). Interestingly, β-propeller folds are believed to have evolved from the Rossmann-fold [229]. Sequence- and fold-based searches using TolB yield many interesting homologues, such as acyl-peptide hydrolases, and Omp85-TolB fusion proteins [72, 214] (SI. 1). In particular, the latter may represent an extant relative of an ancestral TolB progenitor, from which TolB emerged via recombination of the periplasmic domain with a periplasmic secretion signal. The TolB β-propeller is also homologous to the PG-remodelling protein LysM, suggesting that the primordial TolB ancestor may have been involved in septal PG-remodelling even before the emergence of the Tol-Pal system. It is tempting to speculate that modern TolB retains some PG-binding activity, but none has been conclusively demonstrated so far. In some cases, homologues to the TolB β-propeller were found as a fusion protein with the Pal PGB and the CpoB TPR domain, notably in taxa lacking *tol-pal* loci (Fig. 4.5; SI. 1). This protein, PorE, is conserved throughout Bacteroidetes, and may comprise a further speciated Tol-Pal system [221]. Contrary to its classically assumed role within the Type 9 secretion system, PorE may facilitate OM-PG tethering in the Bacteroidetes. No cellular microscopy has been performed to track PorE localisation to date, and the matter is further complicated by the fact that most Bacteroidetes feature multiple PorE paralogues.

Another key observation is that *cpoB* is far more widely distributed than previously suspected, found adjacent to *tol-pal* loci in around 70% of Gracilicute species (and 75% of proteobacterial species) where *tolB* and *pal* were also detected. CpoB plays an enigmatic role at the interface of cell division and cell wall remodelling that remains poorly understood. The broad conservation of *cpoB* and its synteny with the core *tol-pal* genes suggests that the role of CpoB was defined relatively early in the evolution of the Tol-Pal system, potentially even preceding the Tol-Pal system. Loss of *cpoB* is non-lethal in *E. coli* K12, but generates a phenotype with some parallels to deletion of PBP1B [230, 231]. Overall, these data reiterate the importance of this protein as a hub between PG-remodelling and cell division [86, 110]. It is plausible that *cpoB* is present in other (unidentified) species since its low sequence conservation obfuscates detection.

I previously showed that within TolAII, the native α-helix structure is associated with increased force transduction efficiency than PPII-helices, as indicated by faster bacteriocin import and faster recruitment of Pal to the septum [131]. Since the

ancestral TolA is expected to be proline-dominant, what drove the evolutionary transition from PPII-helices to α-helices of TolAII in Proteobacteria? One possible explanation is generation time. Commensal/pathogenic Proteobacteria with rapid growth rates tend to favour α-helices over PPII-helices, suggesting that shorter division times require faster OM-stabilisation during division (chapter three). Because these species tend to be mesothermophilic compared to environmental samples, perhaps their increased TolAII length and α-helicity better tolerates a widened periplasm if OM-fluidity is increased as a function of temperature or changes in osmolarity [232-234]. Those conditions also explain why TolAII is typically longer than the periplasm; in the case of conditions that cause periplasmic widening, TolQRA can still recruit Pal for OM-PG tethering at the septum. Alternatively, perhaps divergent secondary structures better segregate TolA function from TonB. It has been shown, for example, that in *E. coli* both TolQR and TolAII are independently trafficked to the septum by as-yet uncharacterised periplasmic factors, whereas TonB needs no such trafficking [108, 109, 158].

The structural analysis of each TolAII and TonBII also revealed that every sequence without fail, whether from Proteobacteria or wider Gracilicutes, features a disordered region after the TMH (SI. 1-2). Moreover, in the 113 validated TolA sequences from both datasets, 93% feature a predicted β-strand within this disordered region, described as the motor-box [199, 216]. The functional significance of the motor-box in the force transduction mechanism of TonB has been investigated in recent works [199, 216], that suggest it plays a role in the activation of the motor-driven force transduction cycle of TonB. The extensive conservation of this motif within TolA reinforces the view that TolA and TonB utilise a common force transduction mechanism, and hints that the motif, further explored in chapter six, was a feature of their common ancestor.

Overall, these findings suggest that Tol-Pal systems are far more widely conserved amongst Gracilicutes than previously known, and likely originate from a paralogous Ton system. This catalogue of *tol-pal* loci offers an extensive list of targets for further investigation into the diversity of Gracilicute cell envelope physiology. Further, the extensive conservation of the Tol-Pal system bolsters its significance as a potential target for the development of new strategies in the fight against antimicrobial resistance. Further data are needed to understand how they work, such

as the mechanism by which proton flux through the motor generates force, and what is the diversity of biological processes that can be powered by such systems.

5.0 PMF-dependent rearrangements of TolQ-TolR

5.1 Introduction

This chapter features work published in Webby *et al*., and Williams-Jones *et al*., performed and written by myself except where stated otherwise [97, 131].

5.1.1 Identifying TolQR contacts using chemical crosslinking

Up to this point, we have discussed and modelled potential mechanisms that presume TolA utilises torque to generate tensile force. This assumes that TolQR rotates, and that this rotation is directly applied to TolA in the form of torque. However, this is inferred from MotAB homology and is only indirectly evidenced [68, 158]. Previous works have used *in vivo* crosslinking to probe the conformational rearrangements that occur within the TolQR complex during PMF-transduction, predominantly to delineate the topology of the complex. The first used formaldehyde crosslinking to identify interactions formed by TolR, which demonstrated TolR homodimerisation and interactions with TolQ and TolA [235]. They also showed that deletion of the TolR PGB renders it non-functional, suggesting PG-binding is required for PMF-transduction. A truncated mutant of the TolR PGB (residues 60-133) was later demonstrated to bind purified sacculi *in vitro* whereas the full-length PGB did not, suggesting conformational rearrangements occur that regulate this activity *in vivo* [105]. Additional formaldehyde-crosslinking experiments also showed that the TolR PGB forms contacts with TolQ during the PMF-transduction cycle, which may modulate its PG-binding activity [236]. Further-refined crosslinking approaches later utilised cysteine substitution mutants to improve their specificity. The first demonstrated that disulfide bond formation between the PGB of adjacent TolR monomers (via T117C mutants) inactivates TolQR [237]. It was then shown through cysteine mutants adjacent to the essential Asp_{23} residue within the TolR TPH, that the reduced forms are functional while the disulfide-locked TolR homodimers become non-functional, to suggest that relative rotation occurs between the TolQ and TolR subunits [238]. Although the TolR TPHs remained unresolved, two structures of the TolR PGB were solved; one by NMR and one by X-ray crystallography [13, 239]. Together these demonstrated opposing orientations of TolR monomers within each dimer (~180°, Fig. 5.1). Of those two conformations, the crosslinked residue T117 was only proximal in the *E. coli* crystal structure [105]. The other, a truncated form solved by NMR from *H. influenzae*, demonstrated an “open”

conformation [239]. When three *E. coli* TolR constructs were investigated, truncated at different points to mimic those used in the *H. influenzae* work, the strand-swapped conformer was unable to bind PG, whereas deletion of the strand-swapping region (residues 36-60) restores sacculus binding *in vitro* [105]. It was thus suggested that TolR alternates between these two states during force transduction, similar to models proposed for MotB [105, 240].

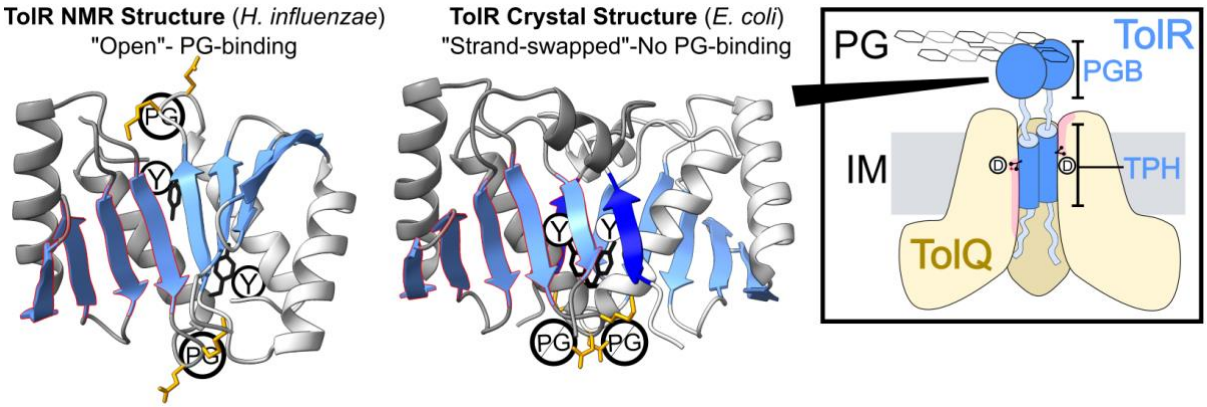


Figure 5.1 | The "Open" and "Strand-swapped" conformations of TolR PGB dimers. A) The open form of the TolR PGB (*left*) was solved by NMR (PDB: 2JWK)[239], while the strand-swapped form (*right*) was crystallised (PDB: 5BY4)[105], revealing an extra β-strand that interacts with the other monomer (*dark blue*), with adjacent subunits flipped ~180° between structures (***β-strands of one monomer outlined in red***). In the open conformation, Y117 residues (*Y*) are too distant for crosslinks to form, while in the strand-swapped form they are adjacent. The proposed PG-binding residues (*circle labelled PG*) are also flipped ~180°, suggesting they are more accessible for PG-binding in the open form but sterically blocked when strand-swapped. Cut-through cartoon of TolQR (*top* right) shows the relative positions of the TolR PGB and TPH, and indicates the position of the conserved D23 residue (*D*) within the TPH.

In this chapter, I aim to further this work, utilising a combination of *in vivo* crosslinking and MS to better understand the precise conformational rearrangements that occur between TolQ and TolR during PMF transduction.

5.2 Results

5.2.1 Co-conservation analyses of TolQ-TolR interface

To identify candidate residues within the TolQR assembly for *in vivo* crosslinking, an adequate structural model is required. The TolQR complex has previously been homology modeled from the cryo-EM structure of ExbBD [13]. By the end of this thesis, the cryo-EM structure of the *E. coli* TolQR complex was published [131], followed shortly after by that of the *Acinetobacter baumannii* [241], but both lack resolution of the TolR PGB due to its flexible linker. In lieu of those structures, the TolQR complex was first modelled using AF2 and relaxed using Rosetta [97]. To validate the model, we used a combination of mutant data and evolutionary co-conservation analysis. The latter used a deep learning model, RaptorX, that aligns a large homologue sequence dataset to two input sequences, then applies statistical methods to detect the residues that are co-conserved between the two proteins indicative of the contact interface [156]. This method is summarised well in Jeon *et al.* [242]. RaptorX was thus used to examine the pairwise co-conservation of TolQ and TolR, to validate the predicted TolQ-TolR interface. The resulting co-conservation matrix was mapped to the TolQR model, where residues of the TolR TPH and TolQ lumen displayed the greatest co-conservation, supporting the topology of the assembly (Fig. 5.2A-B). Analysis of the conserved threonine residues involved in stabilising TolR D23 protonation were found in similar positions to their MotAB/ExbBD homologues, further supporting a shared PMF-transduction mechanism (Fig. 5.2C) [57, 243]. Cysteine point mutants that confer *Δtol* phenotypes were also mapped onto this model, to identify where disulphide formation impinges TolQR rotation [238, 244]. The phenotypes observed were categorised as either a “full” *Δtol* phenotype, where the OM is completely destabilised and colicins are not imported, or a “partial” *Δtol* phenotype, where the OM is slightly destabilised but colicins are still imported. This echoes the spectrum of TolA mutants examined in chapter three, suggesting that the full *Δtol* phenotype prevents rotation absolutely, thereby conferring a destabilised OM and abolishing colicin uptake. Conversely, partial *Δtol* phenotypes likely arise from moderately reduced rotation rates that slightly destabilise the OM due to reduced Pal recruitment rates, but still energise colicin import sufficiently to remain colicin-sensitive.

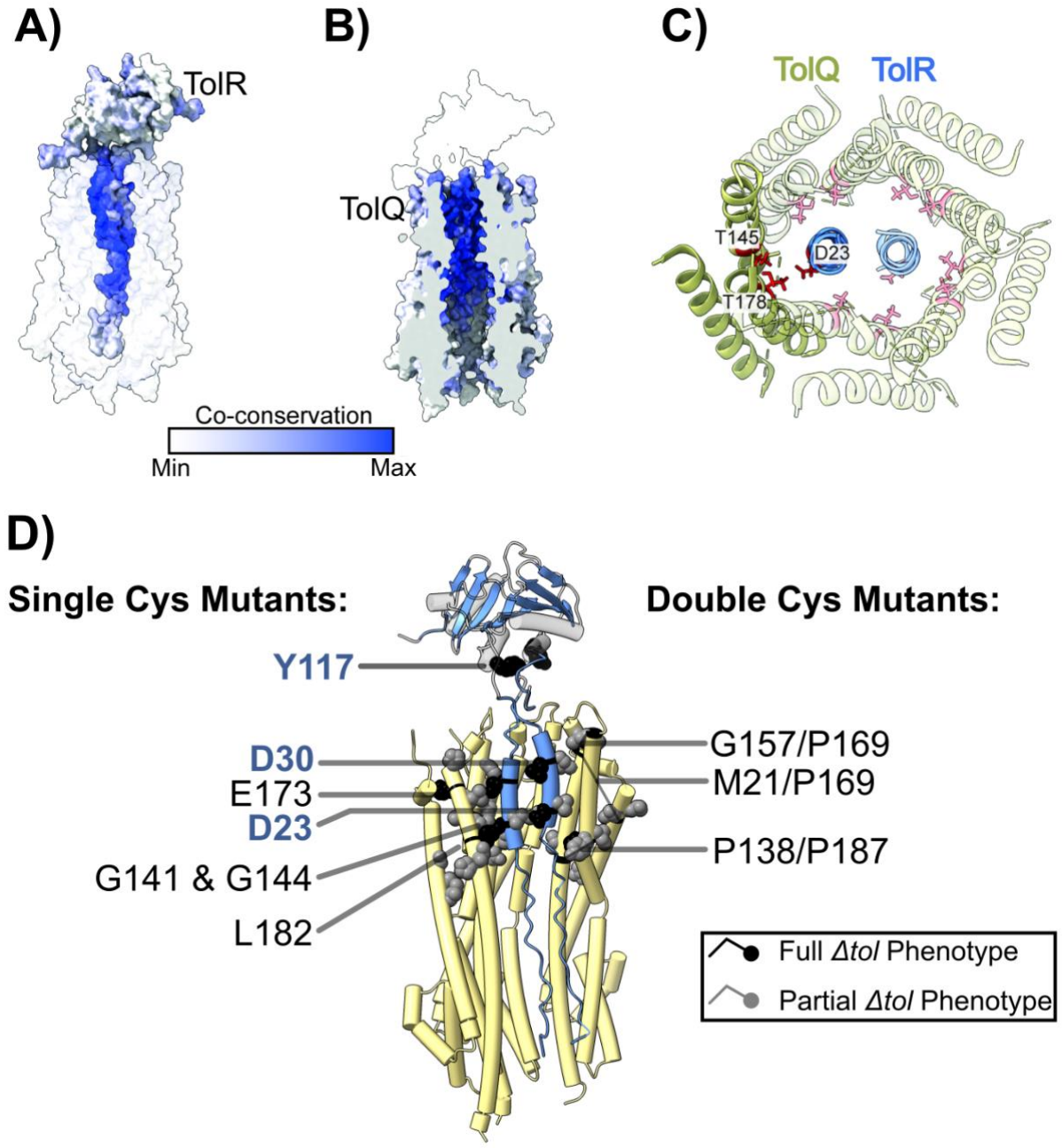


Figure 5.2 | Mapping the TolQR interaction interface. A-B) Co-conservation maxima mapped onto the TolQR model highlight the luminal interface. The TolR co-conservation maxima (*A*) occur on the TPH, whereas TolQ co-conservation maxima occur within the lumen (*B*). The TolR PGB was modelled as its open conformer (*see D)*). **C)** Conserved threonine residues within the transmembrane region of TolQ (T145 and T178) are proposed to stabilise TolR D23 upon protonation [57, 243]. Figure adapted from Webby *et al.* [97]. **D)** Cysteine point mutants that generate *Δtol* phenotypes on the TolQR model [244]. Single cysteine mutants are indicated on the left TolQ protomer (*black labels*) and both TolR protomers (*blue labels*), and double mutants indicated on the right TolQ protomer. Mutant residues that confer a full *Δtol* phenotype (*black residues*) are labelled, while those that confer partial *Δtol* phenotypes (*grey residues*) are unlabelled. Introduction of *Δtol* phenotypes from cysteine mutants suggest proximal crosslinks form at those positions that prevent or reduce rotation.

5.2.2 The TolQR cryo-EM complex

Following this modelling, the *E. coli* TolQR complex was subsequently purified and characterised via cryo-EM, by Dr. Melissa Webby [131]. The complex was shown to form a 5:2 assembly, with great structural similarity to the prediction model (Fig. 5.3). Notably, this structure shows TolR D23 is at the bottom of the TPH, supporting its proposed role as the moderator of proton flux, though insufficient resolution was achieved to determine the true orientation of D23. As predicted, D23 is flanked by the conserved ring of threonine residues that are proposed to stabilise D23 protonation [97]. Given this appropriate structure for modelling TolR site-specific crosslinks , efforts were made to map transient contacts that form during proton flux.

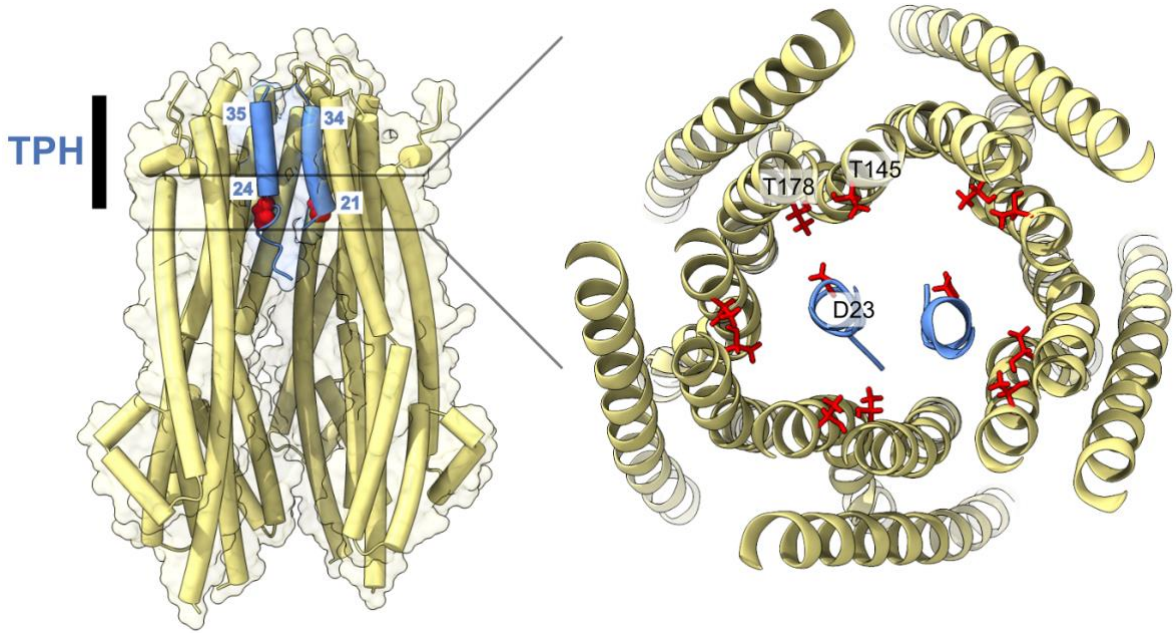


Figure 5.3 | The TolQR complex modelled from cryo-EM data. The TolQR complex (*left*; PDB: 80DT) was similar to the prediction model (*Fig 5.2*), featuring two TolR TPHs within the TolQ pentamer (*TPH*). The TolR residues (*numbered, blue*) comprising the resolved TPH differ slightly between protomers, potentially suggesting some vertical movement. The relative heights of key residues (*red*) are located approximately where expected (*right*). The flexible NTD of TolR could not be resolved, nor could the PGB, due to flexibility conferred by the disordered linker above the TPH.

5.2.3 Designing TolQR mutants for *in vivo* crosslinking

Previous works predominantly utilised formaldehyde and disulphide crosslinking approaches to ascertain the topology of the TolQR complex. To better identify regions that form transient contacts within the complex during force transduction, UV crosslinking was performed with the photoactivatable unnatural amino acid benzoyl-phenylalanine (BPA) [145]. BPA crosslinking is more informative than previous crosslinking approaches, because BPA residues are site-specifically incorporated, then activated by UV exposure in an experimentally controlled manner. The BPA sidechain then forms a ketyl diradical, and proceeds to form irreversible covalent bonds to carbon atoms within 3-4 Å. This specific crosslinking allows precise crosslink mapping *in vivo*, by proteolytically digesting crosslinked proteins and performing MS. BPA mutants were thus designed at positions on TolR, where they reside within the lumen of the TolQ pentamer (Fig. 5.4). Care was taken to substitute bulky hydrophobic TolR residues where possible. With the optimisation of *in vivo* crosslinking and complex purification, the UV-treatment can then be performed in the presence or absence of the PMF. The difference in crosslink profiles between states is thus determined by the conformational rearrangements of PMF transduction.

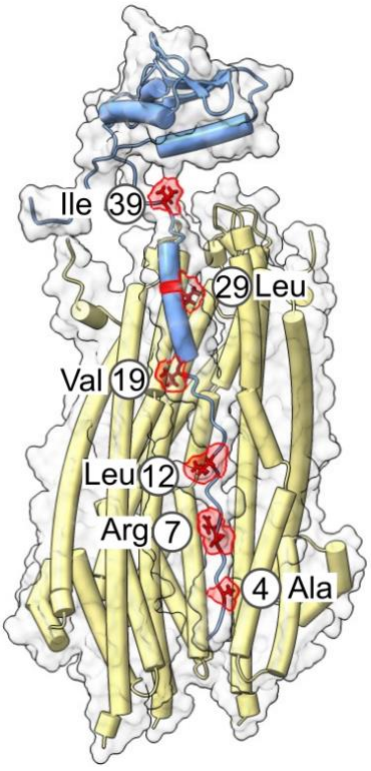


Figure 5.4 | TolR mutants designed to incorporate BPA residues within the TolQ lumen. Native TolR (*blue*) residues are indicated (red), on the cryo-EM model (PDB: 80DT) with full-length TolR docked from Webby *et al.,* [19, 131].

5.2.4 TolQR BPA mutants confer colicin sensitivity

To test whether TolR mutants with BPA incorporated are functional *in vivo*, and therefore representative of WT TolQR, a colicin assay was used as a proxy for force transduction (Fig. 5.5). Competent *ΔtolQR* cells were prepared and transformed with WT *tolQR*, and *tolQR* BPA mutant plasmids with the pEvol BPA plasmid, that encodes a BPA specific tRNA synthetase [145]. A day culture was grown to late exponential phase, before plating in soft agar. Plates were spotted with ColE9 as previously (chapter three). These data suggest that all mutants except $tolR_{V19X}$ encode TolQR complexes capable of PMF transduction *in vivo* (where *X* denotes BPA). Colicin sensitivity in the absence of induction and BPA was unexpected. This was suspected to arise due to promiscuous substrate specificity of the pEvol-encoded aminoacyl-tRNA synthetase, when BPA was absent. Leaky expression of both plasmids was demonstrated, because cells expressing TolQR were sensitive to ColE9 without induction, and the lack of ColE9 sensitivity in the BPA mutants when the pEvol plasmid was not present. A very small amount of killing is observed in *ΔtolQR*, presumably because TolA promiscuously interacts with the ExbBD motor as previously described [160]. These assays nonetheless demonstrated that the TolQR BPA mutants were suitable to proceed with UV-crosslinking.

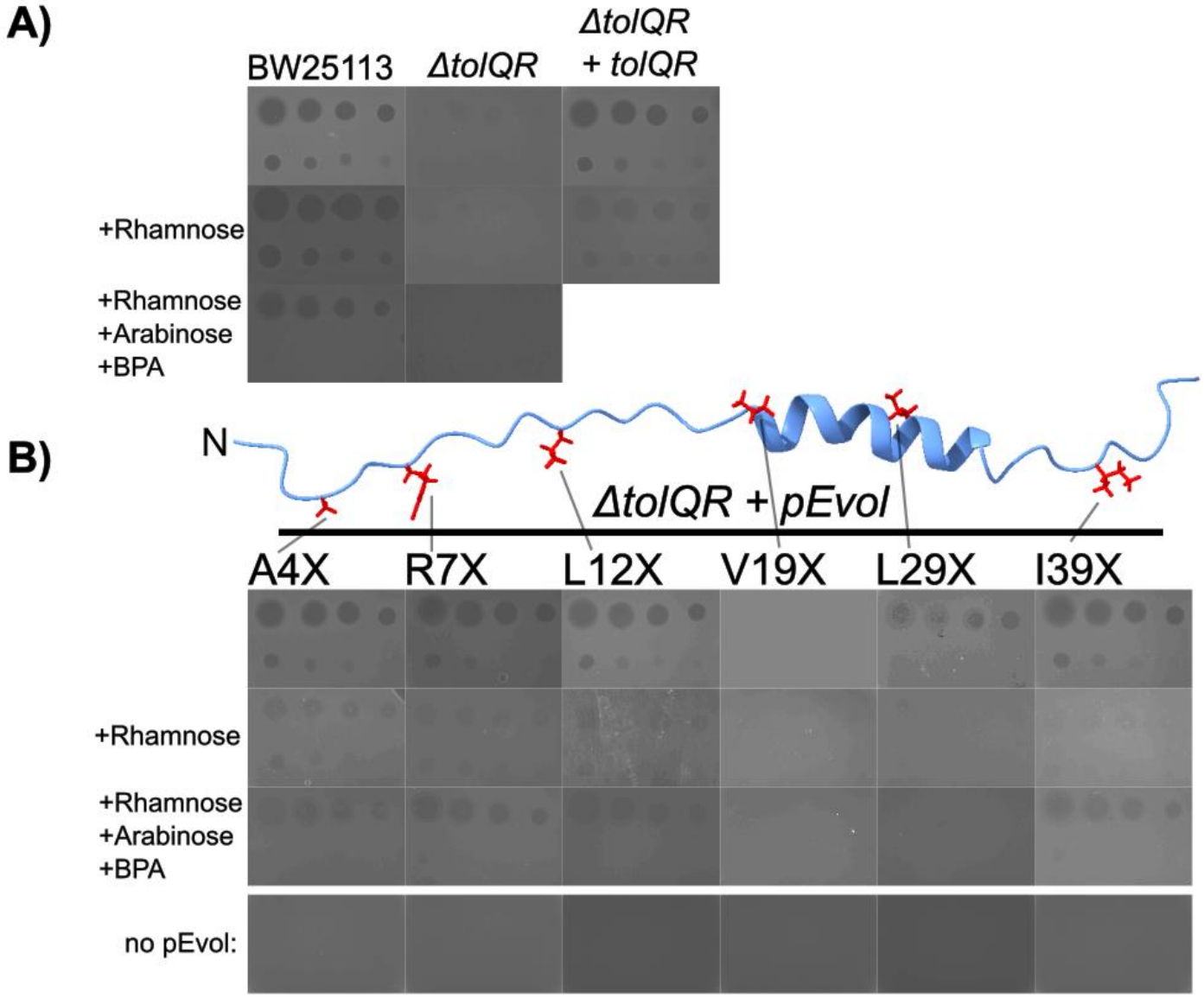


Figure 5.5 | Expression of TolQR BPA mutants elicits ColE9 sensitivity. **A)** Control assays suggest TolQR expression is leaky, because there is ColE9 sensitivity without *tolQR* induction. A very small amount of killing is observed in *ΔtolQR*, likely to due tol TolA promiscuously interacting with ExbBD. **B)** All TolQR BPA mutants confer colicin sensitivity, except for V19X. Leaky TolQR and pEvol expression leads to colicin sensitivity even in uninduced samples lacking BPA supplementation.

5.2.5 Strep-Tactin purification yields pure TolQR

Having shown TolQR BPA mutants are functional, trial purifications were performed with TolQR BPA mutants in *E. coli* BL21 with BPA incorporated during induction [19, 131] (Fig. 5.6). The purified BPA mutants were UV-treated *in vitro* and potential crosslinks were visible, but the presence of contaminants hindered crosslink detection. Size exclusion and Q columns were tested to remove these contaminants, but ultimately unable to deliver sufficient purity and reduced TolQR yield. The *tolQR* construct encodes a Twin-Strep-tag® tag, in addition to its hexahistidine tag [245], thus an affinity purification was performed with the WT TolQR construct using Strep-Tactin resin, which drastically improved TolQR purity compared to Ni^{2+} affinity. The quality of the Strep-Tactin purification was thus deemed sufficient for the purification of crosslinked TolQR complexes in one step. However, the yield of proteins purified from *ΔtolQR* cells was insufficient for crosslink mapping regardless of purification approach, hence BL21 cells were used for expression and purification.

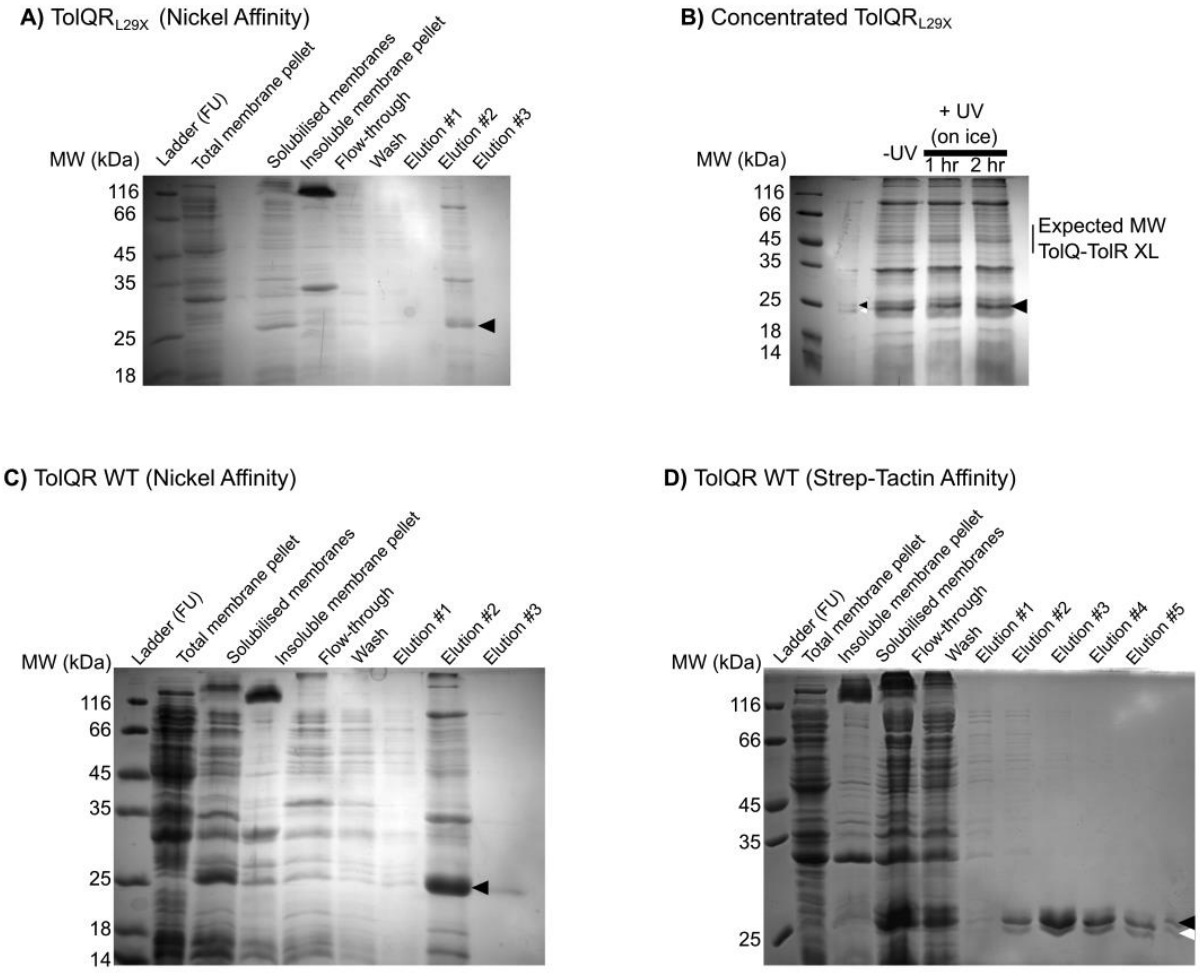


Figure 5.6 | Nickel-affinity purification of TolQR$_{L29X}$ yields insufficient purity for crosslink detection. A) Nickel-affinity purification yields TolQR$_{L29X}$ (*black triangle*) but also a number of contaminants that obfuscate the identification of crosslinked species when concentrated (***B***). **C)** Nickel-affinity purification yields WT TolQR (*black triangle*) also featuring contaminants. **D)** Strep-Tactin affinity purification yields slightly less TolQR (*black/white triangles*), but significantly fewer contaminants.

5.2.6 UV crosslinking solubilised TolQR BPA mutants *in vitro*

After optimising the purification of TolQR BPA mutants, UV-treatment was repeated on solubilised membranes containing TolQR$_{L12X}$, prior to purification by Strep-Tactin affinity. Some high-molecular weight bands appeared as result of UV-treatment, but of insufficient concentration for LC-MS analysis (Fig. 5.7).

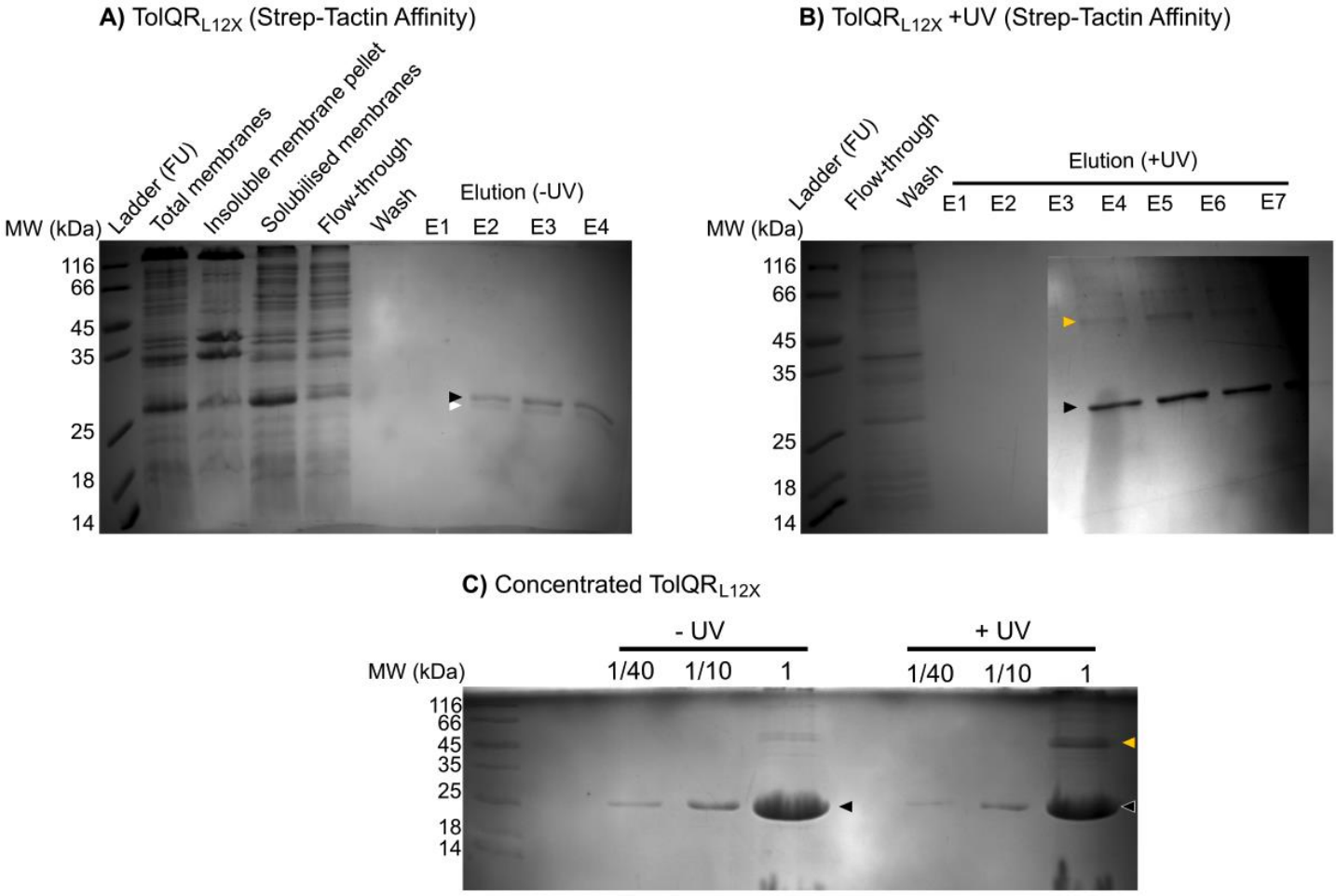


Figure 5.7 | UV-treating the TolQR$_{L12X}$ mutant in solubilised membranes enriches high molecular-weight protein species. Two purifications were performed, one where membranes were not UV-treated (***A***) and one where membranes were UV-treated (***B***). Some faint crosslinked bands were found at around 50 kDa (*contrast enhanced to highlight high-MW bands*). C) Samples were concentrated, demonstrating slight enrichment of crosslinked bands (*~45kDa, yellow triangle*) in the UV-treated sample.

5.2.7 UV crosslinking solubilised TolQR mutants *in vivo*

To optimise *in vivo* crosslinking, live cells expressing TolQR$_{L12X}$ were pelleted and crosslinked for two hours prior to Strep-Tactin-affinity purification (Fig. 5.8). When resolved by SDS-PAGE, high MW bands were present in the UV-treated sample only, suggesting successful crosslinking. This protocol was subsequently applied to the remaining BPA mutants, resolved at their highest concentrations to reveal low-abundance species (Fig. 5.9).

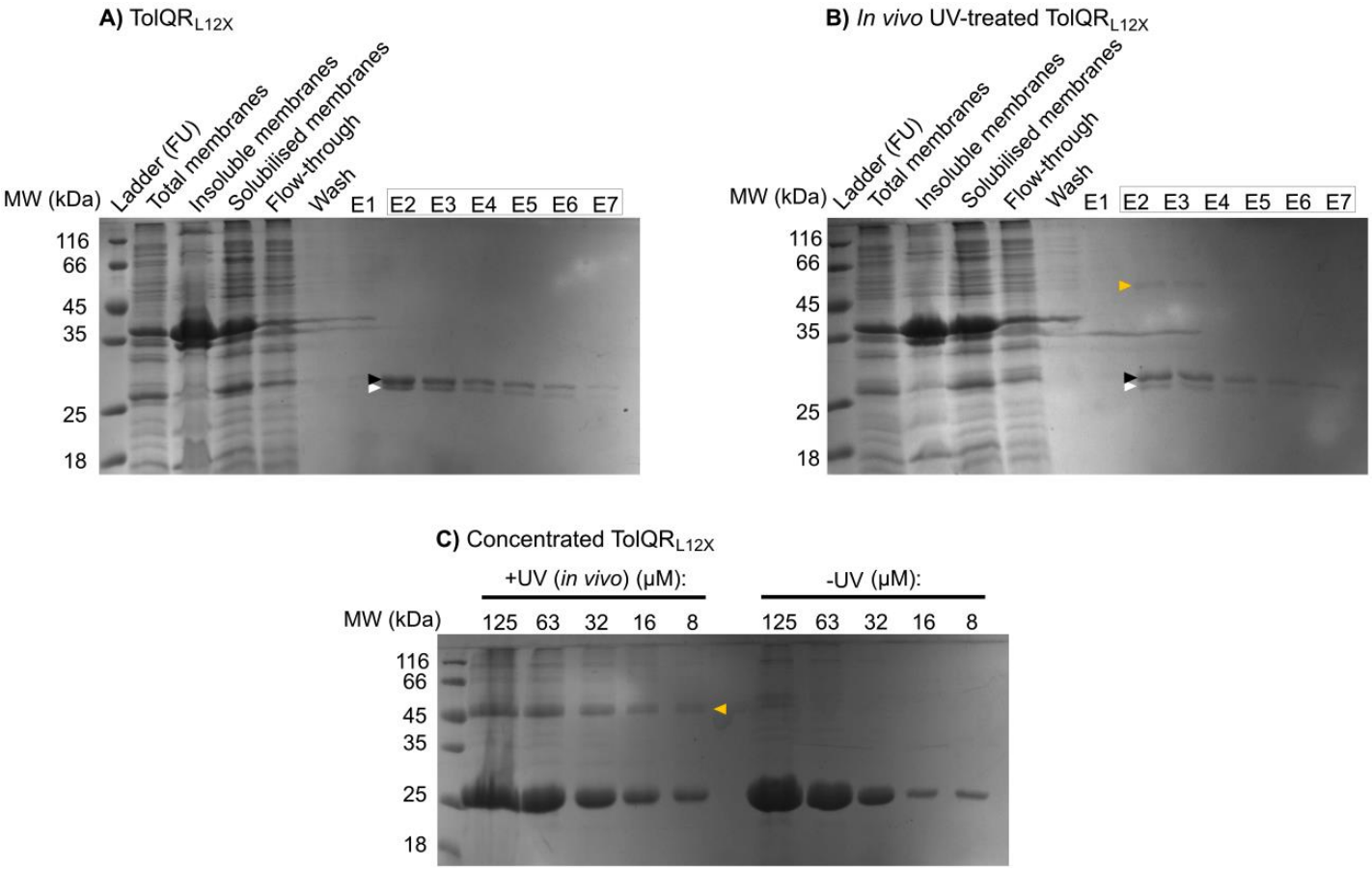


Figure 5.8 | TolQR$_{L12X}$ exhibits *in vivo* UV crosslinking. Two purifications were performed, one with no UV-treatment (***A***) and one where cells were UV-treated (***B***). High-molecular weight protein species (*~50 kDa, yellow triangle*) were visible in the UV-treated sample, suggesting crosslink formation. **C)** Samples were concentrated, demonstrating enrichment of crosslinked proteins in the UV-treated sample.

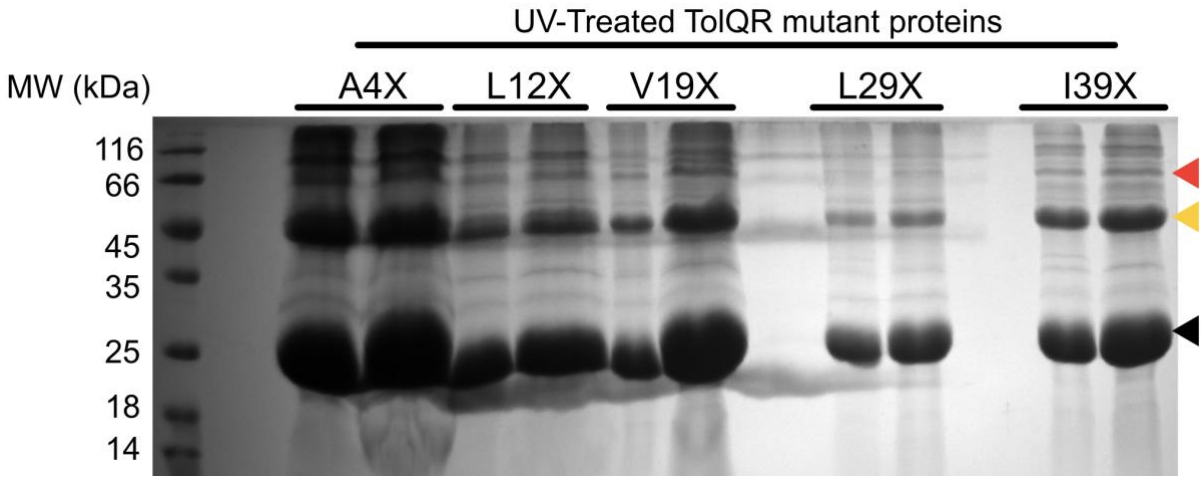


Figure 5.9 | TolQR BPA UV-treated *in vivo* exhibit crosslinking. Each protein was purified by Strep-Tactin-affinity from cells expressing different TolQR BPA mutants that were UV-treated on ice for two hours. These proteins were concentrated as far as possible, then resolved by SDS-PAGE (two lanes per mutant; neat (*right*) and diluted twofold (*left*)). This revealed that each mutant exhibits crosslink formation (*yellow triangle*), in addition to the excess of non-crosslinked TolQ/TolR (*black triangle*). Some higher-order crosslinks may also be present (*red triangle*).

5.2.8 Optimising proteolytic TolQR digests for LC-MS-MS

Preliminary attempts to resolve the BPA crosslinks in complexes were trialled by in-gel digestion using trypsin, before submission for LC-MS-MS. These data confirmed the presence of TolQ and TolR in each band, but the sequence coverage was insufficient for crosslink mapping and could not verify the incorporation of BPA within the peptides. Critically, the TolR NTD is rich in trypsin digestion sequences, generating peptides too short for detection here. Digestion with chymotrypsin and trypsin gave moderate sequence coverage, but not at the regions of interest (Fig. 5.10). To obtain MS data covering the BPA site, different digestions were optimised using wild-type TolR. A chymotrypsin GluC combination generates detectable peptide containing residues 29 and 39 for crosslinking mapping, whilst elastase digestion was required to generate peptides containing residues 12 and 19. No digestion conditions were able to resolve a peptide containing residues 1-10, eliminating these sites from mapping.

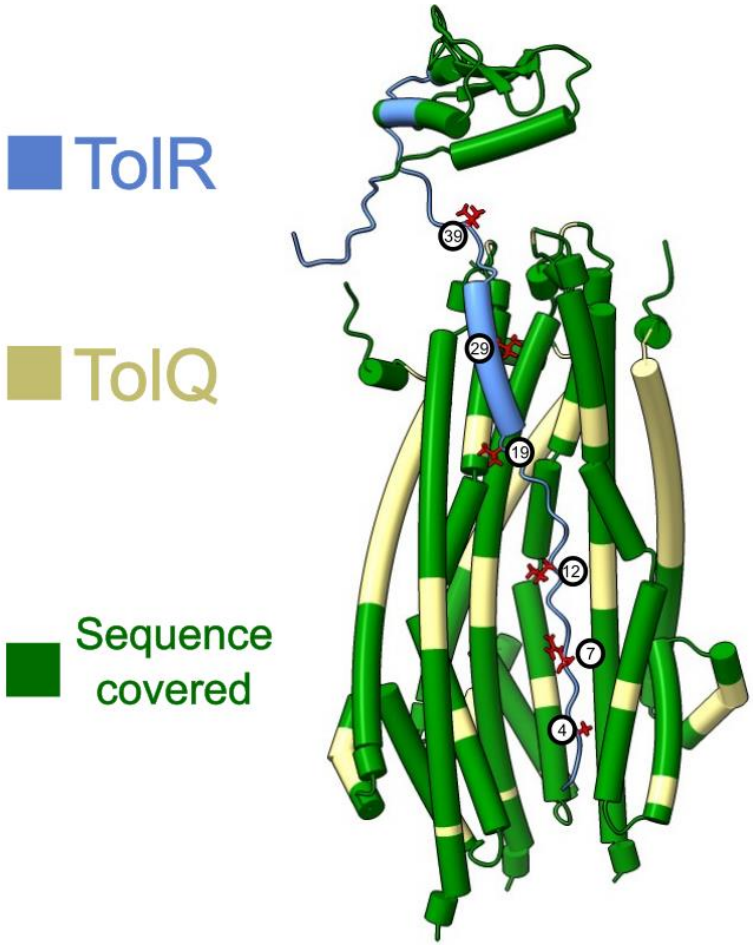


Figure 5.10 | Chymotrypsin-Trypsin double digest fails to resolve the TolR N-terminus. Protein regions detected by MS (*green*) did not include the TolR TPH or NTD. The model shown is the cryo-EM structure (PDB: 80DT) with full-length TolR docked from the Webby *et al.,* AF2 prediction [19, 131].

5.2.9 TolQR I39X forms intermolecular crosslinks

Digestion of crosslinked TolQR$_{I39X}$ samples yielded MS spectra indicating that it forms crosslinks between adjacent TolR$_{I39X}$ subunits. This is in agreement with the TolQR AF2 model, where these sidechains are generally oriented towards one another (Fig. 5.11). These residues are found within the disordered linker, slightly above the TPH, and are thus in a flexible position particularly amenable to self-crosslinking. This demonstrates that this approach is tenable to identify crosslinked residues.

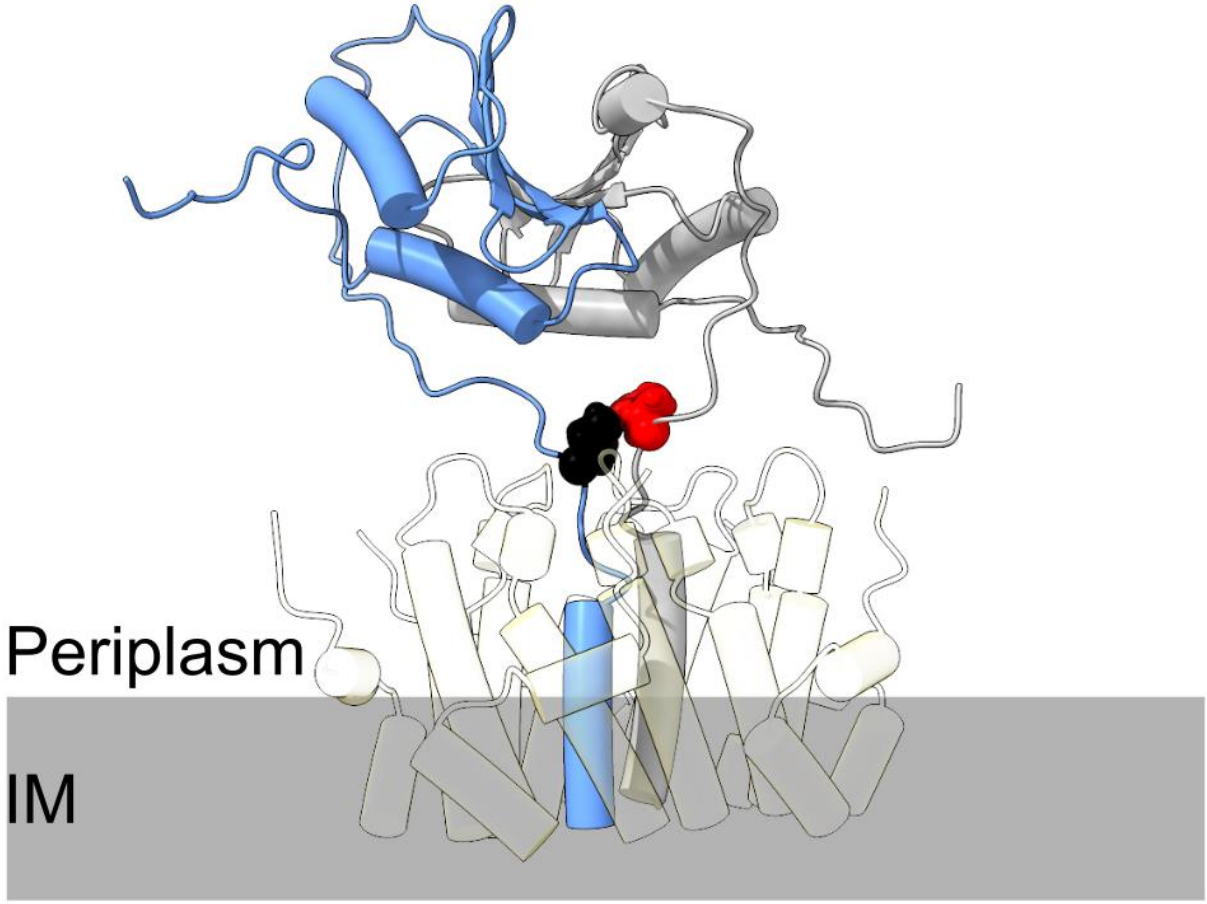


Figure 5.11 | TolR$_{I39X}$ subunits forms crosslinks within the TolQR complex between adjacent TolR$_{I39X}$ sidechains. Crosslinking of adjacent TolR residue 39 sidechains (*red/black spheres*) suggest that these residues are generally oriented towards one another, as predicted. The model shown is the AF2 model from Webby *et al*., featuring the open PGB conformation [97].

5.2.10 Chemical de-energisation alters UV-crosslink formation

Given the optimisation of UV-dependent crosslink formation in live cells, the protocol was thus applied to cells where they were first chemically depolarised by CCCP treatment. Cells were incubated with 1mM CCCP for 40 minutes prior to UV-treatment, which reduced the proclivity for higher order crosslinked species (Fig. 5.12). This suggests that crosslinks form in the energised cells resulting in three main protein bands on the gel. Crosslinks still form in the chemically depolarised cells, but the result is two protein bands on the gel. Additionally, a greater proportion of protein in the CCCP-treated samples was found not to crosslink (~25 kDa). Taken together, this indicates that the higher order crosslinked-species from energised cells reflect conformational intermediates of the TolQR complex. Further optimisation of proteolytic digestion for LC-MS-MS is required to identify the crosslinked residues present.

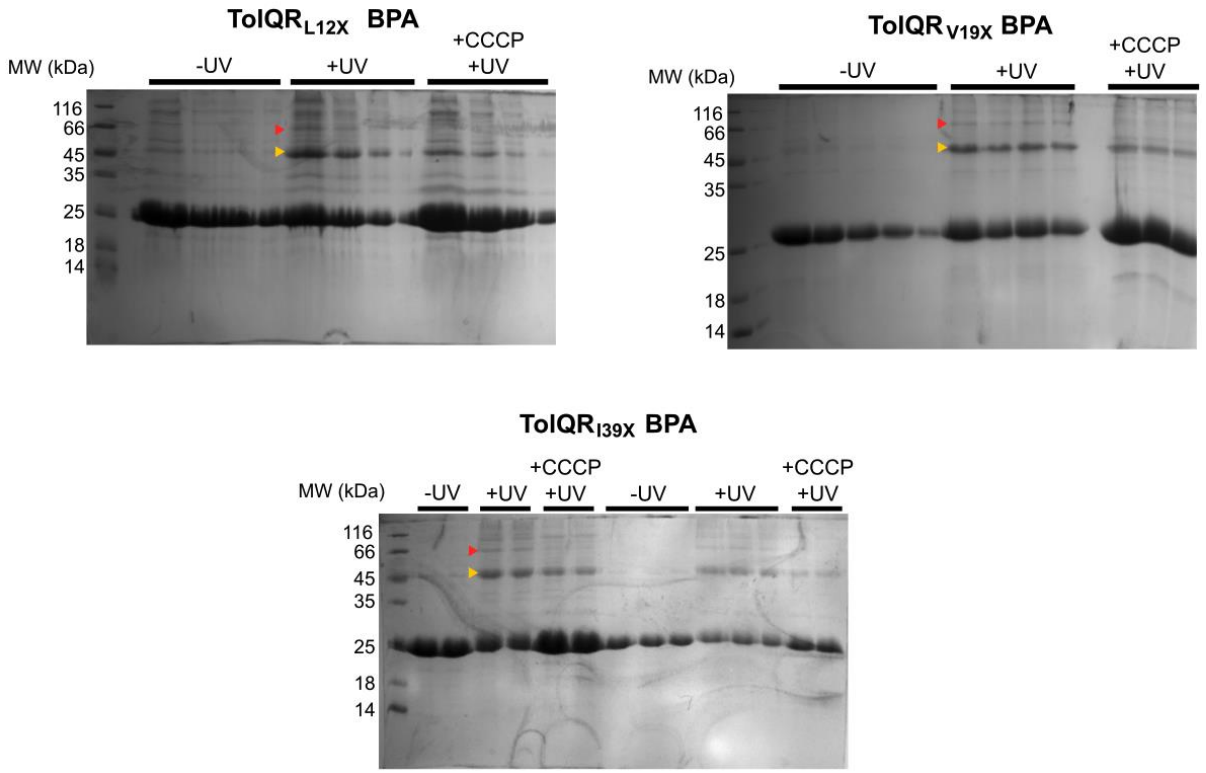


Figure 5.12 | Chemical depolarisation alters crosslink profiles of TolQR BPA mutants. UV-treatment generated crosslinks regardless of energisation state, as shown by CCCP-treated samples (*yellow triangles*). An additional band of low abundance (*red triangle*) was not found in the CCCP-treated samples. A greater proportion of the proteins present in the CCCP-treated samples lacked crosslinks (~ 25 kDa) compared to UV-treated only, similar to the non-crosslinked controls.

5.3 Discussion

5.3.1 Higher-order protein species suggest trimer formation

To elucidate molecular details of the conformational rearrangements that occur within TolQR during PMF transduction, *tag* mutants of TolQR were generated to incorporate BPA. I not only demonstrate that these mutants are functional *in vivo*, but also show that they form UV-dependent crosslinks, suggesting successful BPA incorporation. Moreover, the presence of bands at ~45 kDa are suggestive of crosslinked TolR-TolR (41.2 kDa) or TolR-TolQ species (46.3 kDa). Bands at ~66 kDa may be trimeric species, comprising either two TolR bound to the same TolQ monomer, or a second TolR bound to a TolR that is bound to TolQ (66.9 kDa), but further optimisation proteolysis and LC-MS-MS is required to validate these crosslinks. Moreover, crosslinking efficiency was greater in UV-treated samples that were not treated with CCCP, as indicated by a higher concentration of protein at 25 kDa in CCCP-treated samples. This suggests that during PMF-transduction, these BPA residues are found in positions more amenable to inter-protein crosslink-formation, perhaps because of TolR movement as a result of proton flux. This was perhaps surprising, as one might expect a stalled TolQR to generate more crosslinks, since each protomer is in a static position.

Crosslinked identities are resolved by LC-MS-MS, which has proven difficult due to the problematic nature of TolR digestion, but the identification of TolR39X crosslinks demonstrates the feasibility of this approach. Further work will take the proteins purified here and optimise the digestion protocol to map these crosslinks. The optimisation of BPA crosslinking the TolQR complex can be applied toward many further goals. For example, incorporating BPA residues within the TolR PGB may allow one to determine the conformational rearrangements that occur there during force transduction, and how TolR alternates between open and strand-swapped conformations *in vivo* in response to cellular energisation. Similarly, BPA crosslinking may be used to determine the PG-binding residues of TolR, which I propose are found in a cleft like those of MotB, Pal, and PorE, that sequester mDAP (Fig. 5.13) [119, 120, 246].

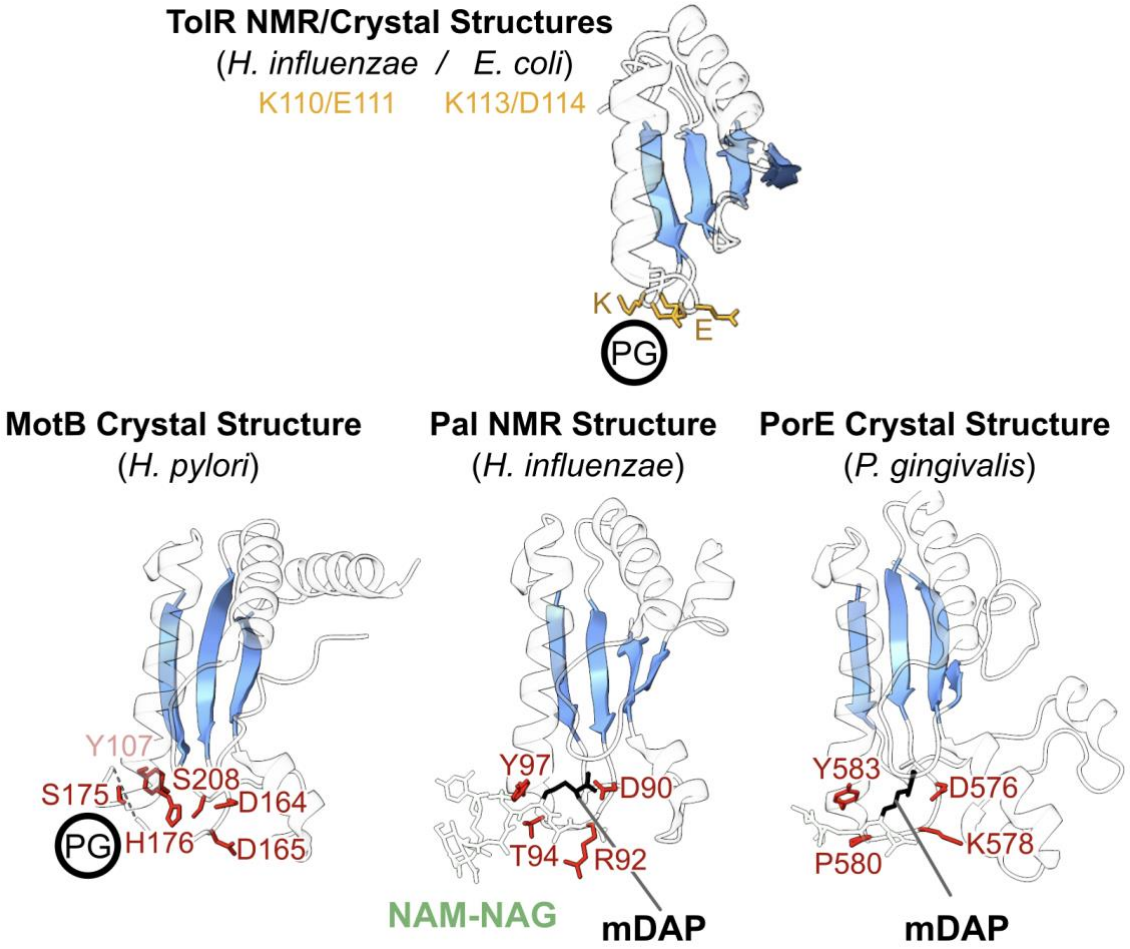


Figure 5.13 | The PG-binding cleft of Pal, MotB, and PorE may be present in TolR. Residues that bind NAM (*red*) are not present in TolR, save for a semi-conserved Glu-Lys pair (*orange*), also present in ExbD. Residues involved in PG-binding (*red*) share some conservation between Pal, MotB and PorE. MotB was solved as a crystal structure with NAM, but is specific to mDap (PDB: 3CYQ) [246]; Pal was solved in complex with a PG fragment by NMR (PDB: 2AIZ) [119]. The equivalent cleft within the PorE crystal structure binds the mDAP residue of a PG fragment (PDB: 6TOP) [247].

Neither ExbD nor TolR have been solved in complex with PG fragments, but both have been demonstrated to bind sacculi [105, 199]. For ExbD, transient binding with sacculi of *E. coli* and *S. aureus*, but not *B. subtilis*, was shown by NMR spectroscopy. *S. aureus* features different peptide stems to *E. coli*, while *B. subtilis* glycans are commonly N-deacetylated [248, 249]. Cleavage of NAM-NAG bonds within sacculi with muramidase abolished this sacculus-binding profile, suggesting that ExbD transiently binds the core glycan chain [199]. If TolR does bind the core glycan chain this may indicate a preference for denuded PG, which is enriched at the cell septum and the target of other division-related proteins [106, 250]. Zinke *et al.*, also propose that transducer binding dissociates the PGB dimer, followed by torque delivered directly to TonB by the ExbD dimer itself; the interactions of TolR-TolA may thus be amenable to BPA crosslinking. The assembly of TolQRA and this possible interaction is explored in the next chapter.

6.0 Interactions of TolQR with TolA

6.1 Introduction

The TolA-TolQ interaction modelling and co-conservation analysis were published in Webby *et al.,* [97]. Work of my own is highlighted accordingly.

6.1.1 PMF-dependent interactions of TolA with TolQR

In the previous chapter, we discussed the conformational rearrangements likely to occur within TolQR that drive rotation of the motor complex. This rotation is assumed to be applied to TolA in the form of torque. In this chapter, I aim to examine how this occurs, starting with how TolA assembles with TolQR motor. In the previous chapter, we modelled the TolQR motor [97]. Here we extended this work, modelling the TolQR-TolA assembly and identifying structural features within TolA that were previously unknown. TolA primarily interacts with TolQ within the transmembrane domain, and it has been indirectly shown that mutations in the TolAI SHLS motif abrogate TolA function [95]. This motif is conserved in TonBI, and its role in mediating interactions with ExbBD are presumed equivalent to those of TolA with TolQR [97, 251]. This is bolstered by evidence that TolA interacts with ExbBD in *ΔtolQR* mutants, and TonB utilises TolQR in *ΔexbBD* mutants, suggesting they share homologous motor-binding mechanisms [160]. The proposed interaction site of our TolQRA model was consistent with previously published mutants and validated by co-conservation analysis, as in chapter five. In doing so we identified the motor-box, a β-strand motif found within TolAII that was predicted to interact with TolR. I thus designed and assayed a collection of TolA mutants to better understand the interaction of TolA not only with TolQ, but also TolR.

6.2 Results

6.2.1 Modelling the TolQ-TolA interaction

Here we used a combination of AF2, molecular docking and Rosetta to model the TolQR complex and its assembly with TolA, and then assessed the degree of co-conservation as estimated by RaptorX [97, 156, 173]. These data highlighted a co-conserved groove along the face of TolQ, complementary to a single face of the TolA TMH (Fig. 6.1). Mutants in the TolA transmembrane SHLS motif tend to confer detergent sensitivity consistent with system dysfunction [95]. The OM-stability of $TolA_{H22P}$ and $TolA_{H22R}$ mutants was greatly reduced by the introduction of these residues that lack aromatic character, but OM-stability was restored by suppressor mutants at the edge of the helical groove ($TolQ_{G26D}$, $TolQ_{I29S}$, and $TolQ_{A30V}$) [95]. In contrast, the $TolA_{H22Y}$ mutant conserves the aromatic character of H22, and retains native levels of OM-stability. In $TolA_{H22Y}$, TolQ suppressor mutants destabilise the OM, suggesting they disrupt π-stacking interactions of $TolA_{H22}$ with $TolQ_{F183}$ and $TolQ_{W32}$.

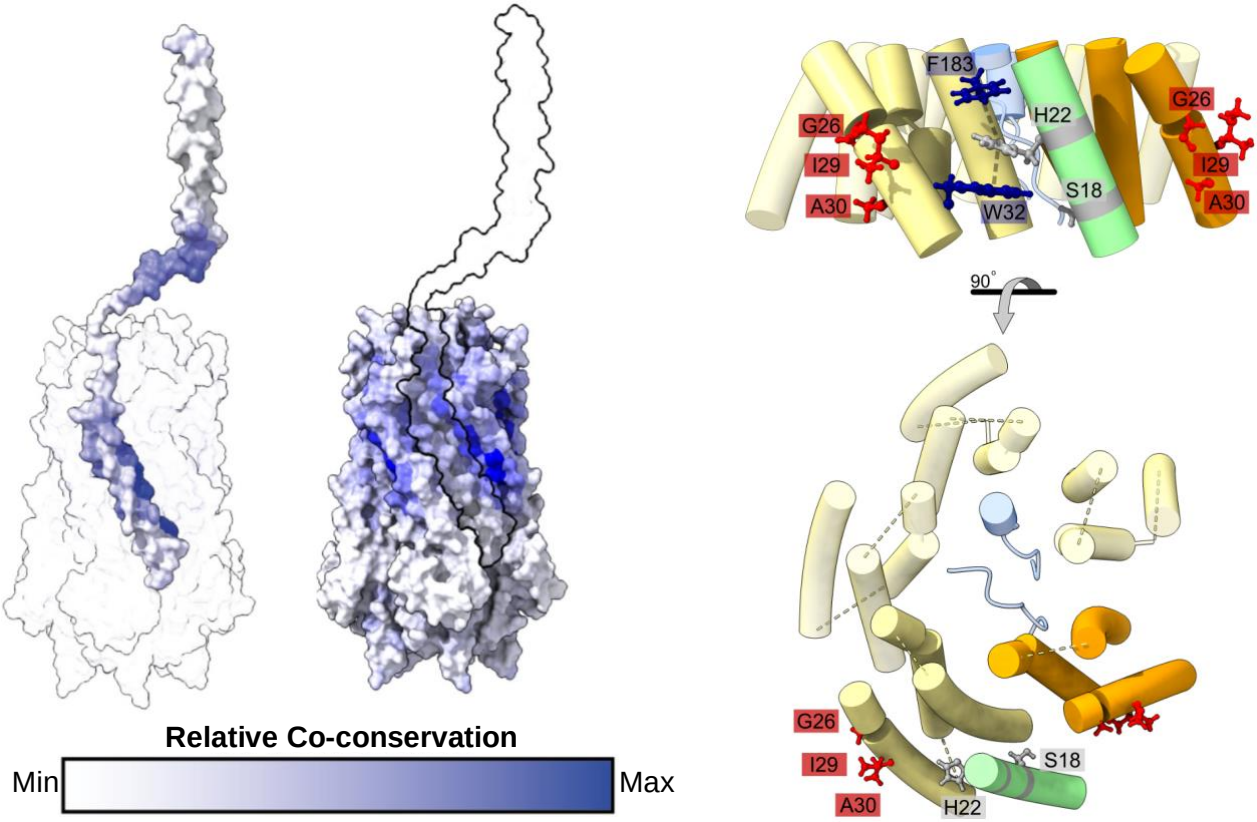


Figure 6.1 | Co-conservation analysis of the putative TolQ-TolA binding site. The TolA TMH was predicted to associate with TolQ helices, at the interface between adjacent TolQ subunits (*left*), as indicated by maximum co-conservation scores of TolQ-TolA from RaptorX output. The groove between the TolQ subunits shows highest co-conservation with the TolA TMH (*black*). The organisation of TolQ helices relative to TolA (*green*) demonstrates packing of TMHs (*right*). The positions of suppressor mutants (red) on TolQ are adjacent to the SHLS residues in this model (*grey*), as well as TolQ residues that may partake in π-stacking (*blue, dashed line*) [95]. Adapted from Webby *et al.* [97].

6.2.2 Conserved TolA TMH His_{22} drives TolA-TolQ interaction

The essential nature of $TolA_{H22}$ belies its assumed role in TMH association, but this interaction has not been clearly demonstrated [95]. TolQ, TolR and TolA have all been shown to independently accumulate at the septum in *ΔtolQRA* cells, but TolAI alone is not recruited to the septum in this strain [108, 109]. To test whether $TolA_{H22}$ drives association with TolQ, fluorescence microscopy was performed (Fig. 6.2). In ΔtolA cells where TolQR is present, GFP-TolAI accumulated at the septum through its interaction with TolQR while GFP-$TolAI_{H22A}$ did not, confirming that the H22A mutant does not associate with TolQ *in vivo*. In the *ΔtolQR* strain, neither TolAI nor $TolAI_{H22A}$ accumulated at mid-cell, confirming that TolQR influences septal trafficking of TolAI.

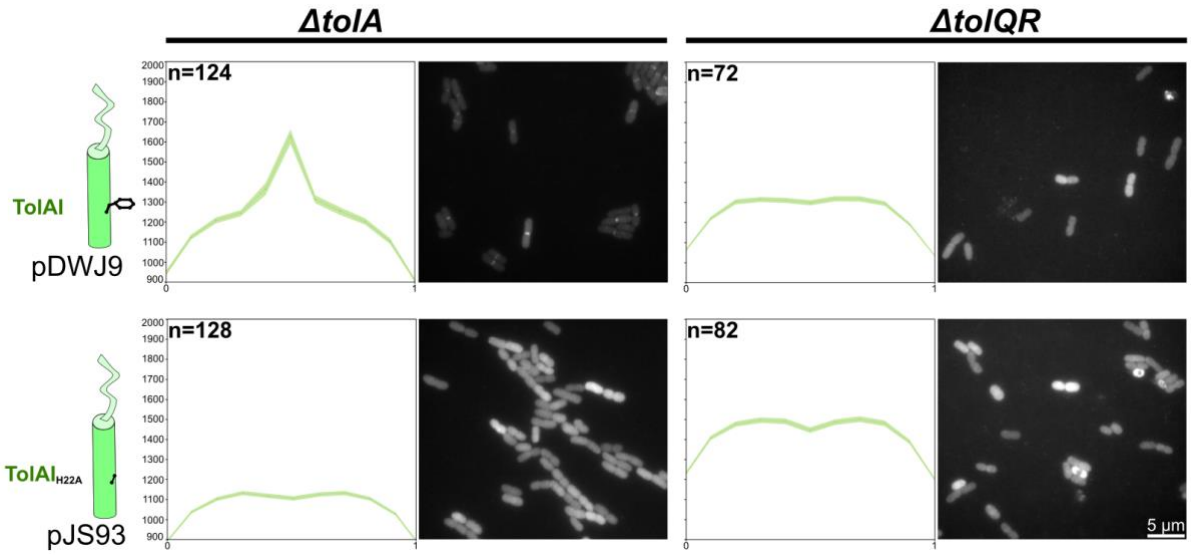


Figure 6.2 | GFP-TolAI is recruited to the septum by interaction with TolQR, driven by $TolA_{H22}$. Competent *ΔtolA* and *ΔtolQR* mutants were transformed with GFP-TolAI and GFP-$TolAI_{H22A}$. In the presence of TolQR, GFP-TolAI migrates to the septum, whereas in the absence of TolQR it does not, suggesting TolQR recruits this domain to the septum. In contrast, GFP-$TolAI_{H22A}$ does not migrate to the septum in either mutant, suggesting that the His_{22} directly interacts with TolQR *in vivo*. Relative fluorescence intensity is indicated on the y-axis (AU). *n* denotes the number of cells being averaged.

6.2.3 A novel motif for transducer-stator interaction

During the co-conservation based validation of the TolQRA model, RaptorX revealed residues within TolA (50-56) that were highly co-conserved with the TolR C-terminal β-strand and to a lesser extent, the other β-strands of the TolR PGB (Fig. 6.3). This co-conservation was subsequently compared with PsiPred 2.0 structural predictions of TolA. This comparison revealed that the co-conserved site on TolA overlaps with the β-strand predicted by PsiPred 2.0, that was not observed in the TolQRA model when TolA was predicted in isolation. A range of additional proteobacterial species were tested by PsiPred and RaptorX, all of which revealed the same co-conserved β-strand within this intrinsically disordered region (IDR) (Fig. 6.4A). This co-conservation analysis was extended to ExbD-TonB, revealing that TonB features a similarly co-conserved β-strand (Fig. 6.4B). As this β-strand is suggested to interact with the dimeric component of both the Tol/Ton motors, it was thus dubbed the "motor-box".

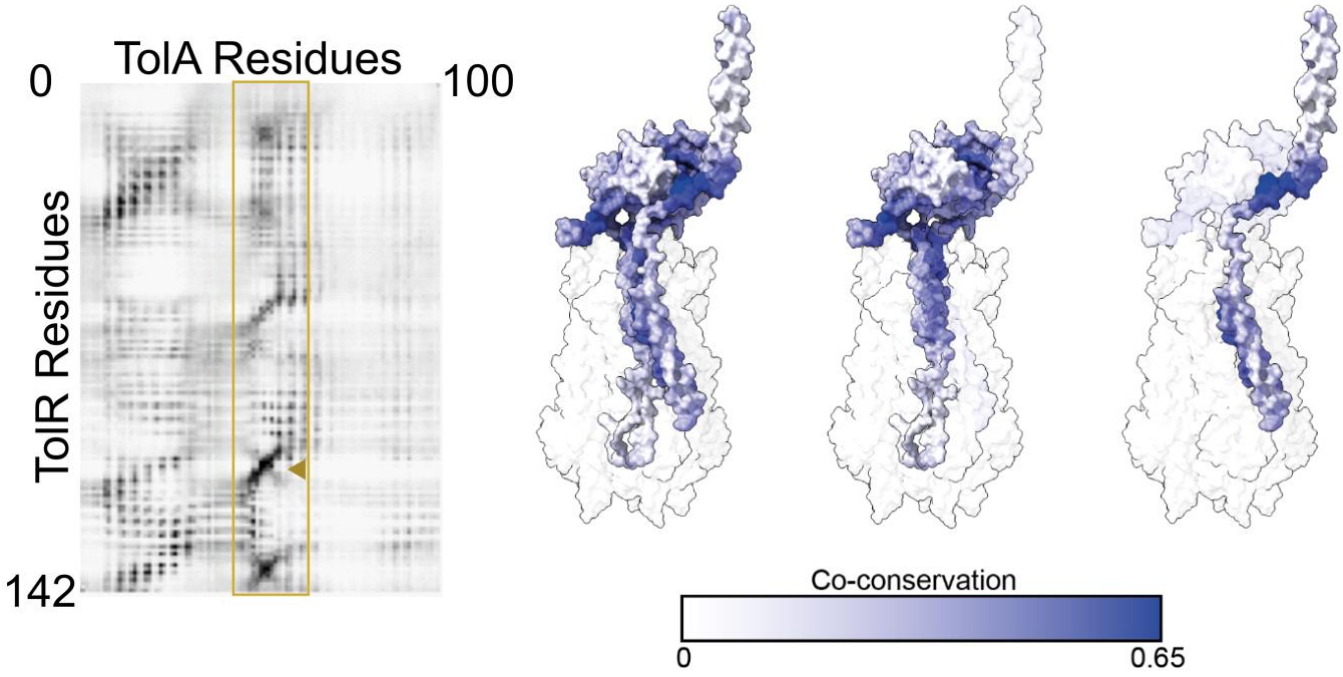


Figure 6.3 | TolR-TolA interactions predicted in periplasmic region between TolR and TolA residues 45-60. RaptorX output of TolR-TolA$_{1\text{-}100}$ (*left*) suggests a short patch of residues of TolA around residue 50 is co-conserved, and therefore likely to interact with TolR [156]. The highest scoring residues of TolR belonged to β-strands comprising the PGB (*brown triangle*), while the highest scoring residues of TolA were found within the disordered region between TolAI and TolAII (*residues 40-60; yellow box*). The highest RaptorX score per residue was mapped onto the TolQRA model, suggesting a plausible interaction site (*right*). The shape of the co-conserved patches are suggestive of antiparallel β-addition, except for the final TolR β-strand which lacks obvious directionality.

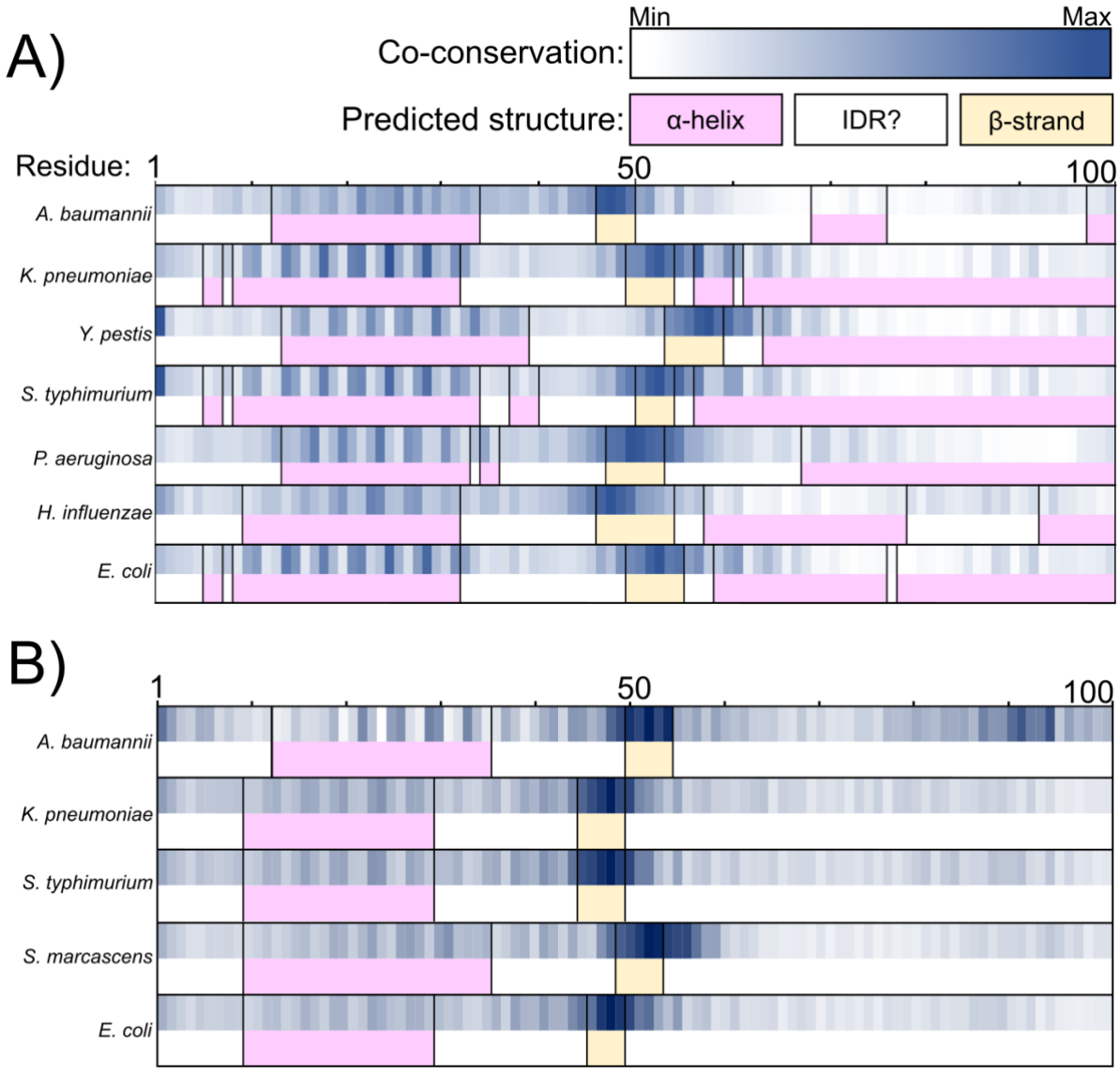


Figure 6.4 | Co-conservation maxima correlate with PsiPred 2.0 β-strand prediction sites. A) TolA sequences from a range of γ-proteobacterial species show a transmembrane helix (*pink*, *left*) followed by an IDR containing a predicted β-strand (*yellow*) followed by domain II α-helices (*pink*, *right*). The disordered region of each of these TolA sequences was found to be the site of maximum co-conservation with TolR from each of these species (*scores indicated by white-blue spectrum*). The transmembrane helix (left) is followed by a disordered region. **B)** TonB sequences from a range of species also possess a predicted β-strand shortly after the transmembrane helix, co-conserved with respective ExbD sequences. PsiPred does not predict polyproline helices, hence there is no distinction between PPII-helices and IDRs.

The motor-box was later identified in 95% of TolA and TonB sequences assessed by PsiPred 3.0 in chapter four (SI. 1-2), as well as each of the TolA/TonB chimeras generated in chapter three, including AIA and BIB. Taken together these data suggest there is a conserved yet enigmatic interaction between the force transducing component and the globular TolR/ExbB PGB, suggesting both Tol and Ton assemblies utilise this β-addition during force transduction. Interestingly, structural prediction with Robetta predicts that *H. influenza* TolA contains a β-hairpin within the co-conserved

region [98]. The co-conservation data for *H. influenzae* TolR-TolA$_{1-100}$ was thus mapped to their Robetta prediction models where the co-conserved region of TolA coincides with a predicted β-hairpin, and the co-conserved region of TolR maps to the β-strands of its PGB (Fig. 6.5).

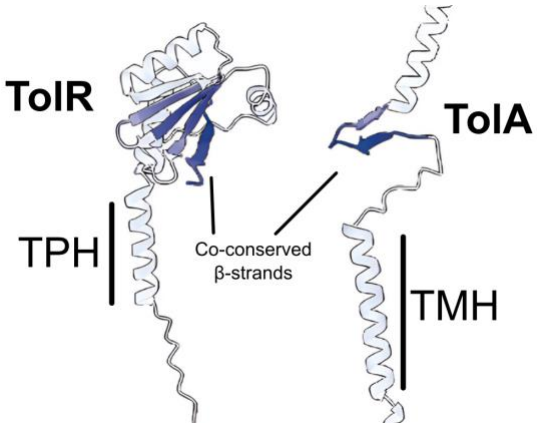


Figure 6.5 | Co-conservation scores of *H. influenzae* TolR-TolA mapped to respective Robetta prediction models. TolR and TolA from *H. influenzae* were structurally predicted in isolation, using Robetta [98]. Notably, Robetta predicted the formation of a β-hairpin in TolA, where one of the β-strands (*darkest blue*) is the co-conserved motor-box motif.

To test whether *E. coli* TolA forms a β-strand when co-predicted with TolR, TolA residues 46-63 were structurally predicted with two TolR monomers using AF2 [71]. The prediction suggested that TolA$_{51-56}$ forms a β-strand that was not predicted in the TolQRA model where TolA was predicted in isolation and then docked onto the TolQR model. This also suggests that the co-conserved β-strand inserts between the TolR C-terminal β-sheet (Fig. 6.6).

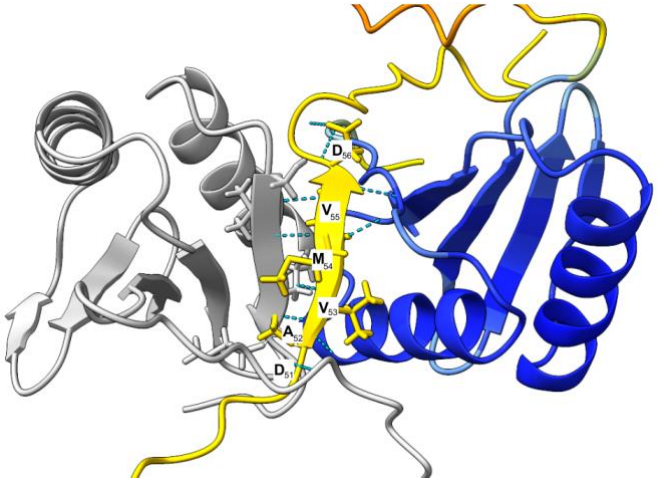


Figure 6.6 | The *E. coli* TolA IDR is predicted by AF2 to interact with TolR via β-addition. TolA (*yellow*, *residues 46-63*) are shown, with motor-box residues labelled (*51-56*). The motor-box is close to the RaptorX and PsiPred predictions (*residues 50-55*), showing backbone interactions with the TolR CTD (*grey*). The other TolR subunit is coloured by AF2 plDDT, as is TolA.

TolR and ExbD are both known to exist in open and strand-swapped conformations (Fig. 5.1, chapter five) [13, 97]. When the same co-prediction was applied to $ExbD_2$-TonB, it became apparent that the TonB motor-box contacts the C-terminal β-sheet of both ExbD monomers. When aligned with the $TolR_2$-TolA motor-box interaction, both are predicted to form a contiguous β-sheet comprising all three proteins where the PGBs are antiparallel, like the strand-swapped models (Fig. 5.1, chapter five) (Fig. 6.7).

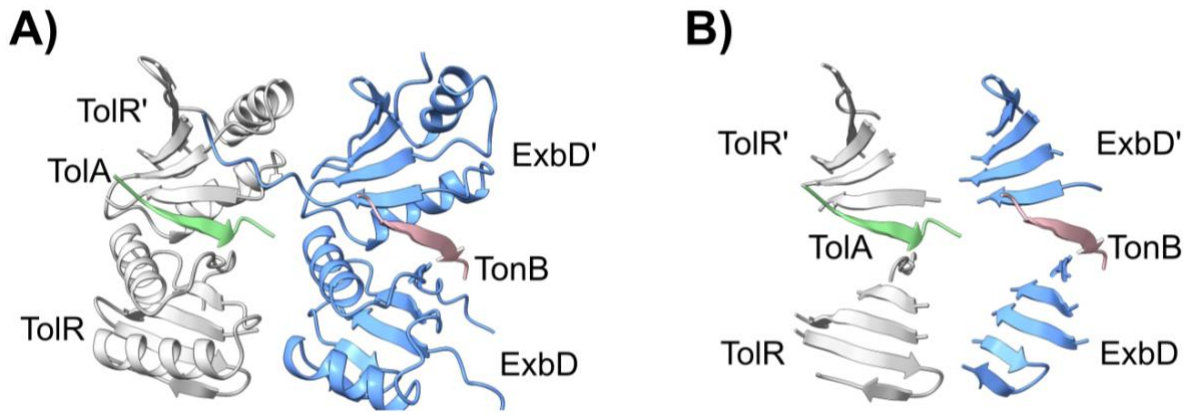


Figure 6.7 | Motor-box predictions suggest formation of a continuous, extended β-sheet. A) The TolA and TonB IDRs are predicted to form β-strands which sandwich between β-sheets of their adjacent PG-binding dimers (*AF2 models*). **B)** Showing only β-strands reveals an extended β-sheet that spans all three proteins.

This structural modelling suggests that the motor-box likely undergoes β-insertion with both PGBs at the dimer interface, thus interacting with the edge of each β-sheet. This connects the dimeric PGBs into a single β-sheet, where TolA/TonB is ostensibly clamped within the centre of respective TolR/ExbD dimers. Upon release of AF3, these motor-transducer assemblies were predicted at full length, which reiterated this prediction of a contiguous β-sheet (Fig. 6.8). Analysis reveals slight sequence similarity between the motor-box motif and the β-strand it displaces, as they both feature residues with a propensity to form a β-strand. In TolA this motif is IDAVMVD, and it displaces a TolR β-strand of sequence ITQSVE. In contrast, the TonB motor box sequence is ISVTMV, which displaces an ExbD β-strand of sequence VDVKVNL.

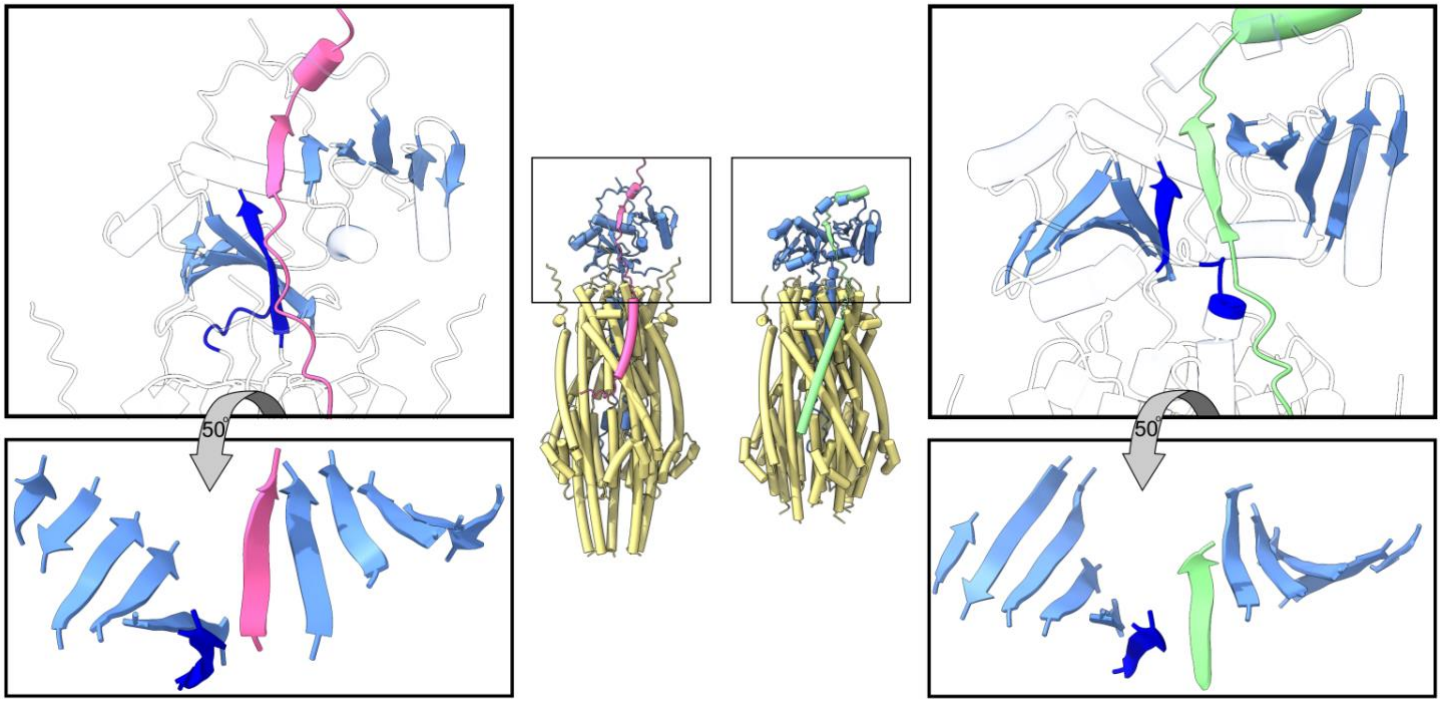


Figure 6.8 | AF3 prediction of Ton and Tol motor-transducer complexes predicts TolA/TonB motor-box threading between single-strand-swapped TolR/ExbD dimers. For both ExbBD-TonB (*left*) and TolQRA (*right*), only one of the two dimeric subunits were predicted to form strand-swapped β-strands (*dark blue β-strands*). The equivalent residues of the other protomer were predominantly disordered (*dark blue random coils*). The relative orientation of the TolA/TonB TMH is predicted perpendicular to the helical groove of the pentamer, rather than inserted as suggested by docking data (*Fig. 6.1*). Similar interactions were predicted in all TolQRA homologue sequences tested, from the species outlined in chapter four.

6.2.4 The motor-box is sequence-agnostic but sensitive to mutation

TolA mutants were generated at the site of the motor-box, to determine whether any could perturb TolA functionality (Fig. 6.9). This included proline scanning, replacement with alanine-rich linkers (12x-Ala, 6X-Ala) glycine-alanine linkers (2x-GAG), charge swapping (2x-GSK), as well as deleting the IDR (ΔIDR), replacing it with an α-helical sequence from domain II (ΔIDR-helix), and replacing it with the IDR from TonB (TonB-IDR). It is well-known that β-sheets form through backbone interactions, via hydrogen bonding, but, the kinked geometry of proline and its lack of a backbone amide proton, should confer an intrinsic capacity to disrupt β-sheet structures [252, 253].

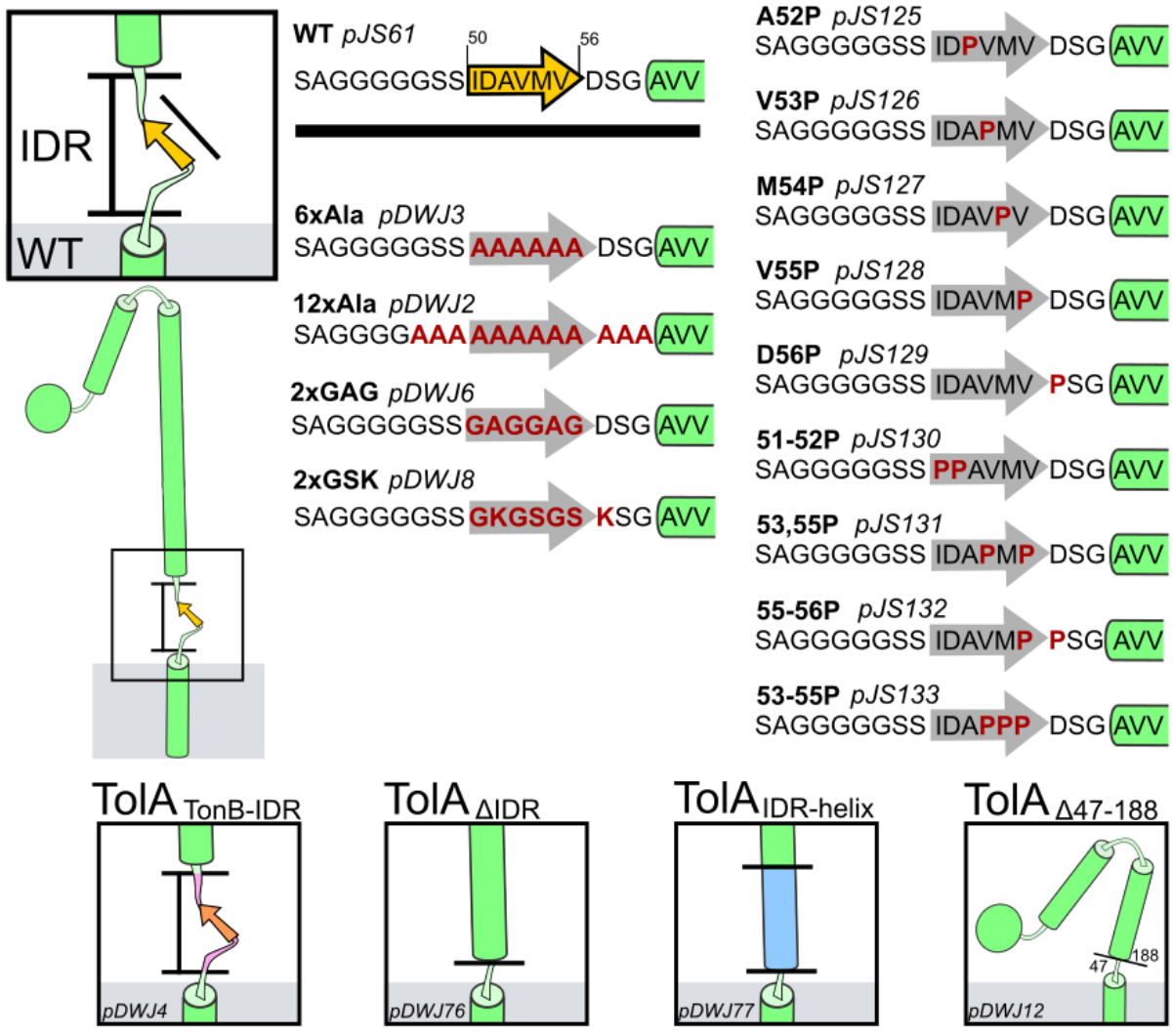


Figure 6.9 | Motor-box mutants designed to perturb TolA function. These were designed to prevent the interaction of the TolA motor-box with TolR through loss/change of secondary structure, loss of charge or strand-breaking.

Of these mutants, all proline point mutants conferred WT OM stability. Substitution of the motor-box with six alanine residues yielded no deleterious effects, whereas twelve alanine residues or two GAG repeats conferred partial OM instability. Swapping the two WT aspartates for lysine residues had no impact on OM stability. Deletion of the entire IDR conferred reduced OM stability, but swapping it for the IDR of TonB conferred WT OM stability. In contrast, deletion of 141 domain II residues in $TolA_{\Delta 47\text{-}188}$ reduces OM stability.

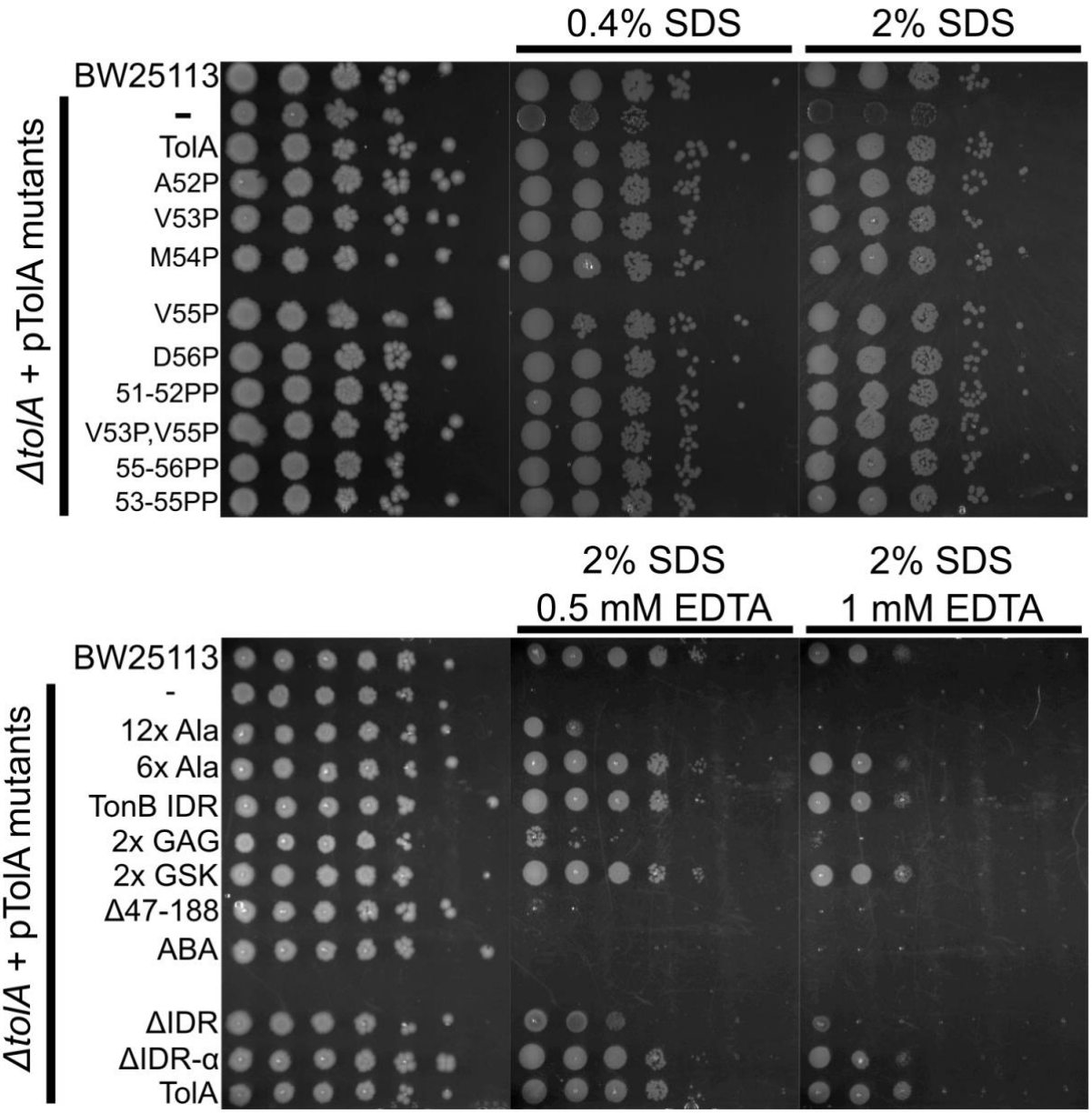


Figure 6.10 | Motor-box mutants vary in OM stabilisation abilities. Cells were spotted on 0%, 0.4% and 2% SDS (v/v) on LB agar plates as previously described (chapter three). EDTA was also tested, to increase the threshold of OM-stability required for growth.

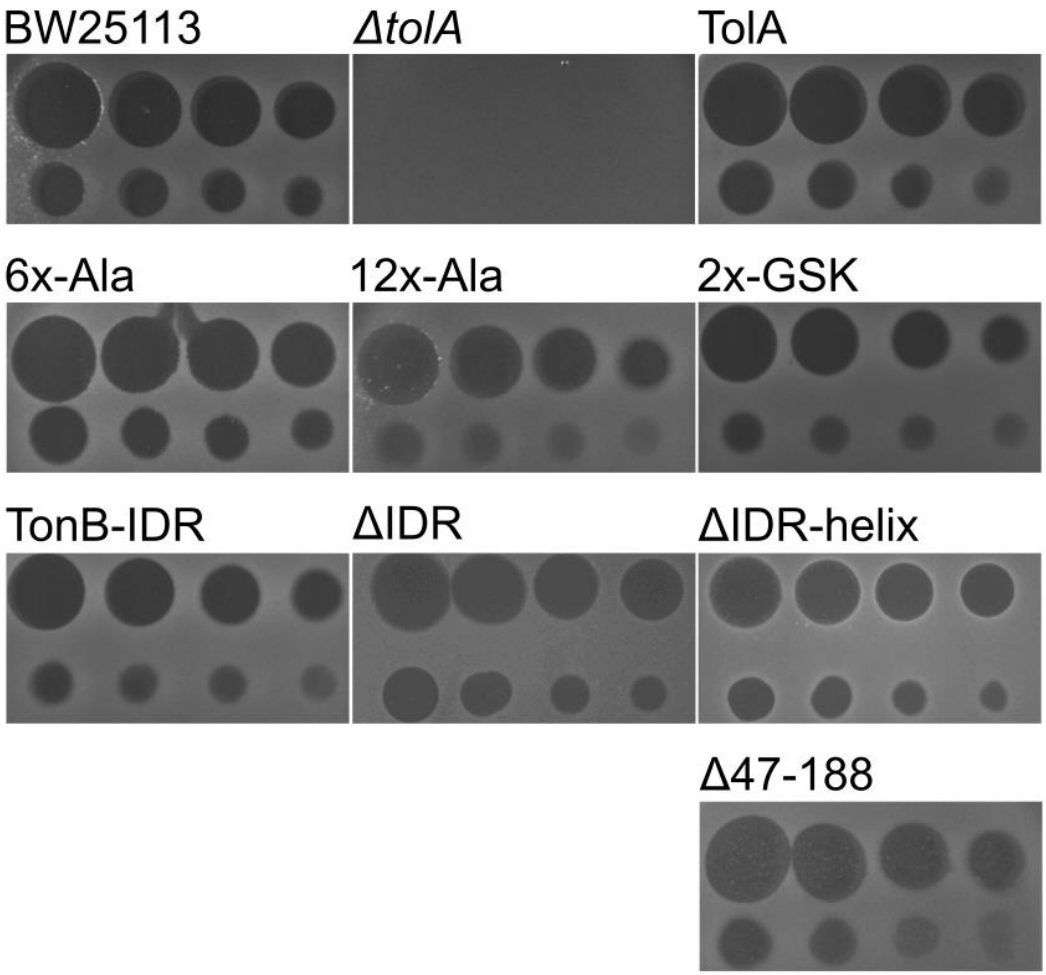


Figure 6.11 | TolA motor-box mutants retain colicin sensitivity. Cells were assayed with ColE9, as described in chapter three. In particular, mutants that gave reduced OM stability (*12x-Ala, ΔIDR, Δ47-188*) were still found to remain colicin-sensitive, but hazier clearance zones suggest a slightly reduced colicin sensitivity.

6.2.5 Modelling suggests the motor-box drives TolR rearrangement

When the full length AF3 prediction is compared with the two solved structures of TolR, it more closely resembles the antiparallel arrangement of the strand-swapped conformer (Fig. 6.12A) [105, 119]. In contrast, only the open conformer was found to bind sacculi previously (Fig. 6.12B), suggesting that the TolQRA AF3 model (Fig. 6.12C) cannot bind PG [105]. Taken together, these models suggest that the motor-box may induce the strand-swapped conformation of TolR, which is no longer able to bind PG. The ExbD dimer was recently solved in complex with the TonB motor-box, again comprising only residues 60-141 (lacking strand-swapping residues), revealing an antiparallel arrangement of ExbD subunits like that expected for a strand-swapped model (Fig. 6.12F) [199]. These residues, in isolation, are equivalent to those that form the open TolR conformer that binds PG (Fig. 6.9A/E)[105, 199, 239]. Moreover, similar ExbD constructs where the strand-swapping residues are deleted, are non-functional [254, 255]. This counterintuitively indicates that strand-swapping is necessary for TolA/TonB function, yet abolishes PG-binding. This is further supported by a pair of ExbD constructs, disulphide crosslinked to maintain either the open or closed conformation, that are both non-functional *in vivo* [199].

These predictions also suggest that the open conformation of TolR is further extended from the IM than the strand-swapped model where the extra β-strands are sequestered, shortening the effective reach of the disordered linker (Fig. 6.8). This is consistent with the previous interpretation that TolR is able to reach the PG layer in its open conformer, but unable to reach the PG layer in its strand-swapped state [97, 105]. This would further suggest that motor-box binding induces both shortening of the TolR linker and strand-swapping, concomitant with loss of PG-binding. Previously, the hydrophobic groove formed by the β-sheet in the open conformation of TolR (Fig. 6.12A) was suggested to be the binding site for PG, though this was never demonstrated [239]. In contrast, structural comparisons with MotB/Pal solved with their PG ligands suggests that binding more likely proceeds via charged residues occurring in a cleft between loops and α-helices (Fig. 5.13). The conformational rearrangements that occur within TolR/ExbD may alter PG affinity by presenting these PG-binding residues at one face of the PGB.

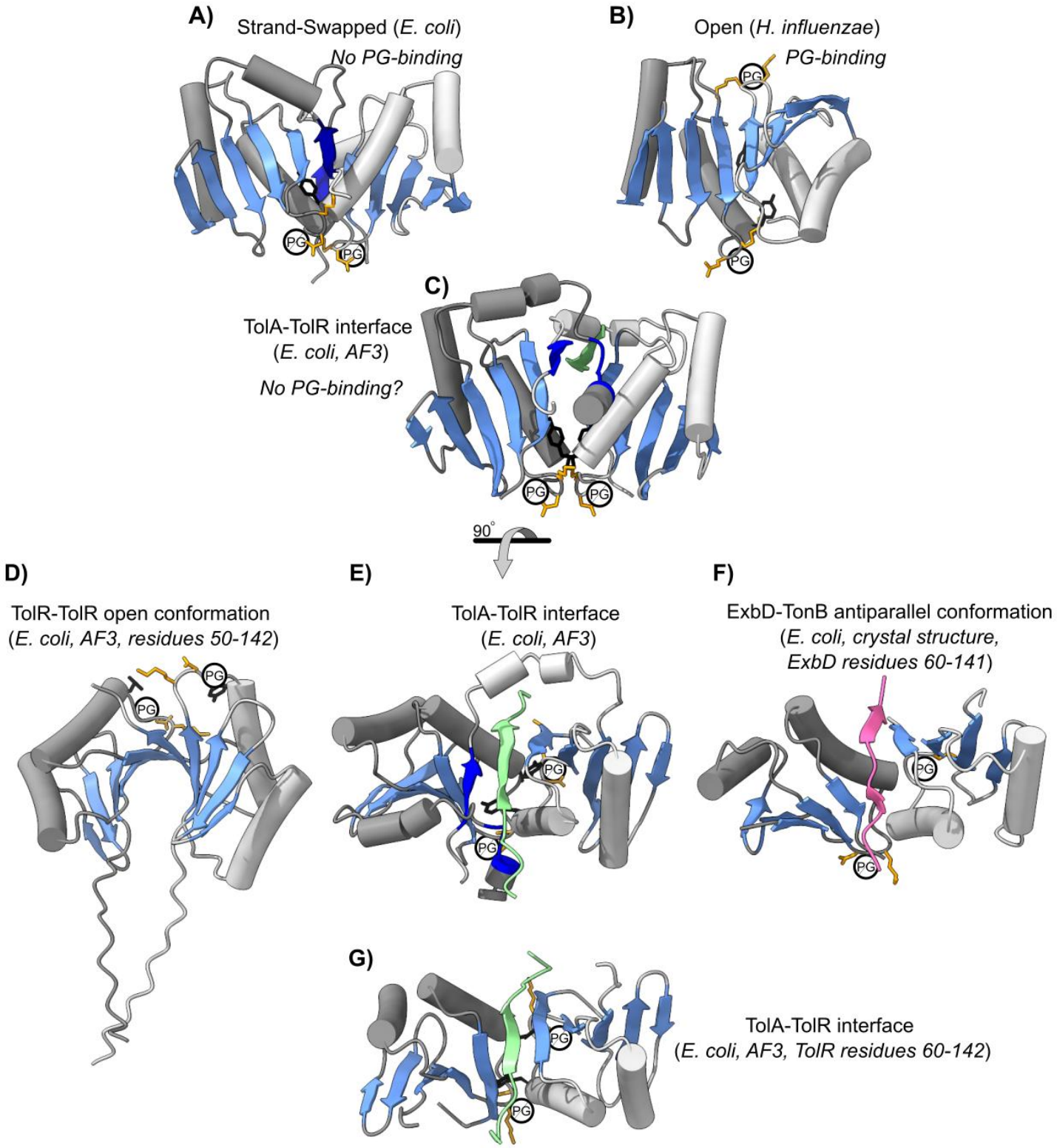


Figure 6.12 | The TolA motor-box may induce TolR strand-swapping. A-B) Resolved structures of TolR dimers (*from Fig. 5.1*) show two distinct orientations, only one of which is able to bind sacculi (*A*). **C)** TolR-TolA AF3 model (*Fig. 6.8*) demonstrates a TolR conformation where only one swapped β-strand is formed (*dark blue*). **D)** Modelling the truncated residues of *E. coli* TolR confers an open conformation, as seen in B. **E)** Co-predicting full length TolR with full length TolA generates a single-strand-swapped TolR model (*taken from Fig. 6.8*). The truncated open TolR model features only residues of the construct demonstrated to bind PG [105], where PG-binding residues (*orange*) are found on opposite sides of the parallel β-sheet (*as seen in* A). **F)** The truncated ExbD-TonB solved structure exhibits an antiparallel ExbBD arrangement in complex with the TonB motor-box (*right*, PDB: 8P9R) [199]. **G)** Co-predicting the open TolR conformer (*same residues as D*) with the TolA IDR yields an antiparallel TolR conformer. Panels D-G are upright, with respect to TolQRA from Fig. 6.8.

6.3 Discussion

To better understand how the TolQRA complex assembles, we generated a prediction model using AF2, where we assumed that a single TolA monomer interacts with TolQR. There are no existing measurements of TolQ expression, but the copy number of TolA is significantly lower than that of TolR, thus it has been proposed that only one TolA molecule interacts with each TolQR motor [19]. TolQ suppressor mutants that rescue $TolA_{H22P}$ and $TolA_{H22R}$ mutants line the transmembrane groove where TolAI docks (Fig. 6.1)[95]. Each of those TolQ point mutants are likely to affect the helix-packing, flexibility and hydrophobicity of TolQ; $TolQ_{G26D}$ introduces a negative charge and reduces flexibility of the outer TolQ helix, whereas $TolQ_{I29S}$ and $TolQ_{A30V}$ introduce polar/apolar sidechains, respectively. I demonstrated here through GFP-TolAI mutants that the conserved $TolA_{H22}$ binds TolQ, which, our model suggests, maybe through π-stacking interactions with $TolQ_{F183}$ and $TolQ_{W32}$ (Fig. 6.1)[97]. Tyrosine residues are also able to partake in π-stacking, thus the $TolA_{H22Y}$ mutant remains functional, but this is disrupted by the TolQ mutants. In contrast, proline and arginine cannot partake in π-stacking, hence $TolA_{H22P}$ and $TolA_{H22R}$ confer the *Δtol* phenotype. The TolQ mutants rescue $TolA_{H22P}$ and $TolA_{H22R}$, likely by permitting alternative interactions orthogonal to π-stacking. In general, these data support the proposed TolQ-TolA interaction site, and explain that the essential role of $TolA_{H22}$ arises due to its aromaticity. While $TolQ_{W32}$ is conserved in $ExbB_{W38}$, $TolQ_{F183}$ is replaced with a valine residue ($ExbB_{V186}$), which is aliphatic but not aromatic, perhaps conferring the preference of ExbB for TonB, where $TonB_{H20}$ is also essential for interaction [256].

Through this modelling, I also discovered a conserved β-strand within TolA that likely interacts with TolR. This motif proved particularly robust, with deleterious OM-instability phenotypes observed only with EDTA supplementation. Over the course of these experiments, multiple studies also identified this motif within TonB, though none demonstrated perturbations *in vivo* by mutagenesis [199, 216]. The conservation of this structural motif suggests that TolA and TonB likely transduce force by fundamentally similar mechanisms, further evidenced by the proline-rich TolAII sequences detected in chapter four and the extensive conservation of this motif in 95% of TolA and TonB sequences tested alike (SI. 1-2). NMR, X-ray crystallography and ITC data suggest that in TonB, this motif binds to ExbD to form the antiparallel ExbD conformation that abolishes PG-binding [199, 216]. It was thus suggested that

rotation of the ExbD-TonB-ExbD trimer within a stationary ExbB complex drives domain II contraction of TonB [199]. Given the many similarities outlined here, this is equally plausible for TolA and TolR. Regardless of whether torque is applied by TolQ or TolR, the fundamental principles of domain II contraction remain similar. Here I demonstrate the biological reduction of function in TolA motor-box mutants (Fig. 6.10), which has yet to be demonstrated for the Ton system. It is unclear how a structural motif can be so widely conserved, as predicted in 95% of TolA and TonB structural predictions (190 total sequences, SI. 1-2), and yet its function remains relatively robust in the face of mutation. It is likely that the motor-box is an induced structure, where proximity of the TolA IDR to the TolR-CTD induces the formation of the motor-box via backbone hydrogen-bonding interactions. β-sheets form hydrogen-bonds through backbone interactions, and are thus difficult to fully abolish. Here, this was attempted through the motor-box proline mutants but insufficient time did not allow for a more exhaustive proline scan of the region.

The paucity of structures for TolQR-TolA and ExbBD-TonB complexes necessitates additional structural work that will further resolve their mechanisms. Foremost is characterising the assembly of the motor-transducer complex, and the conformations of both motor and force transducer in their energised force-applying states. This may be achieved through cryo-ET of live bacterial cells, as explored in chapter 7.

7.0 Cryo-ET of the energised TolQRA assembly

7.1 Introduction

7.1.1 TolQRA energisation may be visualised *in situ* by cryo-ET

Until now we have generally considered the different components of the TolQRA assembly in isolation. Here I sought to integrate the molecular elements of the previous chapters, including the TolQR assembly (chapter five) and its interactions with TolA (chapter six), the PG layer, and the conformational rearrangements that occur during force transduction (chapter three). To do this, I used cryo-ET, which permits molecular resolution of protein complexes *in situ*.

The interaction of both TolA and TonB with their OM components, is one that is mimicked and parasitised by colicins to energise their import [15]. Moreover, this interaction can stall the force transducer in an energised state, through the use of bacteriocins with a modified C-terminal domain that prevents OM-translocation. This has been demonstrated previously through the use of disulphide locked ColE9 variants that fully immobilise TolA in *E. coli*, and pyocin-GFP fusions that are no longer translocation competent in *P. aeruginosa* [195, 257]. Here, we exploit this phenomenon to trap TolQRA complexes in a stalled, energised state in live *E. coli* cells, and image them by cryo-ET. Cryo-ET is an ideal approach for studying these systems, because the system is structurally visualised in its native context at molecular resolution, in the presence of an intact cell wall and PMF.

To achieve a model of the energised TolQRA assembly by cryo-ET, multiple validation steps are required. This includes demonstrating the PMF-dependency of colicin import, colicin–mediated TolA immobilisation, and optimisation of secondary labels to identify those complexes *in situ*. A protocol was designed for labelling minicells, nonviable cells derived from cell division mutants that are significantly smaller than WT cells, because these cells are more appropriately sized for cryo-ET, and feature an intact cell wall and PMF [138].

7.2 Results

7.2.1 *E. coli* strain PB114 is sensitive to Colicins N & E9

To first demonstrate that the minicell strain selected (PB114) is sensitive to colicins and therefore contains functional TolQRA complexes, a killing assay was performed with ColE9 and ColN. These are TolA-dependent bacteriocins that utilise different OMPs for their import and interact differently with TolA (Fig. 7.1A). ColE9 exploits BtuB as its OM receptor and OmpF as its translocon, whereas OmpF is both receptor and translocon for ColN. Further, ColN directly binds TolA, whereas ColE9 binds to TolB, which then forms a ternary complex with TolA. These data demonstrate that the PB114 minicell strain is sensitive to ColE9 and ColN to a similar degree as the parental strain, MG6155 (Fig. 7.1B). Labelling experiments were continued with ColN on the basis that OmpF is at a significantly higher copy number than BtuB in the OM, increasing the number of potential TolQRA assemblies that might be trapped.

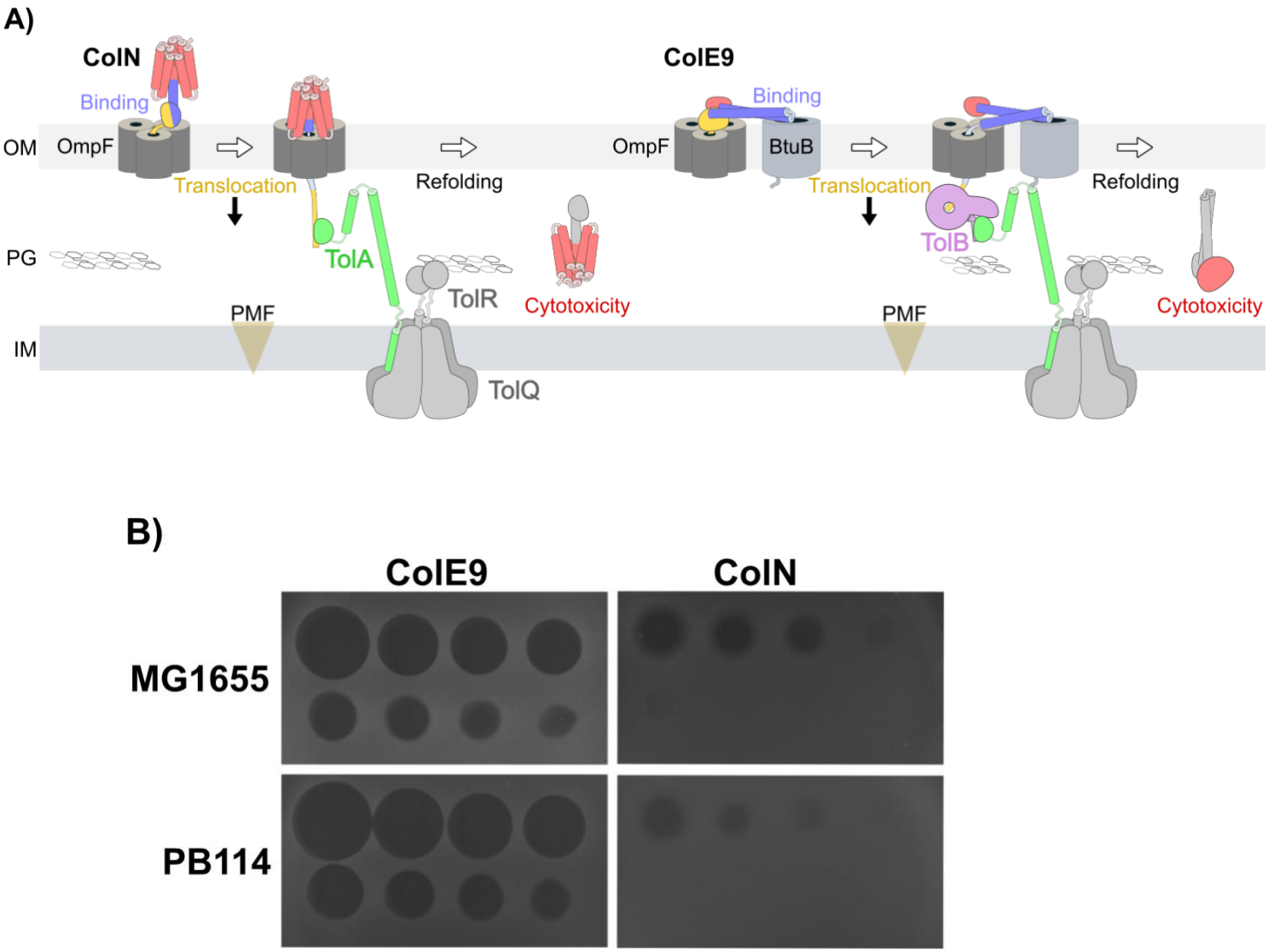


Figure 7.1 | ColN uses OmpF as both receptor and translocon, whereas ColE9 uses BtuB as a receptor and OmpF as a translocon. Both ColN (*left*) and ColE9 (*right*) have three domains involved in their mechanisms. First, is receptor binding (*blue*), where ColN binds OmpF, and ColE9 binds BtuB [258, 259]. Then the N-terminal translocation domain (*yellow*) inserts a disordered N-terminus into the lumen of a neighbouring OmpF lumen, to interact with TolA (ColN) or TolB (ColE9) and energise their import by parasitising TolQRA force transduction [91, 260]. Once fully translocated, the colicin refolds and kills the cell, either by IM pore formation (ColN) or further import to the cytoplasm via FtsH (not shown) and nuclease activity (ColE9) [84, 85]. **B)** Cytotoxicity assays show that *E. coli* PB114 strain is susceptible to colicin toxicity. Mid-log phase lawns of the *minB* minicell strain PB114 and MG1655 were spotted with a fourfold dilution series of purified colicins, beginning with 8 µM. In both cases, the colicins exhibited toxicity at concentrations typically used for fluorescence labelling (~1 µM). These plate-based assays suggest that the PB114 minicell strain is a functional platform for trapping ColN-FPs mid-translocation, since it is capable of translocating colicins.

7.2.2 ColN-GFP labels minicells

Previous work demonstrated that disulphide locked ColE9 stalls TolA diffusion [148, 261]. Here we sought to use ColN fused to fluorophores that render it translocation-incompetent but retain trans-envelope interactions with TolA. Similar work demonstrated that pyocin S5 (a bacteriocin from *Pseudomonas*) cannot fully translocate when fused to GFP [257]. Crucially, TolQRA can only be stalled using a bacteriocin that is translocation competent at the N-terminus, but has a C-terminal plug that remains blocked at the OM. Following the confirmation of colicin sensitivity, PB114 minicells were isolated by differential centrifugation (Fig. 7.2A). These were labelled with ColN-GFP, washed and imaged by fluorescence microscopy (Fig. 7.2B). Having demonstrated that the minicell strain is both sensitive to ColN and can be labelled with a ColN-GFP fusion protein, a translocation-competent fluorescently-labelled ColN variant was also required for validation of import.

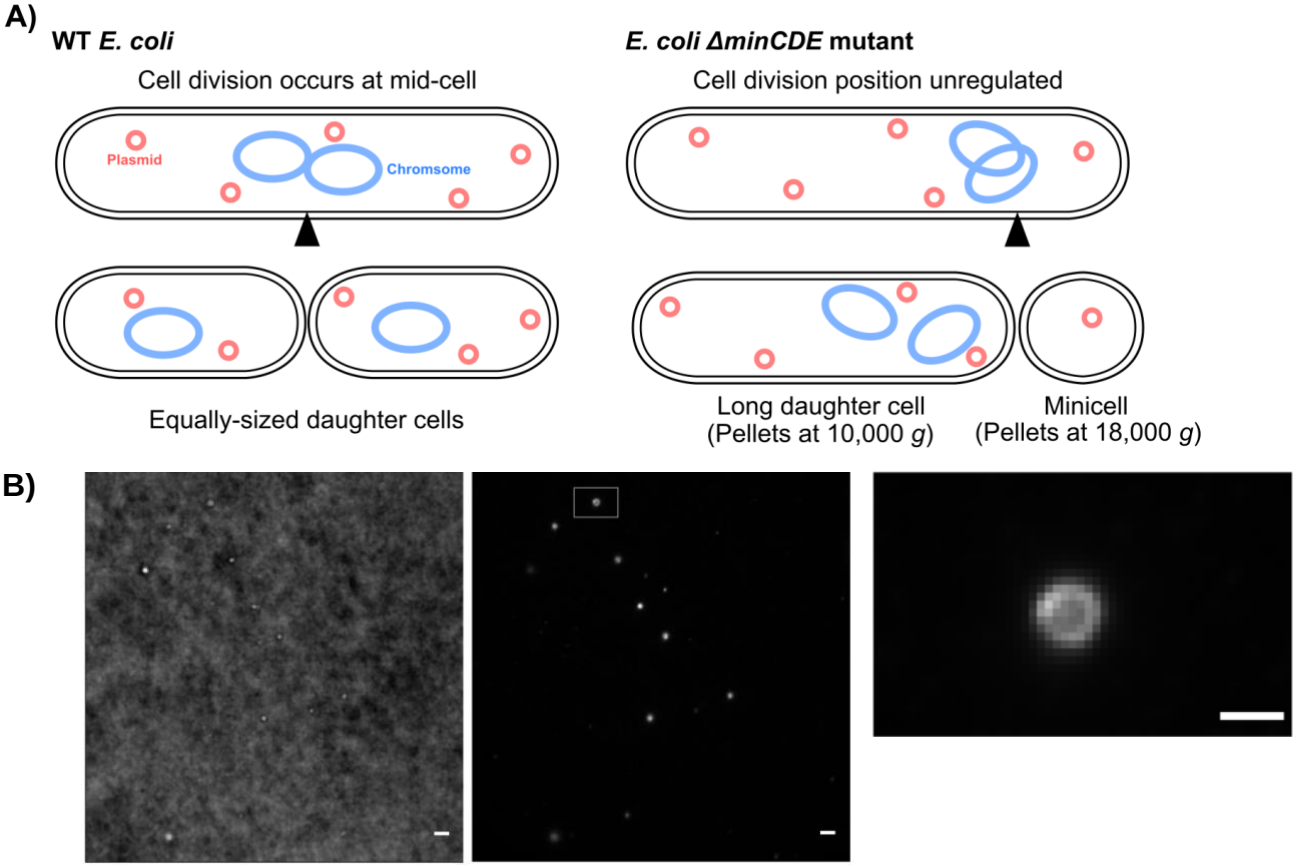


Figure 7.2 | Minicells can be labelled with ColN-GFP. A) Cell division in *ΔminCDE* mutants occurs at a random position along the cell barrel, occasionally leading to polar septation and minicell formation, which can be isolated by differential centrifugation. Minicells are nonviable because they lack chromosomes, but still possess an intact cell wall, PMF, and plasmids. **B)** Transillumination (*left*) and fluorescence (*middle*) show that PB114 minicells can be labelled with ColN-GFP. Zoomed image of a single minicell (*right*) indicates the expected size and morphology. Scale bar = 1 µm.

7.2.3 ColN-mCherry traps TolA-GFP *in situ*

I also sought to orthogonally demonstrate that a ColN-mCherry fusion protein prevents the diffusion of TolA *in situ*, confirming OM-tethering as was previously observed with disulphide-locked ColE9 [195]. ColN-mCherry was selected over ColN-GFP because the TolA reporter protein features an N-terminal GFP fusion. To achieve this, *ΔtolA* cells expressing GFP-TolA were incubated with ColN-mCherry, and the GFP-TolA proteins were excited at 473 nm, after preliminarily bleaching mCherry at 560 nm. The fluorescence micrographs were analysed by signal molecule tracking, demonstrating that ColN-mCherry immobilises GFP-TolA to a similar extent as a formaldehyde-fixed control (Fig. 7.3).

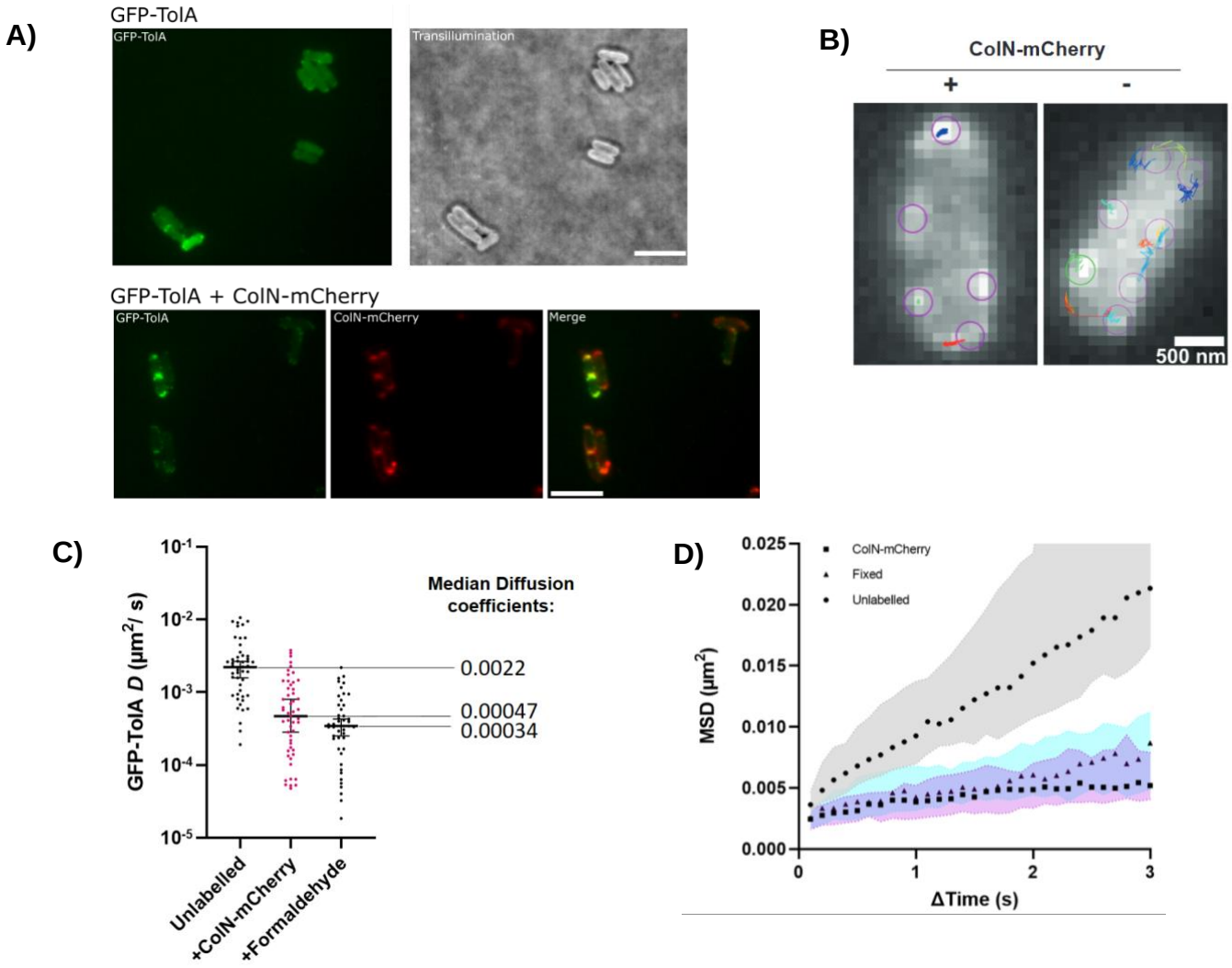


Figure 7.3 | GFP-TolA is immobilised by ColN-mCherry. A) Cells expressing GFP-TolA were labelled with ColN-mCherry and imaged at 473nm. Scale bar, 5 μm. **B)** Representative fluorophore tracks (*coloured by track*) of GFP-TolA ± ColN-mCherry labelling, recorded with Trackmate. **C)** Initial diffusion coefficients from 50 tracks suggest a 7-fold decrease in initial TolA diffusion rate upon incubation with ColN-mCherry. **D)** MSD curves suggest ColN-mCherry restricts TolA diffusion radius.

7.2.4 Fluorescently labelled ColN-AF$_{647}$ is translocation competent

To validate that the ColN N-terminus (present in ColN-GFP) is translocation competent, I designed a ColN$_{1\text{-}186}$-cys mutant, replacing the GFP with a cystiene residue. I then purified and conjugated it to AlexaFluor 647 (AF$_{647}$) via its C-terminal cysteine residue (Fig. 7.4). Unlike ColN-GFP, this protein was expected to be translocation competent because the AF$_{647}$ fluorophore is sufficiently small to pass through the OmpF lumen [183]. This was then used to validate that both the PMF and TolQRA are functional within minicells, by measuring import of ColN-AF.

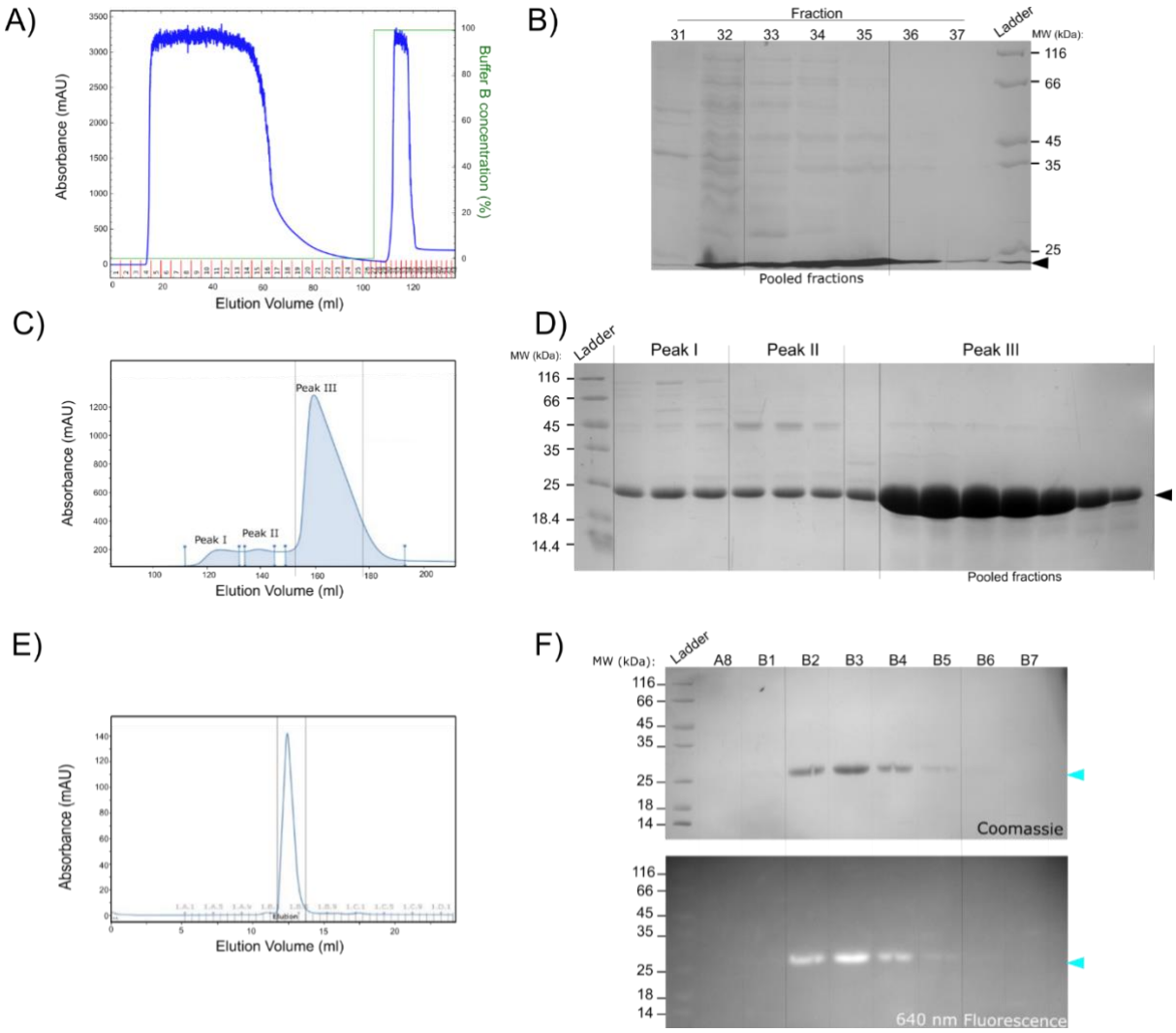


Figure 7.4 | ColN$_{1\text{-}186}$-cys purification and AF$_{647}$ conjugation. A) Nickel affinity purification chromatogram of ColN$_{1\text{-}162}$, resolved by SDS-PAGE (*black triangle*) (**B**). **C)** Gel filtration chromatogram of concentrated ColN$_{1\text{-}162}$, resolved by SDS-PAGE (**D**). **E)** Gel filtration chromatogram of ColN-AF after conjugation. **F)** Fractions (indicated) were diluted into SDS buffer and electrophoresed without boiling, which yielded fluorescent bands at 640 nm (indicated in cyan). Vertical bars indicate fractions which were pooled.

I then isolated PB114 minicells, and labelled them with the purified ColN-AF, before washing with PBS and fixation with formaldehyde. These minicells demonstrated bright peripheral fluorescence at 640 nm excitation (Fig. 7.5). To demonstrate that ColN-AF is imported while ColN-GFP is not, I then performed a trypsin assay, where minicells were labelled, then treated with trypsin to digest and quench surface-bound fluorophores. ColN-GFP was well-digested, suggesting it is present at the surface, whereas ColN-AF fluorescence persisted despite trypsin treatment, indicating import into the minicells and protection from trypsin (Fig. 7.6).

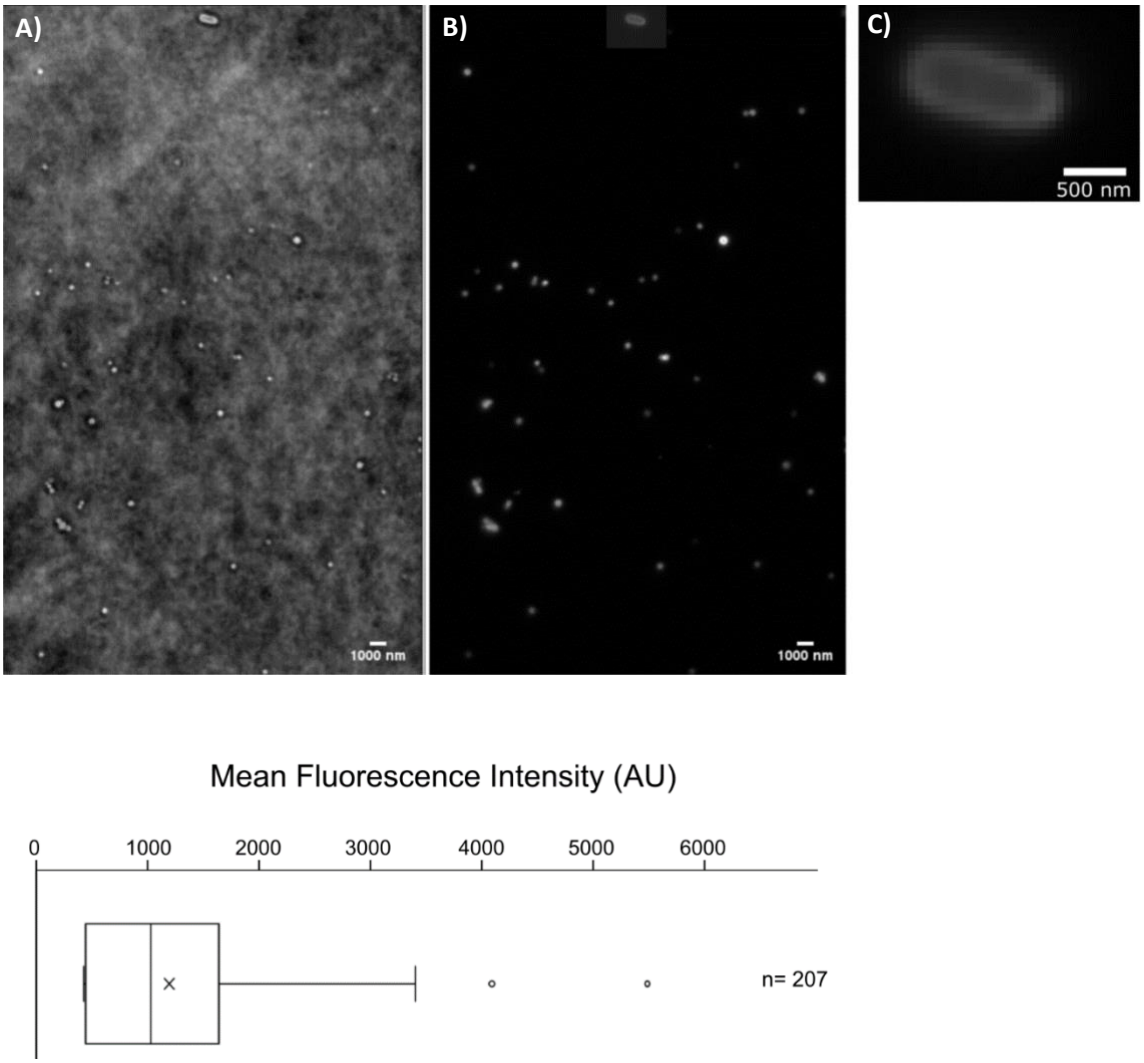


Figure 7.5 | Purified PB114 minicells labelled with ColN-AF$_{647}$. Minicells were imaged under 620 nm transillumination (**A**), and 640 nm fluorescence (**B**). Box in B corresponds with frame of **C**), which depicts a well labelled whole-cell. **D**) The mean fluorescence distribution of labelled minicells.

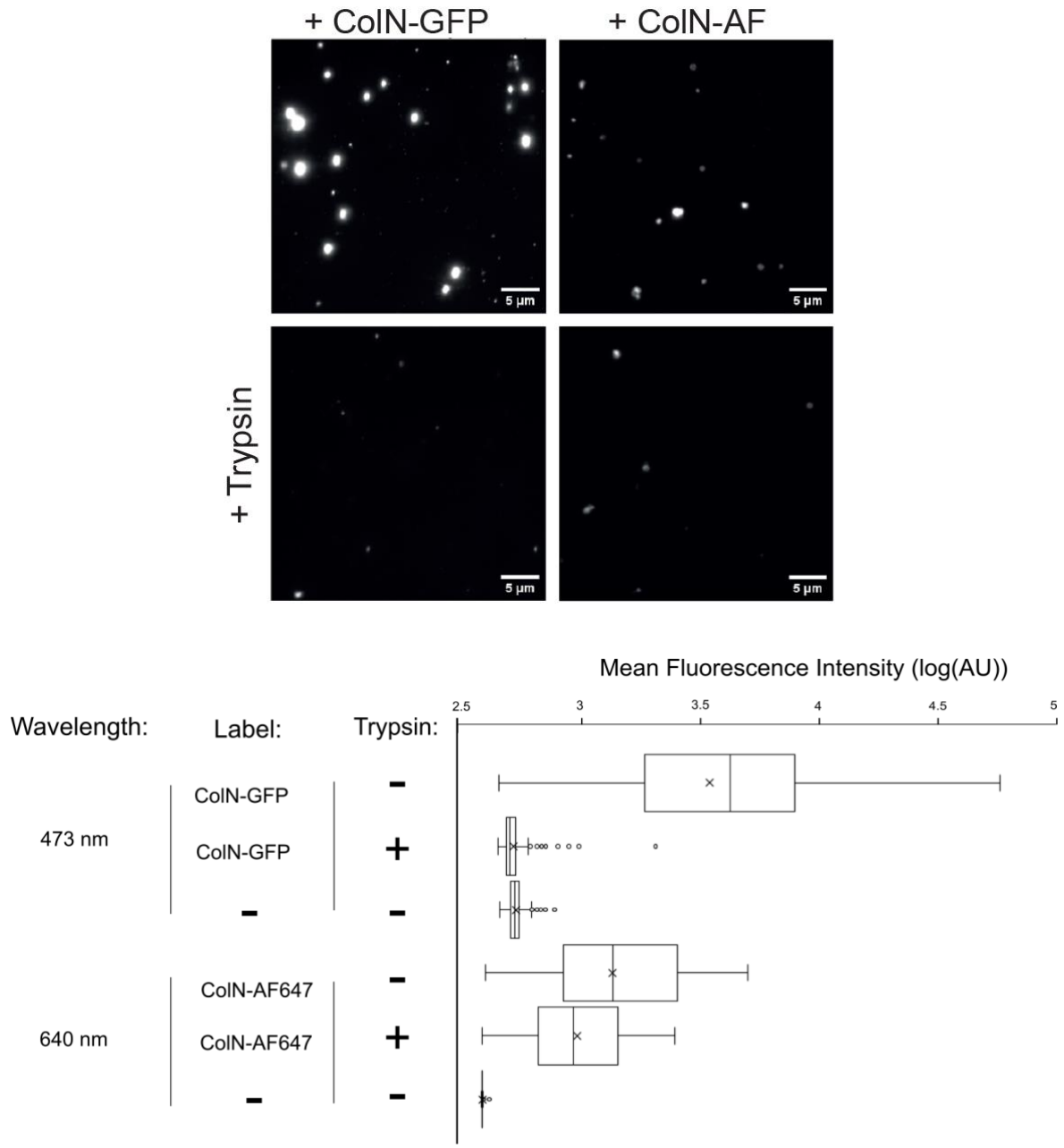


Figure 7.6 | Trypsin assays suggest that ColN-AF$_{647}$ is internalised into minicells while ColN-GFP is not. Minicells were labelled overnight with either ColN-GFP or ColN-AF$_{647}$, then split into two treatment groups, and one half digested with trypsin for 30 minutes at 37° C. Minicells were then fixed and imaged, both under transillumination and with the fluorescence excitation wavelengths indicated above. Fluorescence distributions were quantified appropriately, and compared with nonlabelled control minicells to account for autofluorescence. Fluorescence distributions are displayed on a log scale because GFP is significantly brighter than AF$_{647}$. Overall, these data suggest that the ColN N-terminal sequence is translocation competent in ColN-AF$_{647}$, but the trypsin-accessible GFP-fusion prevents ColN uptake, indicating that a trapped complex is formed.

7.2.5 ColN$_{1-162}$ uptake is Δψ dependent

In order to further demonstrate that colicin import is dependent on an intact PMF, an ionophore assay was performed using nigericin and valinomycin, as described by Francis *et al.,* [183]. These ionophores act as selective carriers of K^+ and H^+ respectively, thus valinomycin collapses membrane potential (Δψ), while nigericin collapses the proton gradient (ΔpH). Cells were first incubated with 1 mM EDTA and the ionophore, in 150 mM KCl in M9 media. Cells were then labelled with ColN-AF$_{647}$, treated with trypsin and imaged by fluorescence microscopy. These data indicate that valinomycin significantly impairs ColN-AF$_{647}$ internalisation, while nigericin does not (Fig. 7.7). This suggests that ColN import requires energisation, predominantly through Δψ rather than ΔpH, in agreement with the work of Francis *et al*. [183].

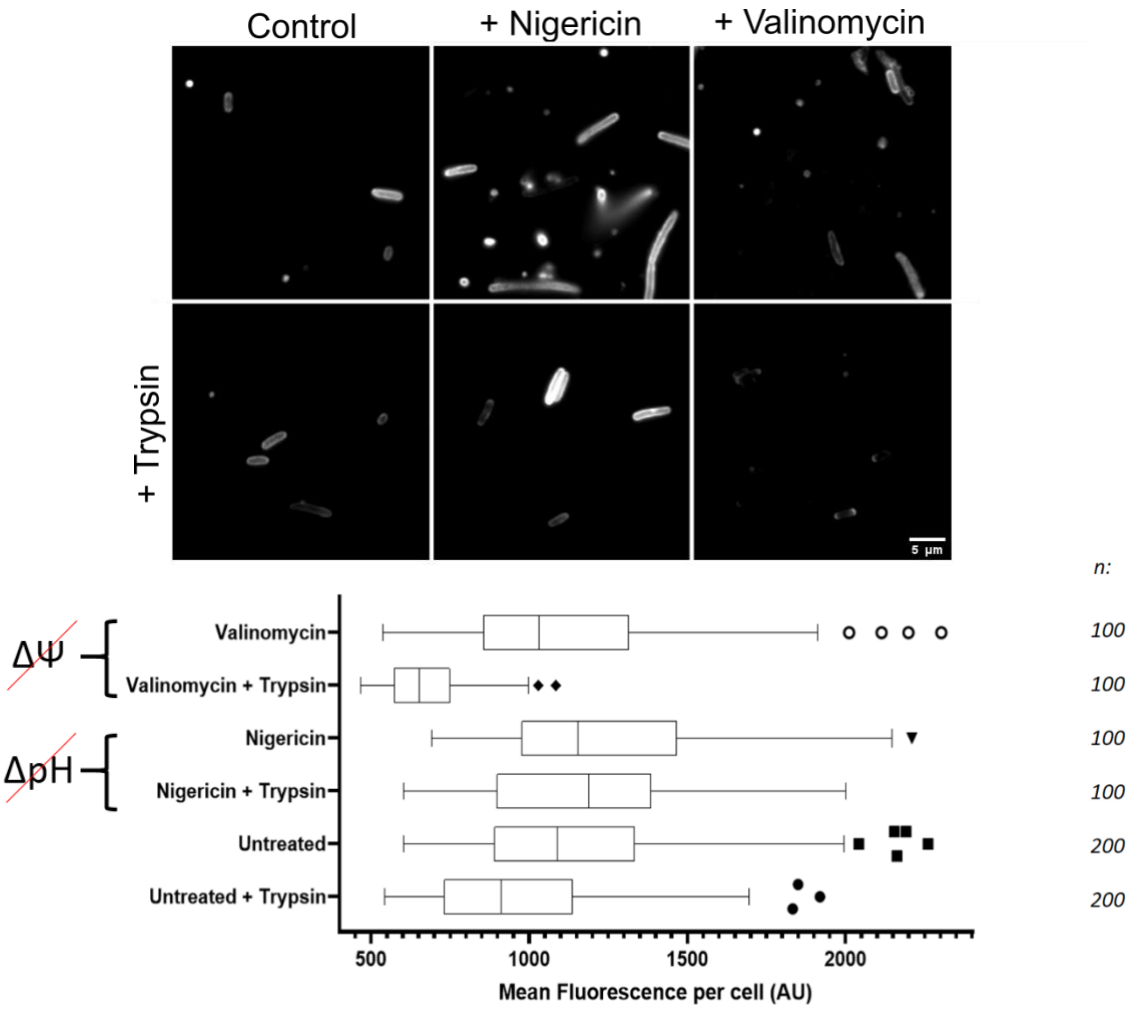


Figure 7.7 | Valinomycin reduces ColN-AF$_{647}$ import into PB114 cells. **A)** Representative images, with equalised background and contrast settings. **B)** Quantification of mean fluorescence per pixel per cell.

7.2.6 Large scale minicell ColN-GFP labelling

Having shown that ColN-GFP labelling captures stalled OMP-TolQRA complexes in minicells, I scaled up the protocol to generate sufficient sample density for electron microscopy. In the example below, 200 ml of cells at OD_{600}: 1 were harvested, and minicells purified. With the increased quantity of minicells, I repeated the ColN-GFP trypsin treatment (Fig. 7.8).

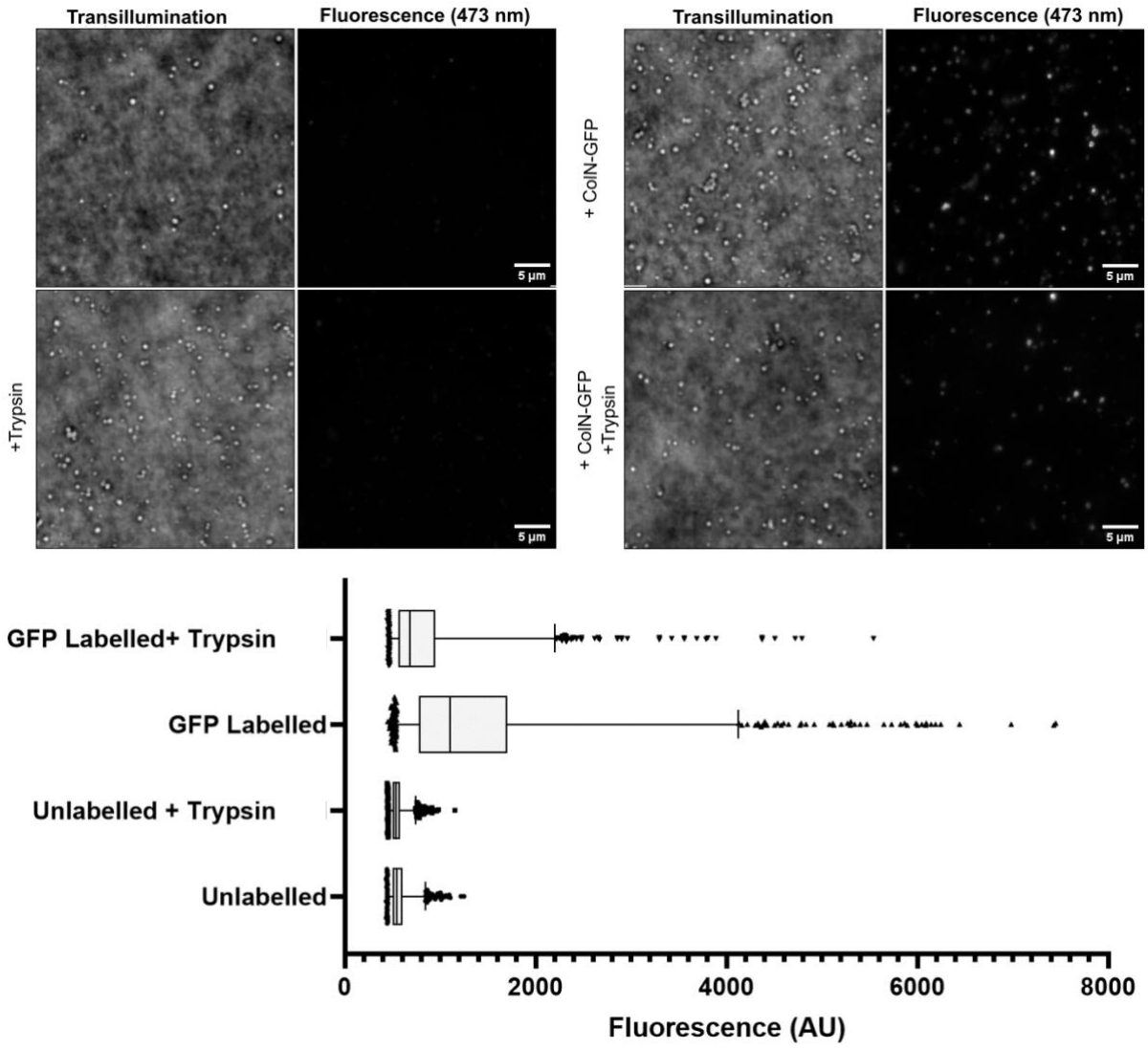


Figure 7.8 | ColN-GFP labelling minicells after upscaling growth in LB. Fluorescence images (*upper panel*), with equal background and contrast settings, show ColN-GFP labelling. Quantification of fluorescence values per minicell (*lower panel*) suggests excess surface labelling with ColN-GFP is quenched by trypsin treatment. N= 2000 minicells per condition.

7.2.7 Gold nanoparticles are used as a secondary label for cryo-ET

The TolQRA assembly is around 200 kDa, which approaches the lower limit of resolution by cryo-ET. As such, a sophisticated labelling system is required to precisely, accurately and specifically detect these complexes in tomograms and to add additional electron density to aid 3D reconstruction. To this end I attempted to secondary label the cell-surface exposed GFP using gold nanoparticles (AuNPs) [262, 263]. These contain functionalised nitrilo-triacetate (NTA) moieties complimentary to the ColN-GFP His-tag. Microscopy confirmed that the minicells remained fluorescently labelled after AuNP secondary labelling (Fig. 7.9). These samples were then imaged by NS-TEM. The presence of AuNPs on the cell surface in the AuNP-only (no bacteriocin) control suggests non-specific labelling. Therefore, a different label was required to locate complexes on cells. These samples were then imaged by NS-TEM, and an AuNP-only labelled control which suggested heterogeneous surface labelling with AuNPs. Presence of AuNPs on the cell surface in the AuNP-only control also suggests non-specific labelling (Fig. 7.10).

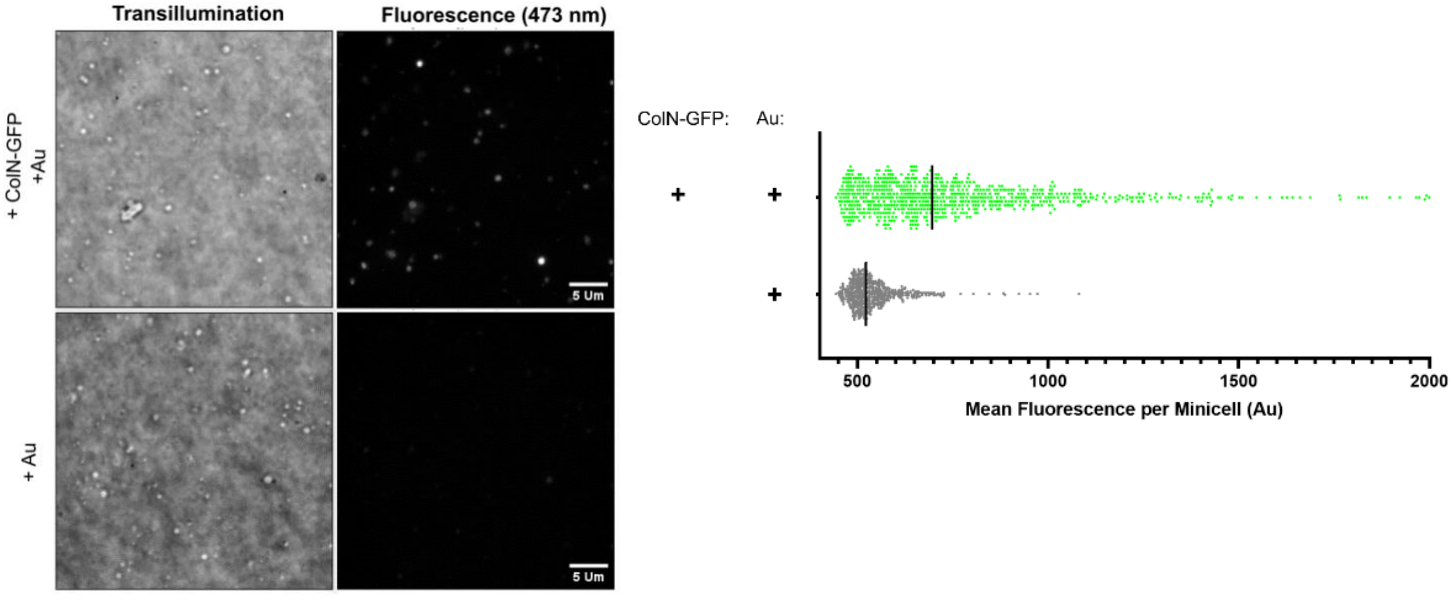


Figure 7.9 | AuNP labelling does not remove ColN-GFP from minicell surface. A) Fluorescence microscopy shows double-labelled and control minicells, with equal background and contrast settings. Some gold aggregates are visible. Cells are fluorescently labelled, compared to the Au-only labelled control. **B)** Quantification of fluorescence distribution suggests double-labelled minicells remain labelled with ColN-GFP.

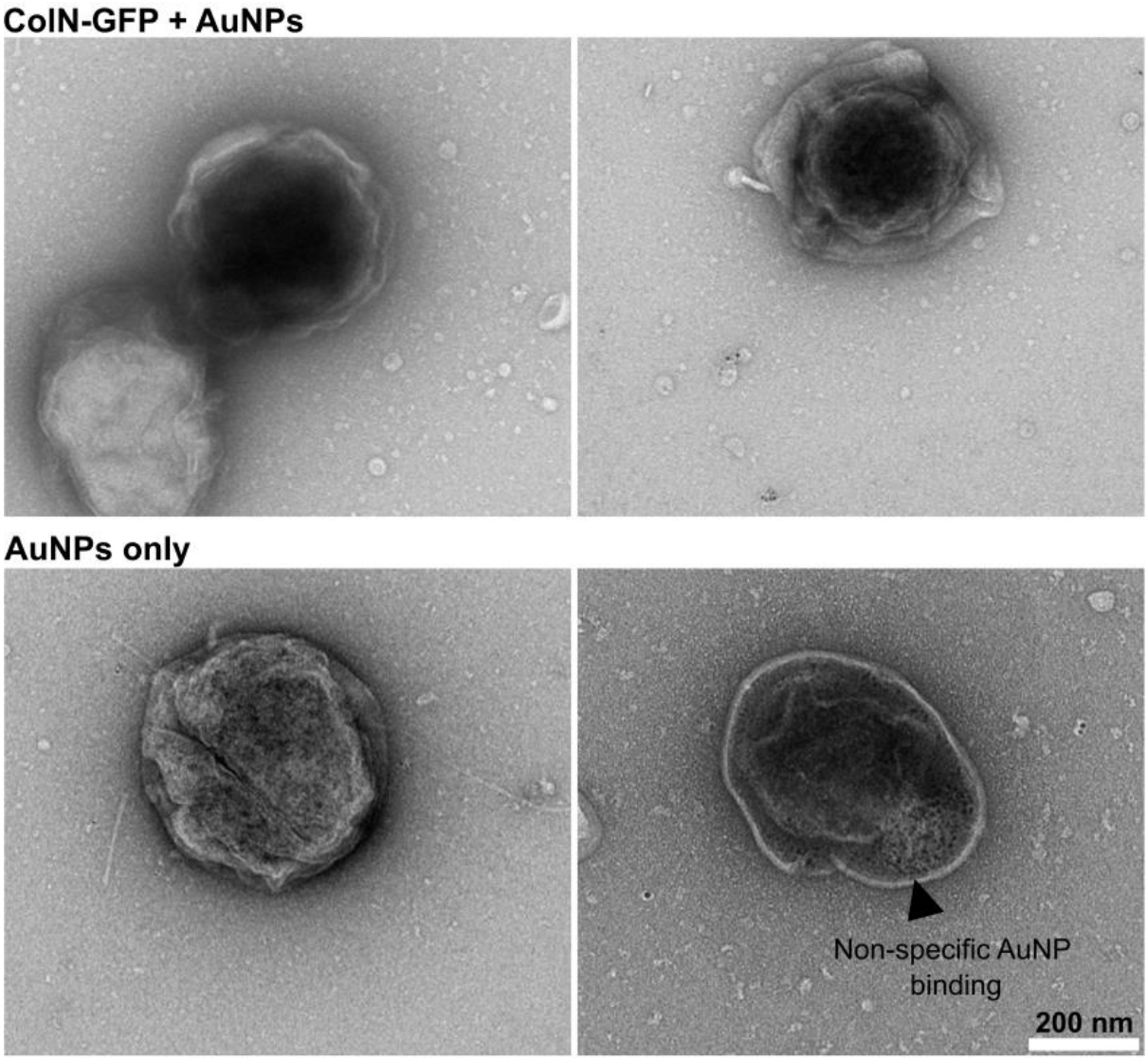


Figure 7.10 | AuNP labels on minicells exhibit heterogeneity. A mixture of heavily labelled and poorly labelled cells are visible in the double-labelled sample (*upper panel*) and the AuNP-only control (*lower panel*). Presence of bound AuNPs in the ColN-GFP-free sample suggests AuNPs bind non-specifically to the cell membrane.

7.2.8 AuNPs label cells heterogeneously

Further efforts were driven toward reducing non-specific binding with AuNPs, by increasing the concentration of bovine serum albumin (BSA) used during secondary labelling. Double-labelled minicells were subsequently imaged by transmission cryo-EM, which generally improved specificity but heterogeneity persisted. Some minicells displayed an appropriate AuNP density, but a low-dosage beam was required due to contrast artefacts arising in regions with high AuNP density (Fig. 7.11). This limits the maximum achievable resolution by cryo-ET.

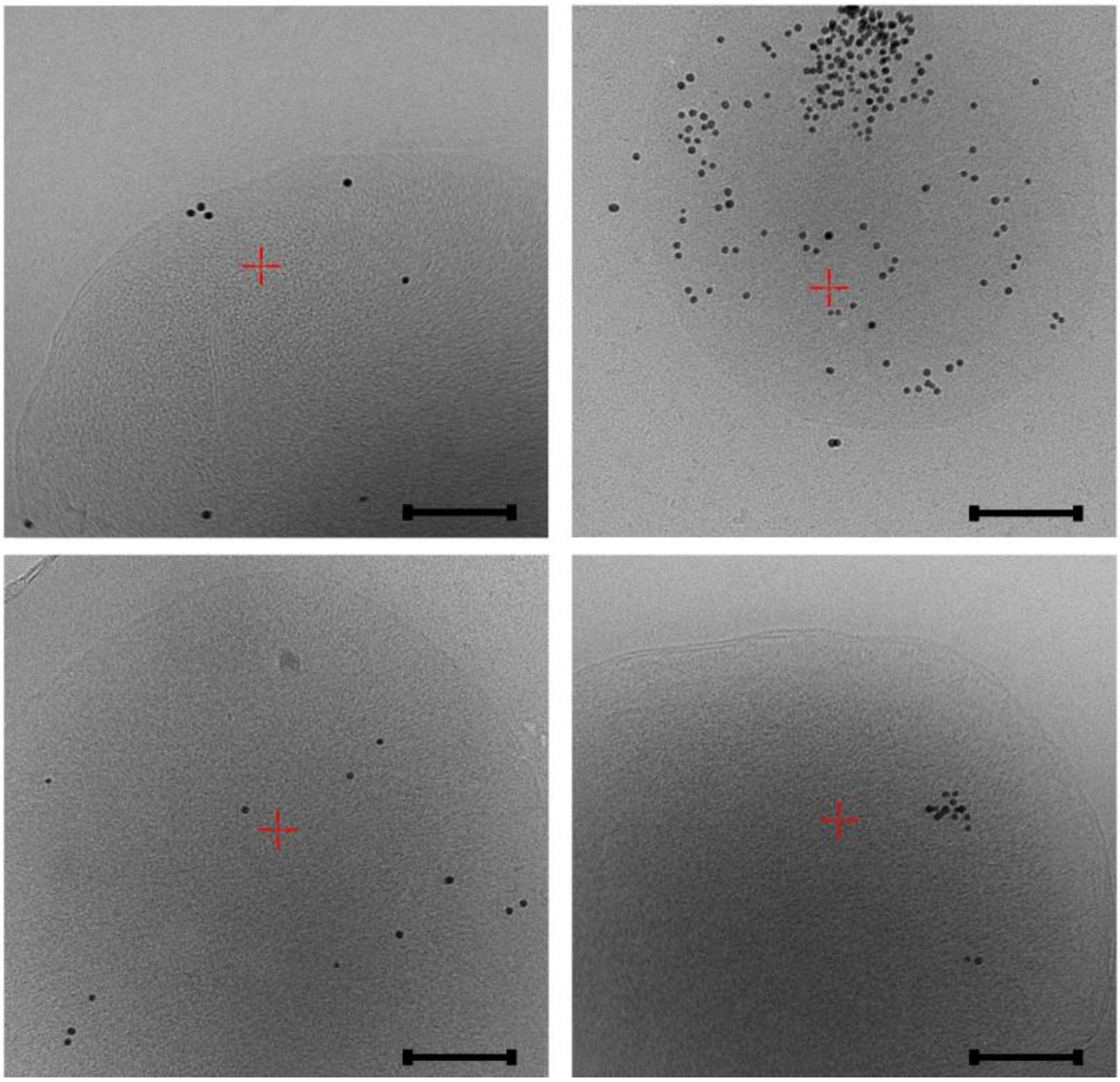


Figure 7.11 | Cryo-TEM shows excess BSA slightly improves AuNP specificity, but contrast artefacts remain. Features such as the IM, periplasm and PG layer are poorly resolved, perhaps due to ice thickness and/or the low dosage beam used to avoid contrast artefacts. Scale bars = 100 nm.

7.2.9 SPOTs are a superior secondary label to AuNPs

Despite adequate sample density it appears that the electron density of AuNPs hinders the maximum achievable resolution due to increased contrast. This may prove problematic during cryo-ET by obscuring the trapped complexes, which are relatively small. A secondary label which circumvents this issue is a DNA signpost origami tag (SPOT) [262, 263]. This is a synthetic DNA nanostructure, with a functionalised RNA aptamer at the base. This allows the SPOT to be targeted to a specific protein, in this case complementary to GFP, thus binding ColN-GFP at the cell surface (Fig. 7.12). To illustrate these contrast artefacts, far greater resolution was seen in samples treated with SPOTs rather than AuNPs (Fig. 7.13).

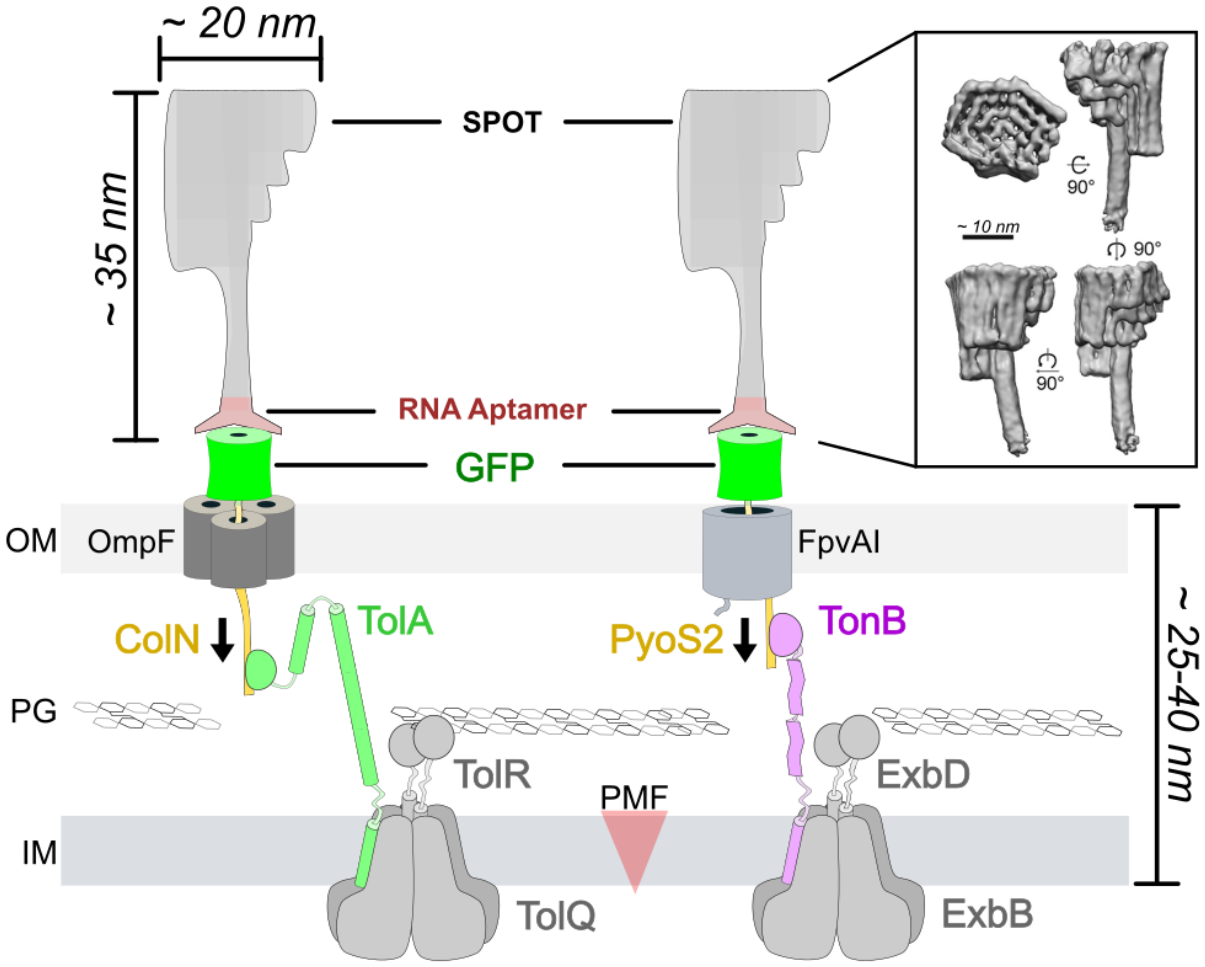


Figure 7.12 | Expected compliment of proteins labelled by SPOTs. This protocol can be applied to both TolA (*left*) and TonB (*right*), depending on the bacteriocin used. The conformation of the stalled complex is currently unknown. SPOTs are asymmetric and directional, aiding detection of labelled targets. Cryo-EM structure of SPOTS shown (*top right*), from Silvester *et al.* [263].

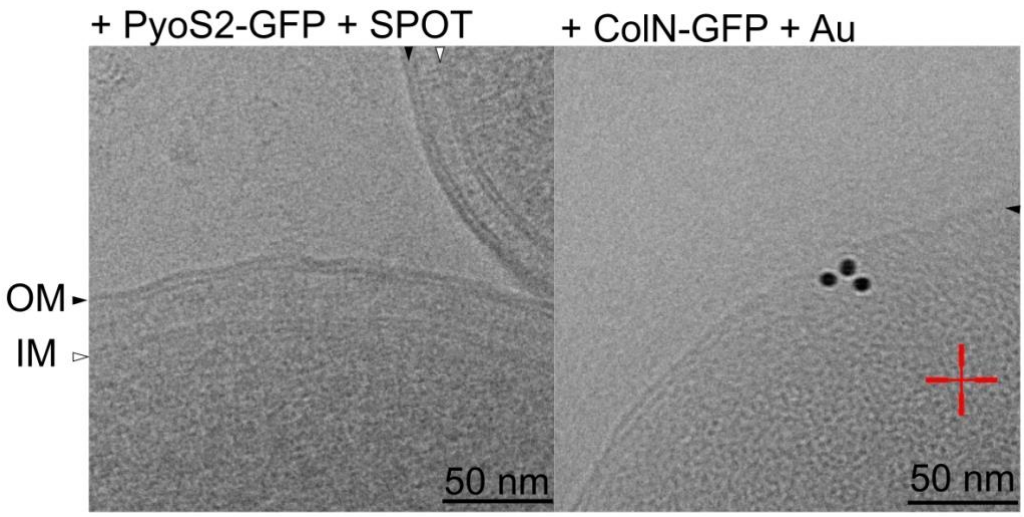


Figure 7.13 | AuNPs impart resolution-limiting contrast artefacts in cryo-TEM. In cells labelled with SPOTs (*top left, Dr. Melissa Webby*), the OM (*black triangle*), IM (*white triangle*) and periplasm are clearly resolved. In AuNP labelled cells (*right*), only the OM can be resolved. Reduced resolution could be due to contrast artefacts arising from AuNP aggregates which were readily observed in most samples.

7.2.10 SPOTs improve cryo-EM resolution over AuNPs

Given the superior EM characteristics of SPOTs over AuNPs, I prepared a PB114 minicell/cell mixture and labelled with a ColN-GFP:SPOT mixture. Here, ColN-GFP was first incubated with SPOTs to label minicells in a single step, reducing the number of washes required. As SPOTs comprise DNA and RNA, the labelling was performed on ice in TAE buffer. Imaging the sample by cryo-EM showed SPOTs located near the surface of cells. However, the density of SPOTs was lower than expected, suggesting poor binding of SPOTs to the surface of cells (Fig. 7.14). Non-functionalised AuNPs were used as fiducials for cryo-ET, but at a much lower concentration than the NTA-functionalised AuNPs used to label ColN-GFP.

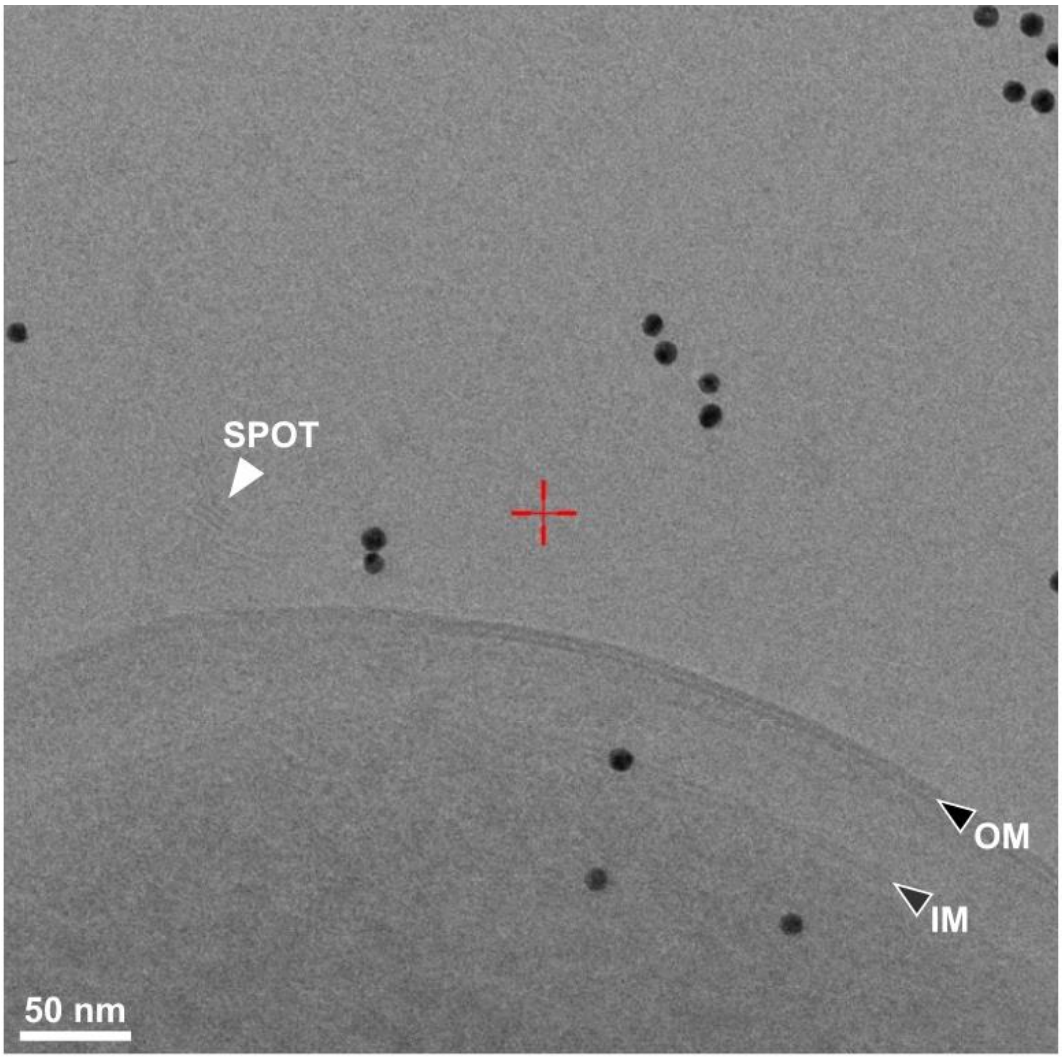


Figure 7.14 | Validation of SPOT-labelling in minicells by cryo-TEM. A SPOT (*white triangle*) was detected on the cell surface. Cell envelope features such as the IM, PG layer, and OM are clearly visible.

7.2.11 Labelling SPOTs in two steps improves efficiency

To improve the quantity of SPOTs per cell, I tested the fluorescence of cells under different labelling conditions, to determine whether the number of surface-bound ColN-GFP or the SPOT labelling conditions were responsible for the inefficient surface display of SPOTS. Labelling cells with ColN-GFP on ice yielded low levels of cellular fluorescence, indicating that low temperature reduces ColN-GFP labelling efficiency (Fig. 7.15). Since ice is required to stabilise the DNA origami structure of the SPOT and is thus essential, a two-step labelling approach was trialled, first labelling with ColN-GFP overnight at room temperature in TAE buffer and subsequently label with SPOTS on ice for 30 minutes. This experiment gave well-labelled cells, but they were plasmolysed due to the TAE buffer (Fig. 7.16).

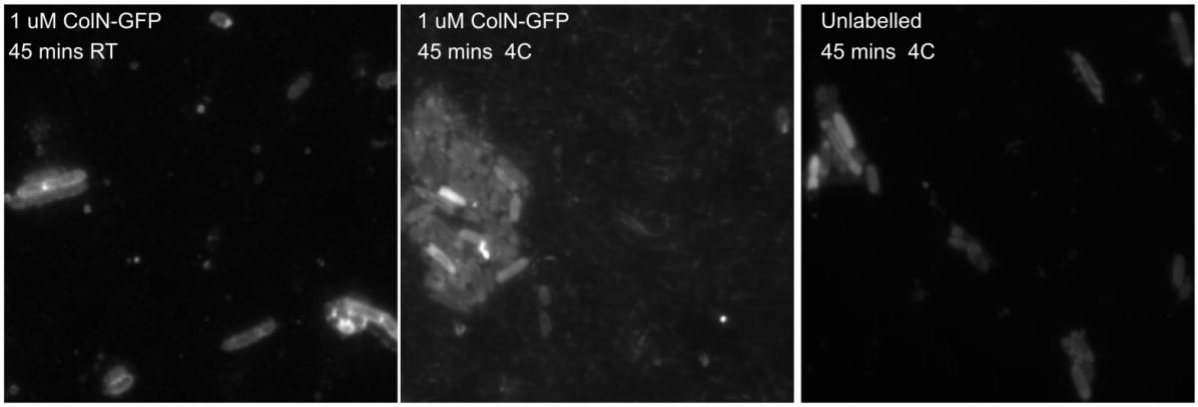


Figure 7.15 | ColN-GFP-labelling is inefficient at low temperatures. Cells were labelled for 45 minutes with 1 µM ColN-GFP at room temperature or on ice, in TAE buffer. Fluorescence micrographs indicate low ColN-GFP labelling efficiency when prepared on ice (*equivalent brightness and contrast settings*).

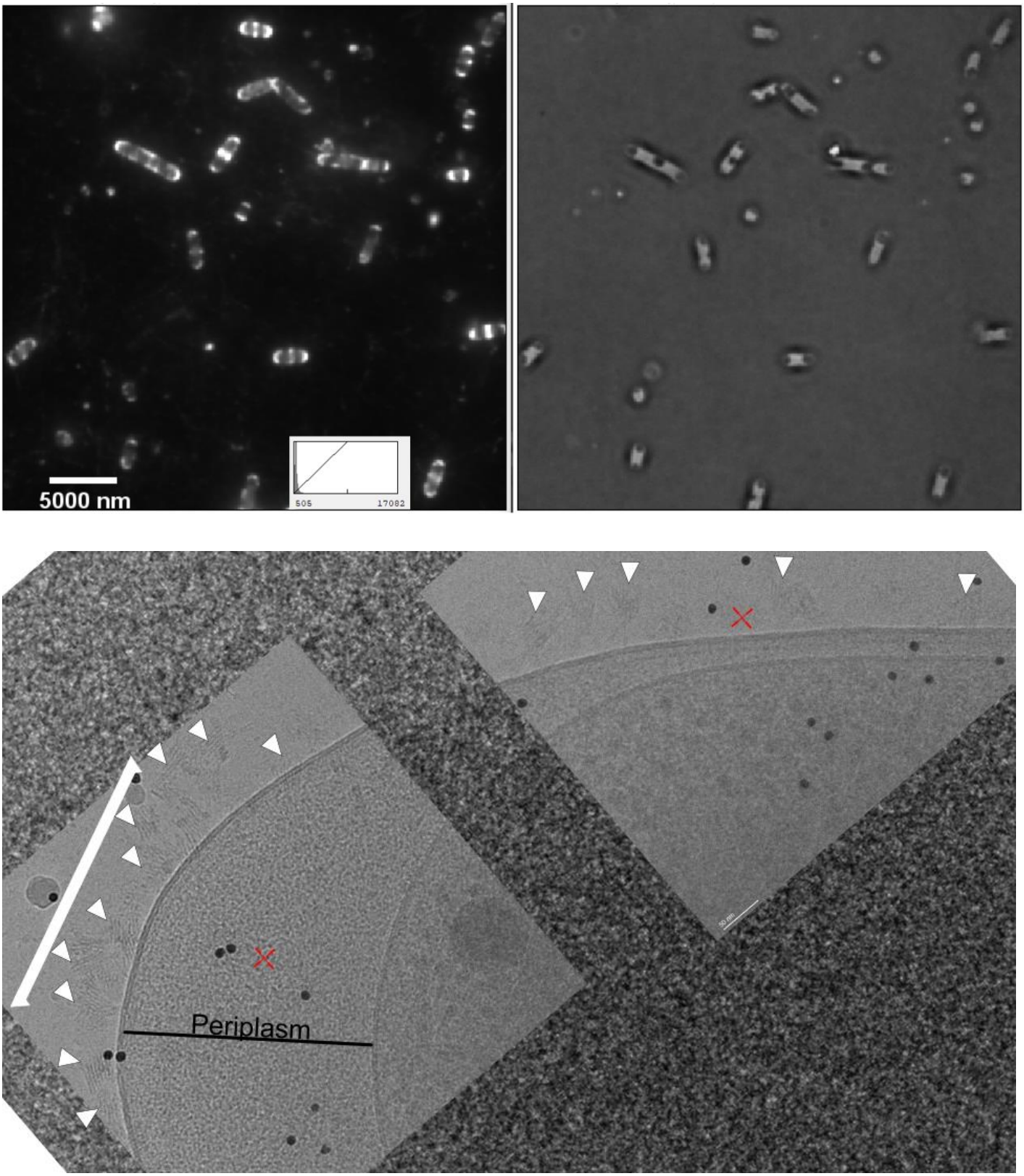


Figure 7.16 | Overnight labelling with ColN-GFP in TAE followed by SPOTs yields well-labelled but plasmolysed cells. Microscopy indicates clustered ColN-GFP distribution along the OM (*top left*). This is similar to the distribution of SPOTs (*white triangles*) as seen by cryo-EM (*lower panel, composite image*). The periplasm was found to be widened at the poles.

7.2.11 Optimised ColN-GFP-SPOT labelled samples

To avoid plasmolysis, minicells were labelled at room temperature with ColN-GFP overnight in LB, then washed and labelled with SPOTs in TAE as optimised. Cryo-TEM showed that these minicells were well-labelled and non-plasmolysed (Fig. 7.17).

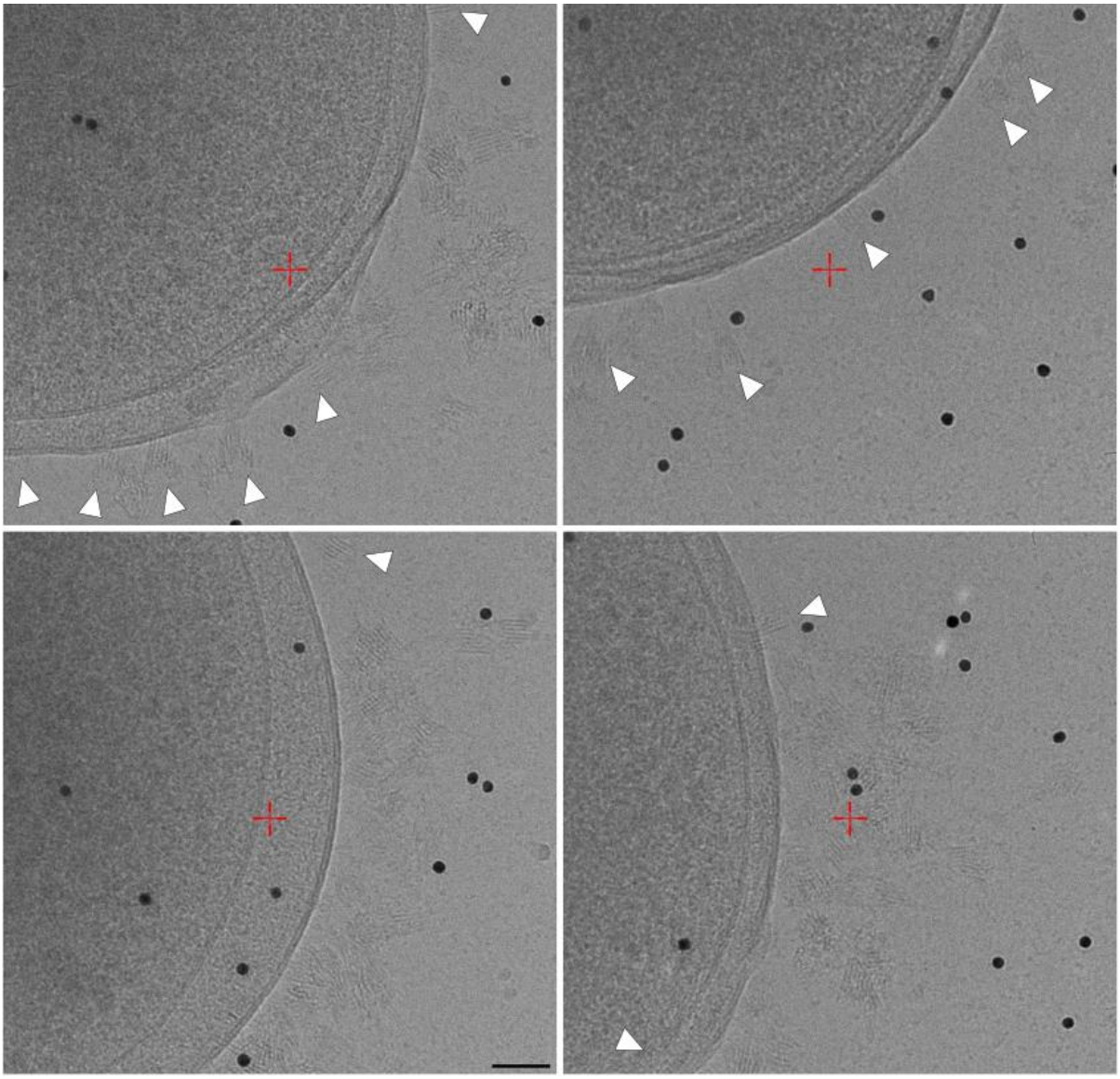


Figure 7.17 | Optimised two-step labelling yields a more ideal SPOT concentration. Single SPOTs appear oriented towards minicell body, though aggregates appear at regions presumably descended from mother cell poles. Scale bar = 50 nm.

7.2.12 SPOT-labelled minicell tilt series collection

We collected a cryo-ET tilt series and used IMOD to reconstruct tomograms from these data. The observed SPOTs were an appropriate distance from the cell surface, oriented towards it, suggesting they are bound to the surface (Fig. 7.18). The collection of further tilt series will be required to enable subtomagram averaging (STA), to structurally resolve the full complement of proteins in molecular detail.

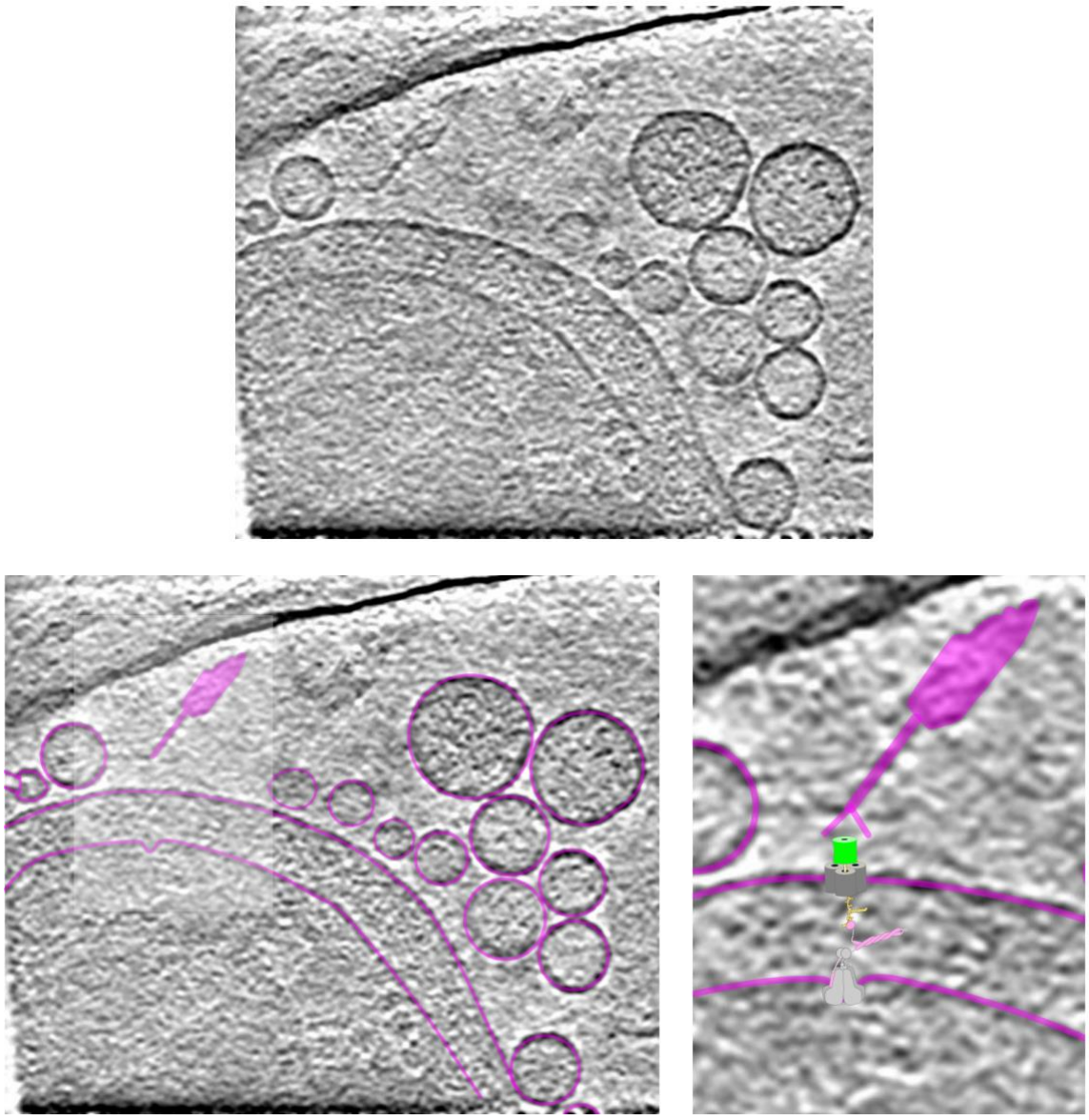


Figure 7.18 | Tomogram of *E. coli* minicell labelled with colicin-GFP and SPOTs. Minicells were labelled with PyoS2-GFP labelled minicells and SPOTs (*by Dr. Melissa Webby, upper panel*). The SPOT is visible at the cell surface, at an appropriate distance to account for the flexibility of the RNA aptamer of the SPOT and the surface GFP domain (*lower panels, manually annotated*).

7.3 Discussion

Here I demonstrate that colicin-GFP fusion proteins trap TolA *in vivo*, and that they can be labelled with SPOTs to detect *trans*-envelope complexes within tomograms. Further work will involve collection of additional tomograms, for subsequent STA. It may then be feasible to detect these complexes in their contracted force-applying state, to identify the conformational rearrangements that occur during force transduction, or more general features such as local OM/PG deformations. This approach bridges the gap between cellular localisation observed by fluorescence microscopy, and molecular resolution. This work is the first use of SPOTs to label individual proteins at the bacterial cell surface, and can be applied to any protein that can be labelled extracellularly. Similarly, this work demonstrates the versatility of minicells, where colicin import demonstrated their intact PMF.

Colicins require force for their import, as demonstrated here and in previous work (chapter three) [183], suggesting that TolQRA/ExbBD-TonB conformations elucidated by this approach are likely similar to those sampled in their native functions of TolB-Pal dissociation and TBDT-plug movements respectively. It is unclear why valinomycin abolished ColN-AF import but nigericin did not, as this suggests that ColN import via TolQRA is primarily driven by charge gradients rather than proton gradients, as shown previously for ColE9 [183]. This is particularly unexpected for a motor proposed to be proton driven; it was previously shown that only a mixture of valinomycin and nigericin together inhibits flagellar rotation via MotAB by dissipating the proton gradient [264]. In *M. xanthus*, nigericin inhibits motility whereas valinomycin does not, suggesting the AglRQS motor is proton-driven [265]. Experimental conditions such as media formulation, particularly KCl concentration, are extremely important for these experiments. For example, at high extracellular K^+ concentrations (such as 150 mM as used here) cells are depolarised, but with low K^+ concentrations, cells become hyperpolarised. Typically, *E. coli* have an internal concentration of 100-200 mM K^+ [266]. In the AglQRS experiments, 100 mM KH_2PO_4 buffer was used, which may be equal or lower than the cellular concentration, leading to hyperpolarisation of *M. xanthus* cells, hence motility after valinomycin treatment. Given their shared essential residues within the motor complex, it was previously assumed that TolQR and ExbBD are also proton driven. To date, no work has explored which aspect of PMF actually drives ExbBD rotation.

8.0 Concluding remarks

Prior to this work, many details of the Tol-Pal system were not known. The general role of the system in stabilising the OM was only recently discovered to operate by specifically recruiting Pal to the mid-cell [42, 93, 140]. Here I demonstrate that septal recruitment of Pal is not only correlated with the continuum of OM stability, but also depends upon the efficiency of TolA in transducing force, as demonstrated by TolA chimeric mutants (chapter three). This finding contextualises previous observations of *tolQR* mutants that confer OM instability, but remain colicin-tolerant, suggesting that these mutants reduce TolQR rotation speed. Thus, these slower mutants deliver less torque to TolA, reducing the rate of Pal recruitment to the septum, similar to observations with ABA and AIA. Moreover, mutants of TolA and TonB with disordered central domains (AIA/BIB) demonstrate colicin import. The disordered nature of their periplasm-spanning domains suggests that torque is delivered at the same time as their OM targets are bound. The force transduction conferred by ordered and disordered group A and B mutants alike also suggests that TolA and TonB likely function with fundamentally similar mechanisms [131].

We also generated the first full-scale model of the TolQRA assembly, revealing that the TolA TMH likely slots into a groove at the exterior face of TolQ, which in combination with mutation data suggest that the essential residue His_{22} likely forms π-stacking interactions (chapter six). I then demonstrate through GFP-TolAI mutants that His_{22} drives assembly of TolQRA *in vivo*. This essential His residue (and the helical groove) is conserved in the Ton system, suggesting a common assembly mechanism. Moreover, this model revealed a cryptic interaction between TolR and TolA via a co-conserved β-strand, also found between ExbD-TonB. The conservation of Tol-Pal and Ton systems throughout the kingdom of bacteria was also catalogued here, as a means to investigate the evolution of TolAII and TonBII secondary structures (chapter four). Previous work noted that TolA and TonB are distant homologues, and here I show that not only does TolAII exhibit proline-rich and mixed domain II structures like TonBII, but also propose a putative origin for TolA from a TonB paralogue, resolving a decades-old question and again supporting a common mechanism of force transduction (Fig. 8.1). This work also highlights the prevalence of CpoB throughout Tol-Pal systems everywhere, suggesting its role in coordinating OM-stabilisation with

cell wall remodeling emerged at the same time as the Tol-Pal system, warranting further investigation.

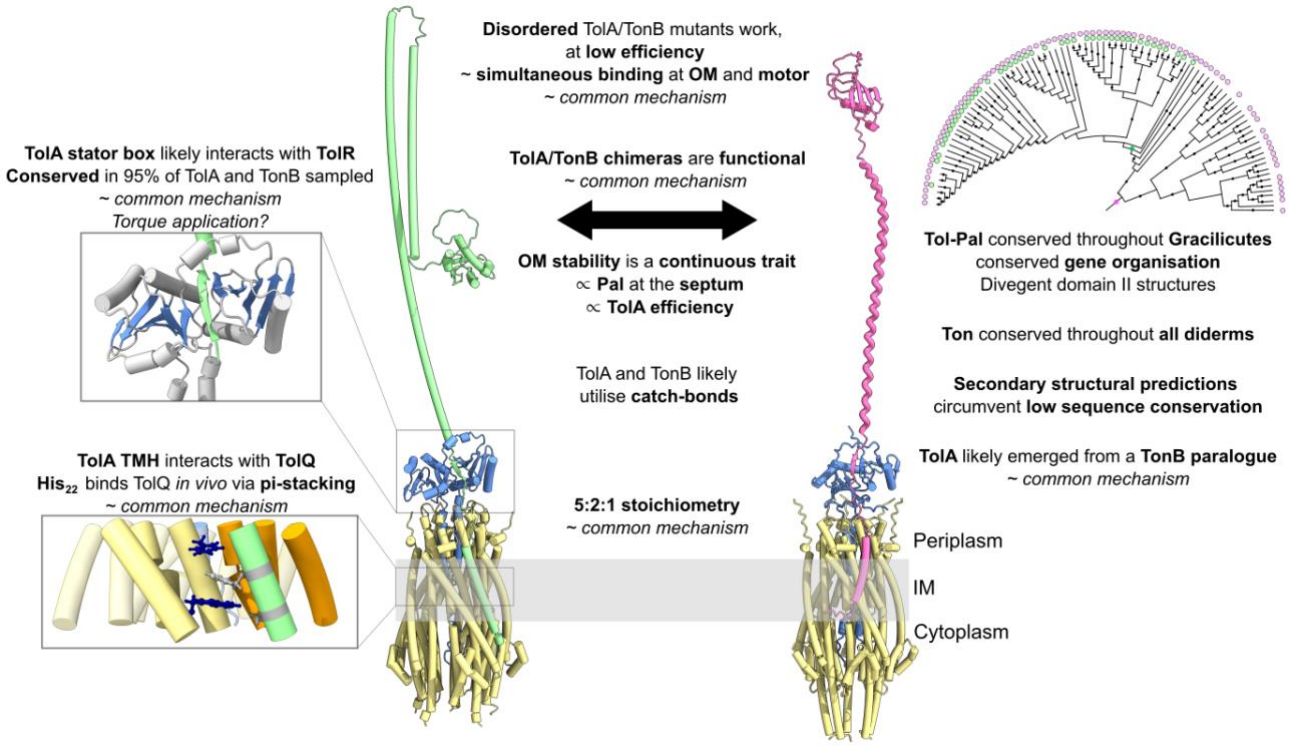


Figure 8.1 | Graphic overview of data from this work that support a common mechanism of force transduction between TolA and TonB.

Despite these advances, many questions remain unresolved. How proton flux drives TolQR rotation, and how this rotation applies tensile force through TolA is another such question, which can be addressed through XL-MS and subsequent MD simulations (chapter five). Relative rotation is strongly implicated in the mechanism of both TolQR and ExbBD, but in absolute terms it is still unknown whether the pentamer rotates around a stationary PG-anchored TolR/ExbD dimer, or whether the TolR/ExbD dimer rotates within a stationary pentamer. Models presented have assumed the former, but the implications of dimer-transducer interactions suggest that force transduction may preclude PG-binding, supporting the latter. Here we also observe OM perturbations arising from deletion and mutation of the motor-box region of TolA, the first biological evidence for the functionality of the motif (chapter six). In combination with ExbD-TonB structural data, this suggests that the motor-box plays a mechanistic role in the Tol-Pal and Ton systems alike. Unlike the Ton system, which is uniformly distributed in *E. coli*, how TolQR and TolA are independently recruited to the divisome also remains to be established, but may be addressed by *in vivo* BPA crosslinking and cryo-ET, both of which were optimised here. I also demonstrate the

use of minicells as viable platform to study cell-envelope processes that require energisation, their amenability to cryo-ET, and the use of next-generation labels for protein identification within tomograms (chapter seven). Through this work we understand the Tol-Pal system better than ever before, and its prevalence throughout bacteria highlights the need for further inquiry into this conserved machinery in the interest of public health, in the age of antimicrobial resistance.

References

1. Muñoz-Gómez, S.A. and A. Roger, *Leaving negative ancestors behind.* eLife, 2016. **5**: p. e20061.
2. Witwinowski, J., et al., *An ancient divide in outer membrane tethering systems in bacteria suggests a mechanism for the diderm-to-monoderm transition.* Nature Microbiology, 2022. **7**(3): p. 411-422.
3. Gram, H.C., *Ueber die isolirte Farbung der Schizomyceten in Schnitt-und Trockenpraparaten.* Fortschritte der Medizin, 1884. **2**: p. 185-89.
4. Oren, A. and G. Garrity, *Valid publication of the names of forty-two phyla of prokaryotes.* International Journal of Systematic Evolutionary Microbiology, 2021. **71**(10): p. 005056.
5. Stackebrandt, E., R. Murray, and H. Trüper, *Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives".* International Journal of Systematic Evolutionary Microbiology, 1988. **38**(3): p. 321-325.
6. Asmar, A.T., et al., *Communication across the bacterial cell envelope depends on the size of the periplasm.* PLOS Biology, 2017. **15**(12): p. e2004303.
7. Briegel, A., et al., *Universal architecture of bacterial chemoreceptor arrays.* Proceedings of the National Academy of Sciences, 2009. **106**(40): p. 17181-17186.
8. Matias, V.R., et al., *Cryo-transmission electron microscopy of frozen-hydrated sections of Escherichia coli and Pseudomonas aeruginosa.* Journal of Bacteriology, 2003. **185**(20): p. 6112-6118.
9. Narita, S.-i., S.-i. Matsuyama, and H. Tokuda, *Lipoprotein trafficking in Escherichia coli.* Archives of Microbiology, 2004. **182**: p. 1-6.
10. Payne, D.J., et al., *Drugs for bad bugs: confronting the challenges of antibacterial discovery.* Nature Reviews Drug Discovery, 2007. **6**(1): p. 29-40.
11. Saxena, D., et al., *Tackling the outer membrane: facilitating compound entry into Gram-negative bacterial pathogens.* npj Antimicrobials and Resistance, 2023. **1**(1): p. 17.
12. Sturgis, J.N., *Organisation and evolution of the tol-pal gene cluster.* Journal of Molecular Microbiology and Biotechnology, 2001. **3**(1): p. 113-22.
13. Szczepaniak, J., C. Press, and C. Kleanthous, *The multifarious roles of Tol-Pal in Gram-negative bacteria.* FEMS Microbiology Reviews, 2020. **44**(4): p. 490-506.
14. Onodera, K., B. Rolfe, and A. Bernstein, *Demonstration of missing membrane proteins in deletion mutants of E. coli K12.* Biochemical and Biophysical Research Communications, 1970. **39**(5): p. 969-975.
15. Kleanthous, C., *Swimming against the tide: progress and challenges in our understanding of colicin translocation.* Nature Reviews Microbiology, 2010. **8**(12): p. 843-8.
16. Davies, J.K. and P. Reeves, *Genetics of resistance to colicins in Escherichia coli K-12: cross-resistance among colicins of group A.* Journal of Bacteriology, 1975. **123**(1): p. 102-17.
17. Davies, J.K. and P. Reeves, *Genetics of resistance to colicins in Escherichia coli K-12: cross-resistance among colicins of group B.* Journal of Bacteriology, 1975. **123**(1): p. 96-101.
18. Benn, G., et al., *Phase separation in the outer membrane of Escherichia coli.* Proceedings of the National Academy of Sciences, 2021. **118**(44): p. e2112237118.
19. Webby, M.N., et al., *Lipids mediate supramolecular outer membrane protein assembly in bacteria.* Science Advances, 2022. **8**(44): p. eadc9566.
20. Jaroslawski, S., et al., *High-resolution architecture of the outer membrane of the Gram-negative bacteria Roseobacter denitrificans.* Molecular Microbiology, 2009. **74**(5): p. 1211-22.

21. Nikaido, H., *Molecular basis of bacterial outer membrane permeability revisited.* Microbiology and Molecular Biology Reviews, 2003. **67**(4): p. 593-656.
22. Marcoux, J., et al., *Mass spectrometry defines the C-terminal dimerization domain and enables modeling of the structure of full-length OmpA.* Structure, 2014. **22**(5): p. 781-790.
23. Iguchi, A., et al., *A complete view of the genetic diversity of the Escherichia coli O-antigen biosynthesis gene cluster.* DNA Research, 2015. **22**(1): p. 101-107.
24. Stenutz, R., A. Weintraub, and G. Widmalm, *The structures of Escherichia coli O-polysaccharide antigens.* FEMS Microbiology Reviews, 2006. **30**(3): p. 382-403.
25. Vergalli, J., et al., *Porins and small-molecule translocation across the outer membrane of Gram-negative bacteria.* Nature Reviews Microbiology, 2020. **18**(3): p. 164-176.
26. Nakae, T., *Identification of the outer membrane protein of E. coli that produces transmembrane channels in reconstituted vesicle membranes.* Biochemical and biophysical research communication, 1976. **71**(3): p. 877-884.
27. Silale, A. and B.v.d. Berg, *TonB-Dependent Transport Across the Bacterial Outer Membrane.* Annual Review of Microbiology, 2023. **77**(1): p. null.
28. Rojas, E.R., et al., *The outer membrane is an essential load-bearing element in Gram-negative bacteria.* Nature, 2018. **559**(7715): p. 617-621.
29. Auer, G.K., *Mechanical genomics identifies diverse modulators of bacterial cell stiffness.* Cell Systems, 2016. **2**.
30. Deghelt, M., et al., *The outer membrane and peptidoglycan layer form a single mechanical device balancing turgor.* BioRxiv, 2023: p. 2023.04. 29.538579.
31. Garde, S., P.K. Chodisetti, and M. Reddy, *Peptidoglycan: structure, synthesis, and regulation.* EcoSal Plus, 2021. **9**(2).
32. Labischinski, H., et al., *On the Secondary and Tertiary Structure of Murein: Low and Medium-Angle X-Ray Evidence against Chitin-Based Conformations of Bacterial Peptidoglycan.* European Journal of Biochemistry, 1979. **95**(1): p. 147-155.
33. Labischinski, H., et al., *Direct proof of a "more-than-single-layered" peptidoglycan architecture of Escherichia coli W7: a neutron small-angle scattering study.* Journal of bacteriology, 1991. **173**(2): p. 751-756.
34. Vollmer, W., D. Blanot, and M.A. De Pedro, *Peptidoglycan structure and architecture.* FEMS Microbiology Reviews, 2008. **32**(2): p. 149-167.
35. Gan, L., S. Chen, and G. Jensen, *Molecular organization of Gram-negative peptidoglycan.* Proceedings of the National Academy of Sciences, 2008. **105**(48): p. 18953-18957.
36. Turner, R.D., W. Vollmer, and S. Foster, *Different walls for rods and balls: the diversity of peptidoglycan.* Molecular Microbiology, 2014. **91**(5): p. 862-874.
37. Verwer, R., et al., *Arrangement of glycan chains in the sacculus of Escherichia coli.* Journal of bacteriology, 1978. **136**(2): p. 723-729.
38. Verwer, R., et al., *Oriented fragmentation of Escherichia coli sacculi by sonication.* Journal of Bacteriology, 1980. **141**(1): p. 327-332.
39. Yao, X., et al., *Thickness and elasticity of Gram-negative murein sacculi measured by atomic force microscopy.* Journal of Bacteriology, 1999. **181**(22): p. 6865-6875.
40. Turner, R.D., et al., *Cell wall elongation mode in Gram-negative bacteria is determined by peptidoglycan architecture.* Nature Communications, 2013. **4**: p. 1496.
41. Boulbitch, A., B. Quinn, and D. Pink, *Elasticity of the rod-shaped Gram-negative Eubacteria.* Physical Review Letters, 2000. **85**(24): p. 5246.
42. Szczepaniak, J., et al., *The lipoprotein Pal stabilises the bacterial outer membrane during constriction by a mobilisation-and-capture mechanism.* Nature Communications, 2020. **11**(1): p. 1305.
43. Mamou, G., et al., *Peptidoglycan maturation controls spatiotemporal organisation of outer membrane proteins in Escherichia coli.* Nature, 2022. **606**(7916): p. 953-959.
44. Hummels, K.R., et al., *Coordination of bacterial cell wall and outer membrane biosynthesis.* Nature, 2023. **615**(7951): p. 300-304.

45. Rowlett, V.W. and W. Margolin, *The bacterial divisome: ready for its close-up.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2015. **370**(1679): p. 20150028.
46. Renner, L.D. and D.B. Weibel, *Cardiolipin microdomains localize to negatively curved regions of Escherichia coli membranes.* Proceedings of the National Academy of Sciences, 2011. **108**(15): p. 6264-6269.
47. Sohlenkamp, C. and O. Geiger, *Bacterial membrane lipids: diversity in structures and pathways.* FEMS Microbiology Reviews, 2016. **40**(1): p. 133-159.
48. McMurry, L.M., A. George, and S. Levy, *Active efflux of chloramphenicol in susceptible Escherichia coli strains and in multiple-antibiotic-resistant (Mar) mutants.* Antimicrobial Agents and Chemotherapy, 1994. **38**(3): p. 542-546.
49. Mitchell, P., *Chemiosmotic coupling in oxidative and photosynthetic phosphorylation.* Biological Reviews, 1966. **41**(3): p. 445-501.
50. Kühlbrandt, W., *Structure and mechanisms of F-type ATP synthases.* Annual Review of Biochemistry, 2019. **88**(1): p. 515-549.
51. Schairer, H., et al., *Topological and functional aspects of the proton conductor, F_0, of the Escherichia coli ATP-synthase.* Bioscience Reports, 1982. **2**: p. 631-639.
52. Martin, F.J., et al., *Ion-driven rotary membrane motors: From structure to function.* Current Opinion in Structural Biology, 2024. **88**: p. 102884.
53. Zgurskaya, H.I. and H. Nikaido, *Bypassing the periplasm: reconstitution of the AcrAB multidrug efflux pump of Escherichia coli.* Proceedings of the National Academy of Sciences, 1999. **96**(13): p. 7190-7195.
54. Mould, R.M. and C. Robinson, *A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane.* Journal of Biological Chemistry, 1991. **266**(19): p. 12189-12193.
55. Arkowitz, R.A. and W. Wickner, *SecD and SecF are required for the proton electrochemical gradient stimulation of preprotein translocation.* The EMBO Journal, 1994. **13**(4): p. 954-963.
56. Cascales, E., R. Lloubes, and J.N. Sturgis, *The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB.* Molecular Microbiology, 2001. **42**(3): p. 795-807.
57. Deme, J.C., et al., *Structures of the stator complex that drives rotation of the bacterial flagellum.* Nature Microbiology, 2020. **5**(12): p. 1553-1564.
58. Santiveri, M., et al., *Structure and Function of Stator Units of the Bacterial Flagellar Motor.* Cell, 2020. **183**(1): p. 244-257.e16.
59. Hu, H., et al., *Structural basis of torque generation in the bi-directional bacterial flagellar motor.* Trends in Biochemical Sciences, 2021.
60. Singh, P.K., et al., *CryoEM structures reveal how the bacterial flagellum rotates and switches direction.* Nature Microbiology, 2024: p. 1-11.
61. Wilhelms, M., et al., *Two redundant sodium-driven stator motor proteins are involved in Aeromonas hydrophila polar flagellum rotation.* Journal of Bacteriology, 2009. **191**(7): p. 2206-2217.
62. Yonekura, K., S. Maki-Yonekura, and M. Homma, *Structure of the flagellar motor protein complex PomAB: implications for the torque-generating conformation.* Journal of Bacteriology, 2011. **193**(15): p. 3863-3870.
63. Luciano, J., et al., *Emergence and Modular Evolution of a Novel Motility Machinery in Bacteria.* PLOS Genetics, 2011. **7**(9): p. e1002268.
64. Hennell James, R., et al., *Structure and mechanism of the proton-driven motor that powers type 9 secretion and gliding motility.* Nature Microbiology, 2021. **6**(2): p. 221-233.
65. Ren, C.-P., et al., *The Flag-2 locus, an ancestral gene cluster, is potentially associated with a novel flagellar system from Escherichia coli.* Journal of Bacteriology, 2005. **187**(4): p. 1430-1440.
66. Sowa, Y. and R. Berry, *Bacterial flagellar motor.* Quarterly Reviews of Biophysics, 2008. **41**(2): p. 103-132.

67. Hendrix, R., *Symmetry mismatch and DNA packaging in large bacteriophages.* Proceedings of the National Academy of Sciences, 1978. **75**(10): p. 4779-4783.
68. Rieu, M., et al., *A new class of biological ion-driven rotary molecular motors with 5:2 symmetry.* Frontiers in Microbiology, 2022. **13**.
69. Taylor, N., et al., *Structure and mechanism of the Zorya anti-phage defense system.* Research Square, 2024.
70. Kirchweger, P., et al., *Structural and functional characterization of SiiA, an auxiliary protein from the SPI4-encoded type 1 secretion system from Salmonella enterica.* Molecular Microbiology, 2019. **112**(5): p. 1403-1422.
71. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
72. Van Kempen, M., et al., *Fast and accurate protein structure search with Foldseek.* Nature Biotechnology, 2024. **42**(2): p. 243-246.
73. Frost, G. and H. Rosenberg, *Relationship between the tonB locus and iron transport in Escherichia coli.* Journal of Bacteriology, 1975. **124**(2): p. 704-712.
74. Garen, A. and T.T. Puck, *The first two steps of the invasion of host cells by bacterial viruses. .* The Journal of experimental medicine, 1951. **94**(3): p. 177.
75. Pressler, U., et al., *Genetics of the iron dicitrate transport system of Escherichia coli.* Journal of bacteriology, 1988. **170**(6): p. 2716-2724.
76. Sean Peacock, R., et al., *The solution structure of the C-terminal domain of TonB and interaction studies with TonB box peptides.* Journal of Molecular Biology, 2005. **345**(5): p. 1185-97.
77. Shultis, D.D., et al., *Outer membrane active transport: structure of the BtuB:TonB complex.* Science, 2006. **312**(5778): p. 1396-9.
78. Larsen, R.A., et al., *Partial suppression of an Escherichia coli TonB transmembrane domain mutation (delta V17) by a missense mutation in ExbB.* Molecular Microbiology, 1994. **13**(4): p. 627-40.
79. Gumbart, J., M.C. Wiener, and E. Tajkhorshid, *Mechanics of Force Propagation in TonB-Dependent Outer Membrane Transport.* Biophysical Journal, 2007. **93**(2): p. 496-504.
80. Ratliff, A.C., S.K. Buchanan, and H. Celia, *The Ton Motor.* Frontiers in Microbiology, 2022. **13**.
81. Reeves, P., *The Bacteriocins*. Molecular Biology, Biochemistry and Biophysics. Vol. 11. 1972, New York: Springer.
82. Hill, C. and I.B. Holland, *Genetic basis of colicin E susceptibility in Escherichia coli. I. Isolation and properties of refractory mutants and the preliminary mapping of their mutations.* Journal of Bacteriology, 1967. **94**(3): p. 677-86.
83. Nagel de Zwaig, R. and S.E. Luria, *Genetics and physiology of colicin-tolerant mutants of Escherichia coli.* Journal of Bacteriology, 1967. **94**(4): p. 1112-23.
84. Walker, D., et al., *The role of electrostatics in colicin nuclease domain translocation into bacterial cells.* Journal of Biological Chemistry, 2007. **282**(43): p. 31389-97.
85. Wilmsen, H., A. Pugsley, and F. Pattus, *Colicin N forms voltage-and pH-dependent channels in planar lipid bilayer membranes.* European Biophysics Journal, 1990. **18**: p. 149-158.
86. Egan, A.J.F., *Bacterial outer membrane constriction.* Molecular Microbiology, 2018.
87. Egan, A.J.F., et al., *Induced conformational changes activate the peptidoglycan synthase PBP1B.* Molecular Microbiology, 2018. **110**(3): p. 335-356.
88. Levengood, S.K., W.F. Beyer, Jr., and R.E. Webster, *TolA: a membrane protein involved in colicin uptake contains an extended helical region.* Proceedings of the National Academy of Sciences, 1991. **88**(14): p. 5939-5943.
89. Muller, M.M., et al., *Membrane topology of the Escherichia coli TolR protein required for cell envelope integrity.* Journal of Bacteriology, 1993. **175**(18): p. 6059-61.
90. Cascales, E., et al., *Pal lipoprotein of Escherichia coli plays a major role in outer membrane integrity.* Journal of Bacteriology, 2002. **184**(3): p. 754-9.

91. Bonsor, D.A., et al., *Allosteric beta-propeller signalling in TolB and its manipulation by translocating colicins.* EMBO Journal, 2009. **28**(18): p. 2846-57.
92. Abergel, C., et al., *Structure of the Escherichia coli TolB protein determined by MAD methods at 1.95 A resolution.* Structure, 1999. **7**(10): p. 1291-300.
93. Petiti, M., et al., *Tol Energy-Driven Localization of Pal and Anchoring to the Peptidoglycan Promote Outer-Membrane Constriction.* Journal of Molecular Biology, 2019.
94. Traub, I., S. Gaisser, and V. Braun, *Activity domains of the TonB protein.* Molecular Microbiology, 1993. **8**(2): p. 409-423.
95. Germon, P., et al., *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, 1998. **180**(24): p. 6433-9.
96. Levengood, S. and R.E. Webster, *Nucleotide sequences of the tolA and tolB genes and localization of their products, components of a multistep translocation system in Escherichia coli.* Journal of bacteriology, 1989. **171**(12): p. 6600-6609.
97. Webby, M.N., et al., *Force-Generation by the Trans-Envelope Tol-Pal System.* Frontiers in Microbiology, 2022. **13**.
98. Kim, D.E., D. Chivian, and D. Baker, *Protein structure prediction and analysis using the Robetta server.* Nucleic Acids Research, 2004. **32**(suppl_2): p. W526-W531.
99. Walburger, A., C. Lazdunski, and Y. Corda, *The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB.* Molecular Microbiology, 2002. **44**(3): p. 695-708.
100. Abramson, J., et al., *Accurate structure prediction of biomolecular interactions with AlphaFold 3.* Nature, 2024: p. 1-3.
101. Kampfenkel, K. and V. Braun, *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, 1993. **175**(14): p. 4485-91.
102. Vianney, A., et al., *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, 1994. **176**(3): p. 822-9.
103. Lazzaroni, J.C., et al., *Transmembrane alpha-helix interactions are required for the functional assembly of the Escherichia coli Tol complex.* Journal of Molecular Biology, 1995. **246**(1): p. 1-7.
104. Cascales, E., et al., *Proton motive force drives the interaction of the inner membrane TolA and outer membrane pal proteins in Escherichia coli.* Molecular Microbiology, 2000. **38**(4): p. 904-15.
105. Wojdyla, J.A., et al., *Structure and function of the Escherichia coli Tol-Pal stator protein TolR.* Journal of Biological Chemistry, 2015. **290**(44): p. 26675-87.
106. Yahashiri, A., M.A. Jorgenson, and D. Weiss, *Bacterial SPOR domains are recruited to septal peptidoglycan by binding to glycan strands that lack stem peptides.* Proceedings of the National Academy of Sciences, 2015. **112**(36): p. 11347-11352.
107. Eick-Helmerich, K. and V. Braun, *Import of biopolymers into Escherichia coli: nucleotide sequences of the exbB and exbD genes are homologous to those of the tolQ and tolR genes, respectively.* Journal of Bacteriology, 1989. **171**(9): p. 5117-26.
108. Hale, C.A., L. Persons, and P.A.J. de Boer, *Recruitment of the TolA protein to cell constriction sites in Escherichia coli via three separate mechanisms, and a critical role for FtsWI activity in recruitment of both TolA and TolQ.* Journal of Bacteriology, 2021: p. Jb0046421.
109. Gerding, M.A., et al., *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, 2007. **63**(4): p. 1008-25.
110. Krachler, A.M., et al., *TolA modulates the oligomeric status of YbgF in the bacterial periplasm.* Journal of Molecular Biology, 2010. **403**(2): p. 270-85.
111. Gray, A.N., et al., *Coordination of peptidoglycan synthesis and outer membrane constriction during Escherichia coli cell division.* eLife, 2015. **4**.

112. Gully, D. and E. Bouveret, *A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in Escherichia coli.* Proteomics, 2006. **6**(1): p. 282-293.
113. Masilamani, R., M.B. Cian, and Z.D. Dalebroux, *Salmonella Tol-Pal reduces outer membrane glycerophospholipid levels for envelope homeostasis and survival during bacteremia.* Infection and Immunity, 2018. **86**(7): p. 10.1128/iai. 00173-18.
114. Rojas, E.R. and K.C. Huang, *Regulation of microbial growth by turgor pressure.* Current Opinion in Microbiology, 2018. **42**: p. 62-70.
115. Braun, V. and K. Rehn, *Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the E. coli cell wall: The specific effect of trypsin on the membrane structure.* European Journal of Biochemistry, 1969. **10**(3): p. 426-38.
116. Braun, V., *Covalent lipoprotein from the outer membrane of Escherichia coli.* Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes, 1975. **415**(3): p. 335-377.
117. Chai, T.-j. and J. Foulds, *Purification of protein A, an outer membrane component missing in Escherichia coli K-12 ompA mutants.* Biochimica Et Biophysica Acta - Protein Structure, 1977. **493**(1): p. 210-215.
118. Mizuno, T., *A novel peptidoglycan-associated lipoprotein found in the cell envelope of Pseudomonas aeruginosa and Escherichia coli.* Journal of Biochemistry, 1979. **86**(4): p. 991-1000.
119. Parsons, L.M., F. Lin, and J. Orban, *Peptidoglycan recognition by Pal, an outer membrane lipoprotein.* Biochemistry, 2006. **45**(7): p. 2122-8.
120. Park, J.S., et al., *Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the Gram-negative bacterial outer membrane.* FASEB Journal, 2012. **26**(1): p. 219-28.
121. Inouye, M., J. Shaw, and C. Shen, *The assembly of a structural lipoprotein in the envelope of Escherichia coli.* Journal of Biological Chemistry, 1972. **247**(24): p. 8154-9.
122. Li, G.W., et al., *Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources.* Cell, 2014. **157**(3): p. 624-35.
123. Koebnik, R., K.P. Locher, and P. Van Gelder, *Structure and function of bacterial outer membrane proteins: barrels in a nutshell.* Molecular Microbiology, 2000. **37**(2): p. 239-253.
124. Sheng, Q., et al., *In situ visualization of Braun's lipoprotein on E. coli sacculi.* Science Advances, 2023. **9**(3): p. eadd8659.
125. Mathelié-Guinlet, M., et al., *Lipoprotein Lpp regulates the mechanical properties of the E. coli cell envelope.* Nature Communications, 2020. **11**(1): p. 1789.
126. Sonntag, I., et al., *Cell envelope and shape of Escherichia coli: multiple mutants missing the outer membrane lipoprotein and other major outer membrane proteins.* Journal of bacteriology, 1978. **136**(1): p. 280-285.
127. de Pedro, M.A., C.G. Grunfelder, and H. Schwarz, *Restricted Mobility of Cell Surface Proteins in the Polar Regions of Escherichia coli.* Journal of Bacteriology, 2004. **186**(9): p. 2594-602.
128. Ursell, T.S., et al., *Analysis of surface protein expression reveals the growth pattern of the Gram-negative outer membrane.* PLoS Computational Biology, 2012. **8**(9): p. e1002680.
129. Rassam, P., et al., *Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria.* Nature, 2015. **523**(7560): p. 333-6.
130. Lloubès, R., et al., *The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity?* Research in Microbiology, 2001. **152**(6): p. 523.
131. Williams-Jones, D.P., et al., *Tunable force transduction through the Escherichia coli cell envelope.* Proceedings of the National Academy of Sciences, 2023. **120**(47): p. e2306707120.

132. Froger, A. and J. Hall, *Transformation of plasmid DNA into E. coli using the heat shock method.* Journal of Visualized Experiments, 2007(6).
133. Dagert, M. and S. Ehrlich, *Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells.* Gene, 1979. **6**(1): p. 23-28.
134. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.* Molecular Systems Biology, 2006. **2**: p. 2006 0008.
135. Meselson, M. and R. Yuan, *DNA Restriction Enzyme from E. coli.* Nature, 1968. **217**(5134): p. 1110-1114.
136. Studier, F.W. and B.A. Moffatt, *Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.* Journal of Molecular Biology, 1986. **189**(1): p. 113-30.
137. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proceedings of the National Academy of Sciences, 2000. **97**(12): p. 6640-5.
138. De Boer, P.A., R.E. Crossley, and L.I. Rothfield, *A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli.* Cell, 1989. **56**(4): p. 641-649.
139. Guyer, M., et al. *Identification of a sex-factor-affinity site in E. coli as γδ*. in *Cold Spring Harbor symposia on quantitative biology*. 1981. Cold Spring Harbor Laboratory Press.
140. Connolley, L., et al., *The quantitative basis for the redistribution of immobile bacterial lipoproteins to division septa.* PLOS Computational Biology, 2022. **17**(12): p. e1009756.
141. Valbuena, F.M., et al., *A photostable monomeric superfolder green fluorescent protein.* Traffic, 2020. **21**(8): p. 534-544.
142. Thorne, H., *Electrophoretic separation of polyoma virus DNA from host cell DNA.* Virology, 1966. **29**(2): p. 234-239.
143. Braha, O., et al., *Designed protein pores as components for biosensors.* Chemistry & Biology, 1997. **4**(7): p. 497-505.
144. Sanger, F. and A. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.* Journal of molecular biology, 1975. **94**(3): p. 441-448.
145. Chin, J.W., et al., *Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli.* Proceedings of the National Academy of Sciences, 2002. **99**(17): p. 11020-4.
146. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
147. Tinevez, J.-Y., et al., *TrackMate: An open and extensible platform for single-particle tracking.* Methods, 2017. **115**: p. 80-90.
148. Inns, P., *The organisation and dynamics of Escherichia coli outer membrane proteins*. 2021, University of Oxford.
149. Buchan, D.W. and D. Jones, *The PSIPRED protein analysis workbench: 20 years on.* Nucleic Acids Research, 2019. **47**(W1): p. W402-W407.
150. O'Brien, K.T., et al., *Prediction of polyproline II secondary structure propensity in proteins.* Royal Society Open Science, 2020. **7**(1): p. 191239.
151. Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nature Methods, 2022. **19**(6): p. 679-682.
152. Benson, D.A., et al., *GenBank.* Nucleic Acids Research, 2012. **41**(D1): p. D36-D42.
153. Sonnhammer, E.L., G. Von Heijne, and A. Krogh. *A hidden Markov model for predicting transmembrane helices in protein sequences*. in *ISMB*. 1998.
154. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.* Journal of Molecular Biology, 2001. **305**(3): p. 567-580.
155. Teufel, F., et al., *SignalP 6.0 predicts all five types of signal peptides using protein language models.* Nature Biotechnology, 2022. **40**(7): p. 1023-1025.

156. Zeng, H., et al., *ComplexContact: a web server for inter-protein contact prediction using deep learning.* Nucleic Acids Research, 2018. **46**(W1): p. W432-W437.
157. Pettersen, E.F., et al., *UCSF ChimeraX: Structure visualization for researchers, educators, and developers.* Protein Science, 2021. **30**(1): p. 70-82.
158. Jordan, L.D., et al., *Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane.* Proceedings of the National Academy of Sciences, 2013. **110**(28): p. 11553-8.
159. Karlsson, M., K. Hannavy, and C.F. Higgins, *A sequence-specific function for the N-terminal signal-like sequence of the TonB protein.* Molecular Microbiology, 1993. **8**(2): p. 379-88.
160. Braun, V. and C. Herrmann, *Evolutionary relationship of uptake systems for biopolymers in Escherichia coli: cross-complementation between the TonB-ExbB-ExbD and the TolA-TolQ-TolR proteins.* Molecular Microbiology, 1993. **8**(2): p. 261-8.
161. Koebnik, R., *The molecular interaction between components of the TonB-ExbBD-dependent and of the TolQRA-dependent bacterial uptake systems.* Molecular Microbiology, 1993. **9**(1): p. 219.
162. Germon, P., et al., *Energy-dependent conformational change in the TolA protein of Escherichia coli involves its N-terminal domain, TolQ, and TolR.* Journal of Bacteriology, 2001. **183**(14): p. 4110-4.
163. Witty, M., et al., *Structure of the periplasmic domain of Pseudomonas aeruginosa TolA: evidence for an evolutionary relationship with the TonB transporter protein.* EMBO Journal, 2002. **21**(16): p. 4207-4218.
164. Derouiche, R., et al., *Circular dichroism and molecular modeling of the E. coli TolA periplasmic domains.* Biospectroscopy, 1999. **5**(3): p. 189-98.
165. Domingo Köhler, S., et al., *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, 2010. **19**(4): p. 625-630.
166. Brewer, S., et al., *Structure and function of X-Pro dipeptide repeats in the TonB proteins of Salmonella typhimurium and Escherichia coli.* Journal of Molecular Biology, 1990. **216**(4): p. 883-95.
167. Evans, J.S., et al., *Sequence-imposed structural constraints in the TonB protein of E. coli.* FEBS Letters, 1986. **208**(2): p. 211-6.
168. Chu, B.C., R.S. Peacock, and H. Vogel, *Bioinformatic analysis of the TonB protein family.* Biometals, 2007. **20**: p. 467-483.
169. Kadner, R.J., *Vitamin B12 transport in Escherichia coli: energy coupling between membranes.* Molecular Microbiology, 1990. **4**(12): p. 2027-33.
170. Cadieux, N. and R.J. Kadner, *Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter.* Proceedings of the National Academy of Sciences, 1999. **96**(19): p. 10673-8.
171. Merianos, H.J., et al., *Substrate-induced exposure of an energy-coupling motif of a membrane transporter.* Nature Structural & Molecular Biology, 2000. **7**(3): p. 205-9.
172. Consortium, U., *UniProt: a worldwide hub of protein knowledge.* Nucleic Acids Research, 2019. **47**(D1): p. D506-D515.
173. Rohl, C.A., et al., *Protein Structure Prediction Using Rosetta*, in *Methods in Enzymology*, B. Ludwig and L.J. Michael, Editors. 2004, Academic Press. p. 66-93.
174. Adzhubei, A.A. and M.J. Sternberg, *Left-handed polyproline II helices commonly occur in globular proteins.* Journal of Molecular Biology, 1993. **229**(2): p. 472-493.
175. Stapley, B.J. and T.P. Creamer, *A survey of left-handed polyproline II helices.* Protein Science, 1999. **8**(3): p. 587-595.
176. Musacchio, A., *How SH3 domains recognize proline.* Advances in Protein Chemistry, 2002. **61**: p. 211-268.
177. Lin, Z., et al., *Evolutionary-scale prediction of atomic-level protein structure with a language model.* Science, 2023. **379**(6637): p. 1123-1130.

178. Okuyama, K., et al., *Structure Analysis of a Collagen Model Polypeptide,(Pro–Pro–Gly) 10.* Bulletin of the Chemical Society of Japan, 1976. **49**(7): p. 1805-1810.
179. Martino, M., et al., *On the occurrence of polyproline II structure in elastin.* Journal of Molecular Structure, 2000. **519**(1-3): p. 173-189.
180. Parry, D.A. and A. North, *Hard α-keratin intermediate filament chains: substructure of the N-and C-terminal domains and the predicted structure and function of the C-terminal domains of type I and type II chains.* Journal of Structural Biology, 1998. **122**(1-2): p. 67-75.
181. Pauling, L. and R.B. Corey, *Atomic coordinates and structure factors for two helical configurations of polypeptide chains.* Proceedings of the National Academy of Sciences, 1951. **37**(5): p. 235-240.
182. Cowan, P.M., S. McGavin, and A. North, *The polypeptide chain configuration of collagen.* Nature, 1955. **176**(4492): p. 1062-1064.
183. Francis, M.-L.R., et al., *Porin threading drives receptor disengagement and establishes active colicin transport through Escherichia coli OmpF.* EMBO Journal, 2021(40): p. e108610.
184. Ricca, R., *The energy spectrum of a twisted flexible string under elastic relaxation.* Journal of Physics A: Mathematical and Theoretical, 1995. **28**(8): p. 2335.
185. Deatherage, B.L., et al., *Biogenesis of bacterial membrane vesicles.* Molecular Microbiology, 2009. **72**(6): p. 1395-1407.
186. Weigand, R.A. and L.I. Rothfield, *Genetic and physiological classification of periplasmic-leaky mutants of Salmonella typhimurium.* Journal of Bacteriology, 1976. **125**(1): p. 340-345.
187. Weigand, R.A., K.D. Vinci, and L.I. Rothfield, *Morphogenesis of the bacterial division septum: a new class of septation-defective mutants.* Proceedings of the National Academy of Sciences, 1976. **73**(6): p. 1882-6.
188. Fung, J.C., et al., *Morphogenesis of the bacterial division septum: identification of potential sites of division in lkyD mutants of Salmonella typhimurium.* Journal of Bacteriology, 1980. **143**(2): p. 1019-1024.
189. Thomas, W.E., V. Vogel, and E. Sokurenko, *Biophysics of catch bonds.* Annual Review of Biophysics, 2008. **37**(1): p. 399-416.
190. Yakovenko, O., et al., *FimH forms catch bonds that are enhanced by mechanical force due to allosteric regulation.* Journal of Biological Chemistry, 2008. **283**(17): p. 11596-11605.
191. Sauer, M.M., et al., *Catch-bond mechanism of the bacterial adhesin FimH.* Nature Communications, 2016. **7**(1): p. 10738.
192. Jonsson, T., C.D. Waldburger, and R.T. Sauer, *Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates.* Biochemistry, 1996. **35**(15): p. 4795-802.
193. Mathelié-Guinlet, M., et al., *Force-clamp spectroscopy identifies a catch bond mechanism in a Gram-positive pathogen.* Nature Communications, 2020. **11**(1): p. 5431.
194. Marshall, B.T., et al., *Direct observation of catch bonds involving cell-adhesion molecules.* Nature, 2003. **423**(6936): p. 190-193.
195. Rassam, P., et al., *Intermembrane crosstalk drives inner-membrane protein organization in Escherichia coli.* Nature Communications, 2018. **9**(1): p. 1082.
196. Tubiana, L., et al., *Topology in soft and biological matter.* Physics Reports, 2024. **1075**: p. 1-137.
197. Carson, M. and C. Bugg, *Algorithm for ribbon models of proteins.* Journal of Molecular Graphics, 1986. **4**(2): p. 121-122.
198. Carson, M., *Ribbon models of macromolecules.* Journal of Molecular Graphics, 1987. **5**(2): p. 103-106.
199. Zinke, M., et al., *Ton motor conformational switch and peptidoglycan role in bacterial nutrient uptake.* Nature Communications, 2024. **15**(1): p. 331.

200. Au, C., *The geometric interpretation of linking number, writhe and twist for a ribbon.* Computer Modeling in Engineering & Sciences, 2008. **29**: p. 151-162.
201. Schendel, S.L., et al., *The TolA protein interacts with colicin E1 differently than with other group A colicins.* Journal of Bacteriology, 1997. **179**(11): p. 3683-90.
202. Gratia, J.P., *Resistance to Colicin B in Escherichia coli. Specificity Relations among Colicins B, I and V and Phage T-4. Genetic Study.* Ann Inst Pasteur (Paris), 1964. **107**: p. SUPPL:132-51.
203. Webster, R.E., *The tol gene products and the import of macromolecules into Escherichia coli.* Molecular Microbiology, 1991. **5**(5): p. 1005-1011.
204. Sun, T.P. and R.E. Webster, *fii, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant tolA and tolB.* Journal of Bacteriology, 1986. **165**(1): p. 107-15.
205. Vianney, A., et al., *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, 1996. **178**(14): p. 4031-4038.
206. Yeh, Y.-C., et al., *The Caulobacter Tol-Pal Complex Is Essential for Outer Membrane Integrity and the Positioning of a Polar Localization Factor.* Journal of Bacteriology, 2010. **192**(19): p. 4847-4858.
207. Dennis, J.J., E.R. Lafontaine, and P.A. Sokol, *Identification and characterization of the tolQRA genes of Pseudomonas aeruginosa.* Journal of Bacteriology, 1996. **178**(24): p. 7059-68.
208. Gaspar, J.A., et al., *Surface expression of O-specific lipopolysaccharide in Escherichia coli requires the function of the TolA protein.* Molecular Microbiology, 2000. **38**(2): p. 262-75.
209. Dubuisson, J.-F., et al., *Tol-Pal proteins are critical cell envelope components of Erwinia chrysanthemi affecting cell morphology and virulence.* Microbiology, 2005. **151**(10): p. 3337-3347.
210. Hirakawa, H., et al., *Roles of the Tol-Pal system in the Type III secretion system and flagella-mediated virulence in enterohemorrhagic Escherichia coli.* Scientific Reports, 2020. **10**(1): p. 15173.
211. Hubloher, J.J., et al., *The Tol-Pal system of Acinetobacter baumannii is important for cell morphology, antibiotic resistance and virulence.* International Microbiology, 2023. **26**(3): p. 543-550.
212. Abdelhamed, H., et al., *Involvement of tolQ and tolR genes in Edwardsiella ictaluri virulence.* Microbial Pathogenesis, 2016. **100**: p. 90-94.
213. Li, Q., et al., *The role of TolA, TolB, and TolR in cell morphology, OMVs production, and virulence of Salmonella Choleraesuis.* AMB Express, 2022. **12**(1): p. 5.
214. Altschul, S.F., et al., *Basic local alignment search tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-10.
215. Sharma, V., et al., *Phylogenomics of the phylum Proteobacteria: resolving the complex relationships.* Current Microbiology, 2022. **79**(8): p. 224.
216. Loll, P.J., et al., *Discovery and structural characterization of the D-box, a conserved TonB motif that couples an inner-membrane motor to outer-membrane transport.* Journal of Biological Chemistry, 2024. **300**(3).
217. Fodje, M. and S. Al-Karadaghi, *Occurrence, conformational features and amino acid propensities for the π-helix.* Protein Engineering, 2002. **15**(5): p. 353-358.
218. Enkhbayar, P., B. Boldgiv, and N. Matsushima, *ω-Helices in Proteins.* The Protein Journal, 2010. **29**: p. 242-249.
219. Armen, R., D.O. Alonso, and V. Daggett, *The role of α-, 310-, and π-helix in helix→ coil transitions.* Protein Science, 2003. **12**(6): p. 1145-1157.
220. Jolley, K.A., J.E. Bray, and M.C. Maiden, *Open-access bacterial population genomics: BIGSdb software, the PubMLST. org website and their applications.* Wellcome Open Research, 2018. **3**.

221. Heath, J.E., et al., *PG1058 is a novel multidomain protein component of the bacterial type IX secretion system.* PloS One, 2016. **11**(10): p. e0164313.
222. Marmon, L., *Elucidating the origin of the ExbBD components of the TonB system through Bayesian inference and maximum-likelihood phylogenies.* Molecular Phylogenetics and Evolution, 2013. **69**(3): p. 674-686.
223. Moi, D., et al., *Structural phylogenetics unravels the evolutionary diversification of communication systems in Gram-positive bacteria and their viruses.* BioRXiv, 2023: p. 2023.09. 19.558401.
224. Amorim, G.C.d., et al., *The structure of HasB reveals a new class of TonB protein fold.* PLoS One, 2013. **8**(3): p. e58964.
225. Lill, Y., et al., *Confined mobility of TonB and FepA in Escherichia coli membranes.* PLoS One, 2016. **11**(12): p. e0160862.
226. Hale, C.A., L. Persons, and P.A.J.d. Boer, *Recruitment of the TolA Protein to Cell Constriction Sites in Escherichia coli via Three Separate Mechanisms, and a Critical Role for FtsWI Activity in Recruitment of both TolA and TolQ.* Journal of Bacteriology, 2022. **204**(1): p. e00464-21.
227. Baccelli, P., et al., *Timing of TolA and TolQ Recruitment at the Septum Depends on the Functionality of the Tol-Pal System.* Journal of Molecular Biology, 2022. **434**(7): p. 167519.
228. Liu, Y. and J. Löwe, *Bacterial and archaeal cytoskeletons.* Current Biology, 2021. **31**(10): p. R542-R546.
229. Qiu, K., N. Ben-Tal, and R. Kolodny, *Similar protein segments shared between domains of different evolutionary lineages.* Protein Science, 2022. **31**(9): p. e4407.
230. Nichols, R.J., et al., *Phenotypic landscape of a bacterial cell.* Cell, 2011. **144**(1): p. 143-156.
231. Ruiz, N., *Filling holes in peptidoglycan biogenesis of Escherichia coli.* Current Opinion in Microbiology, 2016. **34**: p. 1-6.
232. Nenninger, A., et al., *Independent mobility of proteins and lipids in the plasma membrane of Escherichia coli.* Molecular Microbiology, 2014. **92**(5): p. 1142-1153.
233. Los, D.A. and N. Murata, *Membrane fluidity and its roles in the perception of environmental signals.* Biochimica et Biophysica Acta- Biomembranes, 2004. **1666**(1-2): p. 142-157.
234. de Mendoza, D. and M. Pilon, *Control of membrane lipid homeostasis by lipid-bilayer associated sensors: a mechanism conserved from bacteria to humans.* Progress in Lipid Research, 2019. **76**: p. 100996.
235. Journet, L., et al., *Role of TolR N-terminal, central, and C-terminal domains in dimerization and interaction with TolA and tolQ.* Journal of Bacteriology, 1999. **181**(15): p. 4476-84.
236. Goemaere, E.L., E. Cascales, and R. Lloubes, *Mutational analyses define helix organization and key residues of a bacterial membrane energy-transducing complex.* Journal of Molecular Biology, 2007. **366**(5): p. 1424-36.
237. Goemaere, E.L., et al., *Movements of the TolR C-terminal domain depend on TolQR ionizable key residues and regulate activity of the Tol complex.* Journal of Biological Chemistry, 2007. **282**(24): p. 17749-57.
238. Zhang, X.Y., et al., *Mapping the interactions between Escherichia coli tol subunits: rotation of the TolR transmembrane helix.* Journal of Biological Chemistry, 2009. **284**(7): p. 4275-82.
239. Parsons, L.M., A. Grishaev, and A. Bax, *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, 2008. **47**(10): p. 3131-42.
240. O'Neill, J., et al., *Role of the MotB linker in the assembly and activation of the bacterial flagellar motor.* Acta Crystallographica Section D, 2011. **67**(Pt 12): p. 1009-16.

241. Karimullina, E., et al., *Structural architecture of TolQ-TolR inner membrane protein complex from opportunistic pathogen Acinetobacter baumannii.* BioRxiv, 2024: p. 2024.06. 19.599759.
242. Jeon, J., J.-S. Yang, and S. Kim, *Integration of evolutionary features for the identification of functionally important residues in major facilitator superfamily transporters.* PLoS Computational Biology, 2009. **5**(10): p. e1000522.
243. Celia, H., et al., *Cryo-EM structure of the bacterial Ton motor subcomplex ExbB–ExbD provides information on structure and stoichiometry.* Communications Biology, 2019. **2**(1): p. 358.
244. Zhang, X.Y., et al., *Mapping the Interactions between Escherichia coli TolQ Transmembrane Segments.* Journal of Biological Chemistry, 2011. **286**(13): p. 11756-64.
245. Schmidt, T.G., et al., *Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants.* Protein Expression and Purification, 2013. **92**(1): p. 54-61.
246. Roujeinikova, A., *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, 2008. **105**(30): p. 10348-53.
247. Trinh, N.T.T., et al., *Crystal structure of Type IX secretion system PorE C-terminal domain from Porphyromonas gingivalis in complex with a peptidoglycan fragment.* Scientific Reports, 2020. **10**(1): p. 7384.
248. Moynihan, P.J., D. Sychantha, and A. Clarke, *Chemical biology of peptidoglycan acetylation and deacetylation.* Bioorganic Chemistry, 2014. **54**: p. 44-50.
249. Vollmer, W., *Bacterial cell walls: peptidoglycan*, in *Molecular Medical Microbiology*. 2024, Elsevier. p. 45-67.
250. Lyu, Z., et al., *FtsN maintains active septal cell wall synthesis by forming a processive complex with the septum-specific peptidoglycan synthases in E. coli.* Nature Communications, 2022. **13**(1): p. 5751.
251. Larsen, R.A., et al., *His(20) provides the sole functionally significant side chain in the essential TonB transmembrane domain.* Journal of Bacteriology, 2007. **189**(7): p. 2825-33.
252. Joseph, P.R.B., et al., *Proline substitution of dimer interface β-strand residues as a strategy for the design of functional monomeric proteins.* Biophysical Journal, 2013. **105**(6): p. 1491-1501.
253. Li, S.-C., et al., *Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment.* Proceedings of the National Academy of Sciences, 1996. **93**(13): p. 6676-6681.
254. Ollis, A.A., A. Kumar, and K. Postle, *The ExbD Periplasmic Domain Contains Distinct Functional Regions for Two Stages in TonB Energization.* Journal of Bacteriology, 2012. **194**(12): p. 3069-3077.
255. Kopp, D.R. and K. Postle, *The intrinsically disordered region of ExbD is required for signal transduction.* Journal of Bacteriology, 2020. **202**(7): p. 10.1128/jb. 00687-19.
256. Larsen, R.A. and K. Postle, *Conserved residues Ser(16) and His(20) 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, 2001. **276**(11): p. 8111-7.
257. White, P., et al., *Exploitation of an iron transporter for bacterial protein antibiotic import.* Proceedings of the National Academy of Sciences, 2017. **114**(45): p. 12051-12056.
258. Jansen, K.B., et al., *Bifurcated binding of the OmpF receptor underpins import of the bacteriocin colicin N into Escherichia coli.* Journal of Biological Chemistry, 2020. **295**(27): p. 9147-9156.
259. Housden, N.G., et al., *Cell entry mechanism of enzymatic bacterial colicins: Porin recruitment and the thermodynamics of receptor binding.* Proceedings of the National Academy of Sciences, 2005. **102**(39): p. 13849-13854.

260. Raggett, E.M., et al., *Discovery of critical Tol A-binding residues in the bactericidal toxin colicin N: a biophysical approach.* Molecular Microbiology, 1998. **28**(6): p. 1335-43.
261. Rassam, P., *Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria.* Nature, 2015. **523**.
262. Weber, M.S., M. Wojtynek, and O. Medalia, *Cellular and structural studies of eukaryotic cells by cryo-electron tomography.* Cells, 2019. **8**(1): p. 57.
263. Silvester, E., et al., *DNA origami signposts for identifying proteins on cell membranes by electron cryotomography.* Cell, 2021. **184**(4): p. 1110-1121.e16.
264. Manson, M.D., et al., *A protonmotive force drives bacterial flagella.* Proceedings of the National Academy of Sciences, 1977. **74**(7): p. 3060-3064.
265. Sun, M., et al., *Motor-driven intracellular transport powers bacterial gliding motility.* Proceedings of the National Academy of Sciences, 2011. **108**(18): p. 7559-64.
266. Meury, J. and A. Kepes, *The regulation of potassium fluxes for the adjustment and maintenance of potassium levels in Escherichia coli.* European Journal of Biochemistry, 1981. **119**(1): p. 165-170.