Structure and mode of action of the TolA-TolB complex from *Pseudomonas aeruginosa*

Peter Charles Holmes

Linacre College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2016

Structure and mode of action of the TolA-TolB complex from *Pseudomonas aeruginosa*

Peter Charles Holmes

Linacre College, University of Oxford
A thesis submitted for the degree of
Doctor of Philosophy
Michaelmas 2016

Abstract

Protein-protein interactions (PPIs) across the cell envelope of Gram-negative bacteria are critical for mediating signal transduction pathways that underpin cellular homeostasis. The Ton and Tol-Pal systems are two conserved, ancestrally related protein networks that are also required for bacterial pathogenesis. Both Ton and Tol-Pal traverse the periplasm to effect different functions at the outer membrane (OM). Tol-Pal is composed of a homologous complex of three inner membrane proteins, TolQ-TolR-TolA (linked to proton motive force) and two additional periplasmic proteins TolB and Pal. The physiological role of the Tol-Pal system is to stabilise the OM, however the mechanism involved is unknown. TolA is however known to form a crucial protein-protein interaction via its C-terminus with the disordered N-terminus of TolB.

Prior to this thesis, determination of the molecular features underlying a protein-protein complex between TolA and an endogenous binding partner TolB had never been accomplished. In this work, I describe the first structure comprising the TolA-TolB complex from Gram negative bacteria. The structure of this complex was determined from *Pseudomonas aeruginosa* by solution NMR spectroscopy. I determined the interaction between *P. aeruginosa* TolA and a TolB N-terminal peptide to be relatively weak using fluorescence anisotropy. I found that TolB interacts with TolA through an analogous mechanism to that seen in TonB-dependent transporters. Based on these studies and bioinformatics analyses, I hypothesize that the evolutionary resilience of the Tol-Pal system to external pressures is contingent on the preservation of the TolA-TolB interface. Structure-based mutations within the TolA-TolB complex were also evaluated for their effect on *in vivo* function of the Tol-Pal complex and impact on complex formation *in vitro*. Taken together, the results demonstrate that protein networks which transduce energy to the OM through PMF-dependent systems in bacterial cells appear to follow a common β-strand augmentation mechanism.

Table of Contents

List of tables

List of figures

3.11 The Carr-Purcell-Meiboom-Gill (CPMG) field dependencies of residues 228-347 in TolAIII are not consistent with any large conformational exchange contributions

4.1 Two conflicting models of Tol-Pal action in the Gram negative cell envelope

4.2 The TolA-TolB interaction is weak in vitro as determined by fluorescence anisotropy in the presence and absence of a disordered N-terminus

4.3 The TolA-TolB interaction is weak in vitro as determined by solution NMR spectroscopy

4.4 Full assignment of all H-N correlations in a ^{15}N HSQC of TolAIII saturated with $TolB_{22-34}$

4.5 Chemical shift changes in a disordered region of TolAIII as a probe for the binding mechanism of full length TolB

4.6 $TolB_{22-34}$ is purified from the trpLE fusion tag by reverse phase HPLC

4.7 Backbone amide resonances in a ^{15}N HSQC of $TolB_{22-34}$ free and saturated with TolAIII

4.8 ^{15}N, ^{13}C-edited filtered NOESY spectra clearly identify intermolecular NOE constraints between $TolB_{22-34}$ and TolAIII

4.9 NOE contributions made by $TolB_{22-34}$ side chains are observable by ^{15}N and ^{13}C edited-filtered NMR spectroscopy techniques

4.10 Carbon chemical shift changes as a consequence of $TolB_{22-34}$ binding on TolAIII

4.11 Chemical shift mapping and weighted average chemical shift difference ($\Delta\delta av$) for $^{1}H_N$ ^{15}N and $^{13}C\alpha$ chemical shift changes in {^{15}N, ^{13}C}-TolAIII from the free state to saturated with $paTolB_{22-34}$

4.12 Secondary structural changes in TolAIII as determined by TALOS+ measurements suggest a loss of helical character near the C-terminus

4.13 The TolAIII-$TolB_{22-34}$ interaction is a β-strand complementation mechanism involving the binding of $TolB_{22-34}$ to a cleft formed by the exposed face of β-strand 3 in TolAIII

4.14 The structure of $TolB_{22-34}$-TolAIII is a tight ensemble indicating the TolB N-terminus binds using an antiparallel β-complementation mechanism.

4.15 Analysis of NOE contacts determined to be between Leu 345 to residues within the TolB-binding site.

4.16 The binding cleft of TolAIII involves interactions with the first 8 amino acids in $TolB_{22-34}$ and necessitates the rearrangement of helix 4 in TolAIII

4.17 Forward and reverse NOEs from the ^{15}N, ^{13}C NOESY-HSQC experiment and the ^{1}H, ^{13}C edited-filtered NOESY experiment demonstrate clearly the position of peptide sidechains in the interaction interface

4.18 C-terminal TolAIII truncation produces a minimal effect on the dissociation rate constant of TolB for TolA, but does affect the association rate

4.19 The arrangement of aromatic side chains in the hydrophobic core of TolAIII was found to undergo major remodeling during the formation of a complex with $TolB_{22-34}$

5.1 Analysis of aligned N-termini of TolB demonstrates conserved regions preceded by a cleavable Sec/Tat signal.

5.2 The TolB N-terminus is an intrinsically disordered region with a conserved motif

5.3 The TolA-TolB binding interface is strongly conserved in Gram-negative bacteria

5.4 Δ*tolB* cells from the Keio background exhibit a tol phenotype that may be recovered using a plasmid encoding *tolB*

5.5 *tolB* alanine mutations in the [IVL]X[IVL] motif result in abrogation of Tol-Pal activity

5.6 *tolB* alanine mutations in the [IVL]X[IVL] motif result in abrogation of Tol-Pal activity via liquid growth assay

5.7 *P. aeruginosa* TolB suppressor mutant sequences do not out compete a $TolB_{22\text{-}34}$FITC reporter for binding TolAIII

5.8 A structural comparative model of the TolA-TolB interaction in *P. aeruginosa* and *E. coli*

6.1 The new state of knowledge on the TolA and TolB interaction via an intrinsically disordered region

List of abbreviations

Å	Angstroms (10^{-10} metres).
A	Alanine.
ABC transporter	ATP-binding cassette transporter.
AcN	Acetonitrile.
Amp	Ampicillin.
Arg	Arginine.
ARIA	Ambiguous restraints for iterative assignment, an automated NOE assignment and NMR structure calculation.
Asn	Asparagine.
Asp	Aspartic acid.
ATCC	American type culture collection.
ATP	Adenosine triphosphate.
Bam	β-Barrel assembly machinery.
^{13}C	Carbon 13, a rare natural, stable isotope of carbon
^{0}C	Degrees Celsius.
C18	Octadecyl carbon chain bonded silica for reverse phase chromatography.
Chlor	Chloramphenicol.
CNBr	Cyanogen Bromide.
CpoB	Coordinator of peptidoglycan synthesis and outer membrane constriction associated with PBP1B.
C-terminus	Carboxy-terminus.
C or Cys	Cysteine.
D	Aspartic acid.
Da	Dalton.
DANGLE	Dihedral ANgles from Global Likelihood Estimates. A Bayesian inferential prediction method for protein backbone dihedral angles.
DAP	Diaminopimelic acid.
DNA	Deoxyribonucleic acid.
E	Glutamic acid.
ECF	Extracellular fluid.
E.coli	*Escherichia coli*.
EDTA	Ethylenediaminetetracetic acid.
EMBO	European molecular biology organization.
ESI-MS	Electrospray ionization mass spectrometry.
ExbD	Biopolymer transport protein ExbD.
F	Phenylalanine.
Factor Xa or FXa	Serine protease which cleaves after arginine in sequence Ile-(Glu/Asp)-Gly-Arg
FITC	Fluoroscein isothiocyanate.
g	Acceleration due to gravity.
G or Gly	Glycine.
Gln	Glutamine.
Glu	Glutamic acid.
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid.
H or His	Histidine.
HMQC	Heteronuclear multiple quantum coherence.
HNCA	Hyper netted chain approximation.
HNCO	Isocyanic acid.
HSQC	Heteronuclear single quantum coherence.
HPLC	High performance liquid chromatography.
IDP	Intrinsically disordered protein.
IDR	Intrinsically disordered region.
I or Ile	Isoleucine.
Im9	Colicin E9 immunity protein.
IM	Inner cytoplasmic membrane.
IMP	Inner membrane protein.
ITC	Isothermal titration calorimetry.
IPTG	Isopropyl β-D-thiogalactopyranoside.

JNetBurial	Estimation of which residues are likely to form the buried core of a protein
JNetCONF	Confidence estimate for the prediction.
JNetHMM	Hidden markov model profile based prediction.
JNetPRED	Consensus prediction for secondary protein structure from amino acid sequence.
K	Lysine.
K_a	Association constant.
Kan	Kanamycin.
K_D	Dissociation constant.
KDa	Kilodaltons.
KDO	3-deoxy-D-manno-oct-2-ulosonic acid.
K_{obs}	Pseudo first order rate approximation.
LB broth	Lysogeny broth, also described as Luria-Burtani broth.
LC-MS	Liquid chromatography linked to mass spectrometry analysis.
L or Leu	Leucine.
LolABCDE	The Lol lipoprotein-sorting machinery, comprised of five proteins Lol A, B, C, D, and E
LP	Lipoprotein.
LPS	Lipopolysaccharide.
Lys	Lysine.
mA	Milliamperes.
MALDI	Matrix assisted laser desorption ionization mass spectrometry.
MES	2-(N-morpholino)ethane sulphonic acid.
mM	Millimolar (10^{-3} moles per litre).
M	Molar (moles per litre).
M or Met	Methionine.
μL	Microlitres (10^{-6} litres).
μM	Micromolar (10^{-6} moles per litre).
Min	Minutes.
MHz	Megahertz.
ml	millilitres (10^{-3} litres).
MotA	Motility protein A.
MotB	Motility protein B.
MWCO	Molecular weight cut off.
^{15}N	Nitrogen 15, a rare stable isotope of nitrogen with a fractional nuclear spin of of one half
N	Asparagine.
NCBI	National centre for biotechnology information.
Ng	nanogram (10^{-9} grams).
nM	Nanomolar.
nm	Nanometre.
Ni-affinity	Nickel-charged agarose affinity resin that uses nitrilotriacetic acid coupled to cross linked 6% agarose resin.
NHS	N-hydroxysuccinimide
NOE	Nuclear overhauser effect.
NOESY	Nuclear overhauser enhancement spectroscopy.
NMR	Nuclear magnetic resonance.
N-terminus	NH_2 or Amino terminus.
OD	Optical density.
OM	Outer membrane.
OmpA	Outer membrane protein A precursor
OMP's	Outer membrane proteins.
P. aeruginosa	*Pseudomonas aeruginosa*.
Pal	Peptidoglycan associated lipoprotein.
PBP	Penicillin binding proteins.
PBP1b	Bifunctional transglycosylase, also known as peptidoglycan glycosyltransferase.
PCR	Polymerase chain reaction.
PDB	Protein data bank.
PEG	Polyethylene glycol.

PG	Peptidoglycan.
pH	-log of the concentration of H^+ in solution.
Phe	Phenylalanine.
PMF	Proton motive force.
PMSF	Phenylmethyl sulphonyl fluoride.
PPI's	Protein-protein interactions.
P or Pro	Proline.
PSI BLAST	Position specific iterated BLAST (Basic local alignment search
Q	Glutamine.
R	Arginine.
RDC	Residual dipolar coupling.
RCI	Random coil index.
Rmsd	Root mean square deviation.
RNA	Ribonucleic acid.
SDS-PAGE	Polyacrylamide gel electrophoresis performed in the presence of Sodium dodecyl sulphate
SEC-MALS	Size exclusion chromatography multi angle light scattering.
S or Ser	Serine.
SPR	Surface plasmon resonance.
TALOS	Hybrid method for predicting protein backbone torsion angles
Tat	Twin arginine translocation.
TBDT	TonB dependent transporter.
TBE	Buffer solution containing a mixture of Tris base, boric acid and EDTA
tBLASTn	Mode of operation for BLAST that aligns protein sequences to a nucleotide database translated in all six frames
TFA	Trifluoroacetic acid.
T or Thr	Threonine.
TOCSY	Total correlation spectroscopy.
TPR	Tetratricopeptide repeat.
$TolB_{22-34}$	TolB protein.
Tol-Pal system/complex	A complex of proteins found in *Escherichia Coli* that comprises YbgC, TolQ, TolA, TolR, TolB, Pal and YbgF/CpoB proteins.
Ton system	A complex of proteins comprising TonB-ExbB-ExbD.
Tris	2-Amino-2-(hydroxymethyl)propane-1,3-diol. $(HOCH_2)_3NH_2$.
TROSY	Transverse relaxation optimized spectroscopy.
Trp	Tryptophan.
Tyr	Tyrosine.
UV	Ultraviolet.
UV-Vis	UV visible.
V or Val	Valine.
V. cholerae	*Vibrio cholerae*.
W	Tryptophan.
WD 40 repeat	Short structural motif of approximately 40 amino acids, often terminating in a Tryptophan-Aspartic acid (WD) dipeptide.
w/v	Weight per volume.
Y	Tyrosine.
ZZ-exchange spectroscopy	NMR experiment used to measure the kinetics of interconversion between different molecular states.

Acknowledgements

Firstly, I would like to thank my supervisor, Prof. Colin Kleanthous for allowing me to undertake a DPhil project in his laboratory with his supervision. His advice, mentorship, critique, guidance and encouragement throughout the project was strongly appreciated. I would like to thank my thesis advisory panel, Prof. Matthew Higgins and Prof. Simon Newstead, for their advice and suggestions during the progress of my project. I would further like to thank the National Research Council of Canada (NSERC) for their support in the form of an Alexander Graham Bell Canada Graduate Scholarship, as well as the University of Oxford, Department of Biochemistry and the Biotechnology and Biological Sciences Research Council (BBSRC) for their financial support for my project. I would also like to thank the governing body of Linacre College Oxford for their award of an EPA Cephalosporin Write-up Bursary, as well as for election to the Thomas Linacre Studentship and the accompanying financial support.

I would like to extend my thanks to all the members of the Kleanthous group who I had the fortune of working alongside during my project for the support, camaraderie and countless insightful discussions. In particular I want to express my deepest thanks to Dr. Karthik Rajasekar, a colleague and friend with whom I experienced both the highs and lows of this project. Dr. Rajasekar's mentorship throught the course of my DPhil was integral to the success of this study, and I am deeply grateful for his advice, patience and understanding in the progress of teaching me. I would like to thank Dr. Katarina Jansen and Dr. Nick Housden, both for their thorough assistance and advice throughout my project, as well as their outstanding patience in trawling through some awful early drafts. I would like to thank Dr. Renata Kaminska for her support with the molecular biology in this study, as well as Dr. Justyna Wojdyla and Dr. Amar Joshi for their support in efforts to crystallize the TolA-TolB complex. I would like to thank my bench partner Paul White, for his enthusiasm for science, creative whistling and friendship throughout this project and as a travel buddy. I would like to thank Connor Sharp for bioinformatics insight, expertise and the occasional friendly 'over the wall' chat. My thanks to Dr. Patrice Rassam for his friendship, advice and persistence with our fluorescence microscopy attempts.

The NMR within this project would not have been possible without the expert guidance of Prof. Christina Redfield, who has been an excellent mentor, colleague and friend. My thanks as well to Dr. Grigorios Papadakos for late night discussions and insightful critique. I wish to thank

Dr. David Staunton and Dr. Steven Johnson for their assistance with some of the biophysical methods presented in this study. My thanks to Prof. Martin Maiden and Dr. Keith Jolley for their early assistance with efforts to catalogue the *tol-pal* operon in Gram-negative bacteria.

I am indebted to Dr. Jo Claridge for his assistance and mentorship in the development of heterolabelled TolB peptide, as well as general NMR queries. I would like to thank Prof. Elspeth Garman for welcoming me into her discussion group without hesitation, as well as her enduring encouragement and support.

I would like to thank my friends, both in the Department of Biochemistry and in Linacre College, for their company, as inspiration, and friendship. In particular Emmie Readman, Jolyon Faria, Lea Nussbaum and Ron Schweβinger. Thank you to the Oxford Baking Group, as well as the Oxford Board Game Gang, for the wonderful experiences and distractions throughout my project, even when they weren't necessarily at the best of times.

I would like to thank Dr. Ian Robertson, for the time spent under his guidance at the University of Alberta and providing me the inspiration and example to complete a doctorate. I am thoroughly indebted to Prof. Brian Sykes, who has been over the years an outstanding mentor and friend, and without whom I would not be where I am today.

Finally, but certainly not least, I would like to thank my family: my brother Nicholas, my parents Linda and Charles Holmes, and my girlfriend Robyn Millott. Thank you for your continuous support, for encouraging me on whichever path I chose, and for sticking with me through thick and thin across a distance of well over 6500 km and an ocean.

This thesis would not have been possible without you.

Declaration

I declare that all of the work presented in this thesis is my own original work unless explicitly stated in the text.

Peter Charles Holmes, October 2016

1. Introduction

1.1 The Gram-negative bacterial cell envelope

Gram-negative bacteria are a group of bacteria that typically resist the crystal violet stain used in the Gram staining method of bacterial differentiation. Gram-negative bacteria are spread throughout the world in environments that support life. Typical Gram-negative bacteria include the model organism *Escherichia coli*, as well as many pathogenic bacteria, such as *Pseudomonas aeruginosa, Chlamydia trachomatis* and *Yersinia Pestis*. Importantly, several classes of antibiotics target Gram-negative bacteria specifically, including aminoglycosides and carbapenems (Baron *et al*, 1996).

A key hallmark feature of Gram-negative bacteria pertains to their cell envelope, which consists of three concentric layers: an asymmetric outer membrane (OM) bilayer, a symmetric inner cytoplasmic membrane (IM) and a periplasm enclosed in the intervening space (Nikaido & Vaara, 1985; Glauert & Thornley, 1969; Silhavy *et al*, 2010). This envelope provides a wide variety of cellular functions, including energy generation, a barrier to environmental hazards, capacity for immune evasion, as well as the definition of shape and structure. The rigidity of this entire bacterial cell is maintained by a peptidoglycan layer spread throughout the periplasmic space.

The Gram-negative bacterial cell envelope also contains molecular machinery and structures with developmental and reproductive functions. During cell division processes, the cell envelope contains the necessary cytokinetic machinery to uniformly partition itself and the cell into two daughter cells, without compromising the interior machinery of the cell to the external environment. The cytokinetic machinery comprises an assembly of a heterogenous protein molecular machine called the divisome, which after maturation and recruitment of additional molecular complexes involved in the synthesis of peptidoglycan, triggers a process termed septation (Lutkenhaus *et al*, 2012). Septation involves complex coordination of synthetic molecular machinery that spans the entire Gram-negative cell envelope. A diagram of this cell envelope is depicted in Figure 1.1, which illustrates the inner and outer membranes of the cell, the structural peptidoglycan layer that permeates the periplasmic compartment and the proteins and lipoproteins characteristic of this region of the cell in all Gram-negative bacteria.

1.2 The outer membrane of Gram-negative bacteria

The outer membrane of a Gram-negative cell comprises an asymmetric bilayer, approximately 50 Å in thickness. Lipopolysaccharide (LPS) molecules are a major component of the outer membrane that are spread across the outer leaflet and are also endotoxins responsible for the toxicity of Gram-negative bacteria, especially induction of septic shock in humans during Gram-negative bacterial infection (Raetz & Whitfield, 2002). The presence and protection provided by an OM allows Gram-negative bacteria to inhabit and flourish in environments traditionally hostile to bacterial organisms.

LPS contributes greatly to the structural integrity of bacteria, also protecting the membrane from chemical attack. LPS is comprised of three components, an O-antigen, a saccharide core and Lipid A. The primary component responsible for the toxicity of Gram-negative bacteria is Lipid A, which is a glucosamine based hydrophobic anchor: a β-1,6-linked D-glucosamine disaccharide with 5-7 fatty acids protruding into the lipid outer membrane bilayer. The 5-7 β-hydroxyacyl chains anchoring Lipid A to the membrane are usually between 10-16 carbons in length. Lipid A is in particularly high abundance (~10^6) on the outer membrane (Raetz & Whitfield, 2002), and is also vital for the growth of most Gram-negative bacteria. The minimal regions of LPS that are required for a viable outer membrane (OM) comprise Lipid A and 3-deoxy-D-manno-oct-2-ulosonic acid (KDO) present in the LPS inner core (Raetz, 1996).

Gram-negative bacterial cells may grow in the absence of the outer saccharide core and O-antigen, and are viable in a laboratory environment, however these cells deal poorly with antibiotics and environmental stressors. Wild-type bacteria express a non-repeating saccharide 'core' shown in light blue in Figure 1.1, and a distal and lengthy repeating chain of sugars extending up to 40 units from the original core which is the O-antigen. The LPS leaflet is an extremely effective permeability barrier due to the additional rigidity imparted by the quasi-crystalline gel-like structure of the hydrocarbon domains, giving strong lateral interactions between LPS molecules (Nikaido, 2003). This robust arrangement provides protection from threats such as host defense factors, detergents, digestive enzymes, and antimicrobial agents. In addition, outwardly facing saccharide cores and O-antigens endow strong hydrophilicity for improved evasion of any host defenses.

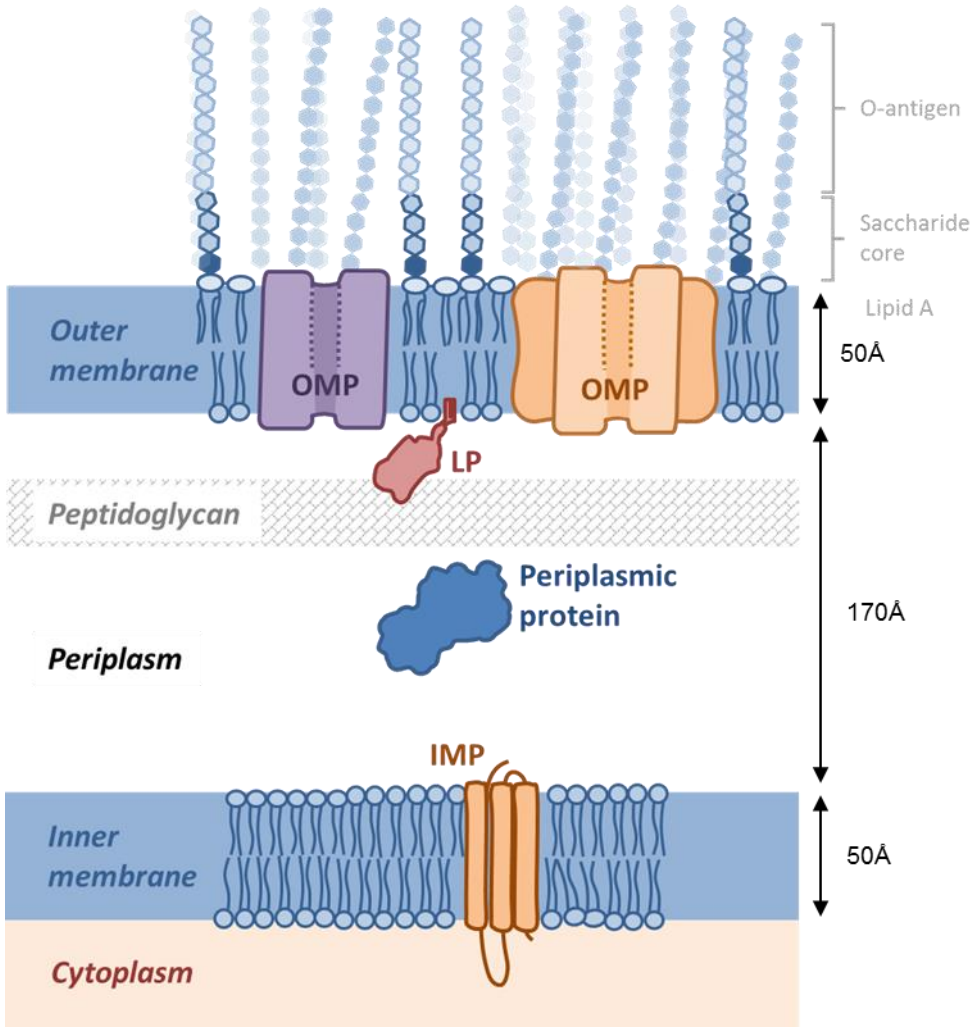


Figure 1.1. **Architecture of the trilaminar Gram-negative cell envelope.** Diagrammatic view of the cell envelope in Gram-negative bacteria. Beginning at the top of the diagram and exterior face of the cell, lipopolysaccharide (LPS) is composed of the following: O-specific polysaccharide, called the O-antigen. The saccharide core, contains an outer core in light blue, and the inner core, in dark blue. The outer core commonly contains glucosamine, galactose and N-acetylglucosamine. The inner core is composed of phosphorylated heptose sugars and ketodeoxyoctulosonate (KDO). Lipid A is indicated as the round oval 'head'. Note the asymmetry of the outer membrane, with glycerophospholipids occupying the inner leaflet and LPS occupying the outer leaflet. Embedded within this asymmetric barrier are an abundance of outer membrane proteins. On the inner leaflet of the outer membrane, lipoproteins are shown (LP). The periplasmic space, approximately 170 Å in length, contains the peptidoglycan layer as well as soluble protein. Inner membrane proteins are also indicated embedded within the symmetric inner membrane. Adapted from (Raetz & Dowhan, 1990)

OM proteins (OMPs) play key roles in effecting entry and exit of molecules past this permeability barrier. These proteins typically possess antiparallel ß-barrel structures, with transport functions such as porins and substrate-specific channels, embedded in the asymmetric OM (Sankaran & Wu, 1994; Nikaido & Vaara, 1985; Nikaido, 2003). The presence of porins makes the OM more permeable than the inner membrane to the entrance and exit of low molecular weight molecules. OMPs such as the ferrichrome-iron receptor of *E. coli* K-12 (FhuA) are multifunctional membrane receptors vital for facilitating the consumption of extracellular

nutrients that are difficult or otherwise impossible to internally synthesize. FhuA uses an energy-dependent relay involving interactions with an inner membrane tethered protein to effect internalization of ferric iron from the environment via capture of iron previously snared by secreted siderophore molecules (Faraldo-Gómez & Sansom, 2003). Other OMPs such as the outer membrane protein C (OmpC) have roles that allow for nutrient diffusion, or function as efflux pumps like the outer membrane channel TolC (Silhavy *et al*, 2010; Murakami *et al*, 2002; Koronakis *et al*, 2000; Nikaido, 2003).

On the inner leaflet of the OM are found lipoproteins (LP). These proteins are embedded into the inner leaflet of the OM via covalently linked lipids attached through their N-terminal moieties and are known to have important structural roles both in tethering the OM to the structural peptidoglycan layer in the periplasm and in OMP biogenesis. Key examples of OM inner-leaflet lipoproteins include Pal (Peptidoglycan-associated-lipoprotein), Braun's lipoprotein (an important structural LP covalently linked to peptidoglycan) and BamB-E (OMP biogenesis pathway); (Silhavy *et al*, 2010). The insertion pathway of LPS is discussed in section 1.6, and the insertion pathway of lipoproteins like Pal is discussed in section 1.8.1.

1.3 The periplasm of Gram-negative bacteria

Bounded by the OM and IM is the periplasm, a highly crowded, oxidizing environment comprising 8-16% of the total cell volume (Dartigalongue *et al*, 2000; Koch, 1998; Wielink & Duine, 1990; Graham *et al*, 1991a). The periplasm is estimated to be 110-250 Å in width, consistent with the 150-170 Å length of trans-periplasmic complexes such as the AcrB/TolC and flagellar motor systems (Graham *et al*, 1991b; Graham & Beveridge, 1991; Murakami *et al*, 2002). This region of the cell is densely packed with protein, and along with periplasmic-facing proteins embedded in the IM and OM contains much of the cell machinery and support structure for the cell.

Compartmentalization of the periplasm away from the cytoplasm allows Gram-negative bacteria to sequester away potentially harmful enzymatic activity, such as RNAse activity, from causing collateral damage in the cytoplasm (Ehrmann, 2007; Silhavy *et al*, 2010). Additionally, compartmentalization results in an electrochemical gradient between the periplasm and the cytoplasm. Energy-generating protein complexes harness this electrochemical gradient across the inner membrane to power the cell machinery (Silhavy *et al*, 2010). The outer bacterial cell membrane prevents proteins from diffusing away from the periplasmic surface of the inner

membrane, so compartmentalization is achieved due to the impermeability of the IM and OM. Key roles for proteins in the periplasm include hydrolytic activities, substrate transport and chemoreceptors involved in the chemotactic response pathway. Proteins destined for the periplasm are expressed in an immature and partially disordered state in the cytoplasm and are targeted to IM import machinery such as the Tat or Sec IM translocation pathways, which then deliver them to the periplasmic space (Driessen *et al*, 2001).

The peptidoglycan layer also permeates the periplasm. Alternating N-acetyl-muramic acid (NAM/MurNAc) and N-acetyl-glucosamine (NAG/GlcNAc) glycan chains weave into a dense, covalently cross-linked mesh-like structure that is stapled to the outer membrane via lipoproteins. Similar to periplasmic proteins, the structural precursors for this assembly are first synthesized in the cytoplasm and must be translocated into the periplasmic compartment. Lipid carrier molecules chaperone nucleotide sugar-linked precursors to the cytoplasmic face of the IM, they are then translocated into the periplasm where they can be assembled into growing peptidoglycan by glycosyl transfer and transpeptidation reactions (Coyette & Van Der Ende, 2008). The peptidoglycan layer then functions as a vital scaffold that maintains cell shape and envelope integrity in the face of fluctuations in osmotic pressures and internal turgor pressure (Coyette & Van Der Ende, 2008). In periods of osmotic stress the IM may collapse against the peptidoglycan layer scaffold and transfers any stress, preventing an appreciable osmotic pressure differential from occurring (Koch, 1998).

The molecular systems governing the synthesis of the vital peptidoglycan layer in the cell envelope therefore represent important targets for antibiotics. β-lactam antibiotics such as penicillins and cephalosporins (which account for over half the known antibiotics produced worldwide);(Madigan, 2012) strongly inhibit the catalysis of the transpeptidation reactions required in bacterial cell wall synthesis. The transpeptidases, known as Penicillin-binding-proteins (PBP), bind tightly to penicillins, consistent with the name, and further stimulate the release of autolysins that degrade the peptidoglycan wall, allowing normal osmotic pressures to cause a failure mode in the cell envelope (Madigan, 2012). Despite this, the OM of Gram-negative bacteria is relatively impermeable to these molecules, thereby restricting the potency of antibiotics that work via this method. Other antibiotics such as moenomycins work via different targets that are peptidoglycan glycosyltransferases, but historically these antibiotics have shown to exhibit suboptimal pharmacokinetic properties (Ostash & Walker, 2010).

1.4 The inner membrane of Gram-negative bacteria

The inner membrane (IM) bilayer is the innermost concentric layer of the bacterial cell envelope, is approximately 50 Å thick and comprises a true phospholipid bilayer (Raetz & Dowhan, 1990). The IM is composed primarily of 70-80% phosphatidylethanolamine, 15-20% phosphatidylglycerol and 5% cardiolipin (Raetz, 1978) along with a variety of transmembrane proteins and lipoproteins inserted into the periplasmic-facing leaflet via an N-terminal lipid moiety (Narita *et al*, 2004; Daley *et al*, 2005; Raetz & Dowhan, 1990). Transmembrane proteins are proteins embedded inside the membrane via a series of one or more trans-membrane helices. These helices traverse the membrane and are typically 50 Å in length (Narita *et al*, 2004). Bacterial cells lack intracellular organelles, so membrane-associated processes that would be performed in eukaryotic structures (such as the endoplasmic reticulum or mitochondria) are instead typically performed in bacteria in the IM. There are exceptions to this, for example metabolic activity within the cytoplasm causes disproportional confinement of cytoplasmic components, allowing for pseudo-compartmentalization (Parry *et al*, 2014). Further, in cyanobacteria carboxysomes are inclusion bodies involved in autotrophic metabolism (Iancu *et al*, 2010; Whitehead *et al*, 2014; Cannon *et al*, 2001). In Gram-negative bacteria lipid and cell envelope biosynthesis, energy production, protein secretion and transport are all typically conducted in the microenvironment of the IM (Silhavy *et al*, 2010).

1.5 Bacterial membrane translocation systems

The Gram-negative cell envelope represents a considerable barrier for the passage of large and small molecules. The bacterial cell therefore possesses a number of translocation and insertion molecular machineries that allow for targeting of specific proteins and molecules to desired regions of the cell envelope. The presence of IM import machinery and OM insertion machinery is vital for cell function and has been well described in a number of key systems.

1.5.1 Sec and Tat secretory pathways

All components in the bacterial cell envelope are synthesized in the cytoplasm or at the cytoplasmic surface of the IM. This requires a translocation system capable of efficiently shuttling important cell components across the IM. The Sec pathway provides the primary IM translocation pathway for Gram-negative proteins. Proteins are expressed in the cytoplasm by ribosomes with a cleavable N-terminal signal sequence (Driessen *et al*, 2001). Protein precursors are targeted

immediately to the IM via either the SecB molecular chaperone or via co-translational targeting by the signal recognition particle (SRP) and its receptor (Driessen *et al*, 2001). Completely unfolded protein is shuttled through the IM SecYEG translocase driven via ATP hydrolytic activity (ATPase) of SecA. Signal peptidase on the periplasmic face of the Sec machinery cleaves the signal peptide after it has begun extrusion through the translocase. The proteolytic site is presented as a loop for cleavage, with the N-terminal signal peptide being threaded back towards the cytoplasm(Beckwith, 2013; Pugsley & Schwartz, 1985; Park *et al*, 2013). After cleavage, mature protein continues to be funneled in unfolded form from N-terminus to C-terminus through the translocase until extrusion, release and refolding occurs on the periplasmic face. Note that OMPs undergo further processing, as detailed in the next section.

The second method that Gram-negative cells use to shuttle proteins across the IM is via the Twin arginine translocation (Tat) secretion pathway. Tat involves a fundamentally different mechanism, whereby proteins are first translocated to the periplasm in a folded configuration. For example, *E. coli* utilises the Tat pathway to deliver proteins which must first receive prosthetic groups in the cytoplasm. Thermophiles and other bacteria that live in harsh environments extensively use the Tat system since the protected environment of the cytoplasm likely provides improved conditions for protein folding compared with anything external to the IM (Silhavy *et al*, 2010). Tat mechanisms are found in both prokaryotic and eukaryotic organisms (Berks *et al*, 2000, 2005).

1.5.2 Bam proteins

Proteins that are destined for insertion in the OM, such as OMPs, are extruded into the periplasm by Sec secretory pathways, but are then immediately protected from degradative enzymes by molecular chaperones which assist transit from the IM to the inner leaflet of the OM. SurA, Skp and DegP all code for proteins that act in parallel to shuttle OMPs to the OM. The majority of OMP assembly pathways show a preference for SurA, whilst a minority fraction show no preference (Silhavy *et al*, 2010; Sklar *et al*, 2007). No OMP has been discovered that preferentially uses the Skp and DegP pathways over SurA. It has therefore been suggested that these redundant pathways are important in the capture of errant OMPs that may have escaped SurA under periods of cell stress (Sklar *et al*, 2007). These periplasmic chaperones deliver OMPs to an assembly site in the OM called the Bam complex (Hagan *et al*, 2011; Silhavy *et al*, 2010). Bam is a highly conserved essential β-barrel protein embedded across the outer membrane that

coordinates the insertion of OMPs into the OM (Kim *et al*, 2012). Bam is composed of BamA (the β-barrel structure embedded in the OM), four lipoproteins (BamB,C,D and E) and a set of large amino-terminal polypeptide-transport-associated (POTRA) domains. These domains are thought to provide a scaffold for the initial β-strand complementation during assembly of the OMP β-barrel, and improve the efficiency of OMP biogenesis inside the Bam complex chamber (Han *et al*, 2016; Hagan *et al*, 2011).

1.6 Trans-envelope systems in bacterial cells

The Sec and Tat translocation machinery are two examples of protein systems that are embedded primarily within the IM. The Bam complex, with which Sec coordinates function, is embedded in the OM. Although molecular chaperones shuttle substrates from the IM to the OM, there was thought to be no direct physical link between the two systems (Silhavy *et al*, 2010), however recently this has been challenged by the suggestion of a supercomplex involving both Bam and Sec proteins spanning the inner and outer membranes (Wang *et al*, 2016). Some protein machinery in the Gram-negative cell is however thought to span the entire periplasmic width, from IM to OM. These trans-envelope systems in Gram-negative bacteria are energized by ATP or an electrochemical gradient and perform a variety of functions in the cell, some of which are described in Figure 1.2. Trans-envelope systems typically fall into the category of an export pathway, an import pathway, or large supramolecular assemblies responsible for cell motility.

1.6.1 Type I, II and III trans-envelope secretion systems

Figure 1.2 depicts the trans-envelope protein complexes in Gram-negative bacteria that use the inner bacterial cell membrane as an energy source for a wide variety of functions. These functions include xenobiotic efflux, secretion of virulence factors, LPS secretion, OM stabilization, nutrient import and cell motility. Secretion systems in Gram-negative bacteria are particularly important for bacterial pathogenesis. In order to attack and defend against host cells, bacteria need a method of recognizing host target cells and delivering virulence factors that modify the metabolism of their targets.

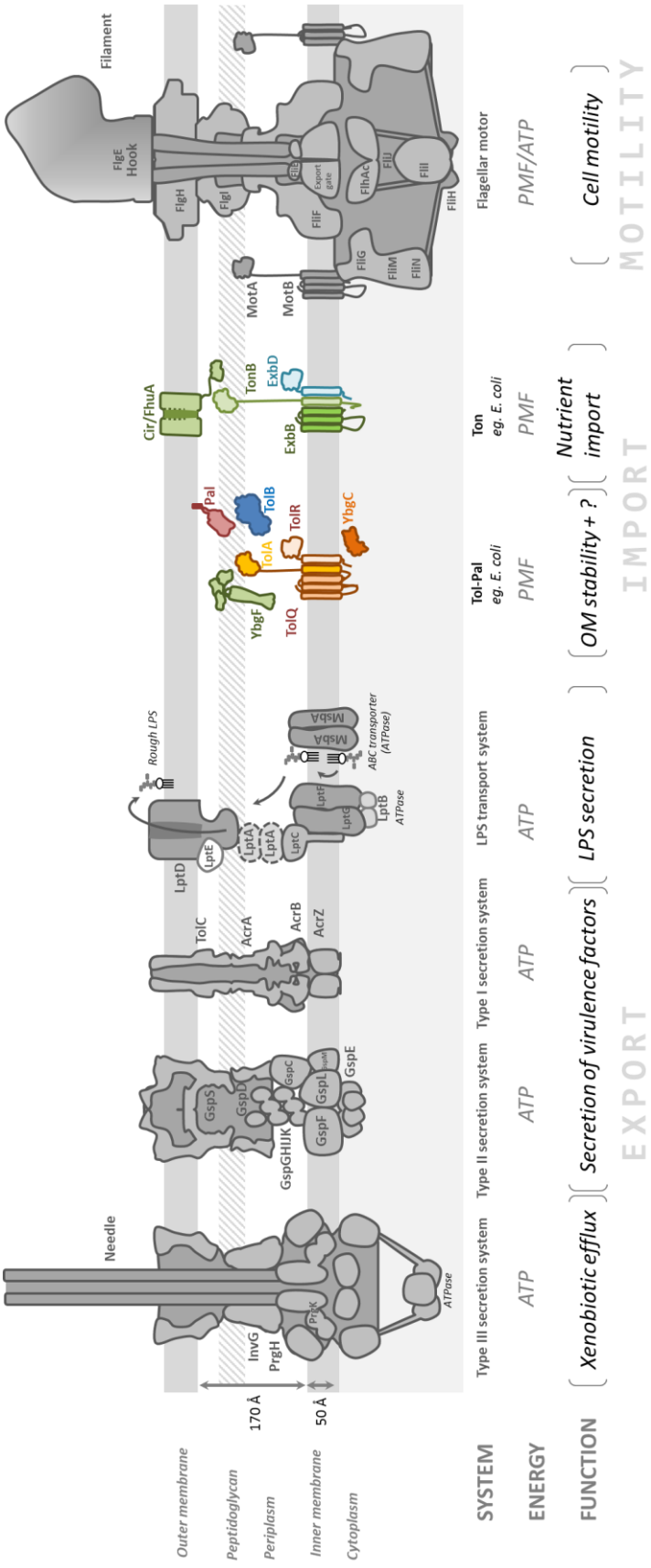


Figure 1.2 Trans-envelope protein complexes in Gram-negative bacteria use the inner membrane as an energy source for a wide variety of functions. Adapted from Silhavy et al. 2010, Nikaido 2003, Parsons et al. 2008, Kleanthous 2010, Ruiz et al. 2009.

There are six types of bacterial secretion systems in Gram-negative bacteria, each with its own unique mode of action. All secretion systems use either ATPase activity or power conversion from the electrochemical gradient across the inner membrane (proton motive force, also termed PMF) to power secretion of substrates from the cytoplasm into the extracellular medium. Type I secretion systems are characterized by possession of an OM protein channel called TolC. They are closely related to resistance-nodulation-division pumps which translocate anitibiotics and small compounds from the cytosol to the extracellular fluid using the PMF. Type I systems use ATP hydrolysis at the IM with an ATP-binding cassette to trigger transfer of the substrate to the extracellular medium (Costa *et al*, 2015).

Type II and type III systems similarly use ATPase activity to power transfer of their substrates to the extracellular space. A principal observation is that all energy-related processes occurring in the periplasm and at the OM of the bacterial cell are powered at the IM. Specifically, this occurs by virtue of the energy captured by protein complex machinery at the IM during the process of ATP hydrolysis or transduction of the proton motive force. The bacterial cell periplasm is devoid of ATP and the OM is similarly devoid of energy-generating processes (Neidhardt & Curtiss, 1996).

1.6.2 LPS secretion

LPS is required for bacterial cell viability in all Gram-negative bacteria with the exception of a few species including *Neisseria meningitides, Moraxella catarrhalis,* and *Acinetobacter baumannii* (van der Ley & Steeghs, 2003; Zhang *et al*, 2013). In all other cases, Lipid A and KDo are the minimal portion of LPS that must be retained for a viable cell. A common laboratory strain which omits the O-antigen from LPS in this way is *E.coli* K12, which has been extensively mutated to retain normal cell function whilst possessing improved mutagenic qualities yet remain non-pathogenic (also termed a “rough” mutant). Recently however, experiments aimed at returning pathogenicity to lab strain K12 by creating a “smooth” wild type strain suggested that O-antigen density is a critical factor in a number of phenotypes, in addition to reducing cell susceptibility to outside threat from antimicrobial bacteriocins and bacteriophage (Browning *et al*, 2013). Bacteriocins and bacteriophages target OM proteins to gain entry into the cell; bacteriocin entry is discussed in section 1.6.6. The LPS transport system is depicted in Figure 1.2 and (in a manner similar to that of secretion systems) utilises an ATP-binding cassette at the IM to power the transport process. LPS precursors are synthesised on the cytosolic leaflet of the IM and then

flipped to the periplasmic leaflet of the IM by the lipid flippase activity of the MsbA multi-drug resistant ABC transporter. In the periplasm, O-antigen polysaccharide is ligated to the LPS core lipid A, catalysed by the enzyme ligase WaaL. A secretory pathway utilising lipopolysaccharide transport (Lpt) proteins then funnels LPS through an OM β-barrel for deposition in the outer leaflet of the OM (Silhavy *et al*, 2010).

1.6.3 Nutrient import

Gram-negative bacteria have evolved systems for acquisition of rare nutrients for the cell. The Ton system, shown in Figure 1.2, is a prominent example and is used for the active transport of iron siderophores, hemin, carbohydrates, vitamin B12 and transition metals (Liao *et al*, 2013; Postle & Kadner, 2003). The Ton system comprises a heterotetrameric protein complex containing TonB, ExbB, ExbD and a variety of OM porins such as OmpF and OmpC. TonB comprises three domains; an N-terminal transmembrane helix which acts as an IM anchor, a periplasm spanning region and a globular C-terminal domain (Karlsson *et al*, 1993; Higgs *et al*, 2002). Energy is funneled to TonB via two partner proteins, ExbB and ExbD.

As depicted in Figure 1.2, TonB binds selectively to the intrinsically disordered region of outer membrane TonB dependent transporters (TBDT) such as Cir or FhuA to effect entry of siderophore-iron or Vitamin B12 (Anwari *et al*, 2010) (Noinaj *et al*, 2010). This process is highly selective and consequently multiple TBDTs exist for a range of siderophore-iron complexes, thus allowing for each entry process to be stereoselective wherein only one type of substrate can be transported by a particular TBDT (Schalk *et al*, 2004; Adams *et al*, 2006). The Ton complex therefore appears to effect the entry of its extracellular substrate into the cell via a distinct but transient interaction between TonB and its receptor.

1.6.4 Outer membrane stability

The Tol-Pal protein complex import system is ancestrally related to the Ton nutrient import system. Tol-Pal comprises three proteins that interact with each other within the IM, termed TolQ, TolR and TolA. The Tol-Pal system also contains two proteins present in the periplasm termed TolB and YgbF/CpoB, as well as an OM-linked lipoprotein, Pal, and a cytosolic protein YbgC of unknown function. The Tol-Pal system, like the Ton system, appears to couple proton motive force across the IM with its physiological role, which is likely stabilization of the OM. However, the molecular mechanisms underlying the different interactions between the proteins

involved in this system are poorly understood and remain to be determined. Tol-Pal is discussed in detail in section 1.7.

1.6.5 Locomotion

The bacterial flagellum represents a third type of trans-envelope protein complex machine. This complex begins as a type III secretion apparatus, but then undergoes transformation via a morphogenetic pathway, involving the addition of a complex series of molecular substructures that result in the creation of a full length flagellum (Macnab, 2003). Although the molecular architecture of the flagellar motor is extremely complex (see Figure 1.2), of particular interest are the roles played by MotA and MotB proteins. These proteins are known to be involved in the transduction of energy from PMF. MotA also performs the role of a peptidoglycan-binding stator. MotA and MotB are orthologous to the ExbD and ExbB proteins in the Ton complex and the TolQ and TolR proteins in the Tol-Pal complex (Cascales *et al*, 2001).

1.6.6 OM transport functions may be subverted

In 1925, the seminal observation of “a remarkable example of antagonism between two strains of *Escherichia coli*” prompted the hypothesis that bacteria kill each other while competing for resources (Gratia, 1925). Later discoveries revealed that *E. coli* and other Gram-negative bacteria secrete antibacterial protein toxins called colicins (a type of bacteriocin), which deliver a cytotoxic domain into cells of nearby competing strains (Kleanthous, 2010). In the late 1960s several types of colicin tolerant gene mutants were observed to adsorb colicin, (Zwaig & Luria, 1969), suggesting that the toxin mechanism relied on breaching the OM and re-purposing of certain host gene products inside the periplasmic compartment. These colicin-tolerant genes were dubbed *tol* genes named after the tolerance their mutants conferred (Webster, 1991; Bourdineaud *et al*, 1989; Sun & Webster, 1987). Mapping of the site that produced the *tol* phenotype revealed the responsible mutations were at 16.8 minutes on the map of the *E.coli* K-12 chromosome (Bernstein *et al*, 1972), and that the flanking genes on the same chromosomal locus produced similar phenotypes. These genes corresponded to a set of genes now known to comprise the Tol-Pal complex.

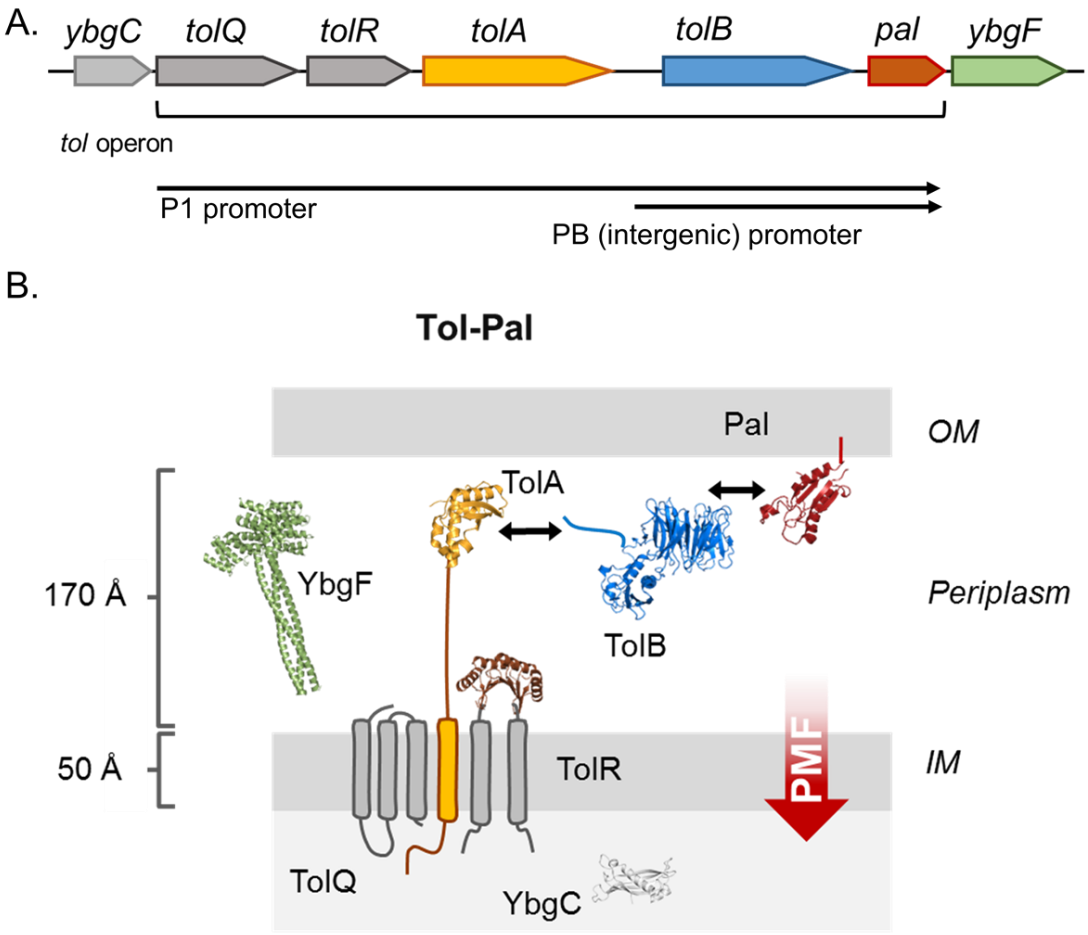


Figure 1.3 The organisation of the Tol-Pal complex and the Tol operon. (A) The Tol-Pal gene cluster is arranged in a continuous region of the genome. The region contains two adjacent operons with a constitutive promoter P1 and an intergenic (iron-regulated) promoter PB (Lafontaine & Sokol, 1998). This results in *ybgC-tolQRA* and *tolB-pal-ybgF* being transcribed independently (Vianney *et al*, 1996; Llamas *et al*, 2000). (B) The structures currently available for Tol-Pal proteins are shown, PDB codes: YbgC 2PZH(Angelini *et al*, 2008), TolA 3QDP(Li *et al*, 2012), TolR 2JWK (Parsons *et al*, 2008), TolB 1CRZ (Abergel *et al*, 1999), Pal 4PWT (Maltseva *et al*), YbgF/CpoB 2WZ7 (Krachler *et al*, 2010a), TPR domain 2XEV (Krachler *et al*, 2010a). The allosteric regulation of TolB by Pal has been reported, indicated by a black arrow (Bonsor *et al*, 2009). The structure of the TolB-TolA interaction is reported for the first time in chapter 4.

Colicins that use the Ton system for entry were later found, and colicins are now classified in groups based on the translocation system that they hijack for entry into the cell. Group A colicins parasitise the Tol-Pal system for cellular entry, whereas Group B colicins parasitise the Ton system (Kleanthous, 2010). Group A colicins include colicin A, colicin E9 and colicin K. The colicins recruit TolA and/or TolB proteins from the Tol-Pal complex for function. In order to deliver a cytotoxic payload, colicins must first attach to a cell, via interaction with an OMP, then locate a power source to drive their own activation, and translocate across the periplasm to destroy the cell from within, via either pore-forming or nuclease activity (Cascales *et al*, 2007). Protein systems such as Tol and Ton, which are already linked to the PMF and span the periplasm represent ideal targets for colicin parasitisation. These systems are also targeted by

bacteriophages using a similar mechanism involving recruitment of PMF-linked protein complexes to drive entry into the cell (Jakob *et al*, 2012).

1.7 The *tol* gene cluster

The Tol gene cluster typically comprises seven genes*: ybgC, tolQ, tolR, tolA, tolB, pal* and *ybgF*/*cpoB*, in an operon pair spanning 6 kilobases, shown in Figure 1.3. In *E.coli,* transcription occurs from two promoters, one of which is upstream of the operon and the other an intergenic region between the *tolA* and *tolB genes.* (Vianney *et al*, 1996; Cascales & Bernadac, 2002; Dennis & Lafontaine, 1996). Mutations to *tol* genes produce *tol* phenotypes, with hypersensitivities to antibiotics and detergents that decrease OM stability and impair membrane integrity (Bernstein *et al*, 1972; Paterson *et al*, 2009). The *tol* phenotype results in the release of periplasmic proteins (Lazdunski *et al*, 1998), production of small OM vesicles (Bernadac *et al*, 1998) and defective O-antigen polymerization (Vinés & Marolda, 2005). *tolA* mutations impair cell division and lead to mucoid phenotypes (Meury & Devilliers, 1999). These phenotypes, that represent OM instability, are conserved in *tol* mutants from a range of Gram-negative bacteria (Llamas *et al*, 2000; Gaspar *et al*, 2000), suggesting Tol-Pal may have a conserved role in membrane stabilization. It is not yet clear whether Tol-Pal stabilizes the OM directly, or if OM stability represents a downstream consequence of Tol-Pal function. The Tol-Pal gene cluster itself makes up a continuous region beginning at 17.3 minutes on the +strand of *E.coli* chromosome (Vianney *et al*, 1996). The organization of this cluster appears to be relatively well conserved amongst a small survey of representative bacterial species, with the exception of a select number of intracellular parasites (Sturgis, 2001). Within the same cluster are however two proteins which do not follow the *tol* naming convention: cytoplasmic YbgC and periplasmic YbgF/CpoB (Muller & Webster, 1997). Deletion of these proteins does not produce a classic *tol* phenotype.

As shown in Figure 1.3, the Tol-Pal cluster is transcribed as two distinct operons. The first transcript transcribes YbgC, TolQ, TolR, and TolA. The second transcript transcribes TolB, Pal and YbgF/CpoB. TolR expression is both tightly regulated and dependent upon prior successful translation of TolQ (Vianney *et al*, 1996) In *E. coli*, expression of TolQRA is induced by the RcsC sensor protein but can also be temperature dependent (Clavel *et al*, 1996). The essentiality of the Tol-Pal gene cluster is evidenced by homologues of Tol-Pal proteins that are found across Gram-negative bacteria, including *B. abortus* (Tibor *et al*, 1994), *P. aeruginosa*

(Dennis *et al*, 1996) and others. Conservation of the Tol-Pal operon is discussed further in section 3.2.1. Mutational analyses in species other than *E. coli* and *V. cholerae* has proven challenging, as *pal* mutants in other organisms are nearly lethal, for example in *P. putida* resulting in highly fragile outer membranes (Llamas *et al*, 2003, 2000).

1.8 Proteins in Tol-Pal

Proteins derived from the *tol* gene cluster are clearly segregated into different parts of the bacterial cell envelope. For example, YbgC is expressed and remains in the cytoplasmic compartment, TolQ TolR and TolA are embedded in the inner membrane via their transmembrane helices (Sturgis, 2001). Immature TolB and CpoB are translocated by the Sec machinery to the periplasm where, after signal sequence cleavage, they become active. Pal is embedded in the inner leaflet of the OM via a lipid moiety. The arrangement and distribution of Tol-Pal proteins can be seen in Figure 1.3. The three inner membrane proteins TolQ, TolR and TolA interact through their transmembrane helices with a respective stoichiometry that has been predicted to be 4-6:2:1, however recent cysteine-scanning analyses of TolQ suggest that there may not be one static stoichiometry but instead highly dynamic complexes involving periodic rearrangments (Zhang *et al*, 2011).

1.8.1 Pal

Peptidoglycan-associated lipoprotein (Pal) is a 13 kDa protein found on the inner leaflet of the OM, with an N-terminal 1,2-palmitoyl-glycerol moiety allowing for membrane inclusion (Mizuno, 1979; Bouveret *et al*, 1999). The structure of Pal has been determined from *E. coli* and *Y.pestis* as well as for its interaction with the periplasmic protein TolB (Bonsor *et al*, 2007; Abergel *et al*, 2001). Pal interacts with the peptidoglycan layer both covalently and non-covalently and in doing so contributes to the overall structural and functional integrity of the OM. Interaction with the peptidoglycan layer via a conserved alpha helical motif allows Pal to function as an OM stabilizer (Koebnik, 1995; Bouveret *et al*, 1999). A similar motif is found also in OmpA and MotB OM proteins.

The Pal-peptidoglycan interaction has been demonstrated by NMR spectroscopy (Parsons *et al*, 2006) which showed that Pal sequesters an NAM precursor molecule *m*-DAP, in a small binding pocket. Pal has also been suggested to bind to TolA via its C-terminal domain, thereby linking the OM to proton motive force (Cascales & Lloubès, 2003). However, a more

recent explanation for this observation was that Pal may interact via an indirect mechanism, explored in more detail in section 1.6.15. Another role for Pal has been suggested in the translocation of surface O-antigens (Godlewska *et al*, 2009).

Recent work with alpha-proteobacteria has shown that Pal is involved in an interaction with the Bam complex. A mechanism underlying this interaction has been proposed wherein Pal acts as a bridge to dock Bam tightly to the peptidoglycan layer and thereby potentially improve substrate transfer from the Sec apparatus to the OM (Anwari *et al*, 2010). In order to reach the OM, Pal must first be expressed in an immature form. After cleavage by signal peptidase II and subsequent acylation, Pal is localized to OM by the LolABCDE protein complex system (Narita *et al*, 2004; Hizukuri *et al*, 2009; Yokota *et al*, 1999). The Lol system works by utilizing ATP hydrolysis at an ABC transporter (termed LolCDE) to extract Pal (and other lipoproteins) from the IM where they have first been inserted, to a soluble periplasmic carrier LolA. In turn, LolA passes Pal to an OM assembly lipoprotein LolB, resulting in insertion of Pal into the inner leaflet of the OM (Silhavy *et al*, 2010). In *E. coli* this targeting is facilitated by a key serine residue at position 2 in the N-terminus of Pal (Gennity, 1991). This signal serine is also conserved in *P. aeruginosa.*

1.8.2 TolB

TolB is a soluble 44 kDa periplasmic protein with a two-domain structure that comprises an N-terminal domain with two α-helices, an antiparallel-strand ß-sheet and a long unstructured loop, and a C-terminal domain with a six-bladed ß ‘propeller’ (Abergel *et al*, 1999). Biochemical data for *E. coli* TolB have been reported wherein TolB appears to interact with Porins, Pal and LPP (major lipoprotein);(Abergel *et al*, 2001; Clavel *et al*, 1998; Bouveret *et al*, 1995). Reported structures of TolB in complex with a binding partner are *E. coli* TolB with colicin E9 domain, TolB with colicin A translocation domain, and TolB with Pal, (Loftus *et al*, 2006; Zhang *et al*, 2010; Bonsor *et al*, 2009). Structural information for *Y.pestis* TolB in complex with Pal have recently been deposited in the Protein Data Bank by Maltseva *et al.* awaiting publication. The tertiary structure of TolB alone and the secondary structure of the TolB N-domain are shown in figure 1.4.

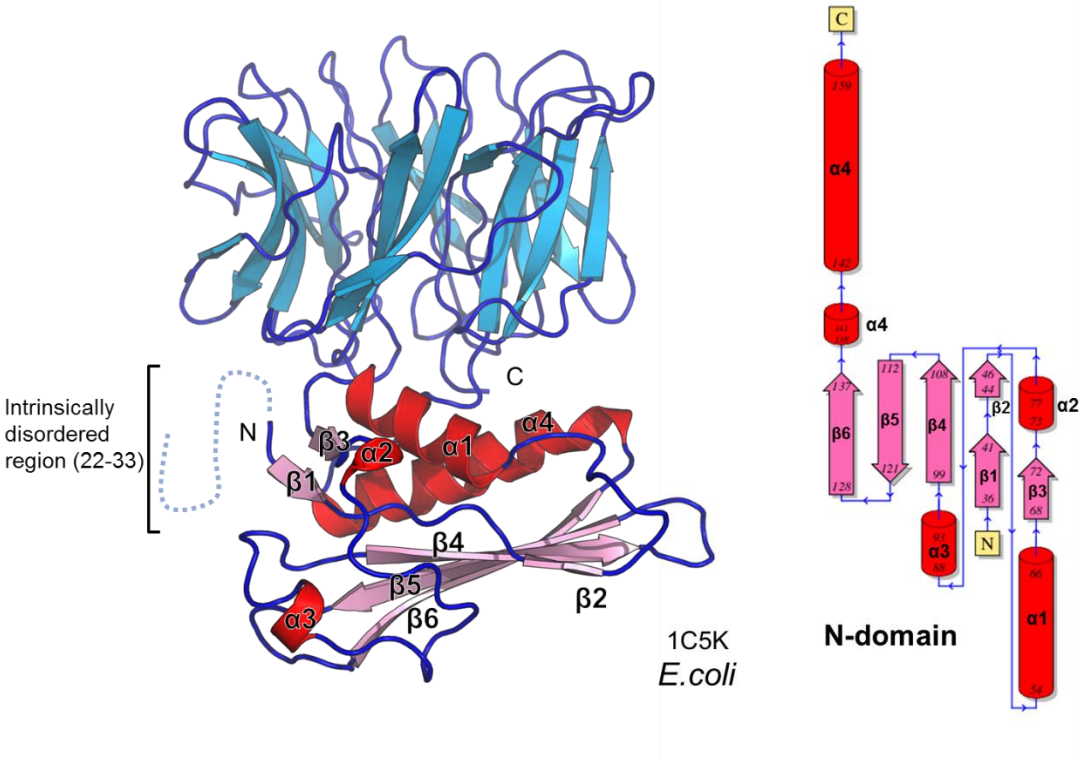


Figure 1.4 Cartoon representation of the structure and organisation of TolB. Shown in blue is the beta-propeller domain, and in red the N-terminal domain, with a characteristic intrinsically disordered region at residues 22-33. Shown is the mature Sec-cleaved form of the protein, missing the N-terminal Sec signal peptide. (Carr *et al*, 2000; Abergel *et al*, 1999). Figures were drawn using PyMol and PDBSum (DeLano, 2002; Laskowski *et al*, 1997; Laskowski, 2009).

Homodimerisation of TolB and the periplasmic domain of TolA have also been suggested (Sturgis, 2001). Based on its known interactions, TolB may have variable roles within the periplasm. For instance, TolB interacts with Pal near the OM (Abergel *et al*, 1999) and the residues critical for this interaction lie just inside the ß-propeller, in the C-terminal domain (Ray *et al*, 2000). If TolB contacts structures of the inner membrane as well, it may indicate that TolB is the keystone in suggested transperiplasmic contact sites, which have been previously attributed to TolA-Pal interactions (Bouveret *et al*, 1995). TolB has therefore been suggested as a binding partner for TolA.

The endogenous interaction between TolB and Pal is subverted by the bacteriocins, protein antibacterial molecules like colicin E9 and colicin A. These invasive proteins hijack TolB and the Tol-Pal system in order to translocate a cytotoxic endonuclease (DNase) domain across the cell envelope to randomly degrade the bacterial genome in the nucleoid. Recruitment of an OM translocon by colicin E9 has been recently reported (Housden *et al*, 2013). In this process an intrinsically disordered region of colicin E9 that can out-compete the endogenous Pal interaction for recruitment of TolB (Papadakos *et al*, 2012) stimulates an order-disorder transition at the N-

terminus of TolB (Bonsor *et al*, 2009). The intrinsically disordered N-terminus of TolB is indicated at residues 22-33 in *E. coli* TolB as depicted in Figure 1.4. By binding to TolB, the colicin gains access to the proton motive force to then trigger cell entry via its interaction with TolA.

1.8.3 TolA

TolA is a 44 kDa protein that comprises three structurally distinct domains (TolAI, II and III) separated by poly-glycine linkers (Levengood *et al*, 1991). The *E. coli* TolA amino-terminal domain I (TolAI) consists of 34 amino acid residues with a 21 residue hydrophobic and inner membrane spanning anchor (Levengood *et al*, 1991), TolAI from *P. aeruginosa* is 37 amino acid residues in length.

E. coli TolAII is approximately 280 amino acid residues in length, containing two poly-glycine repeats and nearly 200 consecutive amino acid residues that contain assorted Lys-Ala-Ala-Ala-(Glu/Asp) repeats. TolAII is thought to span the periplasm (Walburger *et al*, 2002) which would in turn allow the third domain, TolAIII, to be involved in interactions at the bacterial cell outer membrane. If we assume that regions of the TolAII domain are disordered, given a periplasmic width of approximately 170 Å, this model is conceivable. However, circular dichroism experiments and molecular modelling suggest that TolAII may be almost entirely helical and possibly exist in a twisted coiled coil arrangement (Derouiche *et al*, 1999). No structural data are known for TolAII or its binding partners, however TolAII is reported to modulate the oligomeric status of YbgF/CpoB, a protein known to be involved in the coordination of peptidoglycan synthesis(Krachler *et al*, 2010a; Gray *et al*, 2015).

TolAIII acts as a receptor for colicin and phage transport across the periplasm and deletion of TolAII produces a *tol* phenotype and 3000-fold decrease in colicin sensitivity (Li *et al*, 2012; Krachler *et al*, 2010a; Krachler, 2009). This provides strong evidence for the TolAII domain as a 'spacer' that stretches across the periplasm to bring TolAIII in close proximity to the OM. Notably, the length of TolAII varies according to species. In *P. aeruginosa,* TolAII maintains similar heptad repeats, however it is only about 200 amino acid residues in length.

The C-terminal domain of TolA (TolAIII) comprises a folded globular domain approximately 120 amino acid residues in length. Tertiary fold and secondary structure are mostly conserved between *E. coli* and *P. aeruginosa*, despite relatively low sequence identity (20%).

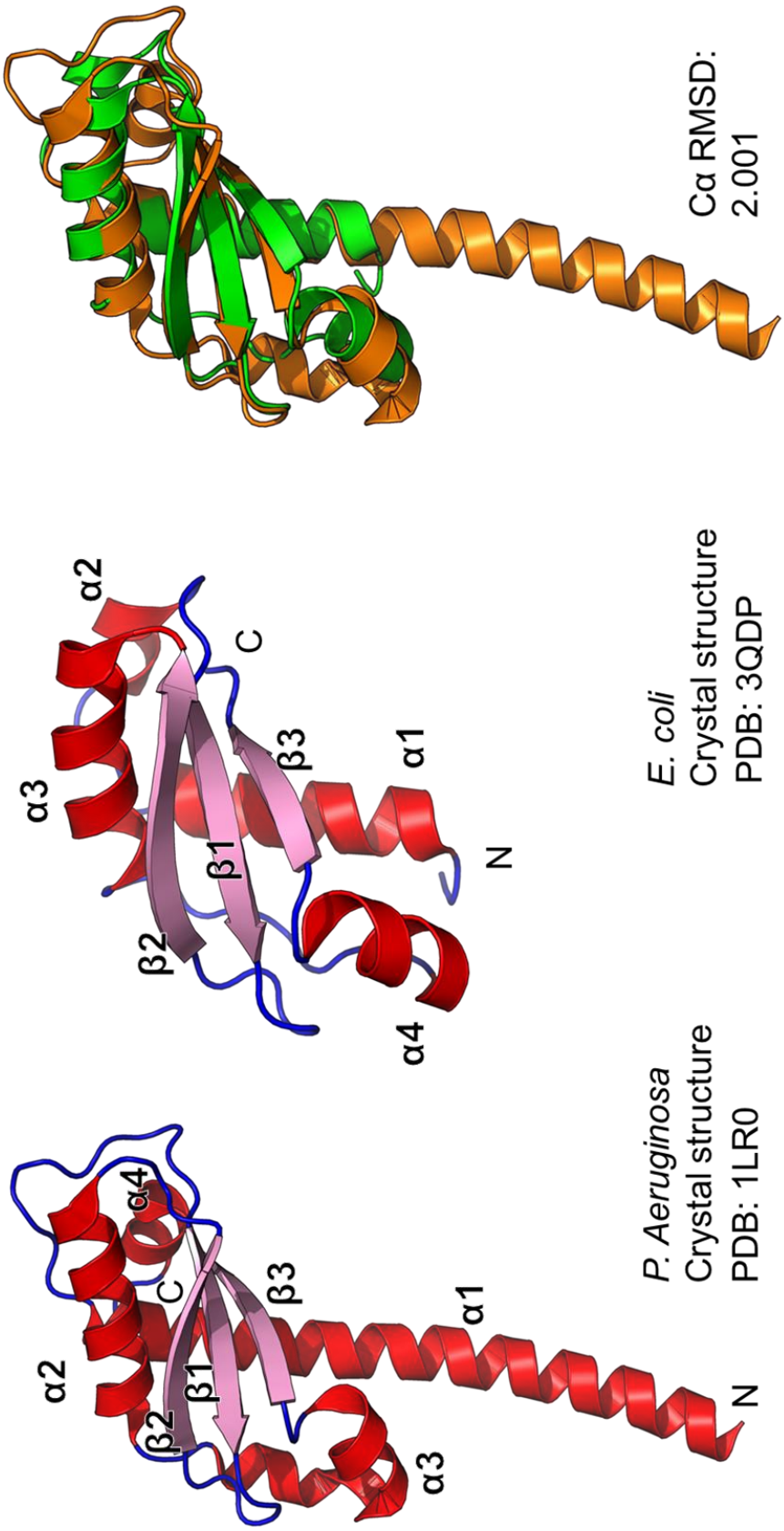


Figure 1.5 Cartoon representations of TolA domain III structures. TolAIII tertiary structure is conserved between *E. coli* and *P. aeruginosa.* Secondary structure is nearly conserved, the position of helix 4 in *P. aeruginosa* is replaced in *E. coli* by helix 2, which does not possess a helix 4 (Witty *et al*, 2002; Li *et al*, 2012)**.** Figures drawn using PyMol (DeLano, 2002).

These structures are depicted in Figure 1.5. The available crystal structures for TolAIII in *E. coli* and *P. aeruginosa* superimpose with a root mean squared deviation (RMSD) of 1.8 Å (Witty *et al*, 2002). In *V. cholera, the* TolAIII domain shares key secondary and tertiary structures as well (Ford *et al*, 2012), with its counterparts from *E. coli* and *P. aeruginosa* and a sequence alignment of a representative array of Gram-negative species suggests that the overall fold of the TolAIII domain may be broadly conserved (please also refer to Chapter 3 for further discussion).

The globular TolAIII domain consists of a single ß-sheet with exposed edges and 4-5 clustered alpha helices and this represents the target for many anti-bacterial proteins, such as colicins and bacteriophage outer coat proteins. Structural data have been reported for the interactions between *E. coli* TolA and Colicin A, as well as for Colicin N and bacteriophage attachment proteins (Hecht *et al*, 2010; Li *et al*, 2012; Ford *et al*, 2012). Interestingly, two key points of interaction on TolAIII have been reported, on either face of the β-sheet via strand complementation. To date these available structures all represent interactions between TolAIII and foreign parasitic molecules and are depicted in Figure 1.6.

It is important to note that the interaction between TolAIII and its endogenous receptor has never been structurally resolved. Indeed, even the identity of a native binding partner for TolAIII has been disputed in the last decade. Cascales et al. have reported on a proton motive force-dependent interaction between TolAIII and Pal (Cascales *et al*, 2000). In these studies the suggestion was made via crosslinking experiments and immunoprecipitation studies that there is a direct interaction between TolAIII and Pal. Subsequently, a conserved polypeptide sequence downstream of the peptidoglycan binding region in Pal was also described as the 'TolA box'. This refers to the region of Pal suggested to bind to TolA, as determined by deletion analysis (Cascales & Lloubès, 2003).

In parallel with these findings, the potential for an interaction specifically between TolAIII and TolB was first suggested (Lloubès *et al*, 2001), then subsequently shown to be required for Tol-Pal function. This conclusion was determined from a yeast two-hybrid screening method but was unable to replicate the results of the 2000 Cascales study that described an interaction between TolAIII and Pal (Walburger *et al*, 2002). Chemical crosslinking experiments also appeared to demonstrate evidence for a physical complex between TolAIII and TolB (Walburger *et al*, 2002).

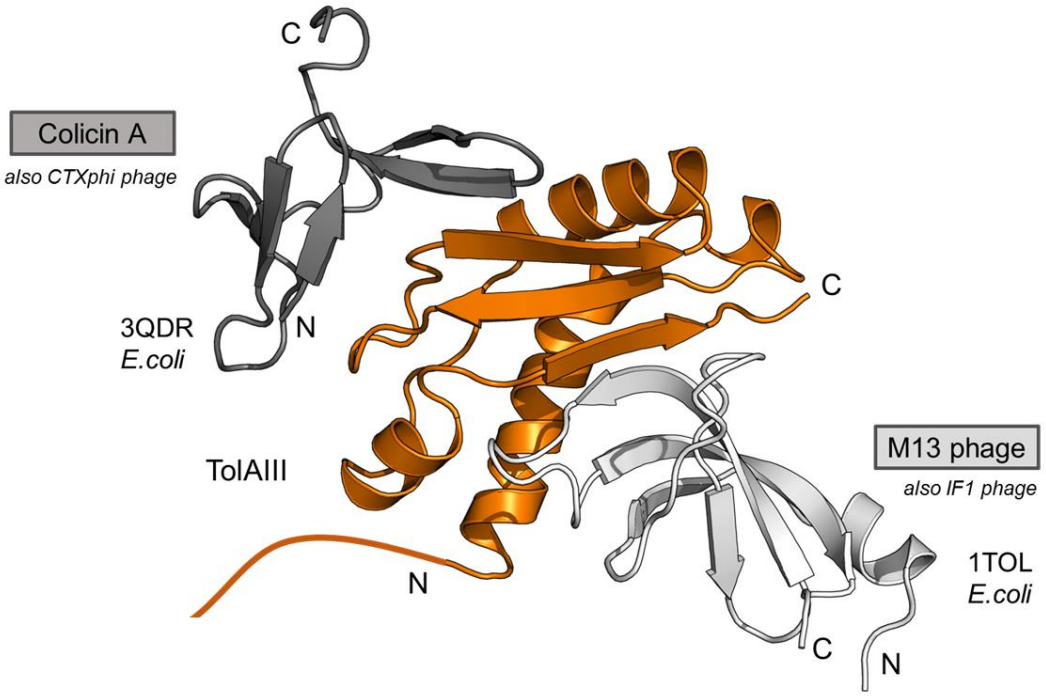


Figure 1.6 TolA is a protein-protein interaction hub in the bacterial periplasm. A cartoon depiction of the known interactions that bacteriocins and bacteriophage proteins make with TolA, both of which recruit TolA to gain entry into the cell. Colicin A uses a parallel beta-sheet complementation mechanism using strand 2 to interact with TolAIII, M13 phage uses a similar mechanism but to strand 3. Neither binding mechanism results in large conformational changes in TolAIII.(Li *et al*, 2012; Lubkowski *et al*, 1999b)

Subsequent to these experiments, the identification of suppressor mutants of the *tolA* domain in *tolb* genes, but not in *pal* genes, from *E. coli* were reported (Dubuisson *et al*, 2002). Then in 2009, a 'TolA binding box' was revealed to be present in TolB. The TolA binding box comprises the 12 amino acid residue intrinsically disordered N-terminus of TolB. In the latter investigation, TolB was shown to act as an intermediary binding protein that could potentially be bound to *either* Pal or to TolAIII. Furthermore, biophysical measurements failed to a direct *in vitro* TolA-Pal interaction (Bonsor *et al*, 2009). Bonsor *et al.* also suggested that the TolB N-terminus undergoes an order-disorder transition between conformational states that may facilitate a signaling mechanism between Pal and TolAIII (Bonsor *et al*, 2009).

1.8.4 TolQ

TolQ, is a 25 kDa IM protein that comprises three membrane-spanning alpha helices (Bourdineaud *et al*, 1989; Kampfenkel & Braun, 1993; Derouiche *et al*, 1995). TolQ, and two

adjacent proteins, TolR and TolAI interact via their trans-membrane helices to form a complex in the IM. (Derouiche *et al*, 1995; Germon *et al*, 2001). This arrangement is shared with the bacterial iron import pathway protein ExbB, with which TolQ is 75% homologous (Kampfenkel & Braun, 1993; Higgs *et al*, 2002). TolQ and ExbB are both energy factories that facilitate processes occurring in the periplasm and at the bacterial cell outer membrane. ExbB, along with its partner protein ExbD, utilise a long tri-domain periplasmic protein called TonB (homologous to TolA) to activate iron siderophore and vitamin B transporters in the outer membrane (Braun & Herrmann, 1993). The ExbB/ExbD complex converts chemical energy from an electrochemical gradient (termed the proton motive force) across the inner membrane to the TonB molecule which transduces this energy to the outer membrane via an as yet unknown mechanism (Postle & Kadner, 2003).

A strong sequence homology reported between ExbB and TolQ exists in the proposed membrane-spanning regions, where it is thought major TolQ functional sites reside. *tolQ* null mutants are tolerant to all group A colicins, suggesting that functional TolQ is essential for the normal action of TolB and TolA, two proteins known to be parasitized by colicins for import. A *tol* mutant containing an alanine to valine substitution (A177V) in the third putative transmembrane helix of TolQ furnishes tolerance to colicin A, but not E1 or filamentous phage f1 (Vianney 1994). Suppressor mutations of TolQ A177V that resuscitate cell integrity and colicin A sensitivity, have been found clustered in the first transmembrane helix of TolR and in the first transmembrane helix of TolQ. Recovery of Tol-Pal function therefore appears directly linked to reinstating an interaction between TolR and TolQ via their transmembrane helices (Lazzaroni *et al*, 1995).

1.8.5 TolR

TolR is a 15 kDa protein anchored in the inner membrane via an N-terminal transmembrane alpha helix. The N-terminus of TolR is exposed to the cytoplasm, and the remainder of the C terminus is exposed to the periplasm (Kampfenkel & Braun, 1993; Derouiche *et al*, 1995). A structure for TolR has been reported from *H.influenza* that omits 61 amino acid residues from the N terminus, nine residues from the C-terminus and which forms a symmetrical dimer (Parsons *et al*, 2008), as shown in Figure 1.3. More recently, it has been reported that intact *E. coli* TolR similarly forms a dimer, however this structure involves strand-swapping between the N and C- termini of each subunit, leading to the suggestion that TolR undergoes structural remodeling to bind peptidoglycan in order to act as a stator (Wojdyla *et al*, 2015). This is similar

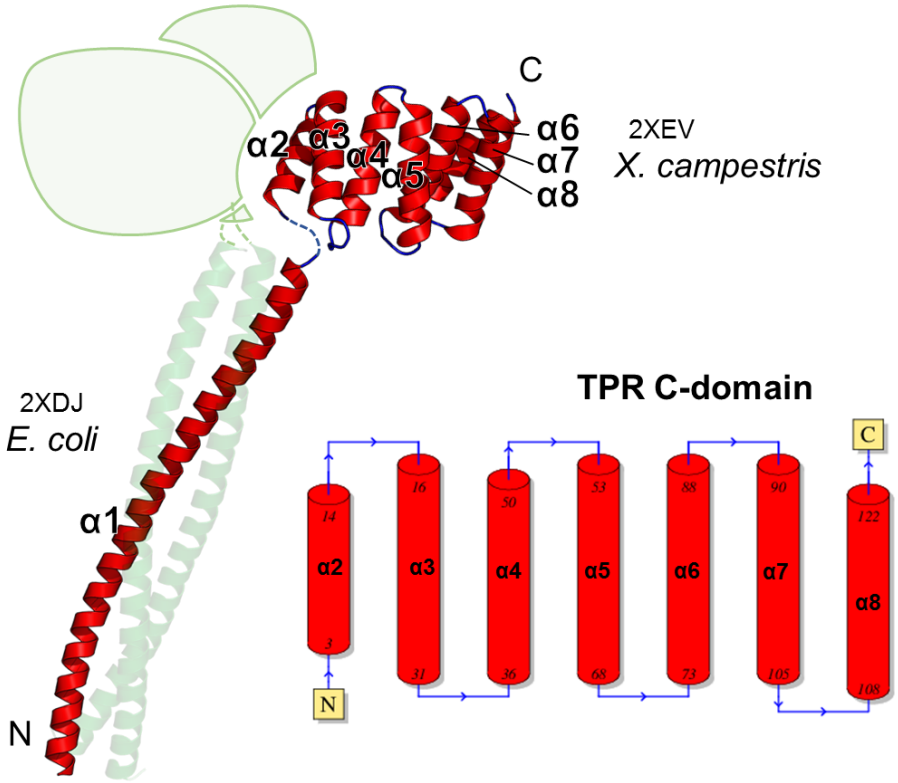


Figure 1.7 Cartoon representation of YbgF/CpoB. Figures generated using Pymol and PDBSum (Laskowski *et al*, 1997; DeLano, 2002). Cartoon depiction shows the trimeric N-terminal domain and the TPR-C-domain monomer. YbgF/CpoB 2WZ7 (Krachler *et al*, 2010a), TPR domain 2XEV (Krachler *et al*, 2010a).

to the mode of action of MotB, and is based on observations of the peptidoglycan-binding activity of *in vitro* mutants that remove these swapped strands (Wojdyla *et al*, 2015). TolQR could therefore function as an energy transmitter, harnessing energy from the proton motive force (Cascales *et al*, 2001) to remodel TolR. Interestingly, cross complementation between Tol and Ton proteins show that TolQ/ExbB and TolR/ExbD are interchangeable (Lloubès *et al*, 2012).

1.8.6 YbgF/CpoB

YbgF is a trimeric periplasmic protein, with subunit monomers 28 kDa in size. YbgF contains an N-terminal coiled coil domain and a C-terminal tetratricopeptide repeat (TPR) domain separated by a flexible linker, as depicted in Figure 1.7. TPR domains spontaneously self-organise into characteristic helix-turn-helix structures and are known to typically assist in the organization of larger protein assemblies (Krachler *et al*, 2010b; Ramarao *et al*, 2001; Eckler *et al*, 2005). Until recently, the function of the *ybgF* gene product was not known, as *ybgF* null mutants produce no discernible *tol* phenotype (Walburger *et al*, 2002).

It has also been reported that YbgF interacts with TolAIII during colicin A import, for an unknown purpose (Walburger *et al*, 2002). Overexpression of YbgF strongly induces the Rcs stress response pathway and it has been proposed that free YbgF has a role in either inhibiting peptidoglycan synthesis or activating peptidoglycan turnover (Krachler, 2009). Krachler *et al.* showed that TolA modulates the oligomeric status of YbgF in the periplasm, suggesting that the function of YbgF is regulated by free TolA acting through destabilization of the YbgF free trimeric state (Krachler *et al*, 2010a). More recently it has been demonstrated that the envelope machines that coordinate septal PG synthesis are themselves coordinated by YbgF, which hereafter is referred to as the molecule CpoB (Coordinator of peptidoglycan synthesis and outer membrane constriction associated with PBP1B) (Gray *et al*, 2015). CpoB localizes to the septum at the onset of constriction, linking both the PG synthetic machinery and the Tol system, and regulates PBP1B activity in response to the energetic state of Tol-Pal. This association between PG synthesis and OM invagination is suggested to selectively modulate PBP1B cross-linking activity, providing a unique mode of PG synthase regulation (Gray *et al*, 2015).

1.8.7 YbgC.

YgbC is a 15 kDa soluble cytoplasmic protein. The structure for YbgC has been determined in *H.pylori*, using X-ray crystallography (Angelini *et al*, 2008). YgbC has distinctly similar molecular architecture to the thioesterase family (Benning *et al*, 1998). It has been observed that YbgC conserves several features needed for thioesterase activity, including the "hot-dog" fold and an active site aspartic acid residue that is critical for function. It has therefore been proposed that YbgC may be involved in phospholipid metabolism (Zhuang *et al*, 2002; Gully & Bouveret, 2006). A 'Hot dog fold' consists of a long α-helix surrounded by a four-stranded antiparallel β-sheet. To date, no interactions between YbgC and the rest of the Tol-Pal complex have been demonstrated.

1.8.8 The divisome in Gram-negative bacteria

The cytokinetic machinery assembled at the septum during cell constriction is termed the divisome: a ring-shaped organelle which carries out the invagination of both IM and OM layers during cell constriction and subsequent division. Core proteins known to be vital in the *E. coli* divisome include cytoplasmic FtsZ and FtsA, integral inner membrane proteins such as ZipA, FtsQ, FtsB, FtsL, Ftsl, FtsN, FtsK and FtsW (Gerding *et al*, 2007a). Polymerisation of FtsZ into a

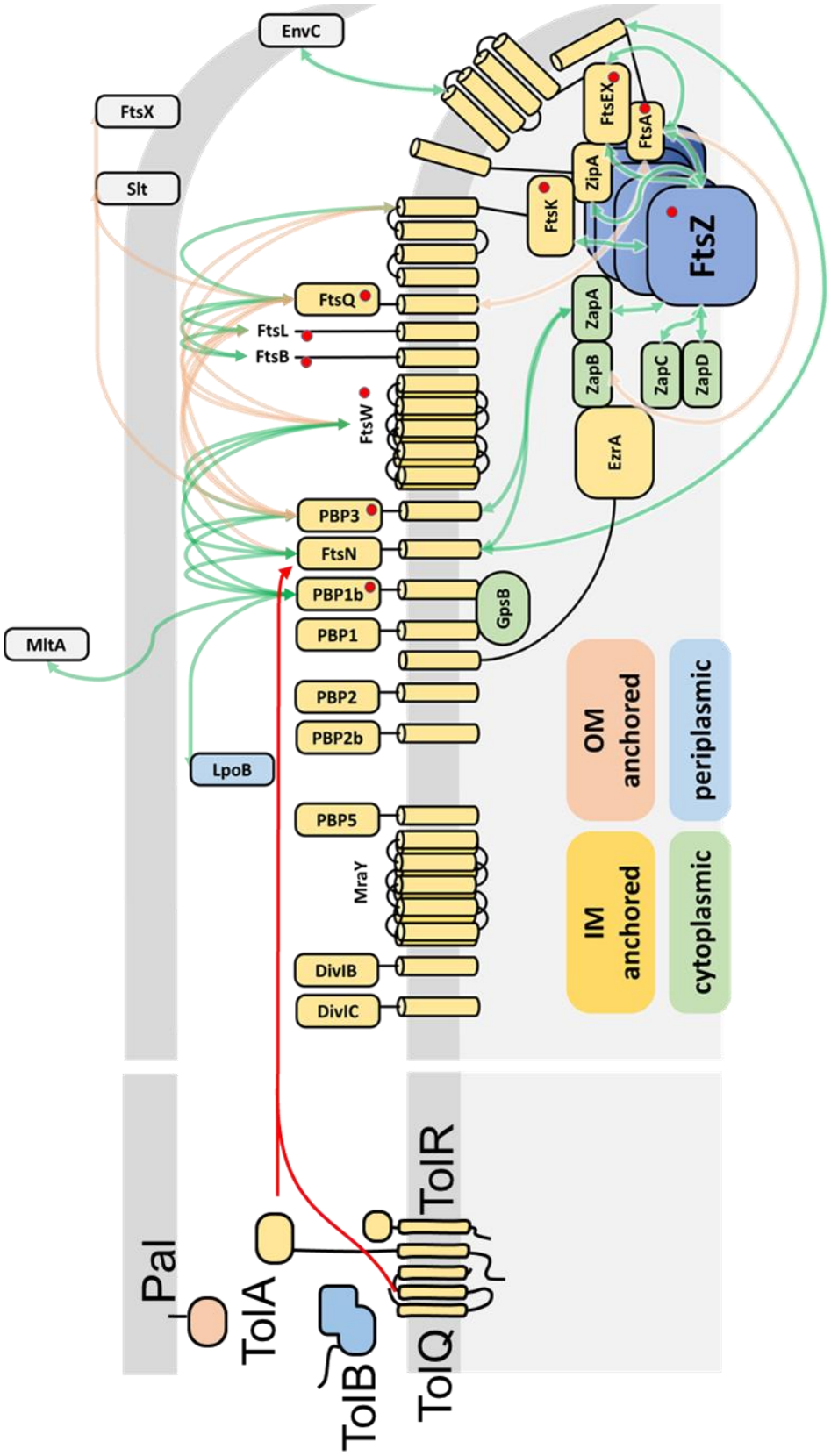


Figure 1.8 **Recruitment of the Tol-Pal complex to the divisome machinery involves interaction with FtsN, a protein with integral interactions to the FtsZ septal ring.** Proteins thought to be essential are marked with a red dot. FtsN is an essential protein, however is not marked as it was recently reported that suppressor mutants of FtsA can compensate for FtsN in *E. coli* (Bernard *et al*, 2007b).(Silhavy *et al*, 2010; Lutkenhaus *et al*, 2012; Bernard *et al*, 2007a; Goley *et al*, 2011; Typas & Sourjik, 2015)

septal ring involves interactions with ZipA and ZapA proteins, the divisome in *E. coli* is represented in Figure 1.8.

The full architecture of the divisome remains to be mapped, however known interactions in Figure 1.8 are shown by green connective arrows, whereas known interactions that may involve a secondary intermediate are shown in pink. The recruitment of Tol-Pal to the divisome machinery occurs through FtsN in *E. coli,* having been demonstrated *in vivo* by fluorescent fusions (Gerding *et al*, 2007a). TolA and TolQ have both been reported to assemble at the septal ring independently, however other Tol proteins require the presence of an intact Tol-Pal complex, suggesting they may be recruited through TolA or TolQ (Gerding *et al*, 2007a). This recruitment is dynamic, and based on the known associations of CpoB with peptidoglycan assembly machinery, may by vital in the rapid assembly of new peptidoglycan at a rapidly constricting division site.

1.9 Thesis Aims and Hypothesis

The Tol-Pal protein complex has been proposed to play a critical role in the structure of the divisome, in mediating peptidoglycan assembly, as well in outer membrane stability. The important capacity for a Tol-Pal complex to be able to span the bacterial cell membrane is strongly supported by the susceptibility of Tol-Pal to antibacterial agents that seek translocation across the bacterial cell membrane. The mechanism of action of the Pal and TolB components of Tol-Pal complex that interact with the peptidoglycan layer, appears to result in a signalling interaction that involves a conformational change of the N-terminus of TolB (Bonsor *et al*, 2009). Furthermore, this conformational change is promoted by colicins to facilitate their entry into bacterial cells, suggesting that it may provide a vital activation step in Tol-Pal function (Bonsor *et al*, 2009).

Despite this, many of the molecular details that underpin the structure and function of the Tol-Pal protein complex remain to be elucidated. Crucially with respect to this thesis, the molecular mechanism and structure underlying the interaction between the N-terminus of TolB and its endogenous partner TolA has never been determined. The principal aims of this work were therefore to firstly determine the molecular architecture and function of the TolA-TolB protein-protein interaction, thereby determining the structure and mechanism by which TolA interacts with TolB. Secondly, to determine if this interaction is broadly conserved amongst Gram-

negative bacteria, and finally, to seek to better understand how the proton motive force may be transduced to the outer membrane of bacterial cells.

A central hypothesis of this study was that one or both of the molecular sites exploited by antibacterial proteins on TolA, and shown in Figure 1.6, provide the molecular sites involved in endogenous interactions with TolB. These protein-protein interaction sites represent important subjects for a detailed structural biological investigation by spectroscopic and biophysical techniques.

2. Materials and Methods

2.1 Microbiology

2.1.1 Bacterial strains and growth conditions

E. coli B-strain (BL21 DE3) cells were used for all protein expression in this study. Cells were grown in Luria-Bertani (LB-Miller) broth (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, MilliQ water, sterilized via autoclave 20 minutes @ 121 °C), or M9 media (0.0477 M Na_2HPO_4, 0.022 M KH_2PO_4, 8.6 mM NaCl, 4 µM $ZnSO_4$, 1 µM $MnCl_2$, 0.7 µM H_3BO_3, 0.7 µM $CuSO_4$, 2 µM $FeCl_3$, 2 mM $MgSO_4$, 0.1 mM, $CaCl_2$, 0.3% (w/v) glucose, 0.1% (w/v) ^{14}N NH_4Cl) for all liquid culture steps, and plated on LB-agar for all solid phase growth steps. Antibiotic selection was performed using ampicillin (amp) at 100 µg/mL, chloramphenicol (chlor) at 34 µg/mL and kanamycin (kan) at 50 µg/mL when indicated (Melford). The bacterial strains used in this work are listed in table 2.1

Strain	Properties	Reference
DH5α	F- *endA1 glnV44 thi-1 recA1 relA1 gyrA96* deoR nupG Φ80dlacZΔM15 Δ(*lacZYA- argF*)U169, *hsdR17*(rK- mK+), λ–	(Meselson & Yuan, 1968)
BL21(DE3)	F- *ompT hsdSB*(rB–, mB–) *gal* dcm (DE3)	(Studier & Moffatt, 1986)
BL21(DE3) pLysS	F- *ompT hsdSB*(rB–, mB–) *gal* dcm (DE3) pLysS (CamR)	(Studier & Moffatt, 1986)
JW5100-1	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), *ΔtolB789::kan*, *λ*$^-$, *rph-1*,*Δ(rhaD-rhaB)568*, *hsdR514*	(Baba *et al*, 2006)
BW25113	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), *λ*$^-$, *rph-1*, *Δ(rhaD-rhaB)568*, *hsdR514*	(Datsenko & Wanner, 2000; Baba *et al*, 2006)**

Table 2.1 **List of *Escherichia coli* bacterial strains used in this study.** (**) BW25113 strain was the parent strain for the Keio collection and represents *wt E. coli* in all *in vivo* experiments.

2.1.2 Vectors and plasmid isolation

The pBAD24 vector was originally purchased from ATCC. The pMMHb plasmid used for pPH01 was kindly provided by Dr. Jo Claridge from Prof. Jason Schnell at the Department of Biochemistry at the University of Oxford. pREN28 kindly provided by Dr. Renata Kaminska from the Prof. Kleanthous group in the Department of Biochemistry at the University of Oxford.

Plasmids were prepared and purified using QIAprep spin miniprep kits (Qiagen). The proteins used in this work are described in the appendix.

2.2 Molecular Biology

2.2.1 Competent cells

Cultures of BL21(DE3) (NEB) and BL21(DE3)pLysS (Promega) cells were grown in LB (10 g/L tryptone, 5 g/L Yeast extract, 10 g/L NaCl, MilliQ water, which was sterilized via autoclave 20 minutes at 121 °C). Cells were brought to log phase (OD_{600} = 0.4-0.6), and pelleted by centrifugation (5,000 *x g*, 4 °C 10 minutes). Cells were subsequently made chemically competent, by suspension in 20 mM Tris, 50 mM $CaCl_2$ pH 8.0. Cells were then pelleted by centrifugation (5,000 *x g*, 4 °C 10 minutes) and resuspended in pH 8.0 20 mM Tris, 50 mM $CaCl_2$, 20% w/v glycerol. Aliquots of competent cells were stored at -80°C.

2.2.2 Transformation of competent cells

Competent cells were thawed on ice, then given 100 ng of target plasmid DNA. Cells were then incubated on ice for 30 minutes, heat shocked at 42 °C for 90 seconds, and incubated on ice an additional 15 minutes. After incubation, 500 µL LB broth was added and cells were then grown with shaking at 37 °C for an additional 30 minutes. Cells were then plated on to LB-agar with appropriate antibiotic and incubated overnight with at 37 °C.

2.2.3 Polymerase chain reaction (PCR)

Gene amplification was performed by PCR, using plasmids or genomic DNA as the template. Primers had restrictions sites flanking the gene of interest and PCR was then performed using an Eppendorf Mastercycler. Polymerases used for PCR were either Pfu Turbo (New England Biosciences, NEB) or a multi enzyme blend (Quikchange II Site Directed Mutagenesis kit, Agilent). In reactions using Pfu Turbo, dNTPs were purchased from NEB, in reactions using Quickchange II, dNTPs were used from the Agilent kit.

2.2.4 DNA primers

DNA primers for gene amplification were synthesized by Eurofins. A full list of primers is in the appendix.

2.2.5 Restriction digests

Restriction digests of plasmids were performed using enzymes purchased from New England Biolabs (NEB). All restriction digests were performed using CutSmart buffer (NEB). BamH1 High-Fidelity (HF) and HindIII HF enzymes were used to minimize star activity. Products were subsequently purified by agarose gel electrophoresis and extracted using a QIAquick Gel extraction kit (Qiagen).

2.2.6 Ligations

Ligations were performed using T4 DNA ligase (NEB). Two types of ligations were used in this study: 1) Ligations where digested PCR products were ligated into an appropriately cut vector. 2) Directly annealing two oligonucleotides to generate an insert with overhangs appropriate to the overhangs of a target digested vector.

Vectors used in double digests were cut to 100% by BamH1/HindIII at a DNA concentration of 101.3 ng/µL and cut overnight at 37 °C. Cut vector was then purified on a 0.7% (w/v) agarose gel, excised with a scalpel and dissolved using a Gel Extraction Kit (Qiagen). PCR products were added to linearized vector in T4 DNA ligase buffer, incubated overnight at room temperature and transformed into chemically competent DH5α cells.

An example ligation reaction was performed as follows: 2 µL 10X T4 DNA Ligase buffer, 50 ng of vector DNA, 37.5 ng of insert DNA, 1 µL T4 DNA ligase, and the addition of nuclease free water to bring the reaction volume to 20 µL. After 60 minutes incubation at room temperature, the T4 DNA ligase was inactivated at 65 °C for 10 minutes, chilled, and 5 µL of the reaction was used to transform aliquots of competent cells and plated on LB-Agar with appropriate antibiotic.

2.2.7 Whole plasmid mutagenesis

Whole plasmid mutagenesis was either through the use of Phusion DNA Polymerase (NEB) or the Stratagene Quikchange kit (Agilent) to generate site directed mutants for use *in vivo* and *in vitro*. Phusion (NEB) and Stratagene Quikchange (Agilent) were used used to alter plasmids to possess mutations in a target gene. A typical PCR reaction mixture was as follows:

Component	Volume
5x *Phusion* HF reaction buffer	10 µL
dNTPs (10 mM)	1 µL
Template DNA (50 ng/µL)	1 µL
Forward primer (10 µM)	2.5 µL
Reverse primer (10 µM)	2.5 µL
Nuclease-free water	32.5 µL
Phusion Polymerase (2.0 U/ µL)	0.5 µL

Table 2.2 **Typical PCR reaction**

Temp (°C)	Time (min:sec)	Description	
98	02:00	Initial denaturation	
98	00:00	Denaturation	15 cycles
60	00:00	Anneal	
72	02:00	Elongation	
72	05:00	Final elongation	
10	∞		

Table 2.3 **Typical PCR cycling parameters.**

2.2.8 Agarose gel electrophoresis

Agarose gels for use in DNA purification and visualization were prepared at 0.7% w/v agarose, with 0.5x TBE buffer (44.5 mM Tris-Borate buffer, 1 mM EDTA at pH 8.3) and microwaved. Before casting, 5 µL of SYBR safe 10,000x concentrated solution (Invitrogen) was added to 50 mL of the agarose mix, then this mixture was poured into a casting apparatus. Gels were then immersed in 0.5x TBE buffer. Samples were added at a ratio of 5:1 with a 6x sample buffer (0.05% w/v bromophenol blue, 6 mM EDTA, 30% v/v glycerol) and loaded. Agarose gels were electrophoresed at a constant 90 V voltage for 35-60 minutes, before visualizing the DNA on a UV transilluminator.

2.2.9 $TolB_{22\text{-}34}$ peptide synthesis

TolB peptide and peptide variants corresponding to the disordered N-terminus of TolB residue $_{22\text{-}34}$ ($TolB_{22\text{-}34}$) were synthesized by thinkpeptides® UK (ProImmune). These sequences are listed in the Appendix. Peptides were delivered as lyophilized powder, and were assessed for full solubility before use. High performance liquid chromatography was used to verify the elution profile of synthetic *P. aeruginosa* $TolB_{22\text{-}34}$ using a Dionex Vydac 250 mm 218TP C18 reverse phase column on a 0-95% (v/v) Acetonitrile (AcN) gradient. Peptides were synthesized with free N-termini, so as to remain comparable with Sec cleaved TolB, and were either amine-blocked, fluorescein isothiocyanate (FITC) labelled via a C-terminal lysine, or unblocked at the C-terminus. Recombinant peptide was also expressed, described in 2.2.13. Not all peptides used in this study were soluble, each was tested for relative solubilty in NMR buffer. Ultimately, as the sequence $TolB_{22\text{-}34}$ proved to be soluble up to 3 mM, and was the closest mimic of the TolB N-terminus, amide-capped $TolB_{22\text{-}34}$ was used in all cases unless otherwise specified.

2.2.10 Expression and purification of Im9-TolAIII fusion and variants

To purify *P. aeruginosa* TolAIII, a fusion protein construct approach was taken, where a $TolAIII_{227-347}$ sequence separated by a FXa cleavage site from a cleavable poly-His-tagged Im9 polypeptide sequence was expressed and cleaved by FXa protease so that TolAIII could be purified by gel filtration.

Competent BL21 DE3 cells were transformed with the expression vector pREN28 pQE-2/Im9. pREN28 encodes a poly-His-tagged Im9 protein followed by a Factor Xa (FXa) cleavage site and *P. aeruginosa* $TolAIII_{224-347}$. Removal of the Im9 His tag was a requirement to keep the size of the construct within a suitable range for an NMR investigation (12 kDa). Fresh transformants were used to inoculate bacterial overnight cultures grown in 50 mL LB with 100 µg/mL amp. Overnight cultures were pelleted at 4,000 x *g,* resuspended in 5 mL fresh sterile LB and used to inoculate 1 L LB media flasks. These expression cultures were shaken at 120 RPM on an Innova 2300 platform shaker (New Brunswick Scientific) at 37 °C for 5 hours to mid-log phase (OD_{600} = 0.6). Cells were then induced with stock 1 M isopropyl ß-D-1-thiogalactopyranoside (IPTG), to a final concentration of 1 mM IPTG in solution.

After 4 hours growth, cells were harvested by centrifugation (4000 x *g* 4 °C and resuspended in ~40 ml 50 mM Tris, 1 mM $MgCl_2$, 2 mM imidazole, 1 mM PMSF (Sigma), 300 mM NaCl, pH 7.5, with protease inhibitor cocktail (Roche cOmplete *Easypack* tablets). Cells were lysed by sonication, on ice using a Sonicator 4000 with a ½ inch stud probe (Misonix) with 3 minutes processing time, 3 second pulses on, with 7 seconds spacing between pulses at an amplitude of 60 %. Cell debris was removed by centrifugation (Beckman JA-25.50, 48,000 x *g*, 4 °C for 30 mins), and the clarified lysate supernatant filtered with a 0.45 µm syringe filter and then further purified by Ni-affinity chromatography: The lysate was loaded onto a freshly charged 5 mL HisTrap FF Ni-affinity column (GE) at 1 mL/min at 4 ºC , and eluted at 2 mL/min. (Buffer A: 50 mM Tris buffer, 50 mM NaCl, 2 mM imidazole brought to pH 7.5, Buffer B: 50 mM Tris, 50 mM NaCl, 500 mM imidazole brought to pH 7.5). The column elution program used was 5 column volumes (CV) of 100% v/v Buffer A, 4 CV of 5% v/v buffer B, then a 5%-50% v/v linear gradient of buffer B over typically 20-30 CV. A 5 CV step at 100% v/v Buffer B was used to elute any residual His-tagged protein. Protein was monitored by UV absorbance at 280 nm. Fractions were checked for the presence of Im9-TolAIII by SDS-PAGE. Fractions were pooled and dialyzed overnight against 4 L 50 mM Tris 100 mM NaCl, 5 mM $CaCl_2$ pH 7.5 at 4 ºC in 8,000 Da molecular

weight cut- off (MWCO) dialysis tubing (Sigma). The sample was concentrated to approximately 5 mg/mL for optimal protease cleavage with MWCO 10,000 spin concentrators (Vivaspin). Dialysis is important because FXa is sensitive to imidazole, and Novagen recommends that protein concentration be greater than 2.5 µM for optimal cleavage.

Im9-TolAIII protein was syringe-filtered at 0.45 µm then cleaved overnight at 4 ºC with 59 µL (150 Units) of Novagen FXa at a ratio of 2 Units per mg of fusion protein in 15 mL. Cleaved protein was purified by Ni-affinity FF His-Trap using the same 5-50% v/v Buffer B gradient as before, with identical Buffer A and Buffer B solutions. Buffer B elutes the fusion partner, as TolAIII no longer binds the resin. Fractions were checked for TolAIII by SDS-PAGE. To ensure no residual protease activity before the dialysis step relevant fractions were pooled and quenched with 1 mM PMSF. Protein was then dialysed 3 times against 2 L of 50 mM Tris, 100 mM NaCl, pH 7.5 at 4 ºC overnight (18 hours) in MWCO 3,500 Da dialysis tubing. MWCO 5,000 Da spin concentrators were used to concentrate the pooled TolAIII, which was syringe-filtered at 0.45 µm and loaded to a room temperature or 4 ºC S75 Superdex 26/60 column, equilibrated in 50 mM Tris, 100 mM NaCl, pH 7.5. Protein was eluted isocratically and the concentration quantified by UV absorbance at 280 nm. Extinction coefficients for TolAIII and Im9-TolAIII were derived by the method of Gill and von Hippel (Gill & von Hippel, 1989), using the Scripps Molecular Weight calculation tool (http://protcalc.sourceforge.net/). After fractionation, samples were checked by SDS-PAGE for TolAIII, pooled, and then concentrated using MWCO 5,000 spin concentrators. The purified protein was snap frozen in liquid nitrogen and stored at -80 °C.

2.2.11 Expression and purification of heterolabelled Im9-TolAIII

Heterolabelled Im9-TolAIII was expressed and purified by a similar method to Im9-TolA, for the use in heteronuclear NMR experiments. Instead of a cell culture growth step in LB, 875 mL cultures were prepared in M9 media. ^{15}N-enriched TolAIII was prepared with ^{15}N-supplemented growth media (0.0477 M Na_2HPO_4, 0.022 M KH_2PO_4, 8.6 mM NaCl, 4 µM $ZnSO_4$, 1 µM $MnCl_2$, 0.7 µM H_3BO_3, 0.7 µM $CuSO_4$, 2 µM $FeCl_3$, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.3% (w/v) glucose, 0.1% (w/v) ^{15}N ammonium chloride). ^{15}N, ^{13}C-enriched TolAIII was prepared using 0.3% (w/v) ^{13}C glucose (Sigma), 0.1% (w/v) ^{15}N ammonium chloride (Sigma) supplemented M9 growth media. Protein was purified as before with Im9-TolAIII.

2.2.12 Expression and purification of TolB

pEC14 expresses *P. aeruginosa* TolB (22-432) with a cleavable C-terminal poly-His tag, without a Sec signal sequence. This protein was produced for use in binding experiments with TolAIII by NMR. Since TolAIII was being used as the heterolabelled reporter protein, TolB was not expressed with either ^{13}C or ^{15}N incorporation. Trial mini-expressions produced TolB protein (identified by SDS-PAGE) after induction with 1mM IPTG in both BL21 DE3 and Rosetta pLysS (DE3) cells, BL21(DE3) cells were used for later purifications.

Competent BL21 (DE3) cells were transformed with the expression vector pEC14. Fresh transformants were used to inoculate bacterial overnight cultures grown in 50 mL LB with 100 µg/mL amp. Overnight cultures were pelleted at 4,000 x *g,* resuspended in 5 mL fresh sterile LB and used to inoculate 1 L LB media flasks at a ratio of 1:100. These expression cultures were shaken at 120 RPM on an Innova 2300 platform shaker (New Brunswick Scientific) at 37 °C to mid-log phase (OD_{600} =0.6). Cells were then induced with a final concentration of 1 mM IPTG and were grown for 4 hours. Cells were lysed as previously described when purifying TolAIII

Cell debris was removed as described for Im9-TolAIII. After the protein solution was loaded on to the column, the column was washed with 5 CV of 100% v/v Buffer A, then a 0%-50% v/v linear gradient of buffer B for 30 CV. A 5 CV step of 100% v/v Buffer B was used to elute any residual proteins bound to the column. Protein was monitored by UV absorbance at 280 nm (tyrosine/tryptophan absorbance). Fractions were checked for the presence of TolB by SDS-PAGE. 1 mL test aliquots of 15 µM TolB were dialyzed 3 times for 3 hours each against Tris, buffer at low (20 mM), medium (50 mM) and high (100 mM) concentrations at pH 8, 7 and 7.5 conditions as well as at 0 mM, 0.05 mM and 0.5 mM NaCl respectively. This was repeated for Phosphate and HEPES buffers to determine the optimum conditions for TolB solubility. The maximum concentration achieved in solution of TolB was 10 µM at 50 mM Tris 100 mM NaCl pH 7.0. Sample stored at -80°C.

2.2.13 Expression and purification of a trpLE fusion protein for TolB peptide synthesis

In collaboration with Jo Claridge in the Prof. Schnell Laboratory (Oxford), I designed a fusion vector for expressing labelled TolB peptide. The pPH03 vector (constructed using pMMHb) incorporates a TrpLE leader sequence followed by a Met residue cyanogen bromide cleavage site, which is then followed by the $TolB_{22-34}$ sequence that corresponds to the 13 amino acid N-

terminal signal sequence for Sec-cleaved *P. aeruginosa* TolB. The TrpLE-TolB$_{22\text{-}34}$ construct was expressed in BL21(DE3). The aim was to produce recombinant TolB$_{22\text{-}34}$ to use in biophysical experiments.

The trp ΔLE 1413 polypeptide is an *E. coli* leader sequence that forces the protein into inclusion bodies (Staley & Kim, 1994). The TrpLE fusion protein expression system has previously been used effectively for the purification of bacterial membrane proteins (Claridge & Schnell, 2012), and was chosen for the purification of TolB$_{22\text{-}34}$ to minimize cell toxicity and protease degradation of the peptide. The pMMHb TrpLE-containing plasmid was digested using BamHI and NdeI overnight at room temperature in 100 µl with plasmid DNA at a concentration of 100 ng/µl. Cut vector was then separated by 0.7% w/v Agarose gel, then excised and purified with a QIAGEN Gel Extraction Kit.

Two overlapping complementary oligonucleotides (forward and reverse) were annealed and ligated directly into restriction digest-cut plasmid. Complementary primers were designed with overhanging ends corresponding to cut restriction sites. Primers were dissolved in sterile autoclaved distilled H_2O to give a final concentration of 100 pmol/µl. 10 µl of each primer were added to a 500 µl tube and heated at 95 °C in hot water bath for 5 min. Reactions were left in a 250 ml plastic beaker filled with 95 °C water and allowed to cool to less than 37 °C, over 1 hour. Reaction tubes were then spun at 10,000 x *g* to collect any condensation. This primer mix was then incubated with polynucleotide kinase in T4 DNA ligase buffer for 1h at 37 °C, to phosphorylate the insert as a prelude to ligation: 2.5 µl T4 DNA Ligase buffer (NEB) and 2.5 µl T4 polynucleotide kinase (NEB) were added.

After ligation, the reaction mixtures were transformed to DH5α cells, plated on LB-agar with appropriate antibiotic and incubated overnight at 37 °C. Colonies were subsequently used for DNA purification and plasmid sequencing.

The pPH01 plasmid was transformed into *E. coli* BL21(DE3) cells, then plated on LB-agar containing kan and incubated overnight at 37 °C. 200mL of bacterial overnight cultures were inoculated with individual colonies from fresh transformants and grown overnight with moderate shaking (150 RPM). Cultures were centrifuged at 2,000 x *g* for 25 minutes at 4 °C and then resuspended into 40 mL of chilled M9 media. Resuspended cells were then used to inoculate large 0.5 L M9 cultures, and grown at 37 °C with moderate shaking. Flasks were induced when

OD_{600} reached 0.6 with final concentration of 1 mM IPTG. Cultures were then grown overnight at 37 °C to a final OD_{600} of approximately 1.4.

0.5 L cultures were pelleted at 5,000 x *g* for 20 minutes, and resuspended in lysis buffer (50 mM Tris pH 8, 100 mM NaCl) using a Dounce homogenizer. These cells were then sonicated on ice for a total of 3 minutes as done with Im9-TolAIII. Inclusion bodies were pelleted at 27,000 x *g* for 15 minutes at 4 ºC. The pellet was resuspended with lysis buffer using the Dounce homogenizer to remove and water-soluble contaminants.

The washed resuspended mix was pelleted using ss34 centrifuge tubes at 27,000 x *g* for 30 minutes at 4 ºC. The pellet was resuspended in 35 mL of guanidine dissolution buffer (6 M guanidine, 25 mM Tris pH 8.0, 5 mM imidazole). Insoluble cell debris was removed through centrifugation of the sample at 18,000 x *g* for 1 h and 30 minutes at 4 ºC. The supernatant containing guanidine-solubilized protein was loaded onto a 5 mL Ni-affinity column (GE HisTrap) equilibrated in 6 M guanidine, 25 mM Tris pH 8.0, 5 mM imidazole. The column eluate was recirculated using a peristaltic pump at 1 mL/min for 1 hour at 4 °C to maximise binding. Protein was eluted from the column stepwise. The column was washed with 6 CV of 5 mM imidazole, 6 M guanidine, 25 mM Tris pH 8.0. Then washed with 5 column volumes of 20 mM imidazole, 6 M guanidine, 25 mM Tris pH 8.0, then 5 CV 50 mM imidazole, 6 M guanidine, 25 mM Tris pH 8.0. The final elution step was with 500 mM imidazole, 6 M guanidine, 25 mM Tris pH 8.0. Before final elution the column was left to rest 10 minutes sitting in the 500 mM imidazole solution before continuing so as to maximise protein separation from the column. Column fractions were subsequently dialyzed against 4 L MilliQ two times over 2 days to precipitate protein. After precipitation, samples were pelleted at 780 x g for 1 h, then decanted to remove 90% of the supernatant, but not allowing the pellet to go dry. The low force centrifugation was used to minimise protein packing for improved efficiency of the subsequent CNBr cleavage step. This was then snap frozen at -80 ºC ready for cleavage by CNBr.

The $TolB_{22\text{-}34}$ peptide was chemically methionine cleaved in a fume hood by dissolving the pellet into 5 mL 70% v/v formic acid. 0.5 g CNBr was then added carefully and dissolved. The reaction vessel was prepared involving a nitrogen feed to evacuate all air from the chamber and the reaction was allowed to continue for 2.5 hours. After cleavage was finished, the quenching process was begun by a tenfold dilution to water followed by flash freezing and then lyophilisation.

The volatile cyanide gas evolves over time in the lyophiliser as the freeze-dry process continues, this step typically took 4 days for completion. A Virtis Benchtop K lyophiliser was used with a Welch 1402 Vacuum pump and acid trap to remove formic acid vapour.

2.2.14 Optimized method of $TolB_{22-34}$ purification by C18 high performance liquid chromatography (HPLC)

After CNBr cleavage, the quenched reaction mixture after lyophilisation contained freeze dried TrpLE and TolB peptide. Since the TrpLE sequence still retained a poly His-tag, to separate the post-CNBr cleavage products, lyophilized reaction mixture was dissolved in 50 mM Tris 100 mM NaCl. This was loaded at 4 ºC into a freshly charged 5 mL HisTrap FF Ni-affinity column (GE) at 1.0 mL/min and was eluted at 2.0 mL/min. Column was eluted on a 0-50% v/v gradient from Buffer A (50 mM Tris buffer at pH 7.5, 100 mM NaCl) to Buffer B (50 mM Tris buffer at pH 7.5, 300 mM NaCl, 500 mM imidazole) at 1 mL/min over 60 minutes. $TolB_{22-34}$ elutes in the flow-through and was subsequently lyophilized for later HPLC purification.

Although this step removes the His-tagged TrpLE protein, the overall yield of the $TolB_{22-34}$ product was determined to be improved by 50% by removing this step as adequate separation of TolB and TrpLE was achieved by reverse phase HPLC. Recombinant peptide purifications were subsequently eluted by separating the CNBr reaction products directly by reverse phase HPLC.

$TolB_{22-34}$ was eluted on a linear gradient of 0-95% v/v AcN with 0.1% v/v TFA using a Dionex 218TP C18 VYDAC column. All experiments were performed using an AKTA Purifier, and no lasting adverse effects were seen on column performance following each run. Each run was preceded and followed by a 0-95% blank injection (MilliQ water, 0.1% v/v TFA) to confirm no protein was retained by the column. All protein was spun down for 10 minutes at 16,000 x *g* to remove any particulates and was loaded using a 2 mL injection loop (rinsed 5-fold with MQ water, 0.1% v/v TFA) after being degassed using a syringe. Cleaved $TolB_{22-34}$ elutes first, around 30% v/v AcN, followed by TrpLE which elutes at approximately 60% v/v AcN. Cleavage efficiency appears routinely near 80-90%. Fractions corresponding to the peptide were identified by running a control elution on an identical gradient using synthetic $TolB_{22-34}$ and subsequently confirmed by mass spectrometry.

2.2.15 NMR Buffer

All samples prepared for use in NMR experiments were, unless otherwise indicated, dialyzed overnight twice against 4 L of 20 mM phosphate 60 mM NaCl pH 7.0 buffer. This buffer composition is hereafter referred to as 'NMR buffer'. Note that all NMR experiments performed included 5% v/v D_2O.

2.2.16 ^{13}C ^{15}N TrpLE and $TolB_{22-34}$ purification

As detailed in section 2.2.9 and 2.2.10, a method for the purification of $TolB_{22-34}$ was determined using a fusion protein (trpLE), BrCN cleavage and subsequent purification by HPLC. Isotopically enriched $TolB_{22-34}$ (^{13}C ^{15}N) for the use in NMR experiments, as detailed in section 4 was subsequently purified by this exact method, with the substitution of ^{15}N ammonium chloride and ^{13}C glucose respectively in M9 media. A total of 1.5 mg doubly enriched peptide was purified, used for heteronuclear NMR spectroscopy. To ensure that the composition of this product was identical to synthetic peptide, and to verify the correct amino acid composition, 1D and 2D NMR spectra were collected on a 300 µL 0.5 mM sample in NMR buffer. Tight clustering between 8-8.8 ppm is characteristic of unfolded peptide, and no significant additional peaks are observed. The small exception to this is that some residues appear to possess small neighbours, which are likely not degradation peaks but rather a small population of cis/trans isomerized peptide. This is to be expected for small disordered peptides in solution.

2.2.17 ^{13}C ^{15}N TrpLE/$TolB_{22-34}$ mass spectrometry analysis for $TolB_{22-34}$

Peptide mass data was collected for the $TolB_{22-34}$ peptides used in NMR experiments. The following samples were analysed in the laboratory of Dr. Shabaz Mohammed (Department of Biochemistry, University of Oxford) for LC-MS analysis, on synthetic, cleaved recombinant, and cleaved/heterolabelled recombinant peptide. Samples injected were 0.8 ng each, freeze dried from MilliQ water and 0.1% v/v TFA overnight and resuspended in water. Peptide masses were determined as follows: Synthetic $TolB_{22-34}$ Peptide: Expected Mass 1314 Da, Observed: 1315.6. Recombinant $TolB_{22-34}$ Peptide: Expected Mass 1314 Da, Observed: 1315.6. Recombinant ^{13}C ^{15}N $TolB_{22-34}$ Peptide: Expected Mass 1385 Da, Observed: 1385.8, estimated degree of isotopic enrichment: 80% or greater.

2.2.18 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was used to resolve protein during and after recombinant expression protocols. Comparison of band position against a molecular weight marker allowed for the visual estimation of protein size to determine the success of recombinant expression. SDS-PAGE were run using running gel matrices of 10%, 14%, and 16% v/v polyacrylamide in 375 mM Tris-HCl, pH 8.8, 0.1% (w/v) SDS. To polymerise the gels a catalyst was used (120 µL 10% (w/v) ammonium persulfate and 20 µL N,N,N',N' –tetramethyl ethylenediamine (TEMED) for a 6 mL gel. Stacking gels were prepared as running gel matrices above, with substitution of 250 mM Tris pH 6.8. 10x running buffer consisted of 3 g Tris, 14.5 g glycine, and 1 g SDS per 1 L of MilliQ water. Gels were immersed in running buffer during operation. The loading buffer used for protein samples contained 200 mM Tris-HCl pH 6.8, 5% v/v mercaptoethanol, 8% w/v SDS, 0.4% w/v bromophenol blue and 40% v/v glycerol. After addition of loading buffer, samples (including the MW ladder) were boiled 4 minutes using a heating block set at 100 °C. A constant current of 30 mA was used for running each gel. Gels were subsequently stained with Coomassie Brilliant Blue R250 and were destained in 10% v/v acetic acid 10% v/v ethanol. Protein ladders for SDS-PAGE were obtained from Fermentas (MW in kDa: 116, 66, 45, 35, 25, 18.4).

2.2.19 Liquid culture *tol* phenotype assay

Liquid culture assays were used to determine the sensitivity of *E. coli tolB* mutants and *wt* strains in the presence of SDS. The outer membrane of Gram-negative bacteria is destabilized upon *tol* deletion and the phenotype can be observed using detergents such as SDS (Bernadac *et al*, 1998). SDS is an anionic denaturing detergent and binds both membrane and non-membrane proteins, which leads to lyses/ruptures of the membrane. In these conditions, cells will display impaired or inhibited growth. The SDS sensitivity of *E. coli* JW5100 cells was determined for *tolb* transformants complemented with *tolB* and *tolB* variants on a pBAD24 vector in comparison with *wt E. coli* strain. Growth of these cells was then compared to *wt* BW25113 strain cells.

LB-agar plates were made with 2% w/v SDS, 2% w/v SDS with 0.02% w/v arabinose (for induction of protein expression), 2% w/v SDS with 0.02% w/v arabinose with antibiotic, and then all of these conditions again in the absence of SDS. Both Keio-strain *wt* BW25113 and JW5100 *tolb*- strains were tested as control comparisons against cells transformed with plasmids encoding TolB and TolB with a poly His tag (pDAB17 and pGP3 respectively). pPH02, pPH03, pPH04, and

pPH04 plasmids encoding single, double, and triple alanine mutants of *tolB* in the conserved [IVL]X[IVL motif were also used to transform JW5100 cells. In all native and mutant transformants of TolB, the signal peptide sequence was present to ensure localisation to the periplasm. Although the pBAD24 vector is arabinose-inducible, this vector shows leaky expression at 37 °C.

All strains were grown overnight in LB with moderate shaking (120 RPM at 37 °C) with appropriate antibiotic. These cultures were used to inoculate 50 mL LB flasks then grown to mid-log phase (OD_{600} 0.45) with moderate shaking (120 RPM at 37 °C). Additional strains were grown where the culture was pre-inoculated with 0.02% w/v arabinose at OD_{600} = 0.15 to test if pre-supplementation of the cells with TolB makes any difference to growth. Aliquots from each growth condition were removed at periodic intervals and OD_{600} was recorded as a function of time.

2.2.20 Agar plate assay

This is an experiment designed to test the stability of the *E. coli* outer membrane, similar to the liquid growth assay. Cells presenting a classic *tol* phenotype will exhibit impaired growth in the presence of SDS or EDTA, increases the sensitivity of bacterial cells to adverse conditions, and accelerates the death rate of starved bacteria (Russell 1969, Strange and Dark 1965). EDTA causes a non-specific increase in cellular permeability, and a concurrent degradation of RNA and nucleotides to nucleosides and bases (Neu et al 1966, 67) as well as chelating away Mg^{+} and $Ca2^{+}$ which disintegrates the outer membrane.

To confirm the *tol* phenotype by a second growth method, strains and plasmid-transformed strains were made as previous with liquid culture assays. Overnight cultures were then used to inoculate 50 mL cultures with appropriate antibiotic and subsequently brought to log phase (OD_{600} = 0.4). After recording this density, aliquots of these cultures were spotted at regular intervals on SDS-containing plates in serial dilutions of spotted cell densities, from 0.002 to 0.000002 using the initial OD_{600} measurement. The rationale for this is that each serial dilution will result in a count of approximately 10-20 cells in the final dilution, which will form discrete colonies. Each SDS plate was poured on a level surface in a biological safety cabinet under sterile conditions and dried at room temperature on the cabinet for 2 h. Plates were incubated at 37 °C overnight.

2.2.21 Size-exclusion chromatography multi-angle laser light scattering (SEC-MALLS/SEC-MALS)

To confirm the monomeric character of TolAIII, this protein was analysed using SEC-MALS in the biophysics suite (Department of Biochemistry, Oxford) and in the Dunn School of Pathology (Oxford). Salt-dependence experiments were run with Dr. David Staunton (Department of Biochemistry, Oxford). Protein was run sequentially at high (200 mM NaCl) and low (50 mM NaCl) salt concentration on a Superdex200 10/300 column equilibrated in phosphate-buffered saline, pH 7.0. Experimental parameters were as follows: Concentration Source: RI, Flow Rate: 0.5 mL/min, Light Scattering Instrument: HELEOS 8, Cell Type: Fused Silica, Wavelength: 662.3 nm, Calibration Constant: 2.8645×10-5 1/(V cm), RI Instrument: rEX, UV Instrument: UV, Solvent: water, Refractive Index: 1.330. Concentration dependence (TolAIII multimerisation) samples were prepared at 1, 5 and 10 mg/mL. Chromatography experiments were run by Dr. Steven Johnson (Dunn School of Pathology). Size exclusion chromatography was performed on a Superdex200 10/300 column (GE Healthcare) equilibrated in Dulbeccos PBS (No Ca/Mg) at 0.4 ml/min. The column was followed in-line by a Dawn Heleos-II light scattering detector (Wyatt Technologies) and an Optilab-Rex refractive index monitor (Wyatt Technologies). Data was analysed using ASTRA 5.3.4.14 (Wyatt Technologies) assuming a dn/dc value of 0.186 ml/g.

2.2.22 Isothermal titration calorimetry (ITC)

ITC was used to initially determine the binding affinity and thermodynamic behaviour of the TolAIII-TolB interaction. Protein was dialysed against appropriate buffer overnight at 4 °C. A typical setup would be protein into 50 mM Tris 50 mM NaCl pH 7.5. Synthetic amide-blocked $TolB_{22-34}$ peptide was dissolved into the same buffer. Particulates were removed by centrifugation (10,000 x *g* for 10 minutes). Sample concentrations were determined by measuring the protein absorbance at 280 nm using a NanoDrop system and UV-Vis. Peptide concentrations were determined by amino acid analysis of a concentrated sample.Typical concetrations involved 3 mM $TolB_{22-34}$ in the syringe, and 0.8 mM TolAIII in the sample cell. An iTC200 (Microcal) was used for all ITC experiments, and experimental data were analysed using built-in Origin software. Typically, a titration involved peptide in the syringe, and dilute protein in the cell at 30 °C. Experiments were conducted using 2 µL injections at 10 minute intervals for 20 injections.

Fitting of ITC data was performed by iterative standard Marquardt least squares to improve the fit of ΔQ_i (heat released from the *i*th injection), K_a , stoichiometry n and enthalpy. In

the case of TolAIII-TolB$_{22-34}$, the signal to noise ratio with the maximal achievable concentration of protein and peptide samples was insufficient to determine a binding constant. The thermodynamic character of the interaction suggests an exothermic association. Subsequent titrations by NMR and fluorescence anisotropy suggest the concetrations in these ITC would need to be 20-50 fold higher to determine a K_d measurement.

2.2.23 Protein crystallization

Protein crystallization trials were set up between TolAIII and TolB$_{22-34}$, in an attempt to obtain a crystal complex. Purified TolAIII was buffer exchanged to 50 mM Tris pH 7.5, 50 mM NaCl, and concentrated to 10 mg/mL. INDEX and PROPLEX screens were set up for 1:1 and 1:4 molar ratios of protein:peptide. Peptide was reconstituted in MilliQ water and freeze dried 2 times to remove residual trifluoroacetic acid (TFA) before use. Protein and peptide solution were then plated using a Mosquito robot at room temperature and 4 °C. The crystals trials were prepared using a hanging droplet setup with concentration by vapour diffusion. At day 3, two conditions produced diffracting crystals: 0.54 µM TolAIII/ 0.54 µM TolB$_{22-34}$, in 50 mM Tris pH 7.5, 50 mM NaCl, and the second at 0.15 M ammonium sulfate, 0.1 M MES, pH 5.5, 25% w/v PEG 4000. Crystals were retrieved with Dr. Justyna Wojdyla and were frozen. Preliminary data collection was performed at Diamond Light Source. However, the preliminary indexing and molecular replacement using the *P. aeruginosa* TolAIII crystal structure *(*PDB:1LR0)(Witty *et al*, 2002) by Dr. Wojdyla suggested the crystals did not contain protein complex. Alternative screening condition around the two conditions were unsuccessful in the generation of any crystals.

2.2.24 Fluorescence anisotropy spectroscopy

To determine a precise binding constant for the TolAIII-TolB$_{22-34}$ interaction, fluorescence anisotropy was used to observe the change in overall anisotropy of a reporter (TolB$_{22-34}$) when in complex with TolAIII. A 96-well black absorbent 100uL plate was used for 40 µL fractions to set up a titration curve, with n=3. Test fluorescence measurements indicated low signal change in test binding experiments between TolAIII and TolB$_{22-34}$. Fluorescence anisotropy was used for all subsequent fluorescence binding assays. In this method polarized light is used to probe the overall anisotropy of a fluorophore, as it will have unequal emission along different axes of polarization. As the anisotropic behaviour changes, the polarization signal changes. This was then quantified in correlation to a titration curve and a binding constant retrieved. All data analysis

performed in SigmaPlot 12.0, using a non-linear regression dynamic curve fit using the quadratic equation:

$$f = \frac{A_T + B_T + k \pm \sqrt{(A_T + B_T + k)^2 - 4A_T B_T}}{2}$$

where: f = Concentration of TolA-TolB, A_T = Total concentration of TolA, B_T = Total concentration of TolB, and k = K_d

A plate reader or fluorescence spectrophotometer can provide polarized light and detect polarized light. Experiments were carried out using a CLARIOstar plate reader, and a Fluoromax-4 spectrofluorimeter (Horiba JobinYvon). Titrations were carried out over appropriate concentration ranges between 0 and 1.2 mM. The fluorophore target was a fluoroscein isothiocyanate-labelled TolB (22-33) peptide (Proimmune), with a molecular weight of 1,832 Da. FITC excites at 495 nm and emits at 519 nm. All experiments conducted in 50 mM Tris 100 mM NaCl pH 7.0. All experiments were performed in light-blocking tubes to minimize any premature quenching. Small bubbles were manually removed by use of a flamed sterile needle. Note that the position of the fluorophore was 4 amino acids away from residues involved in interactions with TolAIII. This site was chosen based on the TolAIII-$TolB_{22-34}$ solution structure so as not to interfere with binding but remain close enough such that a fluorescence change would register.

2.2.25 Stopped-flow Fluorescence anisotropy

Stopped-flow anisotropy was used to assess the on and off rates of TolB for TolAIII. Experiments were conducted on an Applied Photophysics SX20 set up for 1:1 single mixing and thermostated using a circulating water bath at 25 °C. All stopped-flow experiments were carried out in 50 mM Tris pH 7.0, 100 mM NaCl. To observe anisotropy of FITC fluorescence the Stopped-flow apparatus was set up using an SX/FP polarization accessory. FITC was excited at 470 nm and its emission was monitored above 515 nm using cutoff filters at both emission channels in the T-mode. Monochromator entrance and exit slits were set to 2 mm (bandpass = 2 mm x 4.65 = 9,3). A total of 4,000-10,000 data points were collected over each reaction. At least 4 anisotropy traces were collected for each reaction and averaged for each time-dependence fit. All anisotropy traces were fit to a single exponential rate equation. Each reaction concentration was collected in triplicate, using recombinant protein from different protein purification runs. Linear regression analysis was done using SigmaPlot, to fit the dependence of k_{obs} on the concentration of TolAIII.

Titrations were performed at increasing concentration of protein relative to a constant concentration of peptide at 1 µM. Experiments were typically performed with 4-10 second data collection periods. Experiments were also performed with up to a 250 second data collection period, and also at 10 °C, to ensure no aberrations or slow activity were detected.

2.2.26 High performance liquid chromatography (HPLC)

Degassed solutions for HPLC purification were prepared using Buffer A: 5% v/v Acetonitrile, 0.01% v/v TFA, Buffer B: 95% v/v Acetonitrile, 0.01% v/v TFA. All experiments were performed using an AKTA Purifier. Each run was preceded and followed by a blank injection (MilliQ, 0.1% TFA). All protein was spun 10 minutes at 16,000 x *g* to remove any particulates and loaded using a 2 mL injection loop.

2.3 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance experiments were set up to study in detail the structural and biophysical characteristics of the TolA-TolB interaction. Protein used in these experiments was typically 0.5 mM-1.0 mM ^{15}N TolAIII, purified from expressed, cleaved Im9-psTolA3 fusion protein as described previously. All experiments were conducted in NMR buffer as described previously unless otherwise indicated. Protein masses were verified to 2Da accuracy using MALDI Mass Spectroscopy. In sample preparation, all samples first concentrated to 0.5 mM-1.0 mM using a MWCO 5,000 Da spin concentrator (Vivaspin), then dialysed to NMR buffer unless indicated otherwise.

Spectral samples that incorporated synthetic TolB peptide were typically prepared by re-suspending dry lyophilised, peptide in TolAIII previously dialyzed to NMR buffer. Before transfer to a Shigemi tube, D_2O was added to 5% v/v, and the sample was spun at 10,000 x *g* for 5 minutes to ensure no floating particulates. Buffer solutions were prepared fresh before experiments were acquired, and during and after experiemnts, samples were kept at 4 ºC to minimize degradation. Typical sample volumes prepared were 300 µL.

Peptide concentrations could not be quantified by dry weight, so saturation of the complex was determined by the disappearance of peaks characteristic of the unbound state, observed both by 1D ^{1}H and 2D ^{15}N HSQC experiments. Dry peptide was observed to strongly reduce the pH of unbuffered solutions, suggesting the purification process retained some TFA salt as dry weight along with purified peptide. To reach ~95% saturation, for example, 2.9 mg of peptide by

dry weight was added to 300 µL 0.61 mM TolAIII. In all experiments involving peptide, peptide was added to dissolved protein as dry lyophilised powder so as not to cause dilution. Global fitting for representative peaks across titration points of HSQC spectra using the program xcrvfit was calculated (http://www.bionmr.ualberta.ca/bds/software/xcrvfit).

2.3.1 NMR spectroscopy

Experiments were collected on OMEGA 600, 750 and 950 MHz spectrometers, as well as a Bruker 600 MHz spectrometer. All data were collected at 20 °C sample temperatures, All spectrometers were equipped with triple resonance ^{1}H, ^{13}C, ^{15}N probes, and the Bruker 600 MHz instrument has a cryo-probe.

2.3.2 Data processing

All data processing was conducted using NMRPipe (Delaglio *et al*, 1995) either on a Sun computer system or Linux Fedora-based system. Processed spectra were visualized and assigned using the CCPN software suite (Vranken *et al*, 2005). Backbone amides of a TolAIII-x6His construct, (identical to TolAIII but with a C-terminal poly-His tag) with similar spectra to $TolAIII_{227-347}$ were previously assigned to completion by Eoin Cassels and Dr. Karthik Rajasekar (Cassels, 2012). Spectral assignment for TolAIII was completed without automatic transfer of peak assignments, however unique and well resolved amide assignments (approximately 60% of total) provided a useful comparison point in chain assignment. All TolAIII assignments made with assistance from TolAIIIx6His assignments were subsequently cross-verified by the CBCA(CO)NH and CBCANH, which were used to assign the rest of the protein *de novo*. All assignments of $TolB_{22-34}$, TolAIII in complex with amide-capped and uncapped $TolB_{22-34}$, $TolB_{22-34}$ in complex with TolAIII were completed *de novo.*

2.3.3 Titration of $TolB_{22-34}$ to TolAIII, monitored by NMR spectroscopy

To determine the affinity of the TolA TolB complex, a titration of synthetic TolB peptide against TolA was conducted. All samples were prepared in NMR buffer as previously mentioned, at pH 7.6. Protein concentration was determined by absorption at 280 nm. The pH was verified by external minielectrode before each spectral acquisition. Two stock solutions of unsaturated and saturated TolAIII were prepared: 10-points were obtained serially, by removing 15 µL of 0.61 mM TolAIII from a 300 µL solution and replacing with TolB peptide-saturated 0.61 mM TolAIII from the earlier titration point. pH was measured between titration points externally using a

minielectrode and adjusted through serial 1-2 µL additions of stock 0.1 M NaOH, 0.1 M HCl, or 0.5 M NaOH as necessary. All experiments were conducted in a 300 µL capacity Shigemi tube, which was re-used between experiments for consistency.

2.3.4 Titration of TolB to TolAIII, monitored by NMR spectroscopy

For experiments involving the binding of *wt* TolB to TolAIII, 5.0 mL of 14.5 µM TolB was concentrated with a triple washed 5,000 Da MWCO spin concentrator to approx. 10 µM (quantified by UV_{280nm}). 30 µL of 1.1 mM ^{15}N TolAIII was added to 270 µL of concentrated TolB. The sample was prepared in 20 mM phosphate, 100 mM NaCl, pH 7.0 (checked by minielectrode), with D_2O added to 5% v/v. A 14 h HSQC was collected on the 750 MHz spectrometer at 20 °C using a 300 µL capacity Shigemi tube.

2.3.5 Triple resonance assignment

Spectra collected for triple resonance assignment are listed in the appendix. Spectra were assigned to completion through a combination of through-chain and through-space experiments. Samples were prepared as follows: 0.8-1.0 mM ^{15}N or ^{15}N ^{13}C TolAIII, 1-3 mM $TolB_{22-34}$ in NMR buffer and 5% v/v D_2O in a 300 µL Shigemi tube. All experiments collected at 20 °C. Assignments of the main-chain NH, N, Cα, and Cβ resonances were made based on two-dimensional HSQC and three-dimensional CBCANH and CBCA(CO)NH experiments (Muhandiram & Kay, 1994). Carbonyl assignments were made with an HNCO experiment. The NOESY-HSQC experiment relies on the Nuclear Overhauser Enhancement (NOE), which arises from dipolar interactions between spin systems close in space. Using through-bond (TOCSY-HSQC) and through-space (NOESY-HSQC, HSQC-NOESY-HSQC) correlations, the assignment strategy involved constructing a 'map' of proximal spin systems, and using a combined approach of identifying dNN, dαN and dβN NOE effects, sequential spin system chains have been constructed. Comparison of Hα/β resonances with expected random coil chemical shifts helped reduce ambiguity in assignments. To assist with poorly dispersed regions, as well as to provide useful information about sequential backbone connectivities, a 3D HSQC-NOESY-HSQC was used to observe dNN NOEs.

2.3.6 Residual Dipolar Coupling experiments

To determine if the overall fold of the TolAIII molecule in $TolB_{22-34}$-bound and unbound forms was consistent with the *P. aeruginosa* TolAIII molecule, 0.8 mM ^{15}N TolAIII was aligned in

liquid crystal media to perform residual dipolar coupling (RDC) experiments. Liquid crystals were prepared in the presence and absence of $TolB_{22-34}$; two independent solutions were made (the second with 10-fold molar excess of peptide). To make a 300 µL sample, 100 µL of 15% w/v C12E6 PEG with 6.5 µL n-hexanol was added to 200 µL of buffered TolAIII solution. Isotropic and aligned HSQC-IPAP spectra (Ottiger *et al*, 1998) were collected with and without the presence of peptide. Spectra were subsequently processed and peak picked using the program PeakPick 2D. Picked picks were manually assigned by comparison to previously assigned HSQC spectra for the bound and unbound states. Using the program FitAlign (which fits the alignment tensor to experimental data), predicted RDCs for PDB: 1LR0 (Crystal structure of TolAIII) were computed. Using the program JNHCalc, experimental RDCs were taken from spectra picked as previously described. PeakPick 2D, FitAlign and JNHCalc are in-house software (C. Redfield, personal communication).

2.3.7 ZZ-exchange spectroscopy

A ZZ exchange experiment was performed (Farrow *et al*, 1994) to try to characterise slow conformational exchange processes. These data were collected on a Bruker 600 MHz spectrometer. ^{15}N HSQC spectra collected with mixing times were chosen to bracket potential useable times (50 ms, 150 ms, 250 ms, 550 ms). Spectra were recorded over 4 days. Samples were prepared as for TolA-TolB triple resonance experiments. Although these data were not used for rate determination, they provided a useful source of information differentiating bound and unbound TolAIII peaks (see appendix).

2.3.8 Hydrogen-deuterium exchange

After lyophilizing TolAIII previously suspended in 20 mM Phosphate pH 7.0 60 mM NaCl, the same volume of D_2O was added to the sample, mixed briefly and loaded immediately into a dry Shigemi tube and the HSQC spectrum was recorded at 20 °C. This procedure was repeated for peptide-saturated TolAIII with peptide in 4-fold molar excess.

2.3.9 SOFAST HMQC and HSQC of co-concentrated ^{15}N TolAIII with *wt* TolB

A solution comprising 4.25 mL of 14.5 µM *wt* TolB was added to 80 µL of 1.1 mM ^{15}N-TolAIII and co-concentrated in a triple-washed 5,000 Da MWCO spin concentrator (Vivaspin) to approximately 100 µM TolB. Samples were suspended in NMR buffer. A 7 hr HSQC and a 7 hr SOFAST HMQC were performed on a 750 MHz Omega Spectrometer at 20 °C for initial test

scans. In the HMQC two residues at 10.30 and 10.65 ppm (^{1}H) are broadened relative to the HSQC; this peak fall-off is due decreased irradiation of peaks at the edge of the spectrum in the HMQC, not a change in the number of residues in TolAIII. The aim of this experiment was to demonstrate chemical shift changes in TolAIII on binding of full length *wt* TolB.

An additional attempt to concentrate TolB was tried, involving a solution comprising 5.0 mL of 14.5 µM TolB concentrated with a triple washed 5,000 Da MWCO spin concentrator to approximately 100 µM (quantified by UV_{280nm}). 30 µL of 1.1 mM ^{15}N TolAIII was then added to 270 µL of this concentrated TolB sample. A 14 h HSQC was collected in a 750 MHz Omega Spectrometer at 20 °C using a 300 µL capacity Shigemi tube.

2.3.10 Filtered NOE spectroscopy

Isotope-filtered NOESY experiments selectively remove intramolecular NOEs from protons attached to ^{13}C or ^{15}N nuclei to other protons on ^{13}C or ^{15}N nuclei. Remaining NOEs are between the protons of ^{13}C or ^{15}N nuclei and protons on ^{12}C or ^{14}N nuclei. In these experiments 0.8 mM ^{15}N^{13}C TolAIII was prepared in NMR buffer, and added to 3.0 mM synthetic TolB peptide. Most NMR spectra are recorded to select the signals of protons based on the presence or absence of large scalar couplings across one bond to heavy nuclei, such as ^{13}C or ^{15}N (Nietlispach *et al*, 2004). Pulse sequences that remove proton resonances correlated with those ^{13}C or ^{15}N nuclei from the spectra are known as *filtered* (or rejected) experiments. Those that select for these proton resonances are known as *edited (*or separated) experiments (Nietlispach *et al*, 2004). This principle, using a low-pass J-filter (known as *purging*) is used for the 3D ^{13}C,^{15}N isotope-filtered ^{13}C-separated NOESY experiment, which assists in distinguishing intermolecular NOEs from intramolecular NOEs in protein(labelled)-peptide(unlabelled) interactions (Zwahlen *et al*, 1997; Ikura & Bax, 1992).

Intermolecular data between a protein and ligand is already contained within 2D/3D ^{13}C or ^{15}N separated NOESY experiments. Assignments for protein can easily be determined via conventional triple-resonance experiments (CBCA(CO)NH/CBCANH/HNCA etc.). Direct comparison with ^{13}C,^{15}N isotope-filtered ^{13}C-separated NOESY experiments (Zwahlen *et al*, 1997) will distinguish protein-peptide and peptide-peptide NOEs, since all ^{1}H-^{1}H NOEs within ^{13}C,^{15}N TolAIII or between ^{13}C,^{15}N TolAIII and $TolB_{22-34}$ would be removed, allowing for the measurement of distance contacts within $TolB_{22-34}$. Assignments of these peaks were achieved using a 2D

^{13}C,^{15}N filtered TOCSY in tandem with a ^{13}C,^{15}N isotope-filtered ^{1}H-^{1}H NOESY and were used for solving the tertiary structure of the peptide (Ikura & Bax, 1992; Zwahlen *et al*, 1997). Indications of protein conformation could be confirmed using the Hα chemical shift index, which uses chemical shifts to predict secondary structure (Wishart & Sykes, 1994). NOEs from interactions between an unlabelled ligand ($TolB_{22-34}$) and a labelled protein were confirmed using a ^{13}C-edited/filtered HMQC NOESY experiment (Robertson *et al*, 2009; Lee *et al*, 1994).

2.3.11 ARIA structure determination

To solve the structure of TolAIII in complex with $TolB_{22-34}$, ARIA (Habeck *et al*, 2004; Linge *et al*, 2003) was used for the generation of lowest-energy structural ensembles with Dr. Karthik Rajasekar. Backbone torsion angles were estimated from CA, CO, CB, N and HA chemical shifts using the program DANGLE (Cheung *et al*, 2010). Distance restraints were retrieved from NOE assignments made in ^{15}N NOESY-HSQC, ^{13}C ^{15}N NOESY-HSQC and D_2O ^{13}C ^{15}N NOESY-HSQC spectra. Additional NOE assignments were made in H-H NOESY spectra, including ^{15}N and ^{13}C-edited NOESY spectra, however these assignments were all fully corroborated with crosspeaks in the aforementioned spectra and could be transferred, therefore the set of spectra from which distance restraints were derived represent the sum total of all distance restraints determined in the TolA $TolB_{22-34}$ complex. Structure calculation was performed in ARIA 2.2 interfaced to CNS 1.2. NOE distrance restraints were set as ambiguous restraints in early calculations, peak ambiguity was decreased during eight iterations before a final water refinement calculation. Structure was then validated using PROCHECK with Dr. Karthik Rajasekar.

In each cycle a list of violations was generated from the top 20 lowest energy calculated models, and violations were then assessed and resolved. Peaks were only rejected based on established criteria: for example a peak would be rejected for expected spin diffusion if model distance exceeded 6Å for a given NOE and the NOE was weak (single ring contour), if magnetisation transfer was a likely candidate based on the NOE crosspeak having excessive contribution from an non-returning source verified to not be peptide, malformed peaks, or irresolveable peak overlap.

2.4 Bioinformatics

2.4.1 Survey of Gram-negative bacteria for components of the Tol-Pal system using PSI-BLAST and tBLASTn

An improved method for identifying novel and un-annotated sequences of Tol-Pal in bacterial databases, using EMBO software was established. TolA and B sequences were manually obtained using EnsemblBacteria (EMBL software, data from International Nucleotide Databank: INDSC), which grants fast hierarchical navigation from genome to translated protein. Combined prediction of Tat and Sec signal peptides using computed Hidden Markov Models provided a check for the Sec signal on TolB (Bagos *et al,* 2010). Pro/Ala/Lys-rich TolAII sequence and nearby sequences similar to Pal TolQ or –R were all used in identifying the *tol-pal* operon.

6,909 assembled genomes are available in the NCBI database. Previous attempts to categorise Tol-Pal relied on matching sequences with γ-proteobacterial proteins. Given the low sequence similarity however between Tol proteins (15-30%) between species this made thorough documentation of Tol-Pal ubiquity laborious and in some cases impossible. So far no automated approach has been published that would assist with BLAST or HMMER searches to locate Tol-Pal in an accurate but comprehensive manner. This manual method combines targeted manual searches of assembled databases to identify both *tol* and *ton* operons in a genome, the rationale being that if two “Tol-like” genomes are found, the probability of one being Tol-Pal is increased. A “Tol-like” genome encodes for a protein with homology to TolA, two proteins with homologies to TolQ and TolR, and a WD40-repeat containing protein with a clear N-terminal Sec signal sequence.

The NCBI database of fully assembled genomes was partitioned by phylum, subdivided into class and then order. Rounding up, representative species from each order in a ratio corresponding to 1% the total number of available genomes were picked. eg. for Firmicutes, where 1139 genomes are available, 617 of these genomes are from the class Bacilli, so 6 species from Bacilli were searched. TolB, TolQ, TolR and TolA were searched using PSI-BLAST individually, and results were compiled for subsequent analysis for operon conservation across 91 sequences diversely representing all bacteria. The sensitivity of a manual search for Tol-Pal proved necessary to identify the presence or absence of a Tol-Pal genome in these representative species due to the high degree of similarity between Tol, Ton, Exb and Mot proteins. This

database of sequences could in future be used to tailor a broader automated search across the remaining ~2,000 assembled genomes in the NCBI database.

Upon identification of TolA and TolB proteins, PRED-TAT was used to identify TolB Sec-signal and Tat signal peptides using computed Hidden Markov Models (Bagos et al 2010 Bioinformatics.). The subsequent 20 amino acids were aligned to compare the intrinsically disordered region of TolB. Clustal sequence alignment of these sequences in comparison with TolA C-terminal region allowed for the identification of conservation of the TolA-B interaction.

The first 12 amino acids of TolB for a representative spread of TolB across all bacteria were then aligned. TolB alignment by clustalOMEGA reveals a [I/V/L]X[I/V/L] binding motif. Jpred 3 when used to predict the secondary structure elements for the final 50 amino acids of TolAIII shows the position of the third β-strand in TolAIII. A sequence-based alignment of the TolAIII C-terminus consensus sequence with the consensus sequence of the TolB N terminus provides the location of the conserved binding interface on TolA and TolB via conserved residues.

3. Biophysical characterisation of *P. aeruginosa* TolAIII

3.1 Introduction

Although the TolA protein has been studied in a range of Proteobacteria, structural detail for this molecule is restricted to information about the C-terminal globular domain, TolAIII. In isolation, *E. coli* TolAIII has been structurally resolved by crystallography and by solution (Deprez *et al*, 2005; Li *et al*, 2012), and crystallographically for the *P. aeruginosa* variant (Witty *et al*, 2002). Structures for TolAIII have also been reported in complex with non-endogenous binding partners. The most recent is a crystal structure of *V. cholerae* TolAIII in complex with the CTXφ pIII domain of a CTX bacteriophage (Ford *et al*, 2012). A crystal structure for *E. coli* TolAIII in complex with the binding domain from the filamentous phage Ike has been published (Jakob 2012). Crystal structures are also available for the TolAIII interaction with colicin N, colicin A (Li 2012), M13 G3P (Lubkowski 1999), FD G3P, and IF1 G3P protein (Lorenz 2011). The structure of TolAIII in complex with an endogenous binding partner has never been solved.

The globular TolAIII domain consists of a single ß-sheet with exposed edges and 4-5 clustered alpha helices, shown in Chapter 1, Figure 1.6. All known TolAIII structures share these common structural features, despite low sequence identity (20%). TolAII and TolAIII are separated by a poly-Gly linker in *E. coli*. In *P. aeruginosa*, TolA does not possess this linker, therefore in the corresponding crystal structure the N-terminus of the construct begins after the last poly-lysine repeat in TolAII (Witty *et al*, 2002). This means that the *P. aeruginosa* TolAIII construct used by Witty *et al.* contains a small undefined region of TolAII in addition to the globular domain of TolAIII.

This results in a key structural difference between available *E. coli* and *P. aeruginosa* structures. In the *E. coli* TolAIII crystal structure, the N-terminus begins at helix 1 at residue 334. In the *E. coli* NMR solution structure, this construct begins at residue 325, and the first ten amino acids are intrinsically disordered. In *P. aeruginosa* however, a construct is used that includes 30 amino acids before the equivalent point in *E. coli* where helix 1 begins. In the *P. aeruginosa* this region is entirely alpha-helical. It has been suggested this is due to the passing of this region close to a crystallographic axis and the formation of a self-complementary interaction, however subsequent circular dichroic data and X-ray scattering data suggest the area is possibly entirely helical in longer constructs as well (Witty et al., 2002).

The core of the TolAIII molecule is a twisted, open antiparallel kinked β-sheet, containing 3 strands. The hydrophobic core is composed of the side-chains of one face of this sheet, and sidechain contributions from coaxial stripes of helix 1, 2 and 3. The hydrophobic core of the molecule has a volume of approximately 1500 $Å^3$.

3.1.1 The function of TolAIII

In the case of TolAIII, genetics and structure appear closely linked, despite divergent primary sequences. TolAIII is the target for parasitization by colicins and bacteriophage, the structure for which some of which have been solved in complex as described above. Colicins are a type of bacteriocin, a plasmid-encoded bacterial protein that targets and kills susceptible bacteria (Cascales et al., 2007; Zakharov & Cramer, 2002). Colicins are composed of three functional domains, including a translocation domain, a receptor binding domain and a pore-forming or nuclease domain. Integral to colicin function is a separate but co-expressed immunity protein that binds with sub-nM affinity to the colicin protein and prevents action until its dissociation. A common mechanism of killing is as follows: the colicin recognizes the target cell by attachment to an OMP via the receptor-binding domain (R domain). This domain then recruits a second OMP for entry and harnesses the internal cell machinery (the Tol-Pal complex) to provide energy for the import process. In the case of colicin A, or colicin N, this involves threading an intrinsically disordered region of the T-domain through an OMP to contact TolAIII directly via a beta-strand complementation mechanism (Hecht *et al*, 2009; Li *et al*, 2012).

Many filamentous phages also hijack the Tol-Pal complex, and like colicins rely on interaction with TolA to effect entry. The interaction between filamentous phages and TolAIII is via a globular bacteriophage domain, the gene 3 coat protein (g3p), which similarly binds via a beta strand complementation mechanism (Lubkowski *et al*, 1999b). In binding either colicins or phage proteins, TolAIII retains its secondary and tertiary fold, and undergoes no major structutral rearrangements on binding (Lubkowski *et al*, 1999b; Li *et al*, 2012). These proteins are invasive proteins, which use TolA as a means of hijacking the PMF. This provides a strong evolutionary pressure towards the maintenance of consistent protein structure whilst encouraging variance in the primary sequence of the protein. There is no clear rationale for how TolAIII however can transduce signals to or from the OM.

TolA could be assisting in this process via a number of different methods. The similarities between TolQ and TolR and the Mot stator proteins have been highlighted in Chapter 1; it is conceivable that TolAIII mechanically drives the pulling of a colicin or phage inside the cell via rotation of TolA in a windmilling pattern spun around in the same manner as the bacterial flagella. TolAII has also been suggested to fold into a coiled coil motif (Witty et al., 2002), so perhaps the proton motive force drives this pre-translocational folding event which in turn pulls tcolicins or bacteriophage through. Finally, TolA may be acting simply as a protein tether that encourages unidirectional movement by a Brownian ratchet motion.(Simon, Peskin, & Oster, 1992). This would likely be contingent on a series of successively stronger OMP binding sites in the colicin translocation domain, a mechanism that has been described in colicin E9 (Housden *et al*, 2013).

3.2 Aims

Despite large sequence divergence, TolAIII structures bear strong structural homology. A primary aim of this chapter was to determine if the overall fold of TolAIII is conserved in all Gram-negative bacteria possessing a *tol-pal* operon. A further aim was also to establish if the incongruities between *P. aeruginosa* and *E. coli* TolAIII structures were crystallographic artifacts, by using solution NMR to probe the dynamics of the *P. aeruginosa* TolAIII N and C-termini. Additionally, a comprehensive bioinformatics study of the Tol-Pal operon in Gram-negative bacteria was performed to determine the conservation of the TolA and TolB proteins.

3.3 Results

3.3.1 An expanded database of Prokaryotes that express Tol-Pal

The only recent study available for the conservation of the Tol-Pal complex in Gram-negative bacteria was by Sturgis *et al.* (Sturgis, 2001) over a decade ago. Out of a representative set of 44 species, chosen from the genomes available at the time, the *tol-pal* gene cluster was found in 10, most of which were Gammaproteobacteria which were disproportionately represented in the genomes studied (12 of 44). Given the massive expansion of bacterial genomic databases in the last decade, (6,415 complete assembled genomes in the NCBI Prokaryota database as of 2016: http://www.ncbi.nlm.nih.gov/assembly/organism/2/all/) it was an opportune time to revaluate the conservation of *tol* genes and establish: in which species is Tol-Pal a conserved complex?

Annotation of the *tol-pal* cluster is often poor, due to the very poor conservation between even closely related species (Sturgis, 2001). The *tolA* gene in particular is highly divergent, likely reflecting its targeting by bacteriophage and bacteriocins, and annotations often mislabel *tolA* as *tonB* and vice versa. Despite the internal contents of the gene cluster being highly conserved, the overall gene context of the *tol-pal* cluster was suggested to be strongly divergent (Sturgis, 2001). Furthermore, the sequence similarity between *tol* proteins is as low between species as it is homologous to gene clusters such as *motA/motB* and *tonB/exbB/exbD.* This causes a high degree of mis-annotation in all currently annotated genomes. A standard BLAST search for TolA, for example, within a genome may give you protein annotated as *tolA,* but which is actually *tonB.*

Following the suggestion from the Sturgis study, that the contents of the tol-pal cluster were readily conserved between species but the sequence similarity was insufficient to readily identify any members immediately by simple BLAST searches, I used the Microbial Gene Context Viewer (MGcV) web tool to rapidly assess upstream and downstream genes of a target in each assembled genome (Overmars *et al*, 2013; Kerkhoven *et al*, 2004). Each of the available 6415 complete and chrosomal genomes was partitioned by class, order and then by strain. Since an automated protocol would produce as many or greater false hits, each genome was searched manually for the *tol-pal* gene cluster. This many genomes was too large a database to search individually, so 96 genomes were chosen, evenly representative of every major phylum and order in bacteria with fully annotated genomic data, as detailed in chapter 2. To maximize success of finding members of the Tol-Pal complex within a genome, *tolB* was used as the primary search sequence, because of its readily identifiable Sec-signal peptide sequence and WD40 motif. Crucially, neither the Ton complex nor the Mot proteins, with which Tol proteins share homologies, contain a WD40-motif containing protein within their respective gene clusters, making this a key determinant in the identity of any given gene cluster identified in a species in the NCBI complete genome database.

First, starting with *E. coli*, the known sequence for mature (Sec-cleaved) TolB was used to BLAST against all complete genomes available for gammaproteobacteria, both using the MGcV tool and via the NCBI web portal (https://blast.ncbi.nlm.nih.gov/Blast.cgi), for each bacterial order. The genomes used for these searches are annotated by NCBI with known genes from GenBank where possible as well as automated *ab initio* analysis (involving the identification of promoter sequences and start and stop codons). The functional annotation and associated

predicted gene identifiers from the publicly available NCBI database of complete genome, were directly accessed via the MGcV interface which collects them via FTP (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/). Although most searches only required *tolB* as the starting seed sequence to find the *tol-pal* cluster, *tolQ, tolR, tolA* and *tonB* were also used for searches when *tolB* did not produce any matches within a genome. All gene 'hits' for *tol* and *ton* proteins were then arrayed such that their genomic context could be assessed and compared.

Each upstream and downstream gene within an overall context of 15,000 nucleotides from any determined hit were then BLAST searched against known sequences for TolQ, TolR, TolA and Pal proteins in the ancestrally closest class as determined using the iTOL bacterial tree of life (Ciccarelli *et al,* 2006). In addition, the NCBI annotations for each genome was searched for all Tol, Ton and Mot-annotated genes.

To determine the presence of a *tol-pal* gene cluster, I used the criteria of a cluster with a gene encoding a WD-40 motif-containing protein with an N-terminal Sec signal, with upstream and downstream genes encoding for proteins with a sequence similarity of at least >50% to two of either TolQ, TolR, TolA or Pal. Once identified, a *tol-pal* gene cluster for a representative set of species within each phylum was saved, and used as the root for subsequent BLAST searches in the next closest bacterial phylum. An example array of is shown in Figure 3.1, comparing the level of annotation across 8 species spanning proteobacteria and other classes, and demonstrating examples of gene clusters with *tol*-like and *ton*-like genes. In figure 3.1, for example, *C. jejeuni* (Epsilonproteobacteria) possesses a gene encoding a WD40 motif-containing protein with an N-terminal Sec signal and sequence similarity to *E. coli tolB*. Adjacent to this is a gene which, when analysed with BLAST, has a sequence similarity too low to be identified with any *E. coli* gene, however its genomic context strongly suggests this gene encodes a TolA-like molecule. A detailed sequence alignment of *tolA* genes is shown in figure 3.3 and 3.4, with a rationale for how the function of these proteins are conserved in view of the large differences in sequence shown in figure 3.5.

All species investigated were compiled into a database of 91 individually searched representative species across all known bacterial orders for which complete genome data was available. This database of identified *tol-pal* gene clusters could in future be used to tailor broader automated searches for Tol-Pal. The results of these searches are presented in Figure 3.2 and

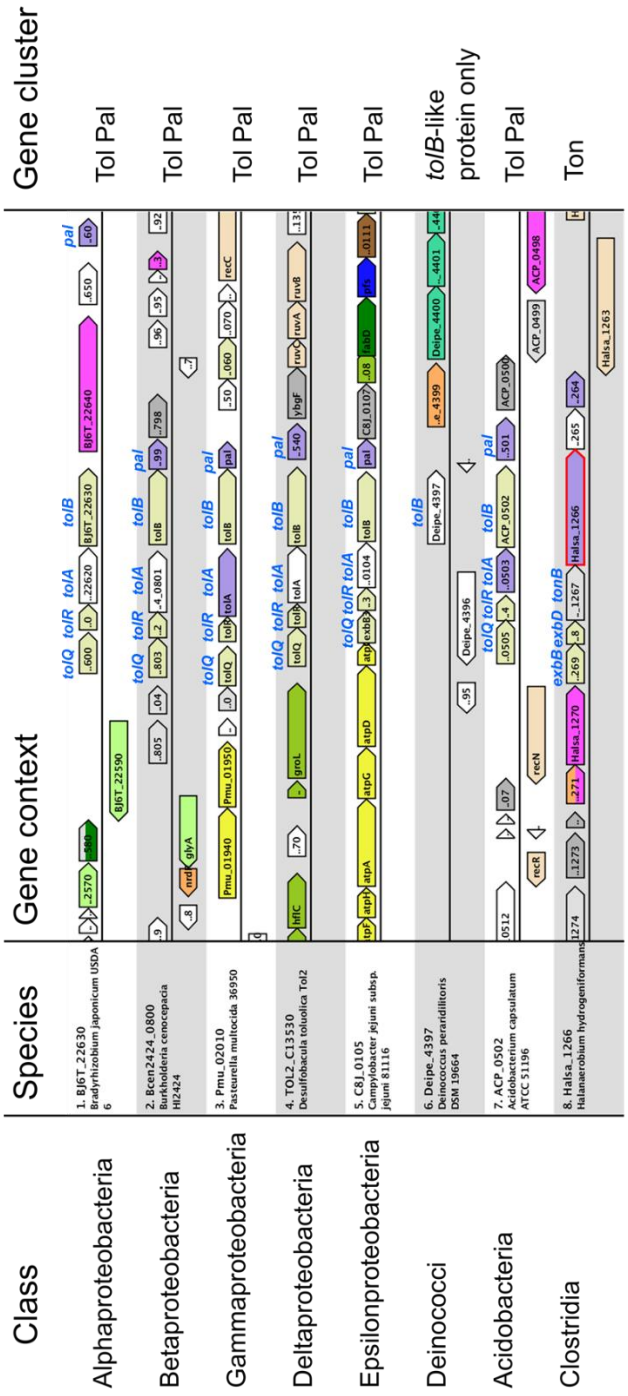


Figure 3.1 **Identification of the *tol-pal* gene cluster within bacterial classes by sequence analysis and by genomic context**. As described in section 3.3.1, genes discovered by BLAST searches and Sec signal analysis were arrayed with their genomic context using the MGcV software (Overmars *et al*, 2013). An example array is demonstrated here. Blue labels indicate the improved or corrected annotation of each gene from this study, determined by sequence analysis of upstream and downstream genes, identification of *tolB* Sec signal regions and secondary structure prediction. For comparison, the *ton* cluster in *H. hydrogeniformans,* from Clostridia (which does not possess any identifiable *tolB* genes) is shown. NCBI gene annotations by comparison are typically thorough in well-studied classes such as Gammaproteobacteria, however in other classes, annotations range from incorrect, (e.g. TolQ in Epsilonproteobacteria annotated as exbB) to incomplete. Deinococci presents an example where a WD40-containing protein with an identifiable Sec signal is predicted to be encoded by the identified gene, but lacks any upstream or downstream genes comparable with *tol-pal.* Clostridia contains genes for proteins with sequence similarity to ExbB, ExbD, TonB, TolQ, TolR and TolA proteins, but lack the genes for a WD40-containing TolB or OmpA-like Pal. This cluster is therefore consistent with a *ton* gene cluster.

the analysis suggests a more nuanced understanding of the range of species in which the *tol-pal* gene cluster is found. Of immediate note is that Tol-Pal is only partially conserved amongst Gram-negative bacteria, within which it is primarily restricted to Proteobacteria and similar species. Most bacteria with trilaminar cell envelopes possess the *tol-pal* cluster, although there are some notable exceptions. Planctomycetes, which until recently were thought to not contain any peptidoglycan in their cell envelope do not have Tol-Pal. With recent findings suggesting that all major species in this phylum possess peptidoglycan layers (Jeske *et al*, 2015; van Teeseling *et al*, 2015), the lack of Tol-Pal is intriguing, and it is unknown which system(s) would replace Tol-Pal function in these organisms.

Tenericutes, thermotogae, and chloroflexi are all atypical Gram-negative bacteria that only possess one membrane (monoderms). These organisms lack a periplasmic compartment and therefore understandably do not use the Tol-Pal system. For most bacteria, either they possess a full Tol-Pal complex, or typically none at all, there were few examples found where a protein with strong sequence similarity to TolB could be found in isolation from the rest of the gene cluster. In *Rhodothermus marinus* in Bacteroidetes for example, a WD-40-like beta propeller repeat-containing protein with a 15 amino-acid predicted N-terminal Sec signal is found (NCBI GI: 345302115). At a different point in the genome, a *ton*-like cluster may be found, but no canonical *tol-pal* cluster. Within the Bacteroidetes/Chlorobi phylum, a complete *tol-pal* cluster is typically found only in the order Chlorobiales. In an expanded search of Bacteroidetes genomes however, 93 of 99 searched contained a gene cluster corresponding to *exbB, exbD* and *tonB*, suggesting the Ton complex has critical function in this organism, being almost completely conserved.

Unexpectedly, evidence of Tol proteins may also be found within some Gram-positive bacteria. In the Phylum Firmicutes, some *Halobacteroides* species possess TolB-like proteins similar to Rhodothermus, also with 23-amino acid long N-terminal Sec signal sequences. No other *tol* proteins could be identified in the Class Clostridia however, so it appears likely that these proteins are a WD-40 repeat containing proteins with strong homology to the TolB sequence. In the Firmicute class Bacilli however the complete genome for *Listeria monocytogenes* possesses a *tol-pal* gene cluster. It is very unlikely for a monoderm Gram-positive bacteria such as Listeria to require the Tol-Pal complex, and the proteins have 60-95% sequence identity with *Vibrio parahaemolyticus tol* genes, hence contamination of this genome information is strongly suspected. No other species with Tol-Pal could be found within Gram-positive bacteria.

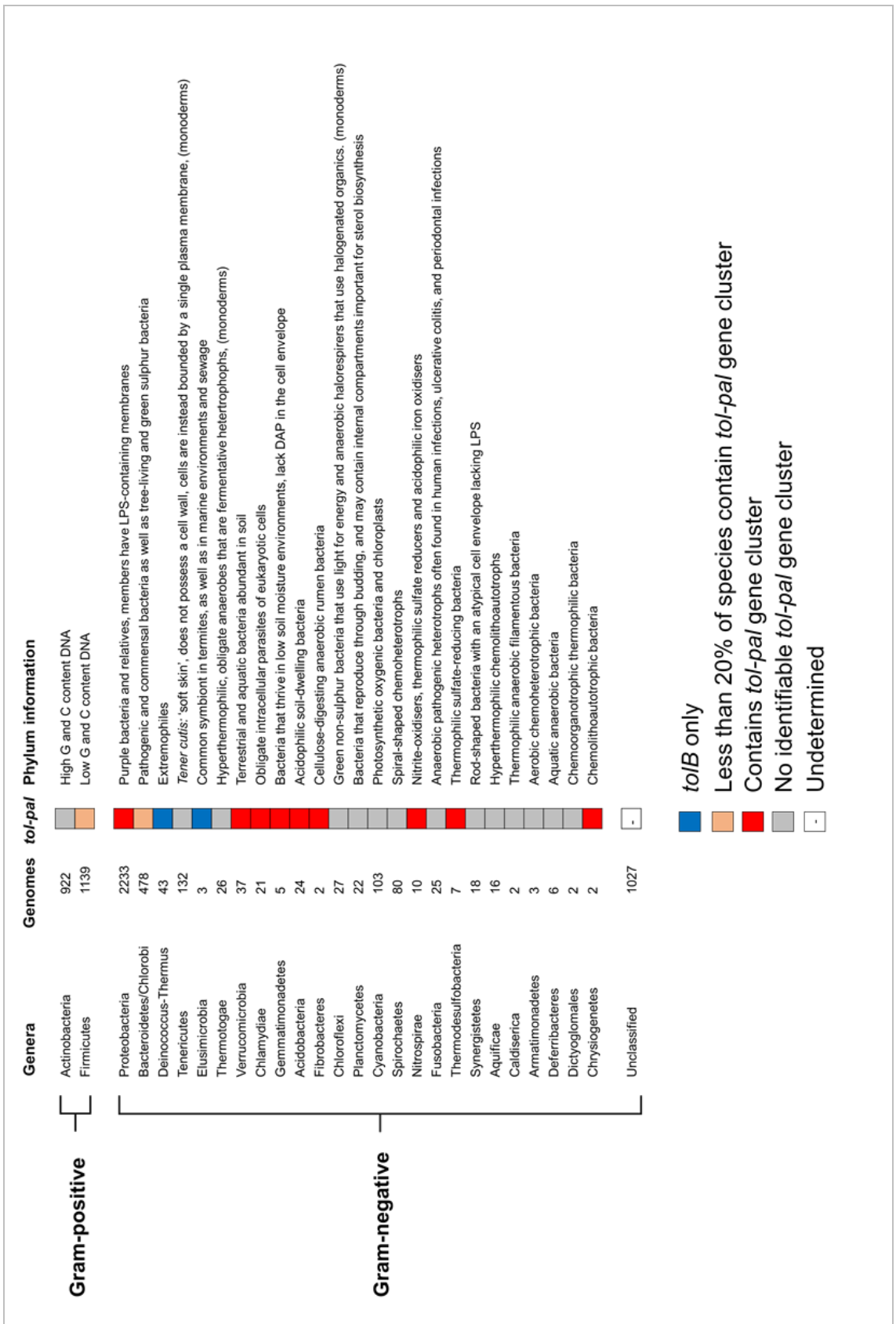


Figure 3.2 **The *tol-pal* gene cluster is conserved in the phylum Proteobacteria, and sparse in the remainder of Prokaryota.** Sequences indicated to contain *tol-pal* genes in bright red mean that the *tol-pal* was discovered in 100% of sequences in these phyla. The *tol-pal* gene cluster appears to be "all or nothing" in terms of inclusion in Gram-negative species. Individual *tolB*-like genes were identified on their own in some species, but no evidence of the rest of a *tol-pal*-like cluster could be identified anywhere in the remainder of the genome. An example of this is demonstrated in Deinococci in figure 3.1.

Elusimicrobia and Deinococcus-Thermus phyla both contained species in which *tolB*-like genes could be found, however no canonical *tol-pal* clusters.

In Chlamydia organisms, which are obligate intracellular parasites, the *tol-pal* cluster lacks genes identifiable as *cpoB* and *ybgC*, however in its place is a gene corresponding to a lytic murein transglycosylase. Acidobacteria similarly lacked a *ybgC* gene, however a *cpoB* gene was conserved.

The concluding observation from this bioinformatics analysis was that the *tol-pal* gene cluster is conserved in Proteobacteria and sparse in the remainder of Prokaryota.

3.3.2 Sequence-based alignment of TolA

TolA sequences are highly divergent, therefore using the complete database of TolA molecules determined from compiling *tol-pal* gene clusters in section 3.2.1, I completed a sequence-based alignment of TolA genes to assess the conservation of the C-terminal globular folded domain, and highlight any conserved structural features. A ClustalOmega alignment of TolA molecules is shown in Figures 3.3 and 3.4. TolAI has a widely variable cytosolic region with a strongly conserved transmembrane helix between positions 50 and 70. No secondary structure information could be predicted for the TolAII region, (not shown) however there is a strong prevalence of 20-40 amino acid long strings of poly proline and poly proline-glutamate sequences in genera like *Rhodobacter* and *Sphingomonas*. Alanine-lysine sequences are also common, however typically in genera like *Pseudomonas or Escherichia* which do not possess long poly-proline regions. In *Pseudomonas* and *Escherichia,* these regions are predicted to have some helical propensity, which is not the case for poly-proline TolAIIs. The functional importance of this difference is not known.

At position 547 in figure 3.4, 58% of sequences have a tyrosine residue, corresponding to the first amino acid in TolAIII whose sidechain makes contacts with tertiary structural elements of the folded globular region. This does not appear to be a vital interaction, as in *P. aeruginosa* this residue is Leu254, and in other species glutamates are also common. Leu254 however makes contacts with Glu320 of helix 3. The next 3 *i*+4 residues are the residues occupying the same coaxial face of helix 1 which have their sidechains buried in the hydrophobic core of the protein. These are all partially conserved (conservation score of >6) and mostly hydrophobic, and are shown in Figure 3.3.

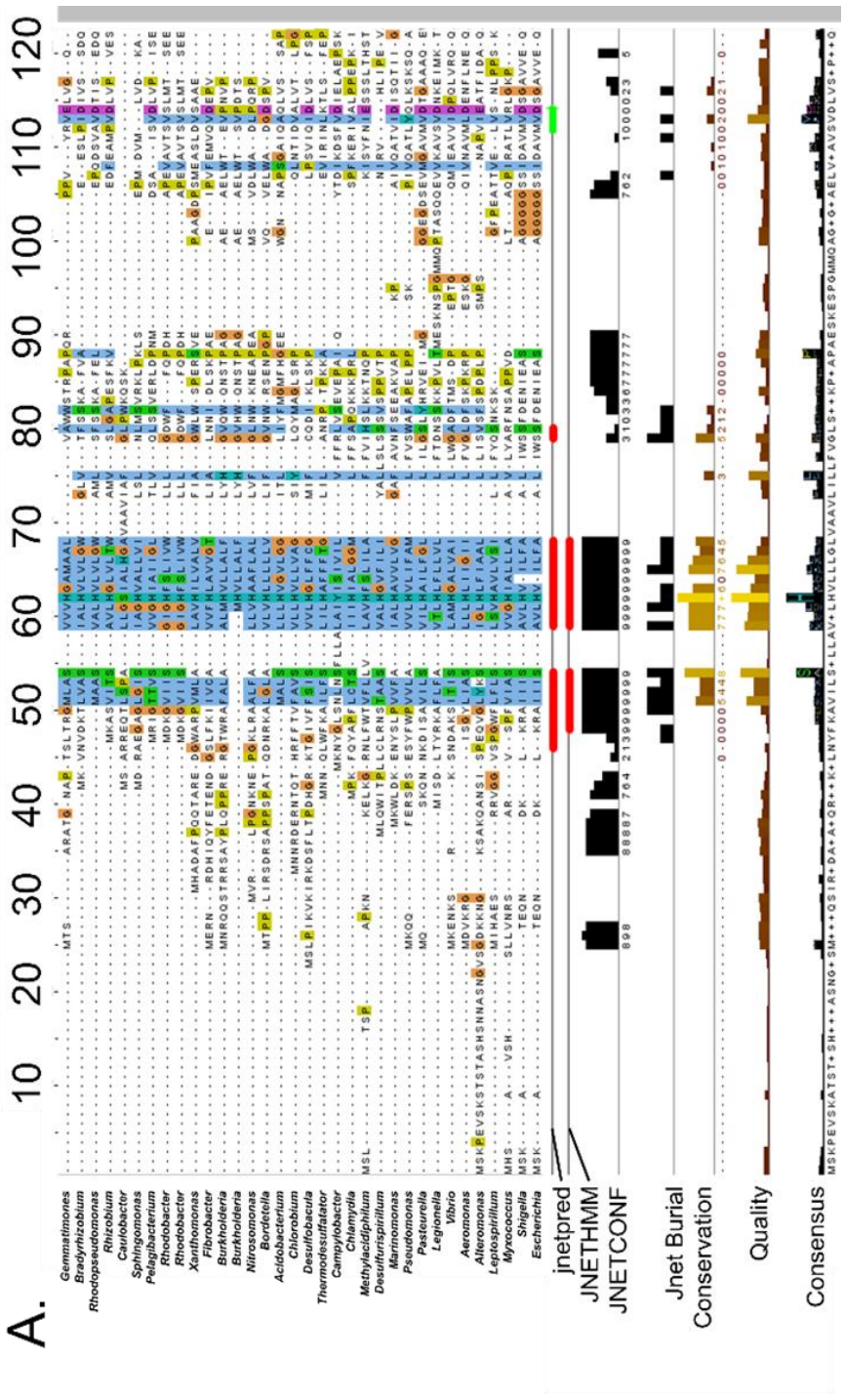


Figure 3.3 **Sequence-based alignment for TolA across Prokaryota reveals conserved secondary structural elements, part A.** The first ~120 residues for TolA are shown for a representative set of species spanning all bacteria. The trans-membrane helix of TolAI is predicted to be strongly conserved at position 50-70. JNetPRED and JNETHMM are consensus predictions for secondary structure, based on a neural network prediction and a hidden markov model profile based prediction respectively. JNetPRED confidence is provided by JNetCONF, a confidence estimate from 0-10 for increasing accuracy. All predicted helices are marked in red, beta-strand marked in green. JNetBurial is an estimation of which residues are likely to form the buried core of a protein. Conservation, quality of conservation, and overall consensus are provided. This figure is paired with figure 3.4, which shows the last ~120 amino acids of TolA. Sequence numbers are given from 0-690, corresponding to the total length of the alignment, each protein variant is typically 300-400 amino acids in length. On the basis of the central region having low sequence complexity, it has not been shown in either figure.

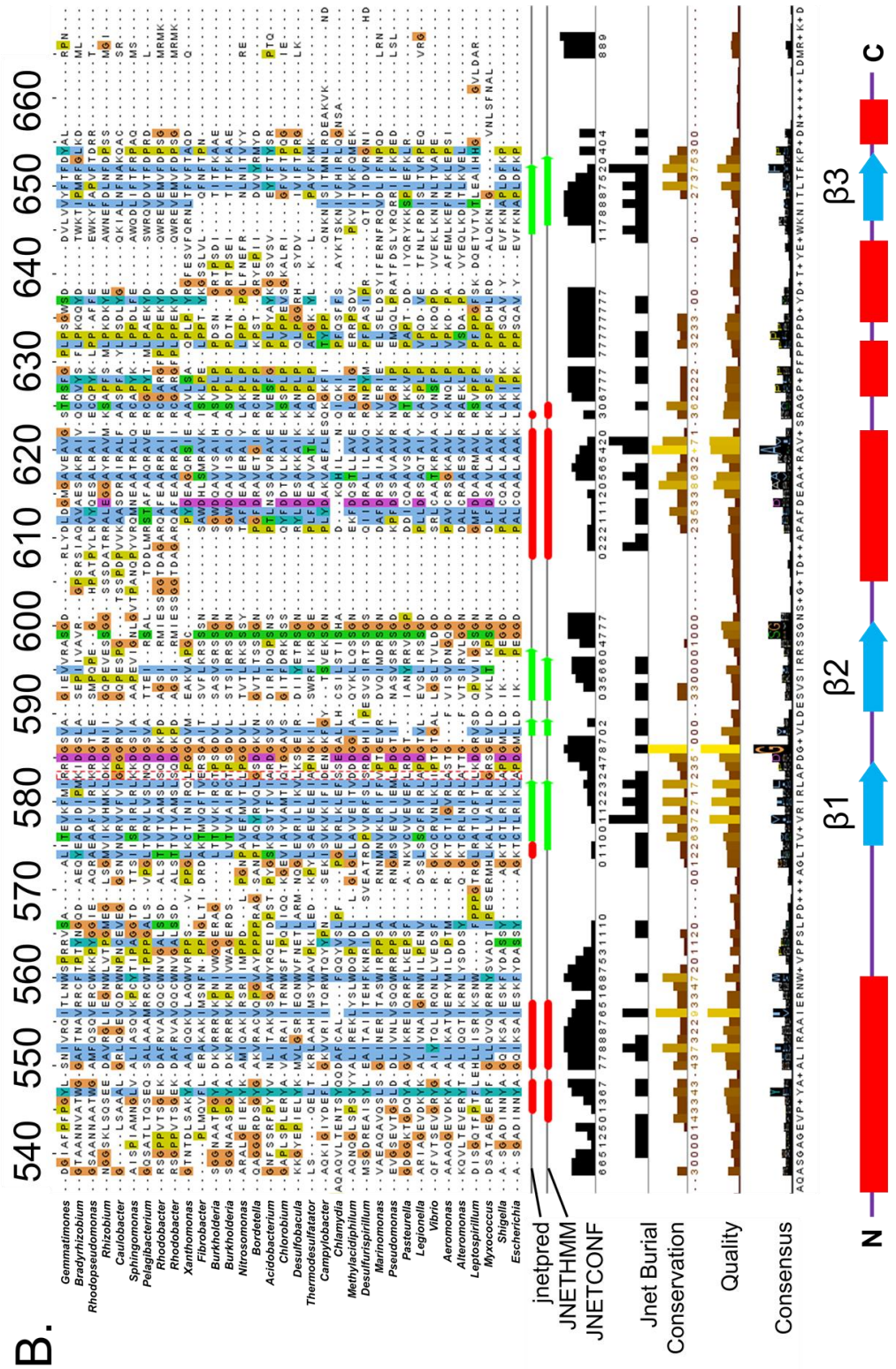


Figure 3.4 **Structure-based alignment for TolA across Prokaryota reveals conserved secondary structural elements, part B.** This figure, paired with figure 3.3 shows the conservation of the last ~120 residues in TolA. This part of the protein comprises TolAIII, and although sequence similarity is poor, some key secondary structural elements can be predicted. Strand 1 and 3 in particular are identifiable by conserved odd residues in each respective strand. The secondary structure for the *P. aeruginosa* TolAIII crystal structure (Witty *et al*, 2002) has been shown in addition to JNetPRED and JNETHMM secondary structural predictions for comparison.

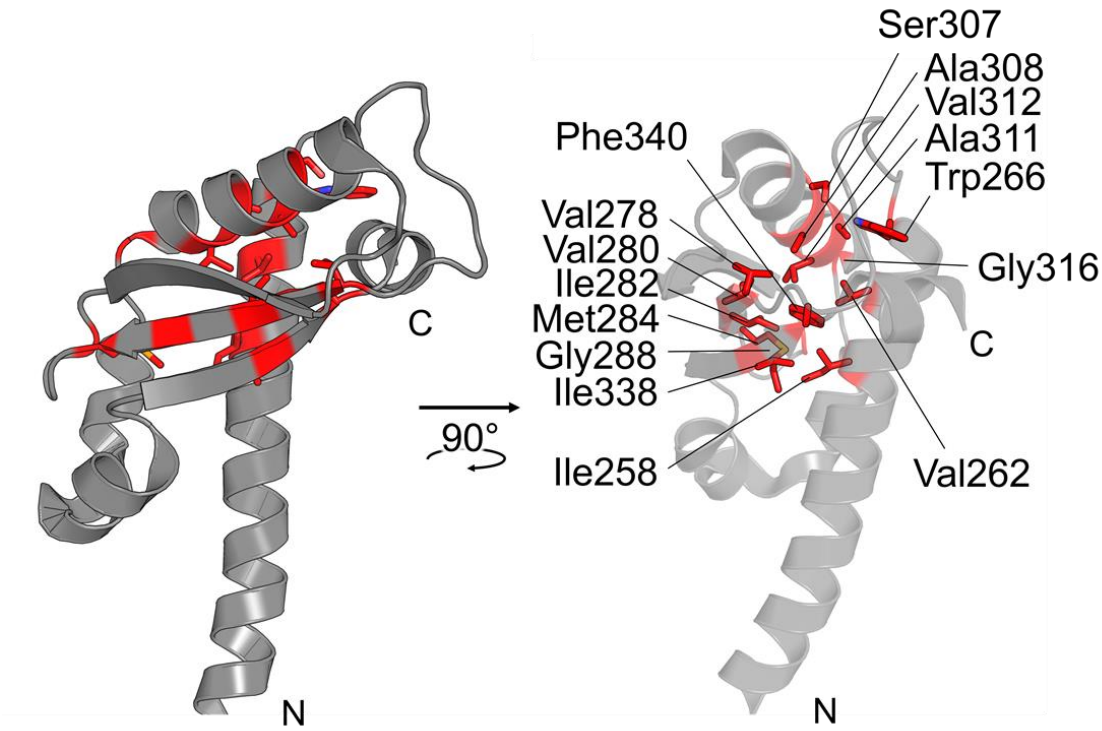


Figure 3.5 **Conserved residues in TolAIII form the hydrophobic core of the protein.** Shown in red are residues with a conservation score of greater than 6 from a sequence based alignment of TolAIII. These colours are superimposed on the *P. aeruginosa* TolAIII crystal structure (Witty *et al*, 2002) and drawn using Pymol (DeLano, 2002).

The major defining structural feature of TolAIII, the antiparallel twisted beta sheet, can be inferred to be nearly fully conserved in all species of TolAIII, as key hydrophobic residues in this sheet are well preserved at positions 574-582 and 648-654. Within these regions are amino acids in the beta sheet whose sidechains face inward and comprise the hydrophobic core of the protein. There is very poor sequence conservation outside of these regions.

3.3.3 Production of *Pseudomonas aeruginosa* TolAIII protein for use in a biophysical investigation

In order to study the physicochemical properties of TolAIII *in vitro,* and prior to a structural determination of TolAIII with its endogenous binding partner TolB in Chapter 4, a suitable expression system for the production of the TolAIII protein needed to be determined. The first successful structural investigation targeting TolAIII was in 1996, when Lubkowski *et al.* determined the structure of *E.coli* TolAIII in complex with the N1 domain of the g3p from Ff bacteriophage (Lubkowski *et al*, 1999b). This crystal structure was determined by a strategy of fusing g3p covalently to TolAIII via the TolAIII N-terminus, and represented the first structural determination of TolAIII in a bound state. Subsequently, the solution structure of *E.coli* TolAIII *in vitro* by NMR spectroscopy in 2005 demonstrated an expression method for a construct that

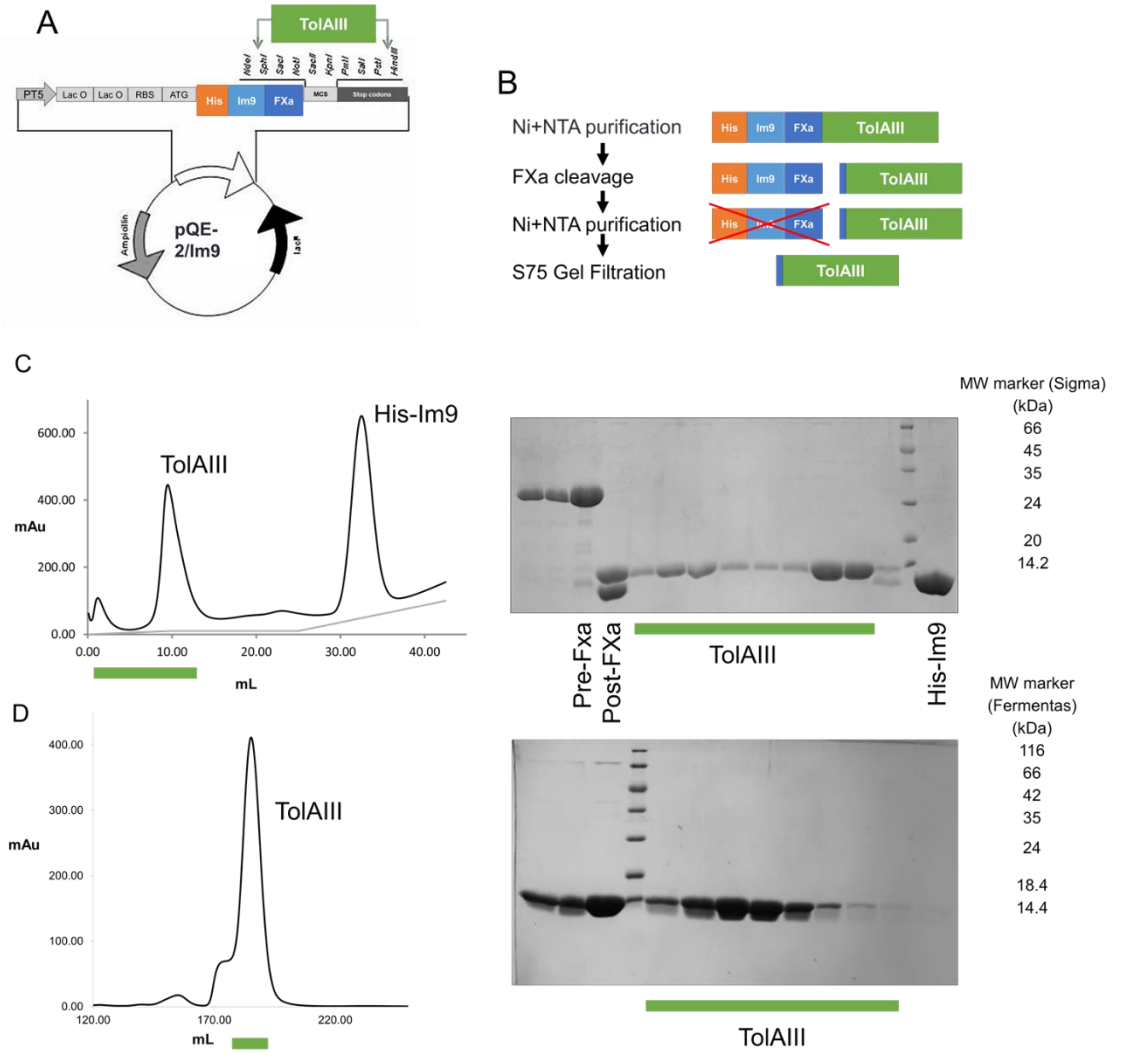


Figure 3.6 **Expression of ^{15}N TolAIII using a high-yield pREN28 pQE-2/Im9 fusion protein vector.** (A) pREN28 pQE-2/Im9 protein fusion vector. (B) General method for optimized TolAIII purification, note that FXa cleavage leaves two residual N-terminal His-Met residues on TolAIII (C) Final Ni-affinity purification step after FXa cleavage shows clear separation of TolAIII from the poly-His tag/Im9 protein. (D) TolAIII purification by gel filtration. SDS-PAGE profile confirms purity, green bar denotes the fractions pooled.

showed minimal structural rearrangement compared to the bound state (Deprez *et al*, 2005). The crystal structure of *P. aeruginosa* TolAIII was determined in 2002 using a construct which necessitated the addition of a C-terminal Histidine tag for purification purposes, as well as a lengthier region of TolA at the N-terminus compared to previous constructs used in *E.coli* studies. The TolAIII crystal structure in *P. aeruginosa* suggested the possibility that the entirety of this N-terminal region was ordered in solution (Witty *et al*, 2002).

Previous investigations in the Kleanthous lab into the use of *E.coli* TolAIII constructs for structural investigation produced protein spectra with signs of degradation and an abnormally large number of backbone resonances in a standard ^{15}N HSQC experiment. Subsequent work

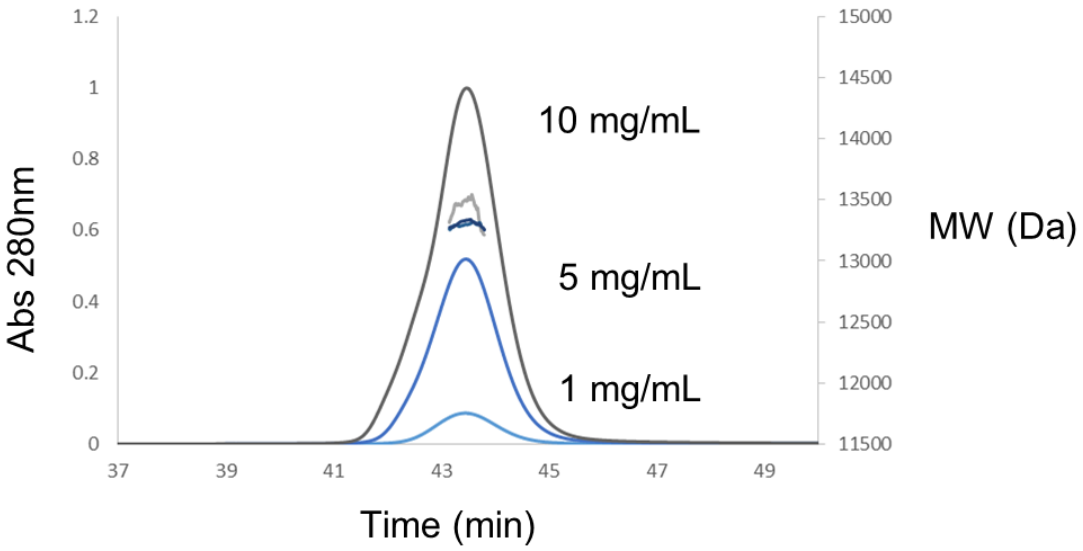


Figure 3.7 **SEC-MALS analysis of TolAIII.** Elution profiles of 1, 5 and 10 mg/mL TolAIII as monitored by the measurement of A_{280nm}. Protein masses were determined by the A_{280nm} refractive index and light scattering data using the protein conjugate analysis within ASTRA software (Wyatt Technologies) and are shown in grey, dark blue and blue respectively. Data were collected and analysed with Dr. Steven Johnson (Dunn Pathology, Oxford).

with a *P. aeruginosa* construct identical to that used in the 2002 crystal structure study (Witty *et al*, 2002) suggested that a peptide representing the N-terminus of TolB caused small chemical shift perturbations near the C-terminus of TolAIII (Cassels, 2012). Furthermore, this construct retained the C-terminal poly-Histidine tag of the original construct used in the 2002 study.

For the purposes of this study, I chose to focus on determining the physicochemical properties of *P. aeruginosa* TolAIII in a construct without a poly-His tag at the C-terminus, so as not to interfere with any potential binding sites. To produce TolAIII, the pQE/2-Im9 expression system was selected for the high yield and N-terminal positioning of the fusion protein tag. An N-terminal tag used for boosting expression and aiding purification would be removed via proteolytic cleavage, with minimal sequence differences to the folded, C-terminal globular region TolAIII. The specific region of TolAIII studied was residues 226-347, and unless otherwise specified, all TolAIII referred to hereafter is this construct.

P. aeruginosa TolAIII residues 224-347 (covalently fused to *E.coli* Im9 protein) containing an N-terminal poly-His tag was purified from BL21-(DE3) cells transformed with pREN28 and grown in LB media. The results of this are shown in Figure 3.6; protein expression and purification steps were as described in Chapter 2. A typical protein yield of 50 mg/(cultured L) was obtained for the pREN28 construct. Factor Xa cleavage of this construct releases TolAIII which was

successfully purified away from Im9 by Ni-affinity chromatography and finished with size-exclusion chromatography. Size exclusion chromatography multi-angle laser light scattering (SEC-MALLS) determined the TolAIII protein was monomeric in solution shown in Figure 3.7 and electrospray ionization mass spectrometry (ESI-MS) verified the molecular weight of TolAIII purified by this method, Expected Mass: 13896 Da, Observed Mass: 13898.6 Da.

3.3.4 Purification of ^{15}N ^{13}C isotope-enriched *P. aeruginosa* TolAIII

With a high-yield expression system for TolAIII determined, I repeated this process with the substitution of minimal (M9) media for LB and the incorporation of ^{15}N and $^{15}N/^{13}C$ sources for the purpose of producing ^{15}N- and $^{15}N^{13}C$-enriched TolAIII. The purification strategy used is detailed in Chapter 2. Isotope-enriched protein was used to record solution-state Nuclear Magnetic Resonance (NMR) spectra, to determine protein fold, regions of flexibility and other structural and dynamic information about the protein for comparison with the TolAIII crystal structure. The aim was to produce a TolAIII protein suitably robust for structural study. This enabled a full structural determination of TolAIII in complex with a binding partner in Chapter 4.

^{15}N-enriched TolAIII was used to collect ^{15}N-HSQC (Heteronuclear single quantum coherence) spectra. The wide chemical shift dispersion, and low number of peaks within the region of the spectrum characteristic of disordered protein, suggested TolAIII was folded in solution. The number of visible peaks corresponded to the expected number of backbone amides. HSQC spectra were collected at 20° C on an Omega 750 MHz magnet in Oxford. This sample was stored at 25°C for a week and test spectra reacquired which confirmed no visible hydrolytic cleavage or appearance of degradation peaks. Standard procedure employed for all subsequent samples was storage at 4 °C. A comparison between TolAIII in the bound and unbound state and an overlay of the corresponding HSQC spectra is discussed in Figure 4.3 in Chapter 4.

3.3.5 *P. aeruginosa* TolAIII is stable at a pH range of 6.95-7.62

In order to colonize the human gastrointestinal tract, organisms like *P. aeruginosa* and *E. coli* must function in a wide pH range. Periplasmic pH range is typically within a range of 6.0 to 7.6, when the pH of the external environment is 5.5-7.5 (Wilks & Slonczewski, 2007). ^{15}N HSQC spectra provide information on increases in regions of disorder and decreases in protein stability reflected in the chemical shift of backbone residues, so the TolAIII protein was tested within a

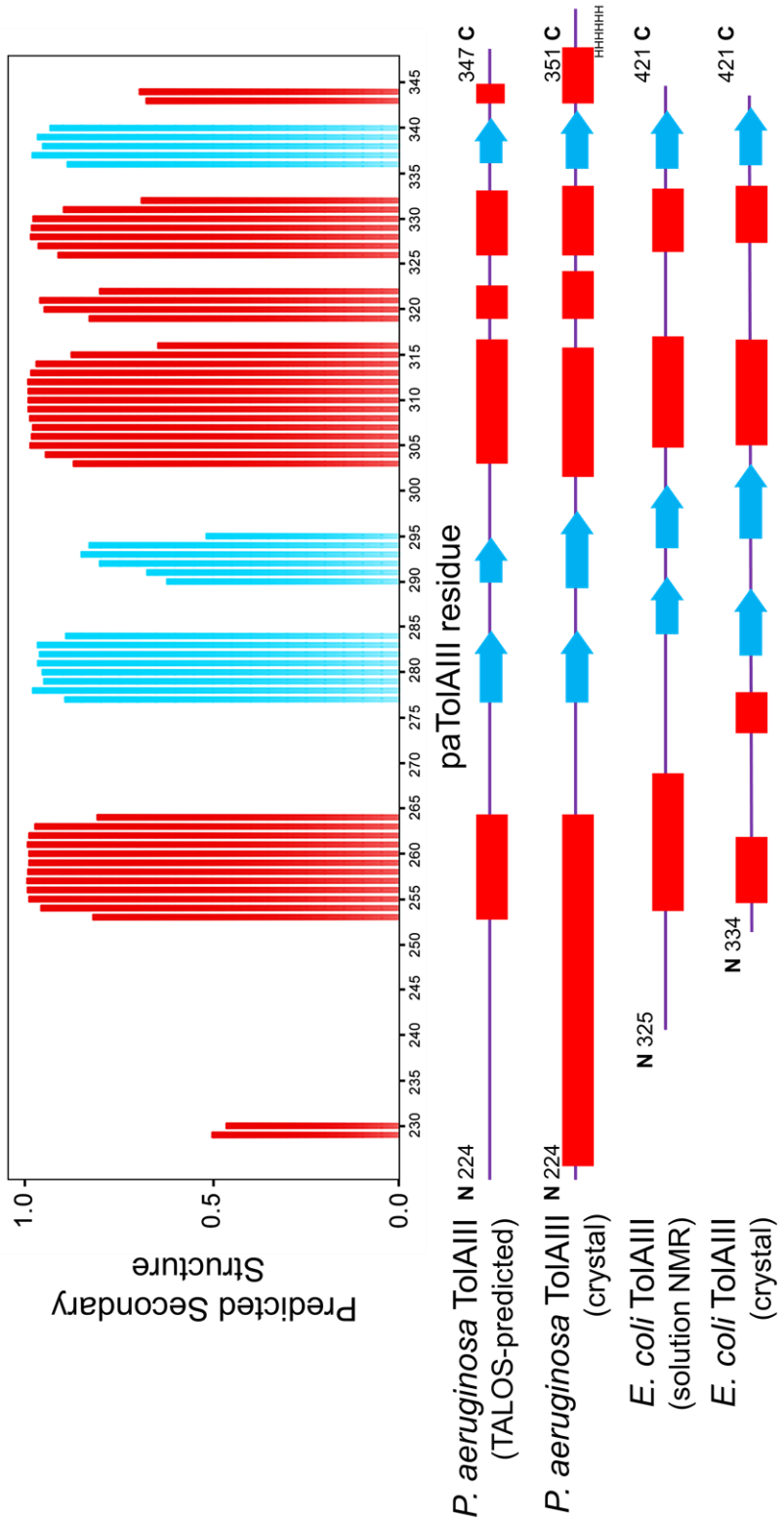


Figure 3.8 **TolAIII secondary structure as predicted by the TALOS+ program in comparison with the secondary structures of solution and crystal states for *P. aeruginosa* and *E. coli*.** TolAIII secondary structure prediction, (blue for β-strand, red for α-helix) is shown as a function of residue number in comparison with the previously determined secondary structures in the *P. aeruginosa* TolAIII crystal structure and *E. coli* TolAIII crystal and NMR structures (Li *et al*, 2012; Witty *et al*, 2002; Deprez *et al*, 2005).

range to assess any effects of pH in a range similar to conditions it may experience in the cell. ^{15}N-enriched TolAIII was used to collect ^{15}N-HSQC spectra above and below the expected physiological pH of 7.0. An overlay of two spectra of TolAIII at pH 6.95 and pH 7.62 in black and red respectively superimposes with minimal (<0.04ppm in the ^{1}H dimension) change in chemical shift for any backbone residues (See Appendix) suggesting that within this pH range there is minimal effect on the overall structure of TolAIII.

3.3.6 *P. aeruginosa* TolAIII chemical shift assignment

NMR spectra were recorded to select the signals of protons with scalar couplings across one bond to ^{13}C and ^{15}N (Nietlispach, Mott, & Stott, 2004). These signals, or 'resonances' correspond to correlations between protons and heavy atoms or between protons. Resonances were then 'assigned': given identities corresponding to the identity of the correlation using conventional triple-resonance experiments (CBCA(CO)NH/CBCANH/HNCA etc.). Representative slices of the assignment process are found in the appendix.

Datasets for HNCO, HNCA, CBCANH, CBCA(CO)NH, HBHA(CO)NH and ^{15}N TOCSY-HSQC, ^{15}N NOESY-HSQC, ^{13}C HSQC and ^{13}C NOESY-HSQC were collected as detailed in Chapter 2. >95% assignment of all backbone resonances was completed. This in combination with full assignment of an HBHA(CO)NH experiment provided a complete set of backbone, Cα, Hα, CO, NH, as well as Cβ and Hβ chemical shifts. Proline chemical shift assignments were determined where possible after completion of *i*+1 shift assignments. ^{15}N NOESY-HSQC, ^{13}C HSQC and ^{13}C NOESY-HSQC were collected in conjunction with HCCH-TOCSY, HCC(CO)NH, (H)CC(CO)NH experiments, providing sidechain assignments for 82% of the protein.

3.3.7 Comparison of the *P. aeruginosa* TolAIII secondary structure to *E. coli* TolAIII in solution

Hα, Hβ, Cα, Cβ, NH, CO chemical shifts for TolAIII as determined by sequential assignment were used to calculate dihedral angles and compute secondary structure using TALOS+ (Shen *et al*, 2009). Since NMR chemical shifts are strongly influenced by regions of local secondary structure and the local chemical environment, TALOS takes the chemical shift data for C, H and N atoms determined in a protein and compares them with a database of proteins to predict backbone torsion angles with an error rate of 2.5%. TALOS uses a neural network to predict protein secondary structure and a Random Coil Index (RCI "order" parameter) (Berjanskii

& Wishart, 2008). This order parameter, termed S^2, indicates regions of predicted protein flexibility, and whilst not direct substitution for measurement of relaxation parameters, has been demonstrated to typically provide accurate estimates for regions of semi-rigid structure and disorder in proteins (Berjanskii & Wishart, 2008). The computed TALOS+ data for the *P. aeruginosa* TolAIII variant suggests secondary structural elements were consistent with those observed in *E. coli* and *P. aeruginosa* crystal structure, with the exception of the presence of an N-terminal helix seen in the *P. aeruginosa* crystal structure. This comparison is shown in Figure 3.9

The aim of the analysis in Figure 3.8 was to establish if the solution NMR data for *P. aeruginosa* TolAIII was more consistent with either the *P. aeruginosa* crystal structure or the *E. coli* NMR or crystal structures. The height of the bars in the secondary TALOS structure plot reflects the probability of each type of secondary structure (red for α-helix, blue for β-strand) based on the neural network calculation and the computed dihedral angles. The predicted 2° structure for *P. aeruginosa* TolAIII computed by TALOS, the 2° structure of the *P. aeruginosa* TolAIII 1LR0 construct (Witty *et al*, 2002) and the 2° structure for *E.coli* TolAIII as determined by solution NMR (Deprez *et al*, 2005) are aligned relative to the 3rd β-strand. Immediately following is the secondary structure as determined in *E.coli* TolAIII via crystallography (Li *et al*, 2012).

Unlike the *P. aeruginosa* TolAIII crystal structure, but similar to *E. coli, P. aeruginosa* TolAIII in solution is disordered at the N-terminus. Helical character in the N-terminal region of *P. aeruginosa* TolAIII as reported by Witty *et al.* is due to crystal-packing forces. This raised concern that the C-terminal helix in TolAIII was also artefactual, possibly stabilised by the presence of a poly-His tag and arrangement in a triskelion-pattern crystal packing unit. The secondary structure predicted for *P. aeruginosa* TolAIII in solution was consistent with a C-terminal helix, however as the prediction is only available for 2-4 residues so it was not on its own conclusive evidence for helix 4.

3.3.8 RCI-predicted order parameters and the $^{15}N\{^1H\}$ NOE (heteronuclear NOE) experiment identify regions of disorder in *P. aeruginosa* TolAIII

As mentioned previously, TolAIII is parasitized by bacteriocin and bacteriophage and one of the two observed binding interfaces in these interactions is adjacent to the C-terminus of TolAIII. With secondary structural prediction suggesting that the N and C-termini of *P. aeruginosa*

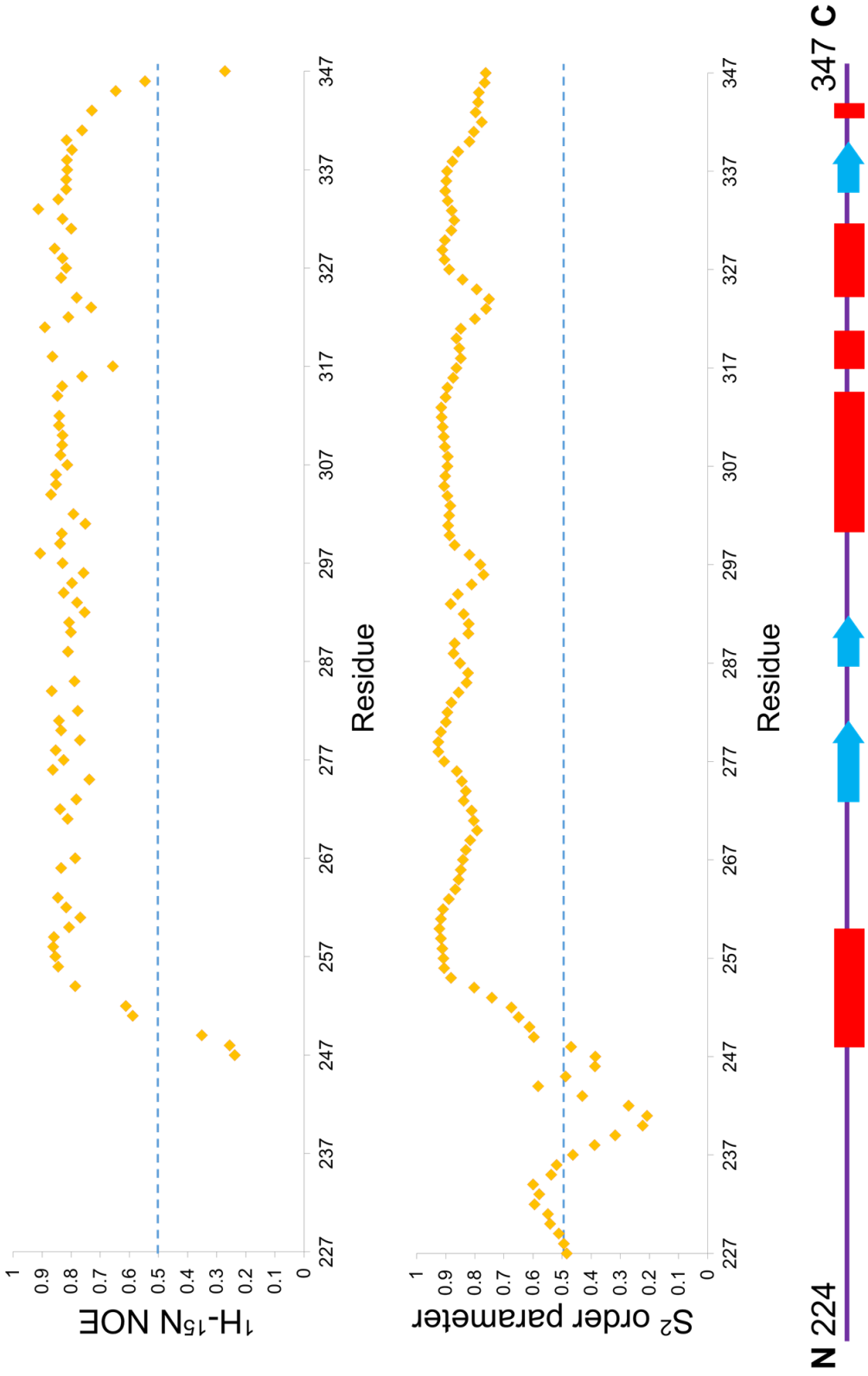


Figure 3.9 **RCI-predicted order parameters and the ^{15}N{1H} NOE (heteronuclear NOE) experiment show the globular region of *P. aeruginosa* TolAIII extends from residue 250-346.** Squared order parameters (S^2) predicted by TALOS+ were compared to measurements of backbone flexibility using a heternuclear NOE experiment. Both the heteronuclear NOE and TALOS+ analysis suggest the N-terminus is disordered from residue 227 to approximately residue 250. The C-terminus of TolAIII appears to possess flexible character for residues 345-347.

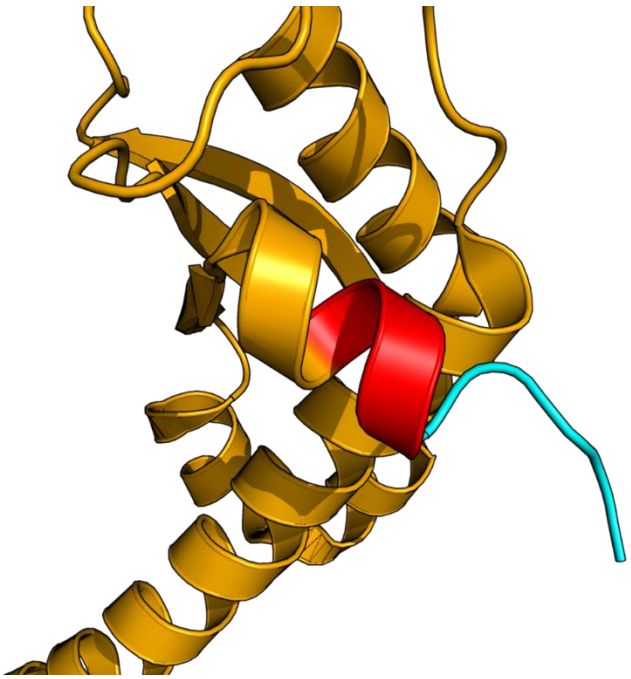


Figure 3.10 **The C-terminus in the *P. aeruginosa* TolAIII crystal structure forms a helix.** Witty *et al*. determined the structure of TolAIII with a C-terminal poly-His tagged construct, (Histidines shown above in teal). This tag is not present in my 227-347 construct of TolAIII, and the data from the heteronuclear NOE suggests that the region highlighted above in red, from 345 to 347 is flexible in solution. (Witty *et al*, 2002)

TolAIII were potentially influenced by crystal packing, it was therefore critical to perform an analysis of TolAIII backbone flexibility using the heteronuclear NOE experiment (Rinaldi, 1983) and compared this with the generalized order parameters predicted by TALOS+. No regions of secondary structure with an RCI greater than 0.6 were observed in the TALOS+ analysis of C, H, and N chemical shift data from assigned TolAIII for the N-terminal region. To confirm this, a direct measurement of the motion of individual N-H bond vectors was made using a $^{15}N\{^{1}H\}$ NOE experiment (Farrow *et al*, 1994). Steady-state NOE values were determined from the ratios of the average intensities of the peaks with and without proton saturation applied before the start of the experiment. Resonance overlap was the major cause of unmeasured peak intensities in NOE spectra. In figure 3.9 the generalized order parameter S^2 as determined by TALOS+ analysis, and the steady-state NOE values as determined by the Heteronuclear NOE experiment were consistent with a region of flexibility in TolAIII corresponding to the first ~20 N-terminal residues. Additionally, the $^{15}N\{^{1}H\}$ NOE experiment further suggests disorder at the C-terminus begins in the final 2-3 residues of TolAIII, and that as shown in figure 3.10, is consistent with the unbound form of *P. aeruginosa* TolAIII possessing only a single turn at the C-terminus instead of an ordered 2-turn helix. The presence of this turn is discussed further in Chapter 4.

3.3.9 Relaxation dispersion spectroscopy does not indicate any hidden conformational states in *P. aeruginosa* TolAIII

I measured relaxation properties of the protein backbone in TolAIII to highlight any potential transient or 'hidden' states corresponding to low-population internal structural rearrangements. *E.coli* TolAIII was been suggested to undergo conformational changes on binding a bacteriophage gene 3 coat protein partner from a pre-existing equilibrium of states or via a conformational fit mechanism, (Deprez *et al*, 2005), however later crystal structure data for free TolAIII suggested bound and unbound TolAIII had similar structures (Li *et al*, 2012). In *P. aeruginosa* TolAIII the N-terminal region has been suggested to be meta-stable (Witty *et al*, 2002). If hidden states existed in TolAIII, and the rate of exchange between these states was slower than chemical shift separation in Hertz, this behaviour would be observable by classic magnetization exchange experiments. If however there is an interconversion in an intermediate exchange regime, Carr-Purcell-Meiboom-Gill (CPMG)-based experiments can be used to extract information about this line-broadened state(Neudecker, Lundstrom, & Kay, 2009). 180 degree pulses applied during the evolution period lead to progressively improved refocusing of magnetization originating on different molecules, in this case on any ground and excited states of TolAIII. This type of refocusing alters the peak intensity visibly if there is an excited state populated as low as ~0.5% (Neudecker et al. BioPhys J 2009) and which typically occurs on the ms timescale.

The change in peak intensity (arbitrary units) as a function of the change in the number of pulses applied during the evolution period have been plotted for every residue in TolAIII in figure 3.11. No large changes were observed in the globular region of TolAIII other than backbone amide intensity changes derived from strong signals of flexible regions. Two residues with large observed changes, Ser331 and Leu 347 are artefactual, due to significant overlap with an adjacent peak and very strong signal. These data are consistent with TolAIII retaining a single structure in solution, and that aside from the disordered region of the N-terminus, are consistent with the crystal structure of TolAIII.

3.4 Discussion

The aim of this chapter was to establish the structural characteristics of *P. aeruginosa* TolAIII in solution and to determine if structural incongruities between crystallographic data for *E. coli* and *P. aeruginosa* TolAIII molecules were also present in solution-state. The second aim was

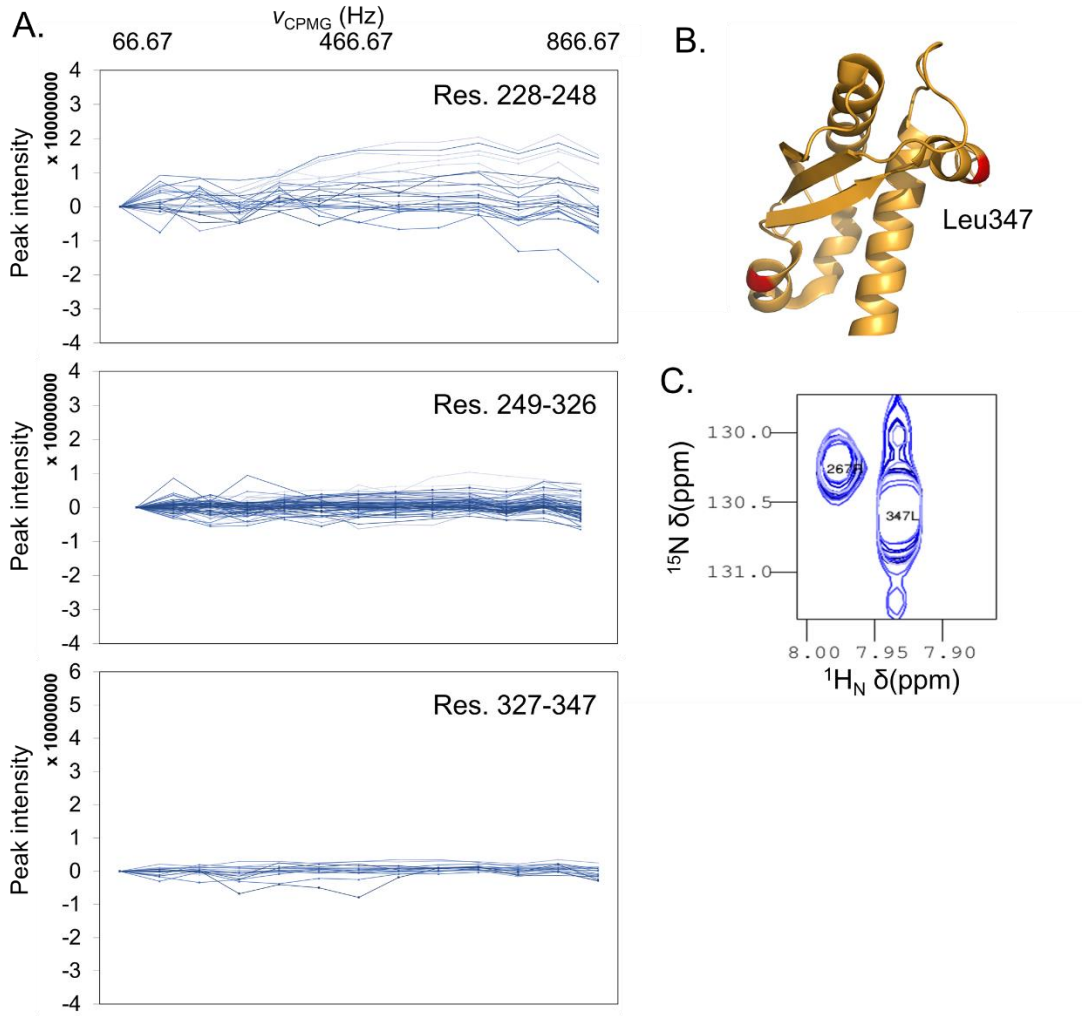


Figure 3.11 **The Carr-Purcell-Meiboom-Gill (CPMG) field dependencies of residues 228-347 in TolAIII are not consistent with any large conformational exchange contributions.** (A) Field dependencies have been shown as a function of the number of 180° pulses within the CPMG pulse train. The C-terminal residue Leu347 (not shown) displays abnormal fluctuations in field dependency, however the strength of this peak due to flexibility makes this behaviour artefactual, this residue is indicated in (B) and a slice of the CPMG spectra displaying residue 347 are shown in (C). Peaks not shown in these traces include those with irresolveable overlap with neighbouring residues. These data suggest that TolAIII does not undergo any micro-millisecond conformational exchanges between structural isomers when free in NMR buffer.

to determine the conservation of the *tol-pal* operon and overall conservation of secondary structure within TolA across all *tol-pal* containing species.

Firstly, a structure-based ClustalOmega sequence alignment of TolAIII molecules identified by BLAST searches of fully annotated genomes revealed that the 2° structure in the folded globular domain of TolAI and TolAIII are strongly conserved. Specifically, the 1st, 2nd and 3rd β-strands in TolAIII are readily identifiable in most Gram-negative strains with a *tol-pal* operon. Notable Gram-negative species without a Tol-Pal operon included intracellular parasites and Gram-negative bacteria with an atypical outer membrane.

Structural data were unavailable for strains other than *V. cholera, E. coli* and *P. aeruginosa* TolAIII molecules. Unlike *E. coli* or *V. cholera* whose polypeptide chains terminate immediately following β-strand 3, TolAIII in *P. aeruginosa* appears to fall under a second class on TolAIII molecules which possess C-terminal structure following β-strand 3. Of concern however, was the effect that crystal packing forces and the presence of a poly-His tag had on this helix.

Crystal contacts, or crystal artifacts refers to interchain contacts that manifest due to the the proximity or arrangement of nearby protein within the unit cell (Dasgupta *et al*, 1997). These interactions are generally small, non-specific and often have little biological relevance (Carugo & Argos, 1997). In the crystal structure *P. aeruginosa,* variations in the secondary structure between *E. coli* and *P. aeruginosa* TolAIII warranted closer examination for the influence of crystal packing. Before this study it was not conclusively known if the C-terminal helix 4 in *P. aeruginosa* was an artefact.

This analysis was performed so that any conformational changes in TolAIII on binding a partner protein could be compared with the unbound state. TolAIII was established as a protein-protein interaction hub in Chapter 1, Figure 1.6., and the C-terminus of TolAIII was in close proximity to one of two known protein-protein interaction sites. It was previously suspected that crystal packing forces may have strongly influenced the *P. aeruginosa* TolAIII structure: Witty *et al.* suggested that the N-terminal region of TolAIII may be metastable, and that the long N-terminal helix 1 observed in the crystal structure may be artefactual (Witty *et al*, 2002). A detailed NMR investigation into regions of flexibility in TolAIII, involving full assignment of the TolAIII molecule *in vitro* further revealed that both the N and C-termini varied from the original TolAIII crystal structure. Of particular importance to this investigation was the structural characteristics of the C-terminus of TolAIII. If in fact the C-terminus was ordered as an alpha helix, the position of the Leu345 sidechain was such that the hydrophobic core of the protein would be occluded. If the C-terminus of TolAIII was fully disordered, this region of the molecule would be exposed to solvent. A detailed examination of the influence of aromatic sidechains and corresponding ring current effects on the chemical shift of Leu345 is covered in Chapter 4 in the context with the TolA-TolB bound state.

Heteronuclear NOE measurements in combination with TALOS+ chemical-shift based structural prediction suggests that the N-terminus from 224-250 in TolAIII is intrinsically

disordered in solution. This is corroborated by the amide chemical shifts for these residues occupying the range where random coil polypeptide chemical shifts are typically observed (8.0-8.5 ^{1}H ppm). Full assignment of the TolAIII molecule allowed for the identification of intramolecular NOE distance restraints between residues in the C-terminus of TolAIII.

The metastability of the N and C-termini of TolAIII were further investigated for any metastable states using CPMG-based relaxation dispersion spectroscopy for the detection of any micro- to millisecond states in conformational exchange. No evidence was observed for states of this kind.

3.5 Conclusions

P. aeruginosa TolAIII has a disordered N-terminus comparable to *E. coli* TolAIII, and no evidence of metastable or hidden states could be found. TolAIII and its 2° structure are mostly conserved across Gram-negative bacteria with a *tol-pal* operon. The C-terminus of TolAIII, may retain partially helical character, and heteronuclear NOE and NOE data suggest this region is structured up to residue Leu345 allowing for occlusion of the hydrophobic core of the protein.

4. Determination of the *P. aeruginosa* TolA-TolB interaction *in vitro*

4.1 Introduction

4.1.1 Protein-protein interactions and intrinsic disorder

Inside the average single cell of *E. coli* there exists at any given point approximately 2 million proteins, representing 55% of the dry weight of the cell, and an estimated 1850 protein species (Neihardt, Umbarger 1996). Individually, proteins represent a crucial component of the cell, but their interaction networks form the backbone of the spatially and temporally regulated chemical reactions that govern all life processes within the cell. Protein-protein and protein-ligand interactions play essential roles in a wide variety of life processes within the bacterial cell, including host-pathogen interactions, cell metabolism, cell maintenance, and reproduction. This chapter investigates the molecular mechanism underlying a key bacterial protein-interaction which involves an intrinsically disordered binding epitope.

Emil Fischer's investigations into the cleavage of β-glucosides and amygdalin to mandelonitrile glucoside via the action of glucosidases, as well as lactose, saccharose, and maltose cleavages led to his formulation of the 'lock and key' principle (Fischer 1895), suggesting unique geometric shapes dictate protein function. When David Phillipps obtained the first 2-ångström structure of hen-egg white lysozyme and the location of the site of attachment of a competitive lysozyme inhibitor, the structure immediately supported previous chemical evidence suggesting amino acids involved in the substrate binding site, and further demonstrated that their orientation and location was likely what produced a catalytic effect. Since 1965, the determination of the three-dimensional structure of a protein became central to the inference of function for these molecules. Since this first determination of the structure of a protein-ligand interaction, 125526 protein structures have been deposited into the Protein Data Bank (Berman *et al*, 2000)

In reality however, not all proteins interact with their partners via lock and key mechanisms. Not all proteins are even fully structured, either in their entirety or even in part. Paramagnetic ion probes and NMR relaxation experiments demonstrated first in chromogranin A that the hydrodynamic properties of this protein suggested it was not a folded globular domain but rather approached “that of a random coil conformation” (Daniels 1978). Intrinsically disordered proteins (IDPs) and proteins with regions of disorder exhibit clear specific functions in the cell

(Uversky 2002). Intrinsically disordered proteins suggest a “new paradigm of coupled binding and folding” whereby an IDR, to perform its function adopts a folded structure upon binding its target. (Sigalov 2016). The frequency of native disorder in eukaryotic organisms is high, with (>30 residue) segments of disorder occuring in 33% of eukaryotic proteins. Conversely, the level of disorder in prokaryotes is low, around 4% (Ward *et al*, 2004). The majority of prokaryotic putatively disordered segments occur in compartmentalized proteins, providing protection from proteolysis. Even within bacteria, the range of proteins with regions of putative disorder varies according to the environment in which that bacteria lives. Thermophiles such as *Thermotoga maritima* for example possess comparatively low (1.8%) numbers of disordered proteins, whereas *Mycobacterium tuberculosis* and other eubacteria are much higher at 7% (Ward *et al*, 2004).

From a biophysical perspective, studying these regions of disorder may be accomplished using a variety of methods. X-ray crystallography, although powerful in the determination of high-resolution structures, identifies intrinsically disordered regions via the absence of significant regions of electron density, or via the stabilisation through the binding of an intermediate which influences the conformational equilibrium of the IDR (Bonsor *et al*, 2009). Solution-based spectroscopic techniques, which do not require the immobilization of protein therefore represent powerful methods for studying interactions and proteins that rely on disorder for function.

4.1.2 The TolA-TolB interaction is a protein-protein interaction involving an intrinsically disordered binding epitope

The first study on the interaction of the TolA and TolB components of the Tol-Pal system by ITC suggested no *in vitro* interaction between these two proteins (Gokce *et al*, 2000). Subsequent work by the yeast two-hybrid method suggested an interaction between TolA and TolB did exist (Walburger *et al*, 2002). The presence of a TolA-TolB interaction has long been contentious, as mentioned in section 3, however a recent discovery suggests that TolB is involved in a conformational selection mechanism involving Pal. TolB alternates between TolA binding-incompetent and TolA-binding competent modes via the sequestration or release of its intrinsically disordered N terminus, $TolB_{22\text{-}33}$ (Bonsor *et al*, 2009). This interaction, shown to be low-affinity interaction between the intrinisically disordered N-terminus of TolB and the folded globular domain TolAIII is of specific interest (Bonsor *et al*, 2009). TolAIII is also the target of filamentous bacteriophages and group A bacteriocins (Davies 1975, Benedetti 1996) to effect an entry mechanism into the cell. The structures of these exogenous proteins provide information on

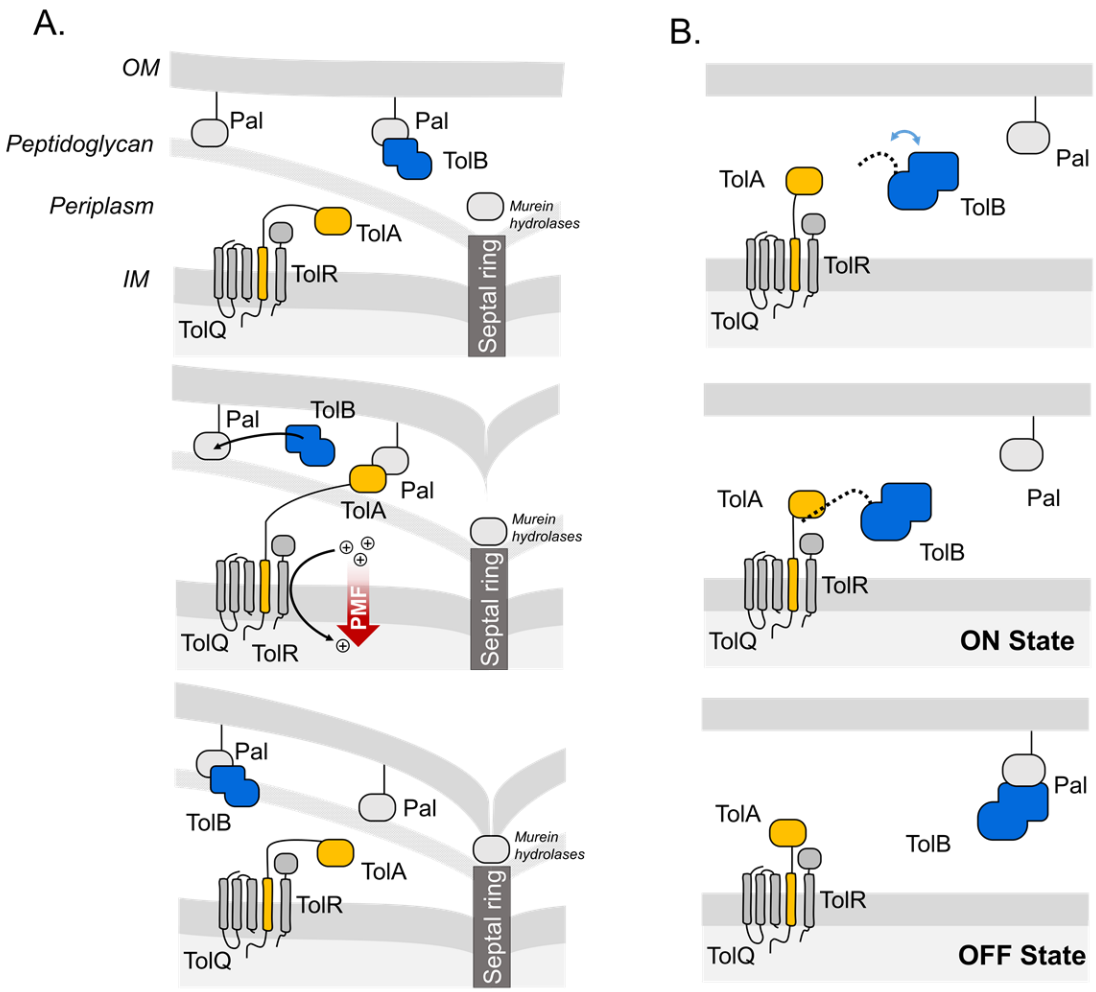


Figure 4.1 **Two conflicting models of Tol-Pal action in the Gram negative cell envelope.** (A) represents a longitudinal section through the septation apparatus and the proposed role of Tol-Pal during cell septation, adapted from (Gerding *et al*, 2007b). (B) represent a similar longitudinal slice, focused on a proposed allosteric signalling mechanism where TolB recruits TolA through an intrinsically disordered N-terminus, adapted from (Bonsor *et al*, 2009). In (A) going from top to bottom, is the first stage of a TolA engagement cycle at the beginning of OM invagination. After energization by the proton motive force funnelled through TolQ and R, TolA assumes an extended conformation to bind Pal at the OM, which causes TolA to 'snap back', disengaging Pal near the peptidoglycan layer such that it may bind and draw the OM inwards towards the septation machinery. In (B) the TolA box of TolB is in conformational equilibrium. When TolB is recuited to TolA, this couples TolB to the PMF via TolA, TolQ and TolR, an 'ON' state. When the OM lipoprotein Pal sequesters TolB, the TolB N-terminal TolA box collapses into a binding cleft between the N and C-domains of TolB. This makes TolB inaccessible to TolA, and is termed the 'OFF' state.

potential binding sites for the endogenous partner TolB, as it is conceivable that they hijack the Tol-Pal system to effect the same function that Tol-Pal normally conducts in the cell, namely the translation of energy from the proton motive force to structural assemblies near or at the outer membrane. The structure of the endogenous interaction with TolB has yet to be described.

Previous studies showed that TolAII-III and TolB fusion proteins grown in a yeast reporter strain would successfully activate a LacZ reporter assay (Walburger *et al*, 2002), suggesting a

TolA and TolB interact. The LacZ reporter assay however does not provide specific information about the binding interface. In the absence of the TolAIII domain, the TolB N-domain would not bind TolA, suggesting that TolAIII was the primary interaction partner (Walburger *et al*, 2002). A mechanism by which TolB interacted with TolAIII was suggested, involving a disorder-order transition in its N-terminal sequence governed by an allosteric regulator, Pal (Bonsor *et al*, 2009). Further to this, the truncation of the TolB N-terminus abrogated binding with TolAIII (Bonsor *et al*, 2009).

4.1.3 The order-disorder hypothesis of the TolA 'box'

TolAIII interacts with the 12 amino-acid N-terminal region of TolB in *E. col,* this region of TolB is referred to as the 'TolA box' (Bonsor *et al*, 2009). In 2000 a direct and energy dependent interaction was proposed between TolA and Pal (Cascales *et al*, 2000). A model was proposed where TolA acted like a molecular crane, involved in sequential engagement-disengagement cycles with Pal near the peptiglycan layer. In this role TolA acted as a 'grabbing device' that, when energized, drew Pal molecules towards the PG layer during septation, drawing the OM inwards as part of the cell constriction machinery (Gerding *et al*, 2007a). This model is outlined in Figure 4.1 (A). In this model, the TolA TolB interaction is absent or inconsequential, as the function of TolB in this model is to compete with peptidoglycan for access to the TolB binding interface of Pal (Gerding *et al*, 2007a). An alternate model suggests that TolB interacts with TolA through its N-terminus allosterically. In this model the TolB N-terminus is sequestered into a binding cleft between the TolB N and C-domains when TolB is bound to Pal in an 'OFF' state, preventing the binding of TolA. When TolB is free, its N-terminus undergoes an order-disorder conformational switch, making it TolA-accessible, an 'ON' state (Bonsor *et al*, 2009). This model is summarised in Figure 4.1 (B).

Before this study it was not known if the binding of TolB to TolA causes a conformational change in TolA. In *E.coli* it was recently shown that the TolA binding epitope in TolB is restricted to the N-terminus by a ^{1}H-^{15}N TROSY-HSQC spectrum; TolAIII binding (and sequestration of the TolB N-terminus) caused the same changes to wild-type TolB as the deletion of the TolA box (Bonsor *et al*, 2009). No other TolAIII-induced conformational changes were reported.

4.2 Aims

The aim for this chapter was to establish the structural mechanism by which TolA interacts with TolB. Specifically to generate structural data that would clarify how TolA might play a role in the disorder-order hypothesis of TolB function, and to elucidate how the PMF might be transmitted via TolA to TolB.

4.3 Results

4.3.1 The *P. aeruginosa* TolA-TolB interaction is weak *in vitro*, as determined by fluorescence anisotropy with FITC-labelled $TolB_{22-34}$

The TolA-TolB interaction is suitable for a structural investigation by NMR and biophysical methods. As determined in Chapter 3, the *P. aeruginosa* isoform of TolAIII is stable under typical NMR buffer conditions, and ideal for study with standard NMR experiments.The aim of this section was to determine a precise measurement fo the binding affinity between TolA and TolB.

The molecular weight range limit for the study of protein interactions by solution NMR is typically in the range of 50 kDa, although more recently, transverse relaxation optimized NMR spectroscopy (TROSY) and methyl-TROSY techniques (Rosenzweig & Kay, 2014; Kay, 2011) have raised the bar in terms of what studies may be performed on large supramolecular protein complexes.

A synthetic peptide representing the N-terminal TolA box of TolB was previously established to recapitulate the TolAIII-TolB interaction in *E.coli* by formaldehyde crosslinking (Cassels, 2012). This complex has comparable thermodynamics and affinity (43 μM for TolB-TolAIII, 40 μM for $TolB_{22-34}$-TolAIII) as determined by isothermal titration calorimetry (Cassels, 2012). A synthetic TolB peptide 13 amino acids long representing the sum total of the N-terminus of TolB in complex with TolAIII is less than 50 kDa and therefore is suitable for NMR structural determination. The *P. aeruginosa* isoform of TolAIII was chosen for this study because the spectra for *E. coli* TolAIII in complex with $TolB_{22-33}$ peptides have been established in previous studies to produce data with HSQC spectra consistent with protein or peptide multimerization (Cassels, 2012; Li *et al*, 2012). Further, *E. coli* $TolB_{22-33}$ is poorly soluble in standard NMR buffer (data not shown). To study the TolA-TolB interaction, *P. aeruginosa* $TolAIII_{224-347}$ was chosen as detailed in Chapter 3, and $TolB_{22-34}$, produced synthetically was chosen to represent the N-terminus of TolB.

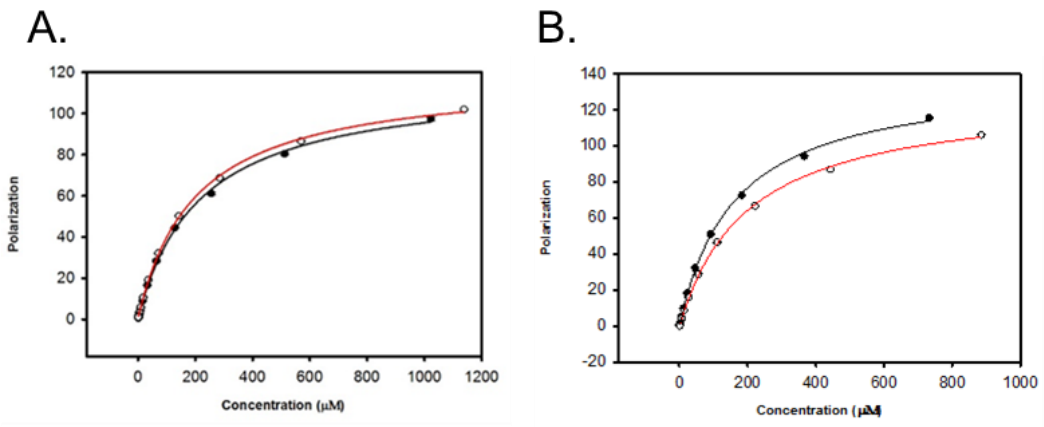


Figure 4.2 **The TolA-TolB interaction is weak *in vitro* by fluorescence anisotropy in the presence and absence of a disordered N-terminus.** (A) Black trace corresponds to $TolA_{224-347}$ in complex with $TolB_{22-34}$FITC, K_D=220±10 μM, n=3, red trace corresponds to $TolA_{244-347}$ (N-terminal truncate) in complex with $TolB_{22-34}$FITC, K_D=200±7 μM, n=3.(B) Black trace corresponds to Im9-$TolA_{244-347}$ in complex with $TolB_{22-34}$FITC, K_D=169±9 μM, n=3, red trace corresponds to Im9-$TolA_{224-347}$ (N-terminal truncate) in complex with $TolB_{22-34}$FITC, K_D=200±4 μM, n=3. All traces conducted at buffer conditions of 50 mM Tris-HCl, 100 mM NaCl, pH 7.0. Data were fit dynamically using a non-linear regression in SigmaPlot.

This peptide was soluble in NMR buffer conditions up to 3 mM and was suitable for a detailed structural investigation.

The *E. coli* TolAIII-$TolB_{22-33}$ interaction is a weak interaction of ~40 μM. To determine the binding affinity for TolAIII-$TolB_{22-34}$ in *P. aeruginosa,* fluorescence anisotropy was chosen to determine a dissociation constant (K_D) for the interaction. I established the affinity of this complex was K_D=220±10 μM, n=3, under buffer conditions of 50 mM Tris-HCl, 100 mM NaCl, pH 7.0. In order to ensure that the disordered region of TolAIII played no role in the TolAIII-$TolB_{22-34}$ interaction, an N-terminal truncation mutant of TolAIII was prepared, omitting the first 20 amino acids of TolAIII. The dissociation constant for this construct was 200±7 μM, n=3, as demonstrated in Figure 4.2 for $TolA_{227-347}$ (black) and $TolA_{244-347}$ (red). Indeed, even the large N-terminal globular domain of Im9 when left uncleaved from the TolAIII-Im9 fusion expression construct bound $TolB_{22-34}$ with an affinity of 200±4 μM, n=3, suggesting that molecules or structure at the N-terminus do not affect the binding mechanism of TolB.

4.3.2 The *P. aeruginosa* TolAIII-$TolB_{22-34}$ interaction is weak as determined by NMR spectroscopy and produces large chemical shift changes in slow exchange

I performed a 2D ^{15}N HSQC titration of synthetic $TolB_{22-34}$ against *P. aeruginosa* $TolA_{227-347}$ to provide a second method of measuring the weak interaction observed *in vitro* between TolA and TolB. An additional aim was to resolve biophysical data regarding the overall fold and dynamics by solution NMR. I established that the TolA-TolB complex produced large chemical

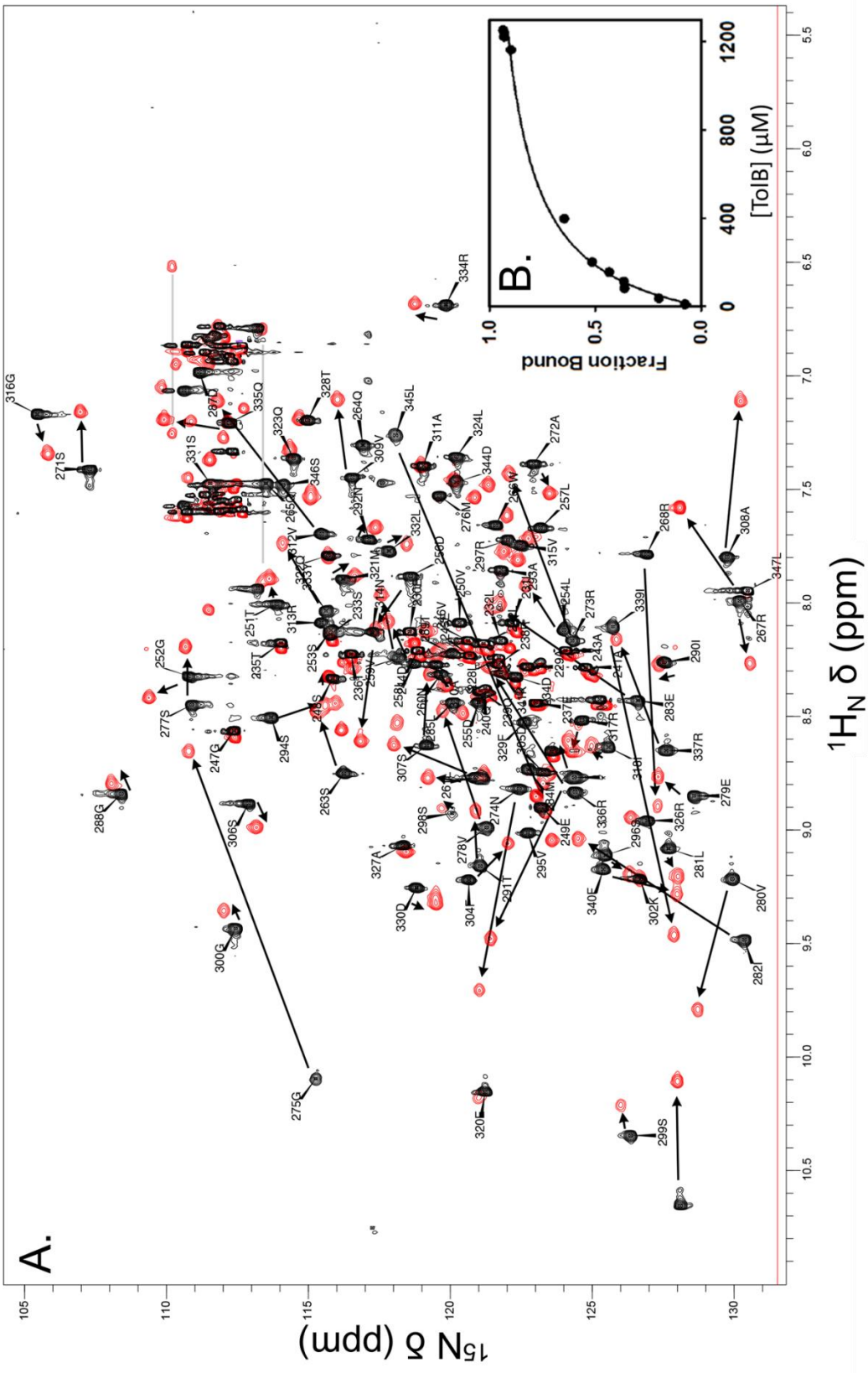


Figure 4.3 **The TolA-TolB interaction is weak *in vitro* as determined by solution NMR.** ^{15}N HSQC spectra of {^{15}N} TolAIII displays large chemical shift changes with the addition of TolB$_{22-34}$. (A) All residues are in slow exchange, and consistent with a 1:1 complex. Unbound TolAIII shown in black, with saturated (>90%) chemical shifts shown in red. Arrows depict the beginning and end points of the titration, both the bound and unbound state was fully assigned. (B) Averaged peak volumes as a function peptide concentration for a titration of TolB$_{22-34}$ into TolAIII from 4 well-resolved peaks. Every peak was in slow exchange, so the sum of the bound and unbound peak volumes for every titration point gives the ratio of bound:free TolAIII.

shift changes in slow exchange, and a global dissociation constant (n=4) of 220 µM +/- 30 µM was determined, shown in Figure 4.3A and B.. Two solutions of 0% bound and ~95% bound TolAIII in 20 mM Phosphate buffer 100 mM NaCl pH 6.9 were prepared. 2D HSQCs were collected for each titration point and pH was monitored regularly externally by microprobe. These data were also consistent with the weak binding affinity observed in section 4.3.1.

The HSQC experiment correlates backbone or sidechain protons with ^{15}N and ^{13}C nuclei in the protein of interest. This means that each cross-peak visible in this spectrum represents a single carbon-proton or nitrogen-proton bond. These correlations are broadly sensitive to changes in the local chemical environment, and so typically when a protein interacts with a ligand, the chemical shifts of these correlations are perturbed in direct relationship to the concentration of the interacting ligand. These interactions may occur on a 'fast' timescale (ps-ns), an 'intermediate' timescale (µs) or a 'slow' timescale (ms-s). Ligand-protein interactions that occur on a fast timescale may be quantified by the change in chemical shift of affected crosspeaks as a function of ligand concentration. Conversely, ligand interactions occurring in slow exchange may also be quantified, by the change in peak volumes representing the bound and unbound states as a function of ligand concentration. These quantifications provide information about ligand stoichiometry, affinity, and may be used to approximate potential binding interfaces, termed chemical shift mapping.

Initial experiments suggested that the TolAIII-$TolB_{22-34}$ interaction involved a significant portion (>80%) of TolAIII amides undergoing large shifts, in slow exchange between bound and unbound species. This necessitated the *de novo* assignment of C, H and N atoms in TolAIII in the bound state. Datasets for HNCO, HNCA, CBCANH, CBCA(CO)NH, HBHA(CO)NH and ^{15}N TOCSY-HSQC, ^{15}N NOESY-HSQC, ^{13}C HSQC and ^{13}C NOESY-HSQC were collected as detailed in Chapter 2. >95% assignment of all backbone resonances were completed. Backbone residues presenting difficulty included the disordered N-terminus and Proline residues. The assignment of backbone amide resonances in TolAIII are shown in figure 4.4. In Figure 4.3A, large translations in the chemical shift of the protein backbone were seen with near-saturation (>90%) of TolAIII by $TolB_{22-34}$.

A striking peculiarity evident from this titration was the large excess of peptide required to approach saturation of TolAIII. This was indicative of a very weak interaction between TolAIII

and $TolB_{22-34}$, as later HSQC spectra for $TolB_{22-34}$ or TolAIII were not consistent with dimerization of either species, nor did any evidence suggest the possibility of a low-population or transient state. This is in agreement with the weak binding constant determined for this interaction. Carbon and proton assignments were subsequently determined, and the full assigned ^{13}C HSQC spectrum is shown in the appendix.

The TolAIII-TolB interaction is weak by both fluorescence anisotropy and solution NMR. This relationship has been determined both by titrating TolA with TolB in excess (NMR) and TolB with TolA in excess (Anisotropy). Furthermore, the bound and unbound spectra for TolAIII suggest unusual kinetics involved in this interaction, as residues across the entire spectra are in slow exchange.

4.3.3 Observable chemical shift changes in a *wt* TolB-TolAIII complex are comparable to a $TolB_{22-34}$-TolAIII complex.

In order to ascertain if the TolB-TolAIII interaction was observable by NMR experiments, and to establish the full length TolB-TolAIII complex was the same as recapitulation with a synthetic peptide, full length TolB was titrated into ^{15}N TolAIII. The results of this titration are shown in figure 4.5. Although the recapitulation of the TolB interaction has been shown with synthetic peptide (Cassels, 2012), it was crucial to ascertain if TolAIII, upon binding to *wt* TolB would produce similar chemical shift changes to synthetic-$TolB_{22-34}$. Maintenance of protein solubility of TolB was of particular difficulty in this experiment and limited the degree of saturation of TolAIII. TolB could not be kept reliably soluble at concentrations exceeding 10 μM. Nonetheless, spectra were collected on a 750 MHz OMEGA spectrometer using conditions as described in Chapter 2. Although the vast majority (>70%) of residues were not observable in the TolB-TolAIII complex, due to the size of the protein-protein interaction at 57 kDa, we were able to observe identical chemical shift changes in the disordered N-terminus up to the transition region at the beginning of helix 1.The diminished anisotropic tumbling of the disordered N-terminus made it ideal for reflecting changes in the folded globular domain of TolAIII. Although the amplitude of chemical shift changes in the N-terminus of TolAIII from binding *wt* TolB were small, observable residues within this region demonstrated chemical shift changes exactly consistent with the chemical shift changes observed in the binding of $TolB_{22-34}$. Examples of this for residues 230 through 248 are shown in figure 4.5. These data are therefore in support of ITC data suggesting the TolA-TolB interaction may be recapitulated by a synthetic peptide (Cassels, 2012).

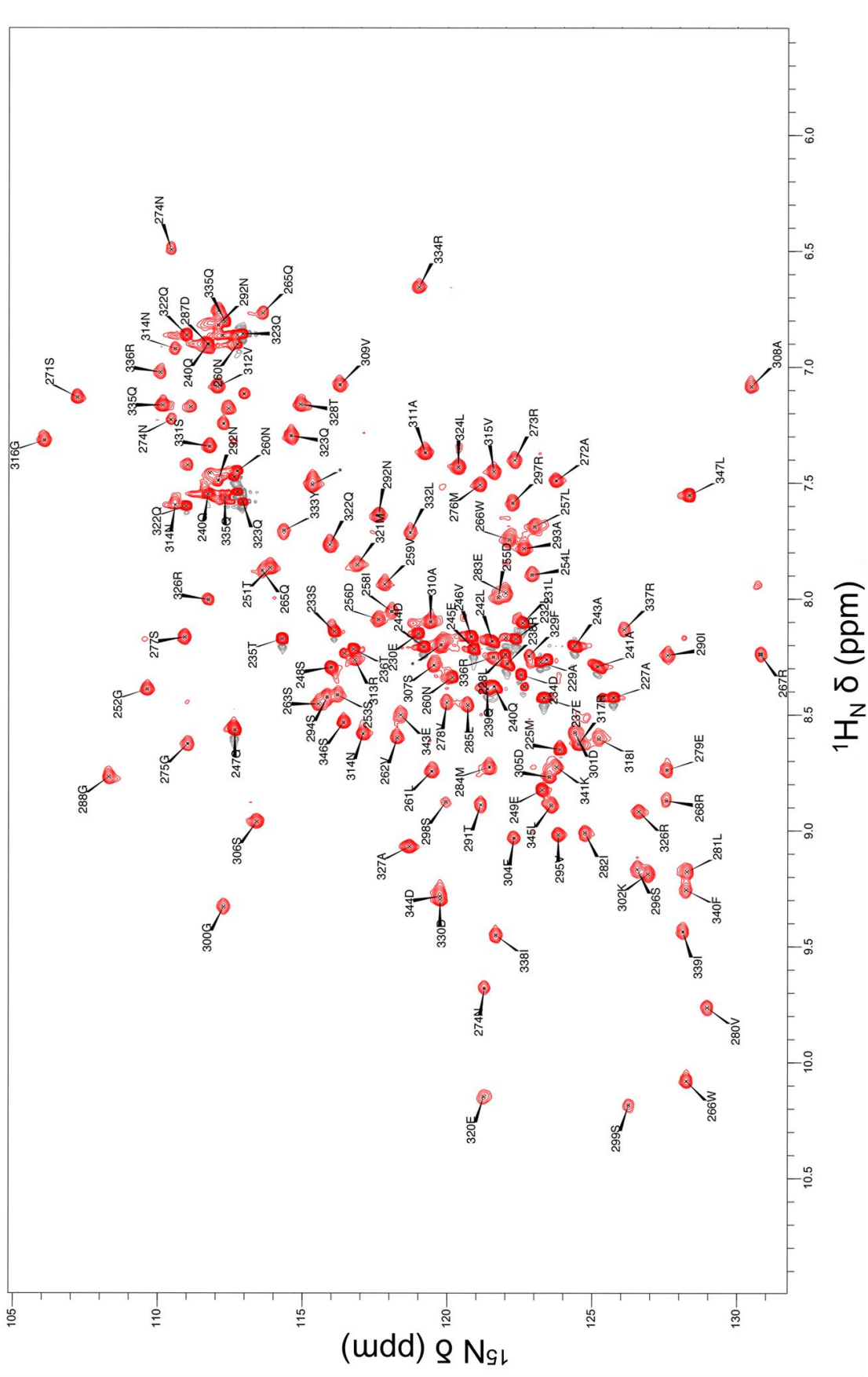


Figure 4.4 **Full assignment of all backbone amide correlations in a ^{15}N HSQC of TolAIII saturated with TolB$_{22\text{-}34}$.** Spectra recorded using 3 mM TolB$_{22\text{-}34}$, 0.6 mM TolAIII, in NMR buffer using a 750 MHz OMEGA spectrometer.

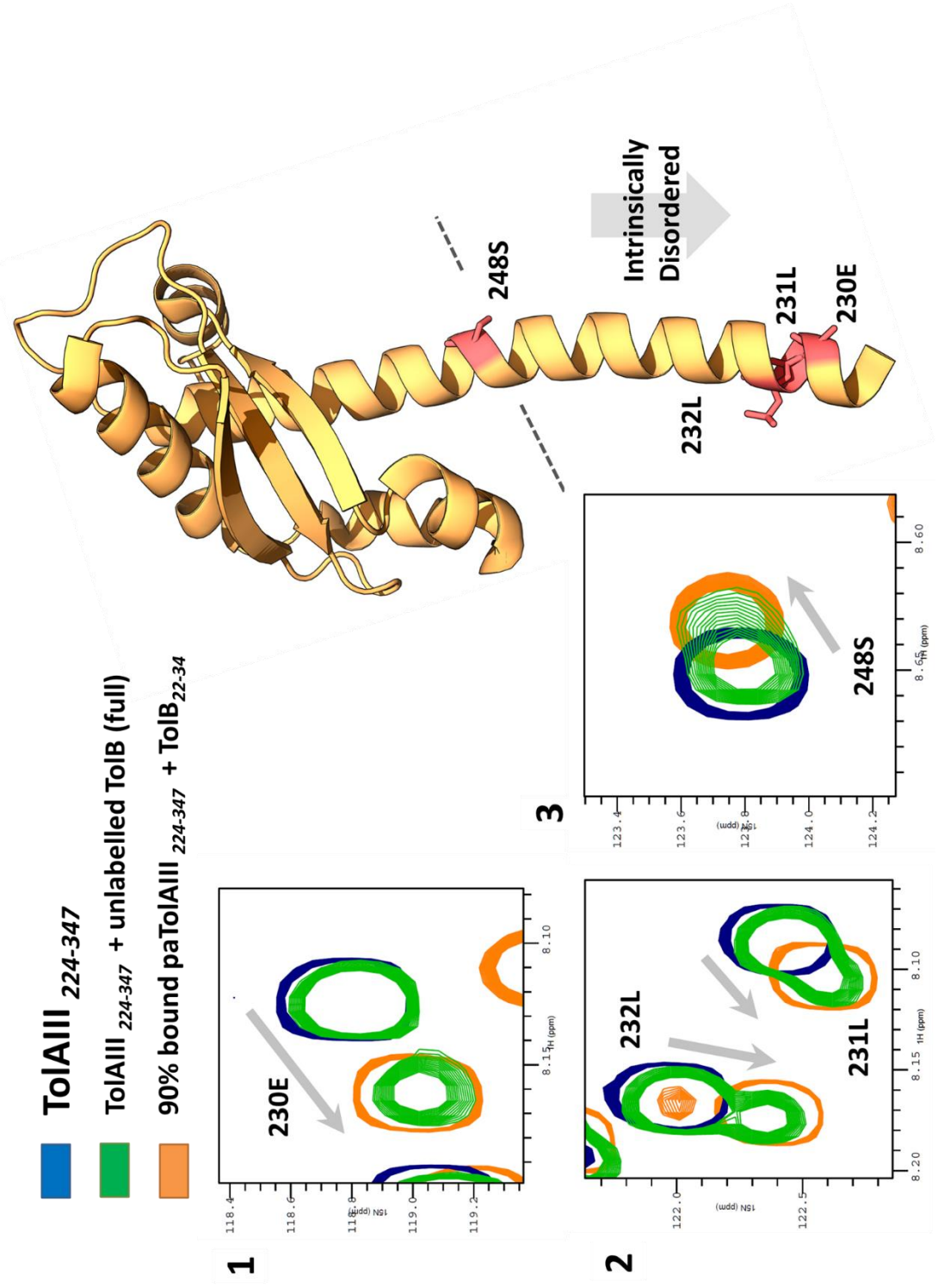


Figure 4.5 **Chemical shift changes in a disordered region of TolAIII as a probe for the binding mechanism of full length TolB.** $TolB_{22-34}$ behaves as full length TolB in NMR titration experiments with TolAIII. Three experiments have been overlaid in boxes 1, 2 and 3. In blue, the positions of 4 exemplar residue backbone amide chemical shifts in TolAIII. In orange, the chemical shifts for those same amides in a $TolB_{22-34}$ saturated state. Shown in green is the titration of 10 μM *wt* TolB into TolAIII. The peak splitting seen in the green spectra indicates that a mixed population of free and bound states occurs, and that the free and bound states have chemical shifts that correspond exactly to the positions of the free and bound states when using TolB peptide. The only observable chemical shifts for the *wt* TolB-bound complex were from the N-terminus to Ser248, likely due to the improved tumbling rate of this disordered region of the spectrum. The chemical shift changes were not large, which is consistent with these residues being disordered in solution, however do suggest that *wt* TolB binds TolAIII in the same manner as $TolB_{22-34}$

4.3.4 TolAIII-TolB$_{22\text{-}34}$ structural refinement required the production of $^{15}N^{13}C$ heterolabelled peptide.

An approach to solving protein-ligand complexes by NMR uses molecular dynamics in restraint-drive docking process to dock ligands (de Vries *et al*, 2010; Stratmann *et al*, 2011). Because of the unusual kinetics of the TolAIII-TolB$_{22\text{-}34}$ complex, and the weak binding regime for TolB$_{22\text{-}34}$, I chose instead to directly resolve intermolecular NOEs in the determination of the TolAIII-TolB$_{22\text{-}34}$ complex. This approach necessitated the purification of $^{15}N^{13}C$ TolB$_{22\text{-}34}$ was required such that standard backbone and sidechain assignment procedure could be carried out in reverse, with a TolAIII-saturated TolB$_{22\text{-}34}$ molecule. In particular, this decision was strongly influenced by difficulties encountered in assigning the unlabelled peptide using edited-filtered experiments (discussed in section 4.3.5) and H-H TOCSY data. This edited-filtered approach, whilst clearly identifying the presence of intermolecular NOEs failed to produce assignments for greater than 60% of the observed backbone residues of unlabelled TolB$_{22\text{-}34}$.

The purification and assignment of doubly enriched TolB$_{22\text{-}34}$ has been described in Chapter 2. Figure 4.6 shows the successful final purification of recombinant ^{13}C ^{15}N TolB$_{22\text{-}34}$ by HPLC. The purity of this peptide is shown in Figure 4.7, Bound (red) and unbound (black) $^{1}H^{15}N$ HSQC spectra for $^{15}N^{13}C$ enriched TolB$_{22\text{-}34}$ are shown superimposed. Unbound TolB$_{22\text{-}34}$ shows characteristically sharp peaks within 8.0-8.5 ppm, typical for a random coil peptide. Upon binding with TolAIII, chemical shift changes result in a disperse spectrum characteristic of a folded peptide with secondary structure and chemical shifts characteristic of β-strand. Although not shown in the trace, contouring of the observation plane lower reveals a minor species of TolB$_{22\text{-}34}$ in solution, consistent with a second species having undergone *cis-trans* isomerization around the proline bond.

Using standard HNCACB, CBCA(CO)NH, TOCSY, ^{15}N NOESY-HSQC and H-H NOESY spectra the bound form of TolB$_{22\text{-}34}$ was assigned to completion for use in identifying intermolecular NOEs from subsequent edited-filtered experiments.

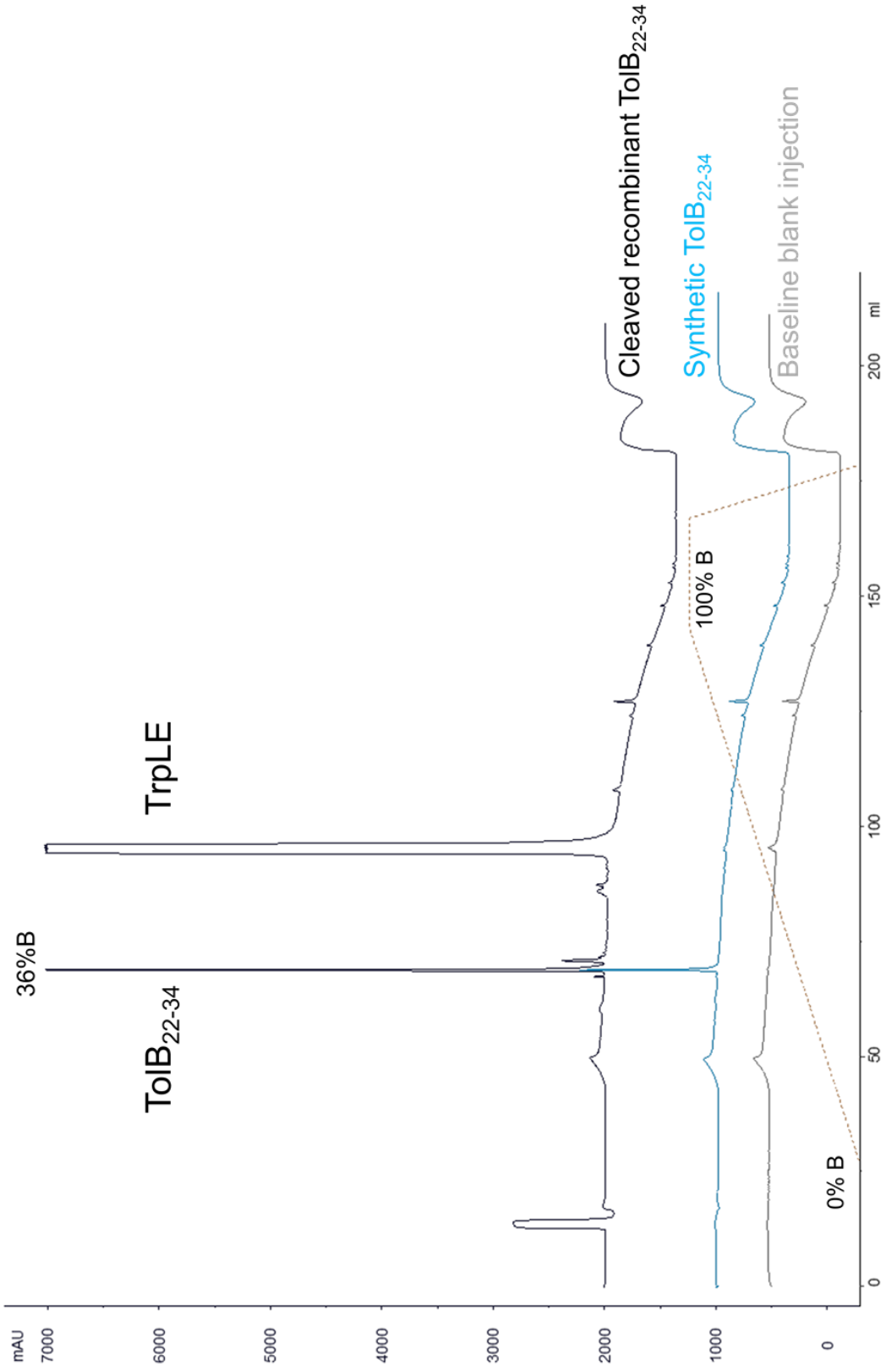


Figure 4.6 **TolB$_{22-34}$ is purified from the trpLE fusion tag by reverse phase HPLC.** Control injections of baseline blank samples and synthetic TolB$_{22-34}$ are consistent with an injection separating TolB$_{22-34}$ from TrpLE. TolB$_{22-34}$ elutes at 36% B (A= 0.1% v/v TFA in MilliQ, B= 95% v/v AcN, 0.1% v/v TFA). The Δ retention volume from synthetic TolB$_{22-34}$ is 0.05 mL. The AcN gradient is shown as a dashed line from 0-100% B. Flow rate was 2.00 mL/min, peptide was detected by absorbance at 205 nm.

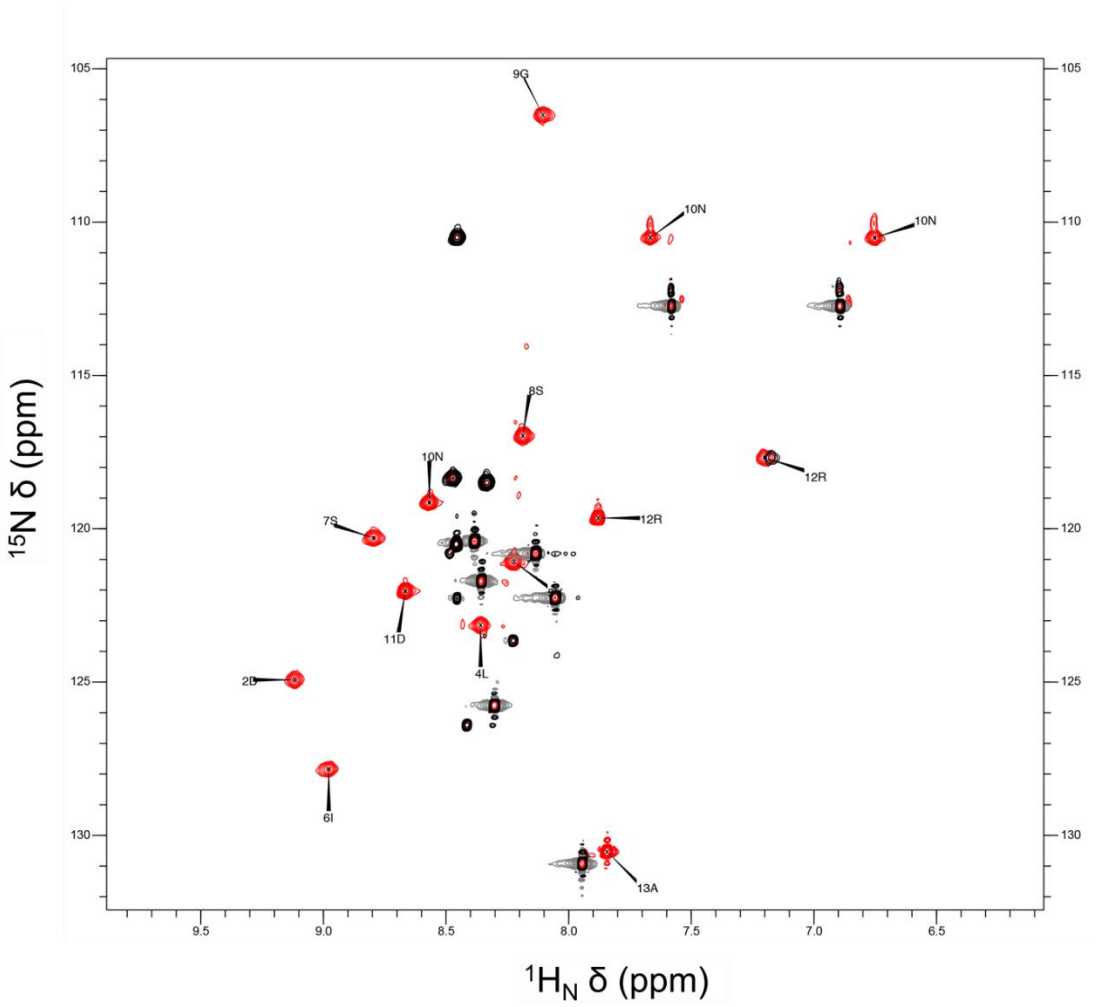


Figure 4.7 **Backbone amide resonances in a ^{15}N HSQC of TolB$_{22-34}$ free and saturated with TolAIII.** Spectra recorded using 200 μM TolB$_{22-34}$, in NMR buffer using a 750 MHz OMEGA spectrometer. In black is free TolB$_{22-34}$, in red the new positions for TolAIII-saturated TolB$_{22-34}$.TolAIII concentration 1.1 mM. All spectra collected in NMR buffer. TolB$_{22-34}$ undergoes a disorder-order transition on binding TolAIII. Chemical shift changes are typical for those seen in a conversion from random-coil to β-strand. TolB$_{22-34}$ was fully assigned in both the bound and unbound states.

4.3.5 ^{15}N ^{13}C-edited filtered NOESY spectra identify intermolecular NOEs and the orientation of the TolB peptide

A challenging barrier to determining the structure of protein-peptide interactions by NMR is the acquisition of sufficient restraints that provide information about the orientation of binding, secondary structures of the smaller ligand and a low internal r.m.s.d. in the ensemble for the smaller species. Data-driven docking approaches have been used to address these issues for similar complexes, (Stratmann *et al*, 2011; de Vries *et al*, 2010). Direct assignment of intermolecular NOEs as opposed to using an automated ambiguous assignment approach for unassigned protein-peptide NOEs in NOESY spectra can be achieved using *filtered* spectra,

which selectively remove crosspeaks corresponding to molecules other than the target, as detailed in Chapter 2. The original intent of these experiments was to provide a method for identifying the orientation, binding interface, and overall solution structure of $TolB_{22-34}$ when bound to TolAIII without having to purify recombinant $^{13}C^{15}N$ enriched TolB peptide. In practice, due to the weak affinity of the TolA-TolB complex, full assignments for unenriched TolAIII could not be determined, thus necessitating the purification of isotopically enriched TolB peptide.

Intermolecular NOEs are measured solely between $^{15}N^{13}C$ TolAIII and non-isotopically enriched $TolB_{22-34}$. Figure 4.8A shows a ^{15}N-edited/filtered NOESY spectrum corresponding to the NOEs between TolB sidechain protons and TolAIII amide protons. Figure 4.8B shows a ^{13}C-edited/filtered NOESY spectrum corresponding to the NOEs between TolB sidechain protons and TolAIII ^{13}C-bound protons. Intermolecular NOEs assigned in these spectra were cross-verified for the reverse corresponding NOE in 3D NOESY-HSQC spectra using the same sample. All NOEs that could be verified in this manner were then used for structural calculations.

Once a full assignment for bound TolB peptide was determined, clear structural details could be resolved and the identity of intermolecular NOEs were then used to determine the orientation of the $TolB_{22-34}$ peptide in complex with TolAIII. Specific and unique NOEs, determined in the edited-filtered experiments were cross-verified with return NOEs in $^{13}C^{15}N$ NOESY HSQC experiments, and a detailed understanding of the sidechains involved in the TolA-TolB binding interface was obtained. In Figure 4.9 the NOE contributions made by TolB sidechains have been plotted against an early structure calculation performed in ARIA 1.2 and clearly show that NOEs are either weak, medium or strong, and correspond roughly to the overall distance of each restraint.

4.3.6 Combined chemical shifts indicate the position of the TolAIII-$TolB_{22-34}$ interface

$TolB_{22-34}$ perturbed residues across the entire TolAIII protein, so no immediate binding site is apparent from amide shifts or Cα shifts alone. A weighted average of C, N and H chemical shift changes however allowed us to posit that the interaction interface was positioned near the C-terminal β-strand 3, based on the large chemical shift changes mapped to this region of TolAIII, as seen in figure 4.10. With full assignment of $^{1}H^{15}N$ and $^{1}H^{13}C$ HSQC spectra, chemical shift changes in the carbon-proton spectra were analysed for the effect of $TolB_{22-34}$ on TolAIII. The Cα chemical shift is noted for its sensitivity to the secondary structure of a protein, whereby regions

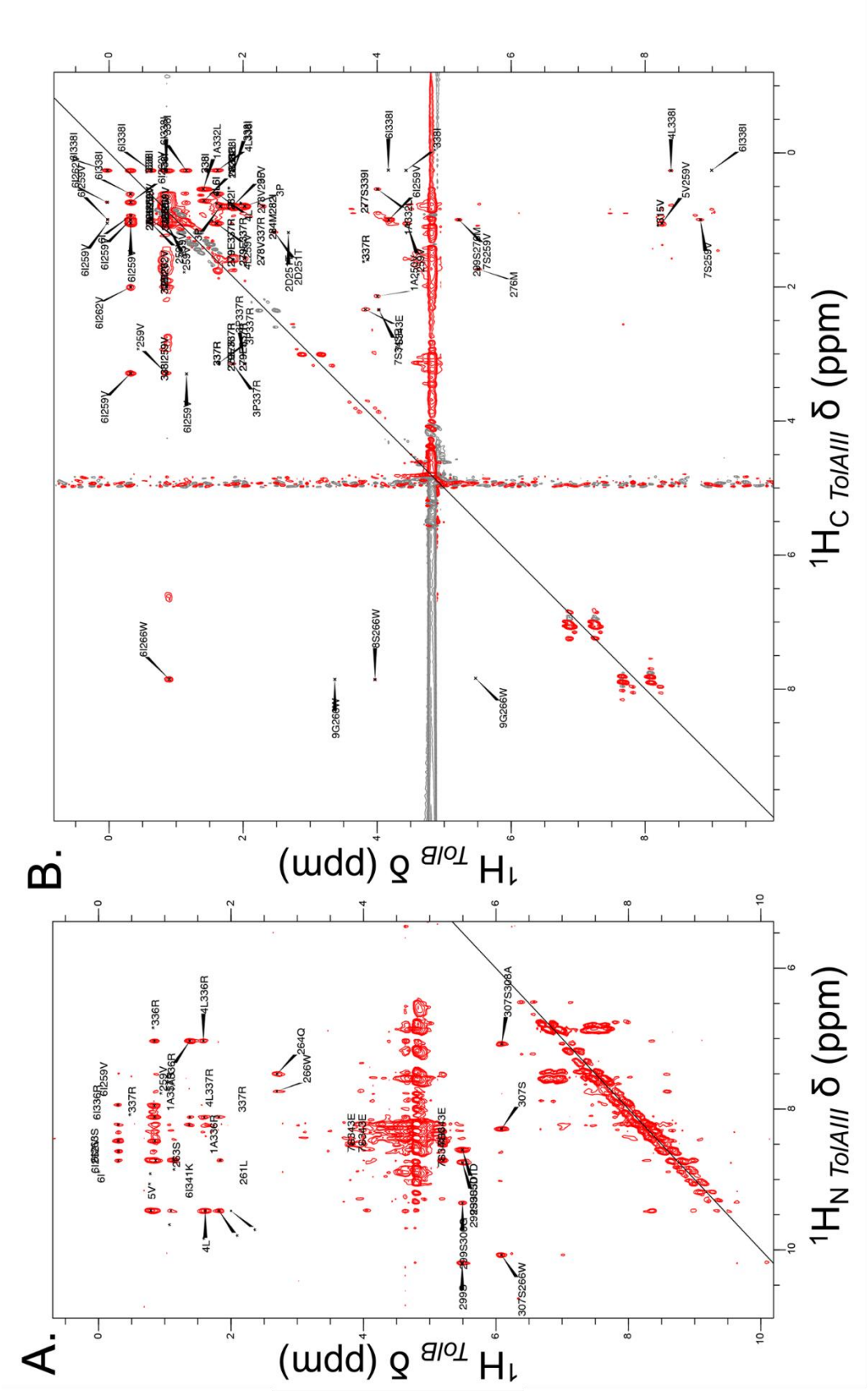


Figure 4.8 **^{15}N ,^{13}C-edited filtered NOESY spectra identify intermolecular NOE constraints between TolB$_{22-34}$ and ^{13}C^{15}N TolAIII.** 2 mM TolB$_{22-34}$, in complex with 1.1 mM TolAIII. Intermolecular NOEs are measured between ^{15}N^{13}C TolAIII and non-isotopically labelled TolB$_{22-34}$ (A) shows a ^{15}N-edited/filtered NOESY spectrum corresponding to the NOEs between TolB protons and TolAIII amide protons. (B) shows a ^{15}C-edited/filtered NOESY spectrum corresponding to the NOEs between TolB protons and TolAIII ^{13}C-bound protons.

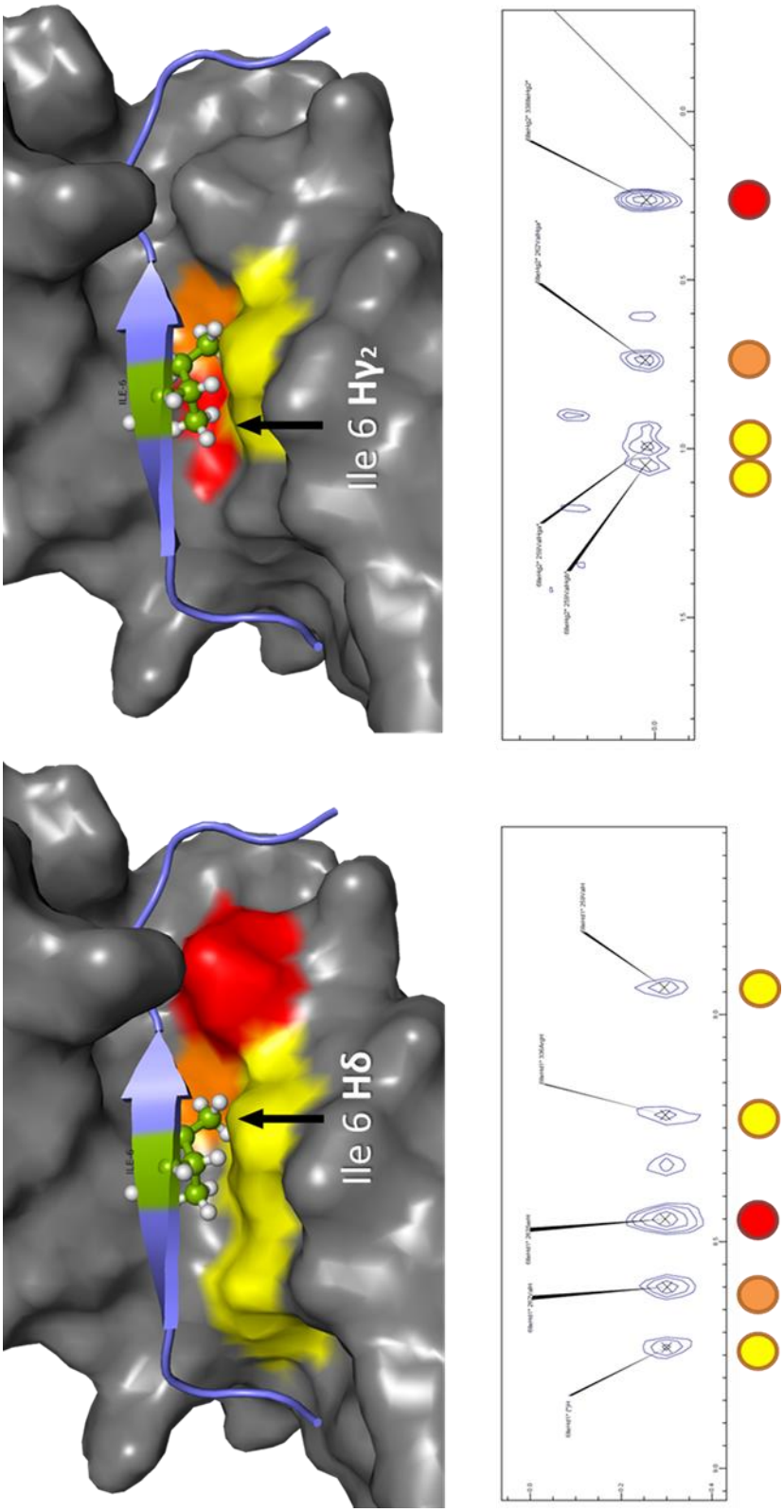


Figure 4.9 **NOE contributions made by TolB$_{22-34}$ sidechains are observable by ^{15}N and ^{13}C edited-filtered NMR spectroscopy techniques.** Slices of the ^{15}N-edited NOESY experiment in figure 4.8 have been shown with a cartoon depiction of the TolB peptide in complex with TolAIII. Strong NOEs, such as the NOE between Ile6 Hδ and TolAIII have been colour mapped according to the colour-coded circle on the surface of the TolAIII protein.

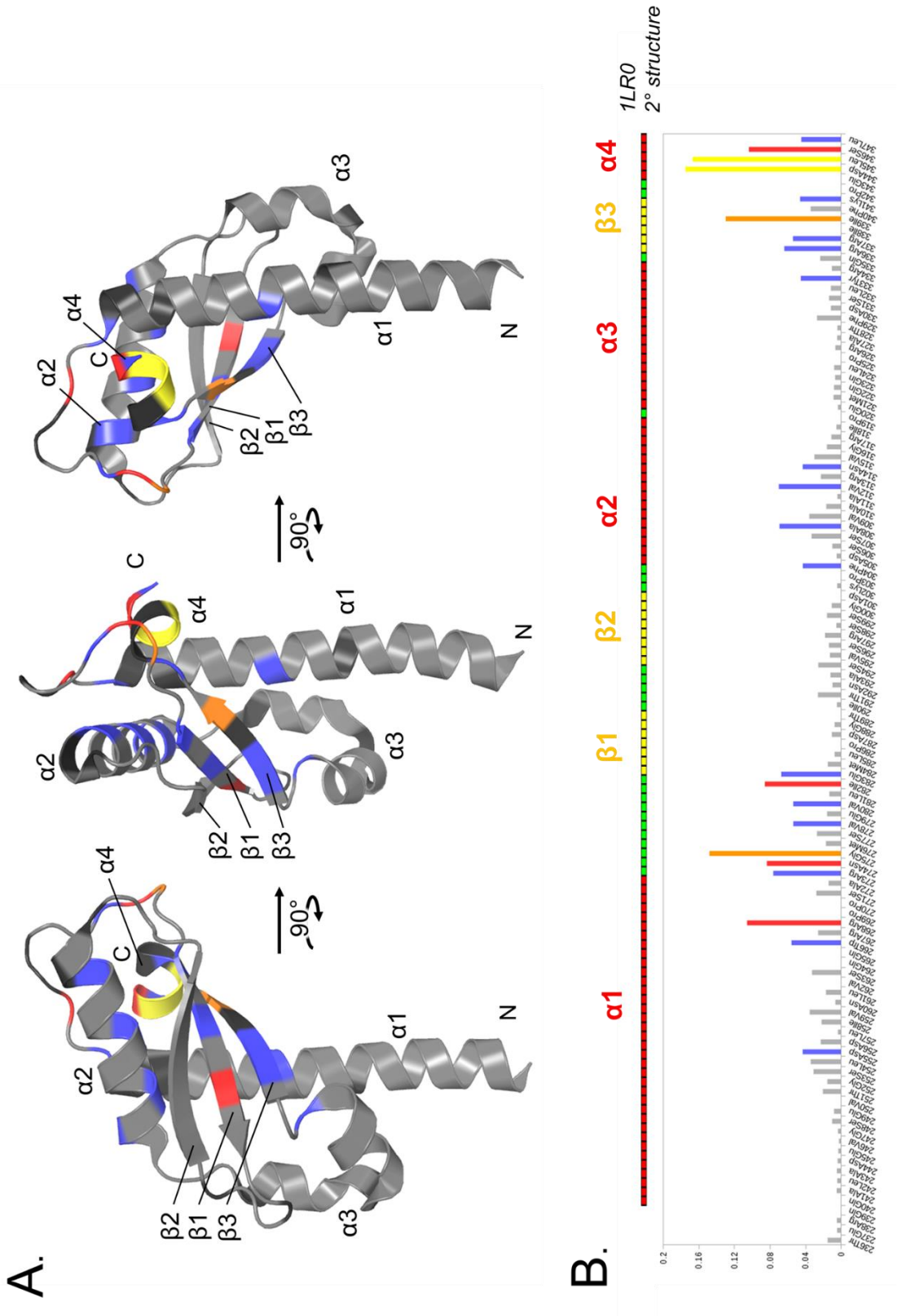


Figure 4.10 **Chemical shift mapping and weighted average chemical shift difference (Δδav) for $^{1}H_{N}$ ^{15}N and $^{13}C\alpha$ chemical shift changes in {^{15}N, ^{13}C}-TolAIII from the free state to saturated with $TolB_{22-34}$.** A) Residues 236-347, shown from the crystal structure 1LR0 have been depicted with colours superimposed corresponding to the average combined chemical shift for these residues. Residues N-terminal to residue 237 have not been shown as the region is intrinsically disordered. To account for the intrinsic differences in chemical shift dispersion of $^{1}H_{N}$ ^{15}N and $^{13}C\alpha$ nuclei, Δδav was defined using the equation $(\Delta\delta av) = \{[(\Delta\delta^{1}HN/3.5)^{2} + (\Delta\delta^{15}N/30)^{2} + (\Delta\delta^{13}C\alpha/20)^{2}]/3\}^{1/2}$ (Pellecchia *et al*, 2000).

of α-helical structure tend to have on average downfield shifted carbon chemical shifts, and β-sheet has upfield shifted carbon chemical shifts (Mielke & Krishnan, 2009). In the TolAIII-$TolB_{22-34}$ interaction, we can observe that areas with major perturbations in the Cα and Cβ correspond to the C-terminal end of helix 1, β-strand 1, a residue in helix 2, and a series of residues in the C-terminus of TolAIII corresponding to β-strand 3 and helix 4, as shown in figure 4.4. In order to obtain a more detailed understanding of the interaction site, the solution structure of TolAIII in complex with $TolB_{22-34}$ was determined.

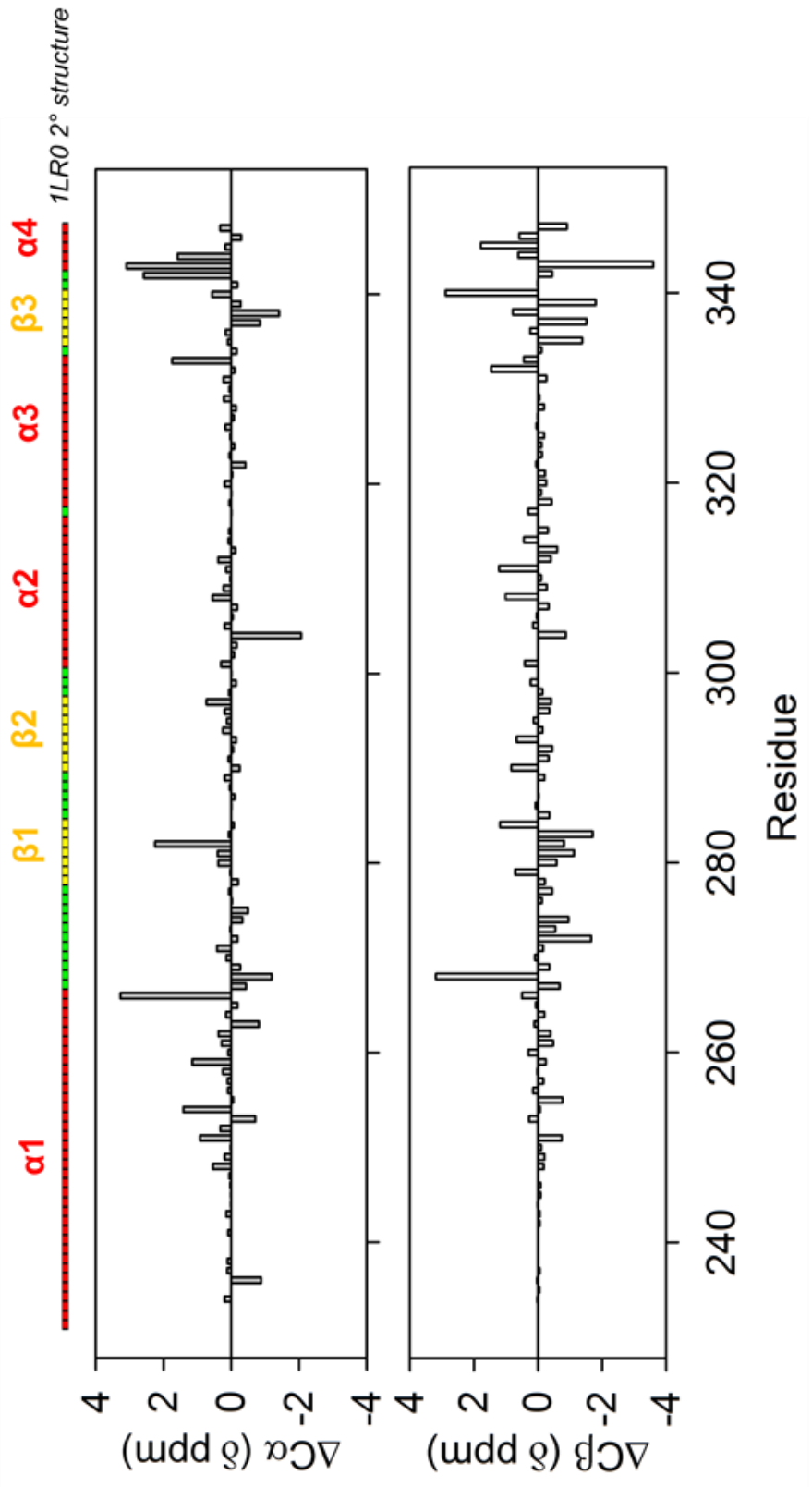


Figure 4.11 **Carbon chemical shift changes in TolAIII as a consequence of TolB$_{22-34}$ binding.** This figure depicts a comparison between the secondary structure of the *P. aeruginosa* TolAIII crystal structure and the major carbon chemical shift changes in TolAIII. The largest chemical shift perturbations are located at the C-terminus of TolA. Specifically centered around residue 343, 344 and 345. An additional region of perturbation occurs at the end of helix 1. Topologically, these two regions both correspond to the area of the protein surround the C-terminus of TolAIII.

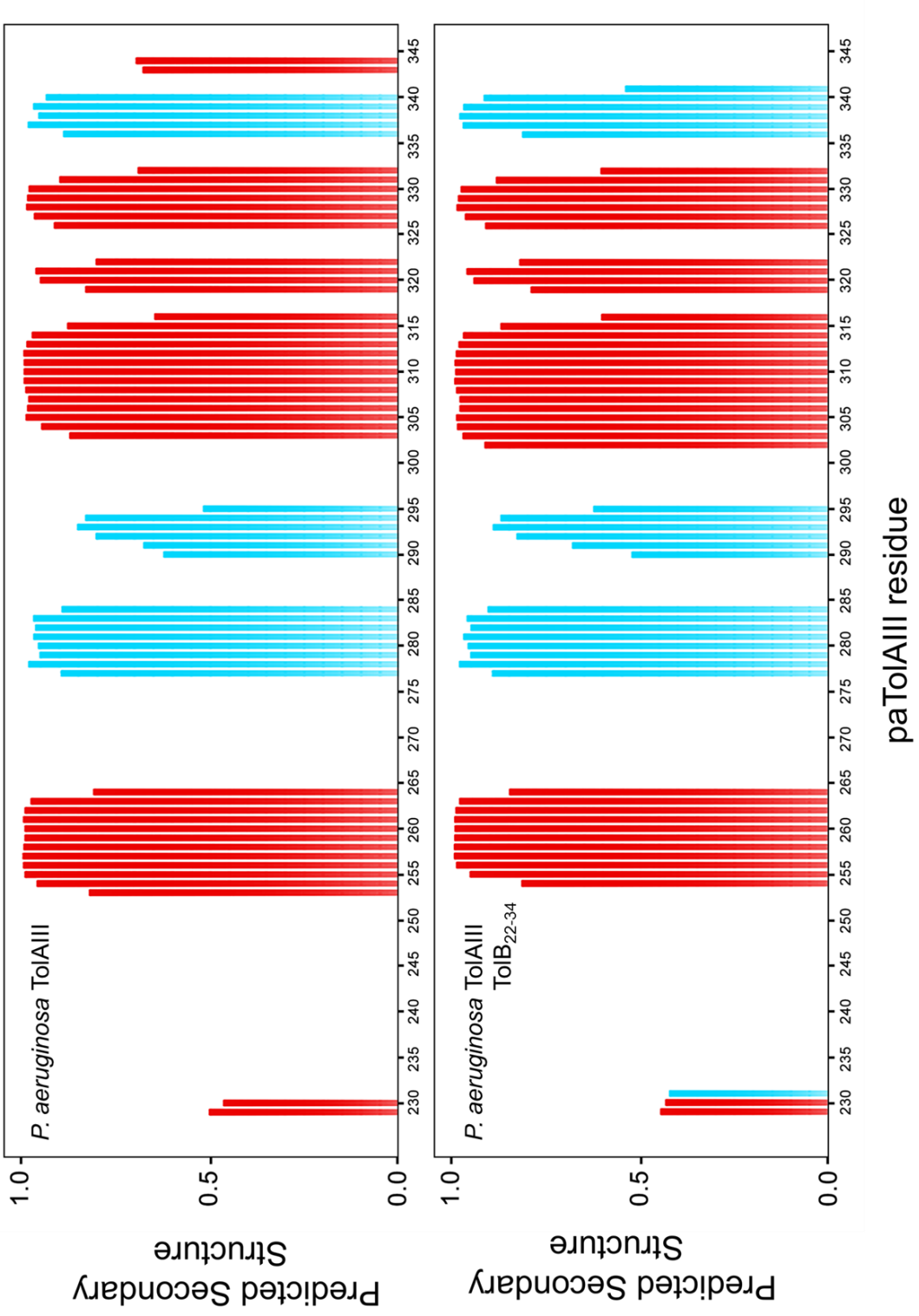


Figure 4.12 **Secondary structural changes in TolAIII as determined by TALOS+ measurements suggest a loss of helical character near the C-terminus after binding TolB.** TALOS+ predicts secondary structure from backbone chemical shift data. These data have been compared for both the bound and unbound form of TolAIII to show that overall secondary structure is predicted to be conserved on TolB binding, with a small rearrangement at the C-terminus.

4.3.7 Full structural determination of the $TolB_{22-34}$-TolAIII complex

To understand the mechanism of the TolA-TolB interaction in *P. aeruginosa,* the structure of TolAIII in complex with $TolB_{22-34}$ was determined by solution NMR. The determination of distance restraints was completed by the assignment of NOE spectra, and with the completion of Hydrogen-Deuterium exchange experiments to determine regions of solvent accessibility and hydrogen bonding patterns. Dihedral angle restraints were determined using the chemical shifts derived from the program DANGLE (Cheung *et al*, 2010). The chemical shifts corresponding to the backbone atoms of TolAIII were assigned using $^{1}H^{15}N$ HSQC, $^{1}H^{13}C$ HSQC, CBCA(CO)NH, CBCANH, HNCA and HNCACO spectra. The chemical shifts corresponding to the sidechains of TolAIII were assigned using 3D HCCCONH, CCCONH, 3D TOCSY-HSQC, 3D ^{15}N NOESY-HSQC, and 3D ^{13}C NOESY-HSQC spectra. ^{15}N-edited NOESY and ^{13}C-edited NOESY spectra were used for the identification of protein-peptide NOEs. TOCSY, CBCANH, CBCA(CO)NH, HNCO, HNCACO, HNCACB, $^{1}H^{15}N$ HSQC and $^{1}H^{13}C$ HSQC were used for the full assignment of $TolB_{22-34}$ bound and unbound to TolAIII. Distance restraints for the TolAIII-$TolB_{22-34}$ structure calculation were derived from 3D ^{15}N NOESY-HSQC, 3D ^{13}C NOESY-HSQC in H_2O, 3D ^{13}C NOESY-HSQC in D_2O, and 2D NOESY (aromatic) spectra. These restraints were used in the calculation of a water-refined ensemble, with Dr. Karthik Rajasekar (Department of Biochemistry, University of Oxford)

In total, 5095 NOE assignments were used in structural calculation, of which 4479 were unambiguous assignments. The pairwise root mean squared deviation (r.m.s.d.) between the top 20 structures (water-refined) was 1.1 +/- 0.2 Å. 87.8% of residues were in favoured positions and the remaining 12.2 % in additional allowed regions by Ramachandran analysis. The lowest energy structure of this ensemble is shown in figure 4.14A, in addition to the ensemble of 20 lowest energy structures in figure 4.14B. The overall fold for TolAIII is essentially structurally similar to what has been previously been described for TolAIII (Witty *et al*, 2002). There is a twisted antiparallel β-sheet with two well-defined α-helices and one kinked helix. Helix 1 is half the length of helix 1 in the TolAIII crystal structure; the N-terminal segment of the protein is entirely disordered and due to complete rotational freedom for clarity has been omitted from figure 4.13 and 4.14. The binding site of $TolB_{22-34}$ is to a cleft on the exposed β-strand 3. $TolB_{22-34}$ forms a parallel β-strand, which is well defined and positioned with precision in the ensemble by 76

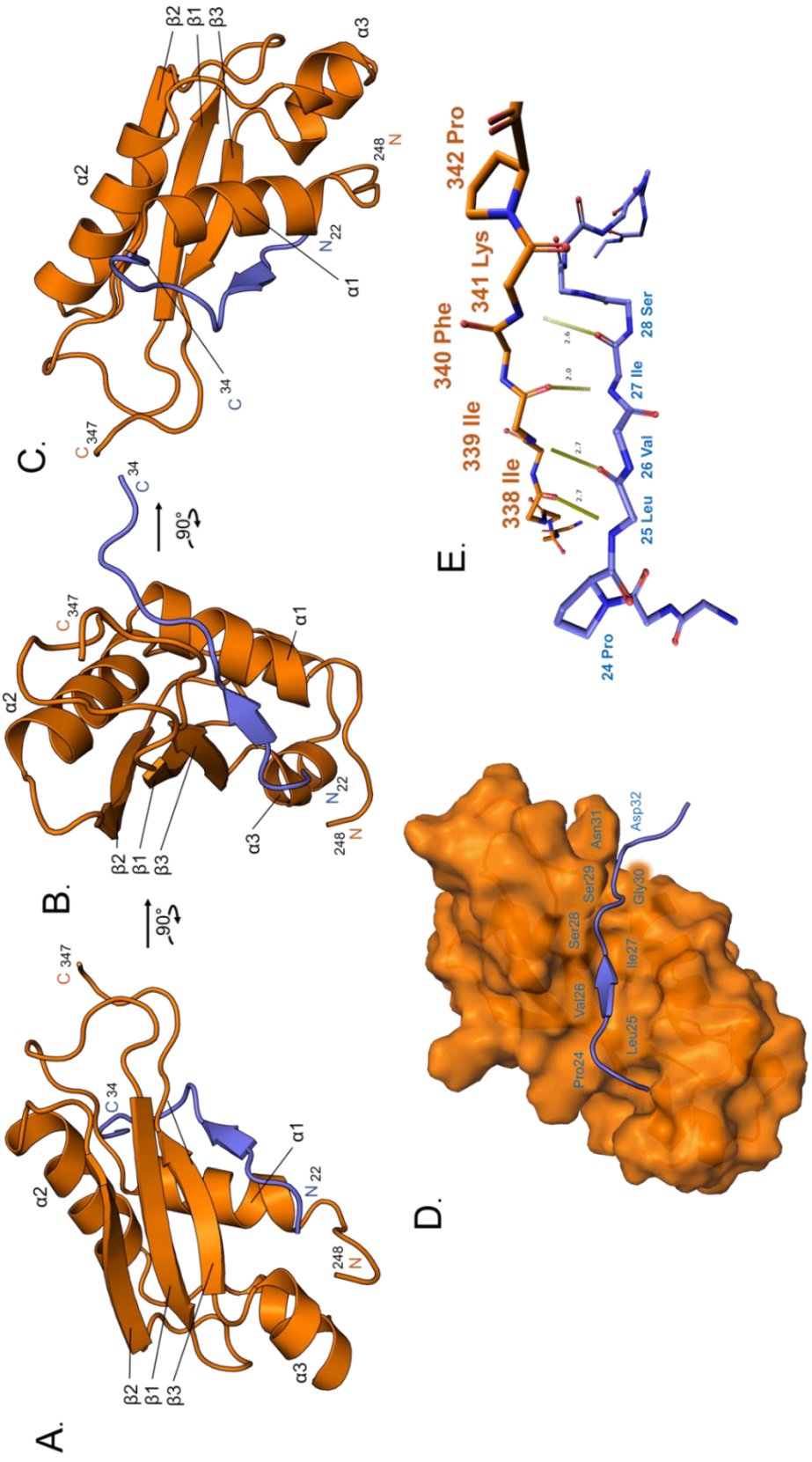


Figure 4.13 **The TolAIII-TolB$_{22\text{-}34}$ interaction is a β-strand complementation mechanism involving the binding of TolB$_{22\text{-}34}$ to a cleft formed by the exposed face of β-strand 3 in TolAIII.** (A, B and C) depict three rotations of the novel TolAIII-TolB$_{22\text{-}34}$ structure. (D) depicts the overall surface of TolAIII in the TolB-bound state, demonstrating that TolB binds to a cleft corresponding to the third β strand in TolAIII. The TolA-TolB interaction mechanism involves parallel β-strand complementation between three paired residues. The conservation of this mechanism is discussed in Chapter 5.

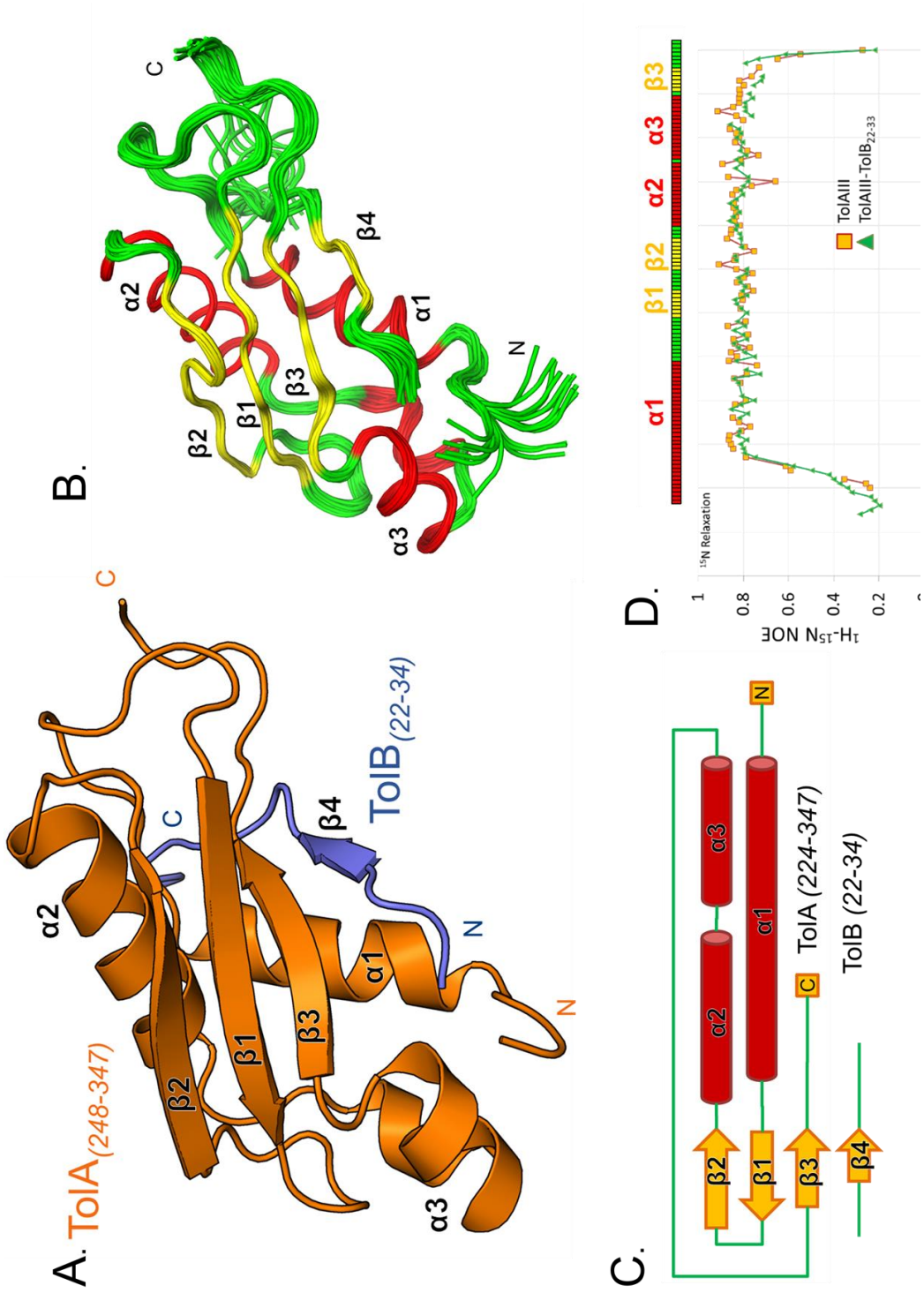


Figure 4.14 **The structure of $TolB_{22-34}$-TolAIII is a tight ensemble indicating the TolB N-terminus binds using an antiparallel β-complementation mechanism.** (A) Depicts in cartoon formt he lowest energy structure from the 20-structure ensemble calculated by Aria. (B) Depicts the overlay of the top 20 structural ensemble for the final calculation. (C) A secondary structure diagram of the TolAIII-$TolB_{22-34}$ complex, indicating the parallel beta-strand arrangement of $TolB_{22-34}$ (D) Overlay of heteronuclear NOE data for both the free and bound states of TolAIII suggest no major changes in the overall flexibility of TolAIII after binding.

NMR distance and dihedral constraints	
Distance constraints	
Total NOE	5095
Unambiguous	4479
Ambiguous	616
Intraresidue	2081
Inter-residue	
Sequential (|*i* – *j* | = 1)	1018
Short (|*i* – *j* | = 2-3)	483
Medium-range (|*i* – *j* | = 4-5)	143
Long-range (|*i* – *j* | > 5)	678
Inter-chain	76
Hydrogen bonds	31
Total dihedral angle restraints	
φ	76
ψ	76
Structure statistics	
Violations (mean ± s.d.)	
Distance constraints (Å)	0.048 ± 0.002
Dihedral angle constraints (°)	1.0 ± 0.1
Number of dihedral angle violation (> 5°)	1.6 ± 0.9
Number of distance constraint violation (> 0.5 Å)	0.7 ± 0.7
Deviations from idealized geometry	
Bond lengths (Å)	$0.0011 \pm 2 \times 10^{-4}$
Bond angles (°)	0.99 ± 0.02
Impropers (°)	2.32 ± 0.09
Average pairwise r.m.s. deviation[a] (Å)	
Heavy[b]	1.0 ± 0.2
Backbone[b]	0.44 ± 0.07
Ramachandran analysis[b] (%)	
Residues in most favored regions	87.8
Residues in additional allowed regions	12.2
Residues in generously allowed regions	0.0
Residues in disallowed regions	0.0

Table 4.1 **NMR and refinement statistics for the *P. aeruginosa* TolAIII-TolB$_{22-34}$ complex.** a. Pairwise r.m.s.d was calculated for the top 20 structures.
b. Statistics applied to residues 3-9 & 254-340

inter-molecular distance constraints, however this causes disruption to the overall secondary structure of TolAIII which loses its C-terminal helix 4. An overlay of Heteronuclear NOE data for the bound and unbound state of TolAIII suggests no major changes in amide flexibility (Figure 4.14D).

4.3.8 Analysis of the $TolB_{22-34}$ binding interface reveals $TolB_{22-34}$ binds to an exposed 'cleft'.

Of the first 6 residues in $TolB_{22-34}$ that interact with TolAIII, 4 appear to play a key role in the formation of the parallel β-strand. A 180° view of the $TolB_{22-34}$ binding site in Figure 4.13A,B,and C shows the binding interface is formed from the exposed strand of β3 and a coaxial stripe from helix 1. The large carbon chemical shift changes observed in β-strand 1, as well as the combined chemical shift changes in Figure 4.10 correspond to the point of interaction with TolB. $TolB_{22-34}$ interacts with TolAIII along a deep binding cleft, indicated in figure 4.13D. The major perturbation to the TolAIII structure is an 'opening' of the binding interface that allows or is caused by the interaction with the $TolB_{22-34}$ molecule. Between the new structure of bound TolAIII and the TolAIII crystal structure backbone atom alignment has an r.m.s.d. of 1.77 Å for resides 244-347 (omitting the N-terminus).

4.3.9 The $TolB_{22-34}$ binding site on TolAIII is occluded by the Leu 345

A structurally aligned overlay of the TolAIII C-terminus between bound and unbound states reveals a clear perturbation to the helix 4 suggested by the unbound TolAIII crystal structure. The $TolB_{22-34}$ molecule and the C-terminal helix 4 from TolAIII are positioned in the same place and so are sterically incompatible. As discussed in figure 3.9 of Chapter 3, the length of the C-terminal helix 4 is expected to be reduced to a single turn in comparison with 2 turns of the TolAIII crystal structure. Crucially however, Leu345 (occupying a position in close proximity to the hydrophobic core) remains in a position occluding the TolB binding site. Chemical shift analysis and NOE data have been shown in figure 4.15 for Leu345, along with a comparison of the positions of Leu 345 in TolB-bound and unbound states. Of note, the proximity of Val262 $H\gamma_1$ to nearby aromatics gives it a unique ^{1}H chemical shift, making the Leu345-Val262 NOE (and return NOE) simple to identify. Val262 and Ser263 both possess clear return NOEs in their corresponding ^{13}C NOESY-HSQC and ^{15}N NOESY-HSQC spectra (not shown). A return Trp266 NOE is ambiguous due to the poor sensitivity of collected aromatic NOESY spectra, however Trp266Hδ1 occupies a unique chemical shift and thus the NOE in 4.15C has no other possible

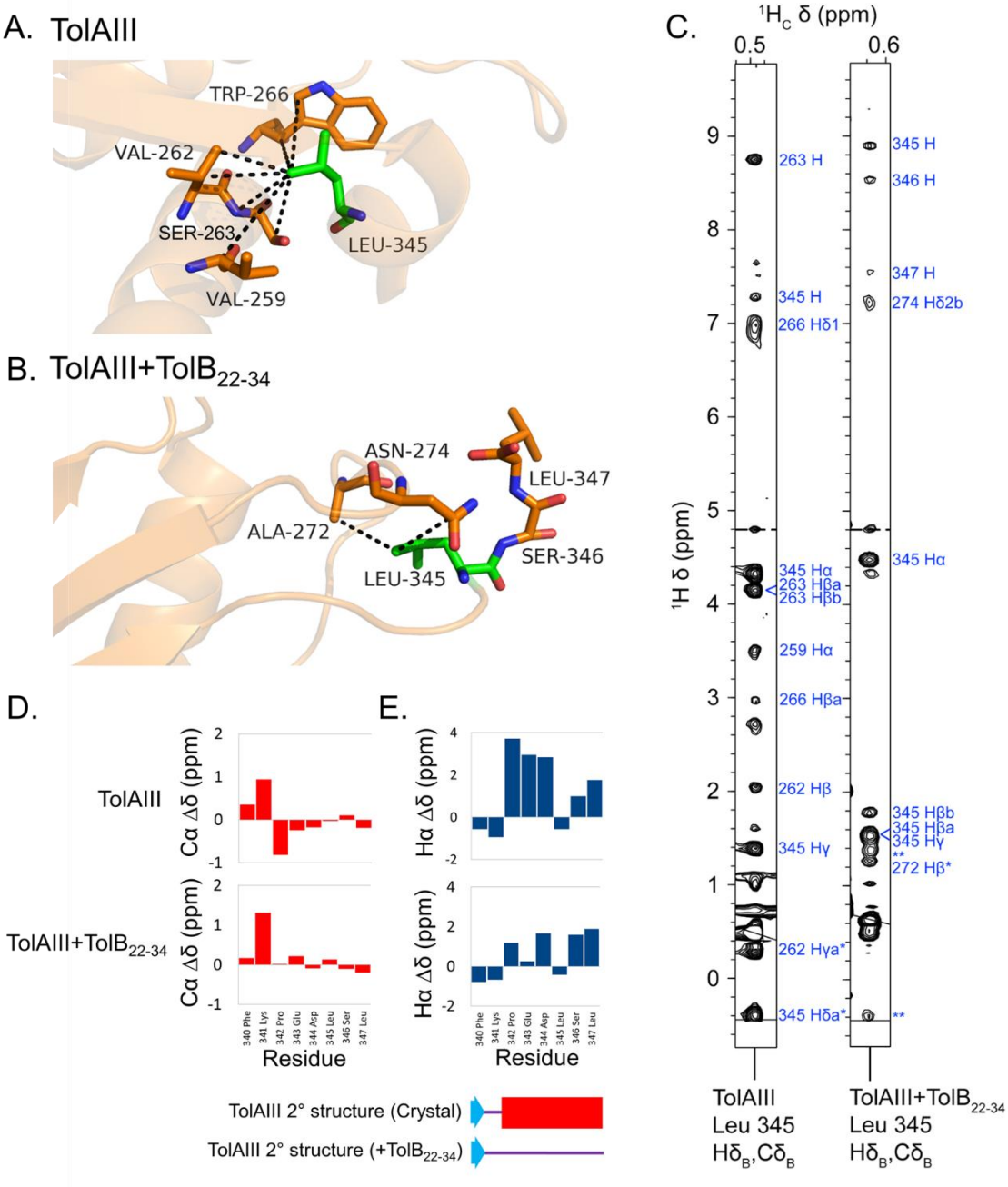


Figure 4.15 **Analysis of NOE contacts determined between Leu345 to residues within the TolB-binding site.** (A) Intramolecular NOEs between Leu345, Trp266, Val259, Val262 and Ser263 suggest that in the unbound TolAIII state, the C-terminus of TolAIII occludes this region. These NOE contacts are displayed on the crystal structure of TolAIII (Witty *et al*, 2002) (B) On TolB binding, Leu345 is displaced, displayed here with the NMR structure of $TolB_{22-34}$-bound TolAIII. New NOE contacts to Ala272 and Asn274 appear and contacts to the TolB-binding site disappear. (C) NOEs from ^{13}C NOESY-HSQC spectra of $^{15}N,^{13}C$ TolAIII are shown alone and in complex, corresponding to residues identified in 4.15A and B. Crosspeaks between Leu345 $H\delta_B$ and aromatic, backbone and sidechain protons are labelled in blue. (D) Chemical shift differences (Δδ ppm) between observed Cα chemical shifts from random coil values. (E) Chemical shift differences (Δδ ppm) between observed Hα chemical shifts from random coil values.

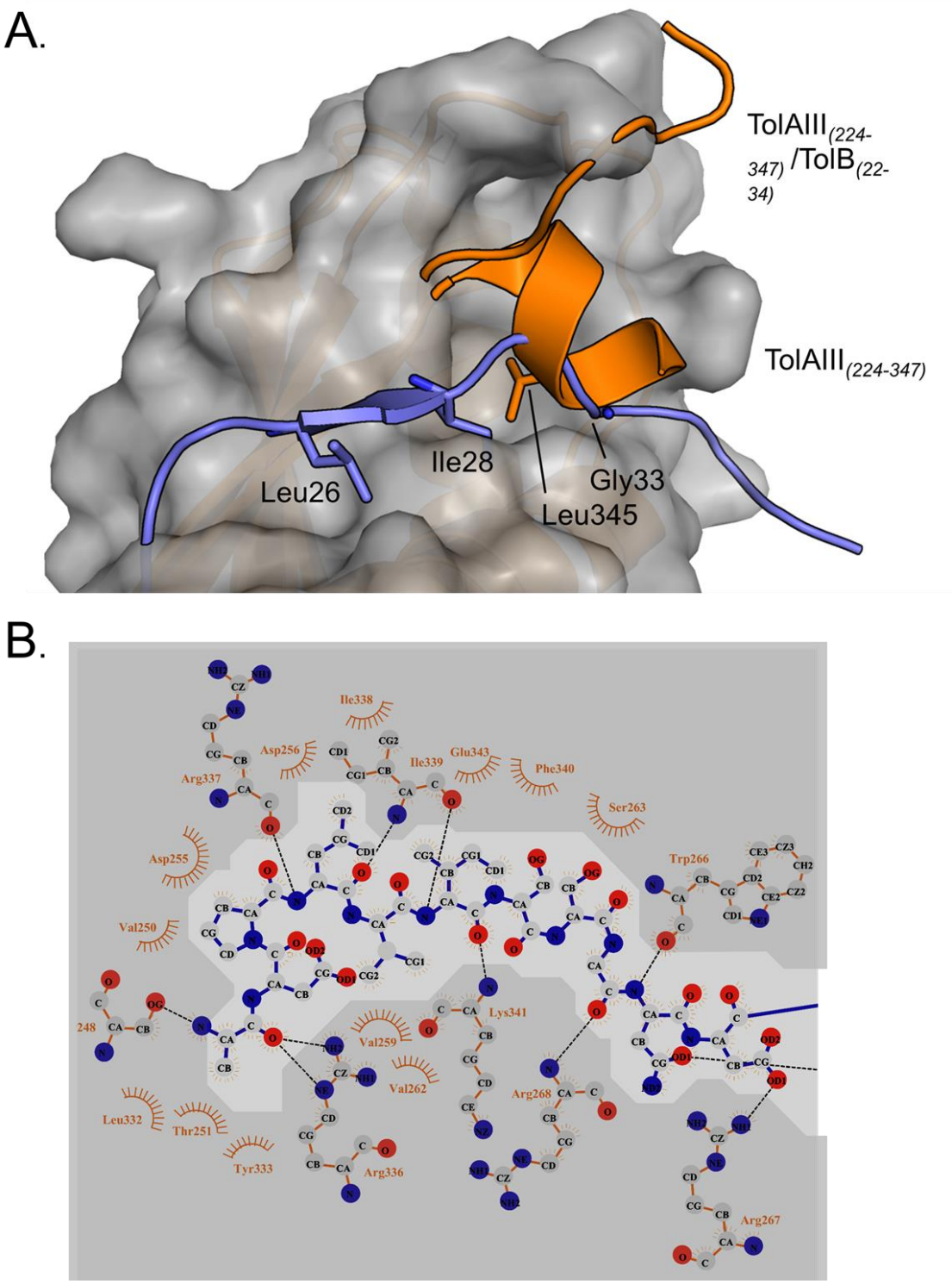


Figure 4.16 **The binding cleft of TolAIII involves interactions with the first 8 amino acids in TolB$_{22-34}$, and necessitates the rearrangement of the C-terminus in TolAIII.** (A) Depicted here as a cartoon, TolAIII has been overlaid three times. The first, in grey, as a surface plot, that omits the last 6 amino acids from the C-terminus for the surface calculation. This is intended to show the 'canyon' that TolB$_{22-34}$ occupies in the TolA-TolB complex. The second and third cartoon overlays of TolAIII, in orange, depict the change in the C-terminus between the bound and unbound states of TolAIII. In Figure 4.15, Leu 345 partially occludes the binding cleft and has been shown here that in the bound state, the TolB peptide would sterically clash with this amino acid if the C-terminus did not assume a novel conformation. (B) Depicts the range of intermolecular contacts made between TolA and TolB, as depicted with the program LigPlot+ (Laskowski & Swindells, 2011)

contribution. These NOEs are lost upon TolB binding. Residues 345-347 undergo minor chemical shift changes, but residues 342, 343 and 344 Hα resonances undergo downfield shifts and Cα resonances undergo upfield shifts, consistent with a loss of helical character. The lack of chemical shift changes in residues 346-347 is consistent with heteronuclear NOE data discussed in Chapter 3 suggesting flexibility in these residues. These data support the conclusion that although the C-terminus is not as helical as the crystal structure, it must undergo a conformational change as Leu345 must be displaced from the TolB-binding site upon binding $TolB_{22-34}$.

Figure 4.16 shows a surface plot of unbound TolAIII omitting the last 4 residues, superimposed in orange with the bound and unbound forms of TolAIII in cartoon form, as well as the position of bound $TolB_{22-34}$. The hypothesis was then posited that $TolB_{22-34}$ either A) causes the displacement of Leu345 via an "induced fit" model, or B) requires the displacement of Leu345 for binding, a "conformational selection" mechanism. Figure 4.16B shows the residue interaction partners for $TolB_{22-34}$, and shows that 8, and potentially up to 10 amino acids from the TolB N-terminus may be involved in interactions with TolAIII. Of further note is that the length of the β-strand complementation appears restricted by the position of Proline 342, as seen in figure 4.13E. The solution structure for *E.coli* in complex with bacteriophage protein and bacteriocin binding partners have been previously described and the TolB interaction site is the same interaction site as used by g3p in the bacteriophage translocation pathway (Lubkowski *et al*, 1999b).

4.3.10 TolAIII-$TolB_{22-34}$ hydrophobic interactions and the displacement of the C-terminus of TolAIII are explained by key NOE restraints determined at the core of the TolAIII-$TolB_{22-34}$ interface.

Ile6 in the TolB peptide (Ile27 in TolB) is a remarkably buried residue of the TolB peptide at the interface of $TolB_{22-34}$ and TolAIII. This residue projects nearly 4 Å deep into the hydrophobic core of TolAIII, as seen in figure 4.16B. Typical examples of the NOE restraints that are made between TolB and TolAIII are depicted in figure 4.17, which illustrates the ^{15}N ^{13}C NOESY-HSQC experiment in blue, with the ^{1}H ^{13}C edited-filtered NOESY shown in red. These spectra have been overlaid for comparison. Of specific interest, and marked with an asterisk and red arrow, is a set of crosspeaks depicting the forward and reverse contribution from each experiment. The blue crosspeak shows the NOE made between ^{15}N ^{13}C TolAIII to unlabelled $TolB_{22-34}$, whereas the red crosspeak is the NOE from unlabelled $TolB_{22-34}$ to ^{15}N ^{13}C TolAIII. The use of these restraints give convincing precision to our understanding of the position of these sidechains in the interface. The

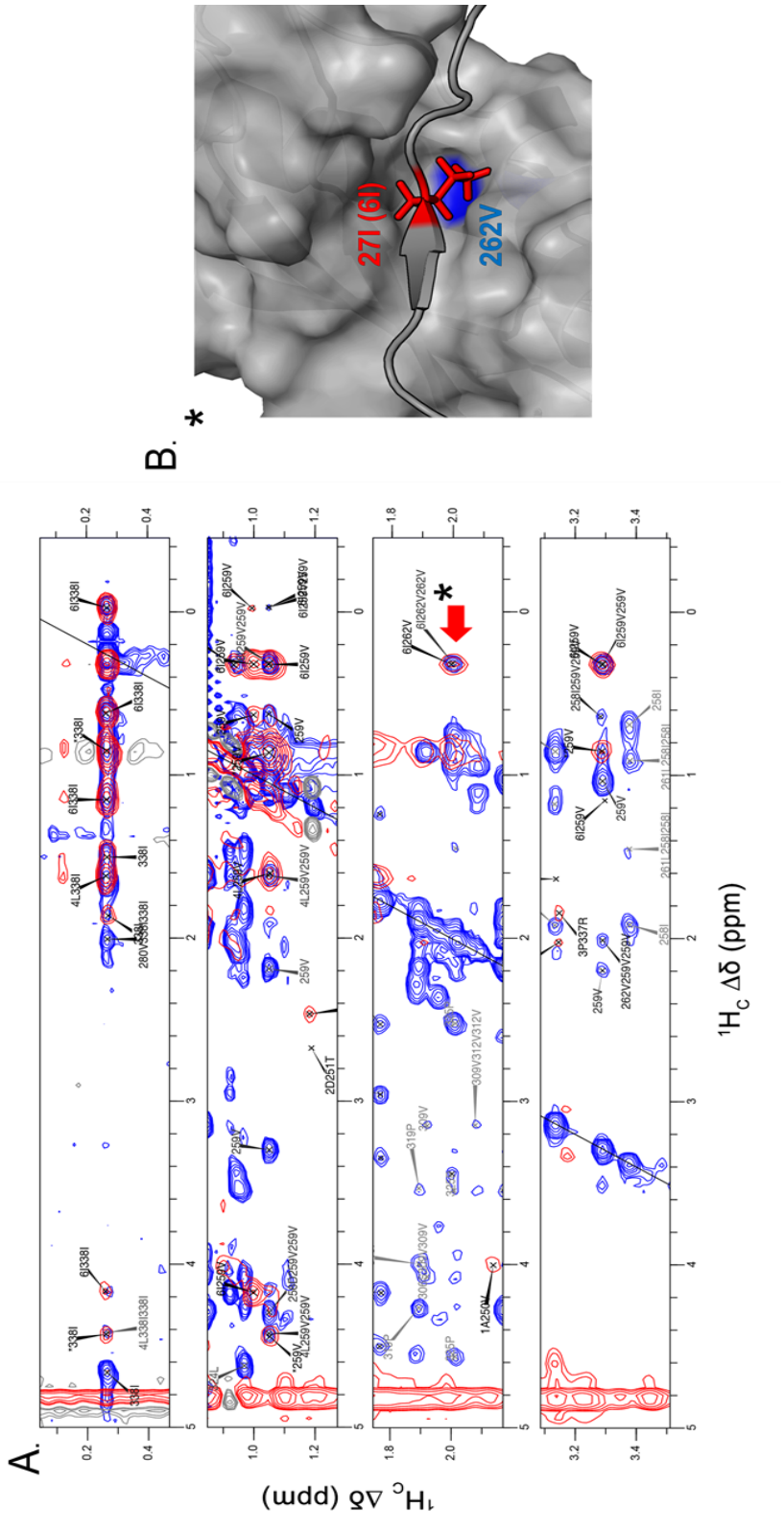


Figure 4.17 **Forward and reverse NOEs from the ^{15}N ^{13}C NOESY-HSQC experiment and the ^{1}H ^{13}C edited-filtered NOESY experiment demonstrate clearly the position of peptide sidechains in the interaction interface**. (A) strips from the ^{13}C NOESY HSQC are shown in red, overlaid with the corresponding strips from the ^{13}C edited-filtered NOESY spectrum. This figure is intended to show that forward and reverse intermolecular NOEs between protein and peptide are clearly visible betweent he two spectra. An example NOE has been shown, with a large red arrow, between Ile27 and Val262, depicted in the cartoon in (B). This residue sidechain penetrated 3-4 Å into the hydrophobic core of TolAIII. Between these two experiments, 76 restraints between protein and peptide were determined, allowing for a precise determination of the conformation of $TolB_{22-34}$ in the binding cleft.

full assignments of these NOEs as described in chapter 2 allow for complete confidence in both the orientation of the peptide in binding, and the number of residues in contact with TolAIII.

4.3.11 The TolAIII C-terminus in *P. aeruginosa* modulates TolB binding affinity

The influence of the C-terminal residues of TolAIII on $TolB_{22-34}$ was investigated by fluorescence anisotropy. Two mutants, $TolAIII_{224-344}$ and $TolAIII_{224-345}$ that substituted in an early stop codon were used to assess the influence of the TolAIII C-terminal helix 4 on TolB binding. Fluorescence anisotropy determined that a C-terminal truncate relative to *wt* TolAIII had a three-fold amplification in binding affinity. This suggests the C-terminus represents a barrier to association for $TolB_{22-34}$. $TolAIII_{224-344}$ in complex with $TolB_{22-34}$ demonstrated a dissociation constant of 86 +/- 3 µM, n=3. *Wt* TolAIII in complex with $TolB_{22-34}$ in the simultaneous experiment demonstrated a dissociation constant of 250 +/- 10 µM, n=3.

4.3.12 The kinetics of the TolAIII-TolB interaction is modulated by the TolAIII C-terminus.

The effect of *P. aeruginosa* $TolB_{22-34}$ on TolAIII was dissected kinetically to better understand the physicochemical properties governing the paradoxic situation of NMR titration data in slow chemical exchange, paired with a weak dissociation constant. To measure the rate of association directly, and rates of dissociation by competition, two species of $TolB_{22-34}$ were used: Fluoroscein isothiocyanate-tagged $TolB_{22-34}$ as detailed in section 2, and a competitor *wt* $TolB_{22-34}$ peptide. The time course of the change in anisotropic character of this labelled molecule was monitored as a function of TolAIII concentration. TolAIII was stoichiometrically mixed in a 1:1 volume ratio with increasing concentrations of TolAIII (100-500 µM) in a stopped-flow apparatus. An increase in the anisotropic character of $TolB_{22-34}$ was monitored as a function of time for each independent experiment, as $TolB_{22-34}$ was bound by excess TolAIII in solution.

The time-dependent association of $TolB_{22-34}$ with TolAIII produces large and clear signal changes observable on TolAIII binding via this method. The k_{obs} is then derived from averaged fluorescence traces and fitted with a single exponential. The on and off rates determined for TolA and $TolB_{22-34}$ demonstrate truncation of the C-terminal helix produces a 3-fold increase in association, as shown in figure 4.18. $TolAIII_{227-347}$-$TolB_{22-34}$ has an on rate of 0.74×10^4 +/- 0.02×10^4 $M^{-1}s^{-1}$, and an off rate of 0.9 +/- 0.1 s^{-1}. $TolAIII_{224-344}$-$TolB_{22-34}$ has an on rate of 2.38×10^4 +/- 0.04×10^4 $M^{-1}s^{-1}$, and an off rate of 1.4 +/- 0.1 s^{-1}. This is commensurate with the experimentally determined dissociation constants for TolAIII-$TolB_{22-34}$ by NMR and fluorescence anisotropy.

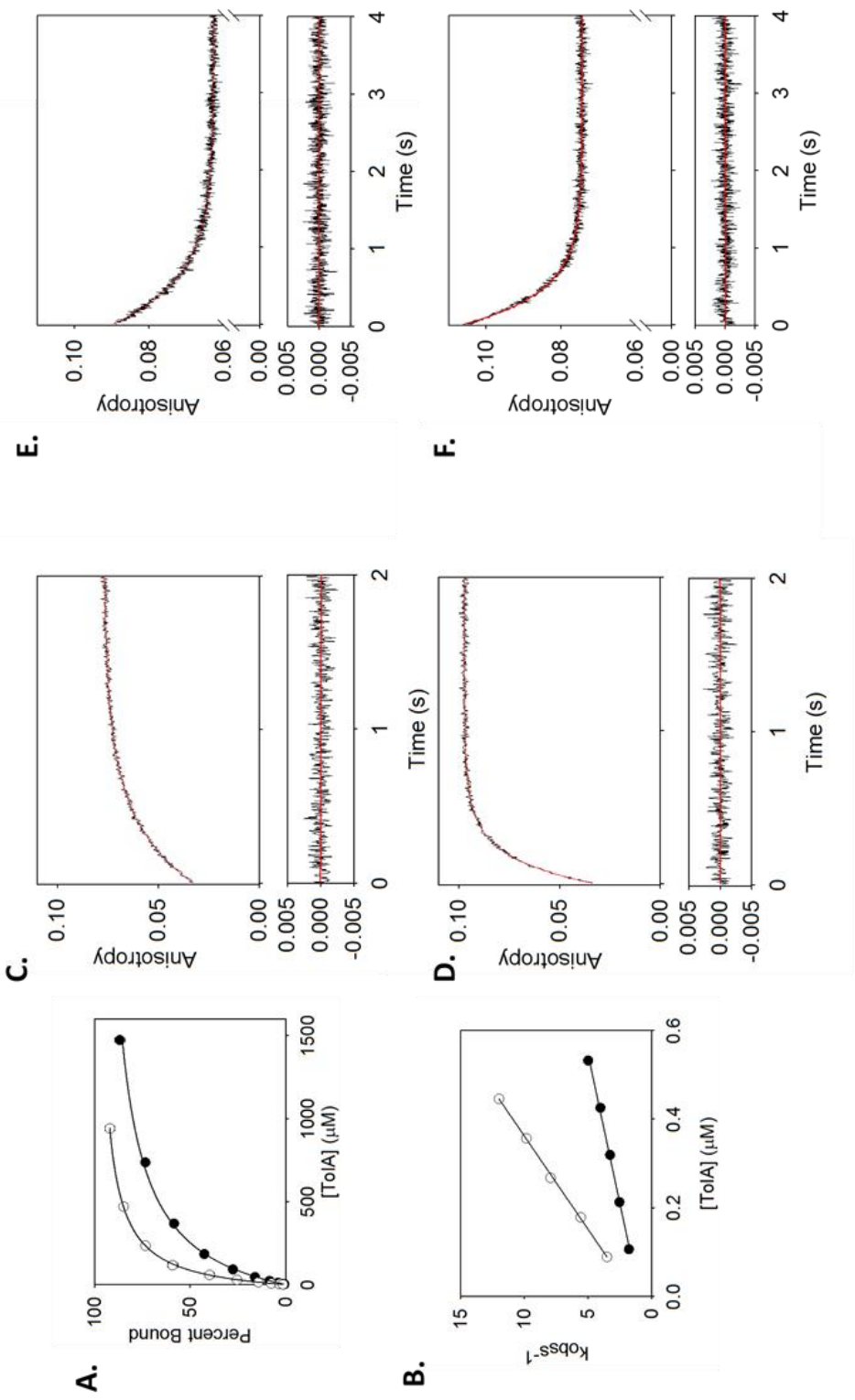


Figure 4.18 **Kinetics of the TolAIII-TolB interaction are modulated by the TolA C-terminus** (A) *In vitro* C-terminal mutants of TolA augment binding affinity 3-fold. TolAIII$_{224-344}$ -TolB$_{22-34}$ has a K_d=86 +/- 3 μM, n=3 whereas TolAIII$_{224-347}$-TolB$_{22-34}$ has a K_d=250 +/- 10 μM, n=3. (B) TolAIII-TolB$_{22-34}$FITC association was monitored by stopped-flow fluorescence anisotropy in 50 mM Tris 100 mM NaCl pH 7.0 at 25 °C. Time dependence of the fluorescence polarization of FITC at 460 nm under pseudo-first-order conditions with TolB$_{22-34}$ was monitored, and a linear regression (n=3) plots for the association of full length TolAII and C-terminal truncated TolAIII indicates truncation of the C-terminal helix has minimal effect on the dissociation rate constants of TolB for TolA however does produce a 3-fold increase in association. Error bars are shown, however are too small to see in plots A and B. plots C-D and E-F show the association and dissociation competition curves for TolB-FITC in complex with truncated and full length TolAIII respectively. Association and dissociation rates are presented in 4.3.13.

4.4 Discussion

The primary aim of this Chapter was to determine the structure of the TolA-TolB complex in *P. aeruginosa.* This included establishing the molecular mechanism of this interaction, as well as providing a detailed assessment of the underlying biophysics and thermodynamics. Further aims sought to analyze of how TolB may interact or be influenced by the PMF via TolA.

This work determines the first structure of TolAIII in complex with its endogenous partner, the intrinsically disordered N-terminal TolA box from TolB. This structure demonstrates for the first time that the TolA-TolB complex is similar but fundamentally different from the interactions that parasitic bacteriocins or bacteriophage make with TolA. Whereas Bacteriophage and bacteriocin mechnisms involve the formation of antiparallel β-strands with either β1 or β3 in TolAIII, my finding is that TolB forms a parallel β-strand with β3 in TolAIII. The interface of the interaction comprises hydrophobic amino acids, a parallel β-strand, four amino acids in length and an overall binding interface involving the first 8-9 residues in TolB which occupy a cleft in TolAIII.

Protein-protein interactions which occur through β-strand addition fall into three main classes: (1) β-augmentation, involving addition of a strand to an exposed acceptor edge of a β-sheet. (2) β-strand insertions and fold complementations, where acceptor domains are not stable folds in the absence of their ligand and (3) the formation of two-stranded β sheets, or 'β-zippers' (Remaut & Waksman, 2006). The TolA-TolB interaction is a classic parallel β-augmentation mechanism.

The purification of ^{13}C ^{15}N doubly enriched $TolB_{22-34}$ was essential to enable successful structural determination of this interaction. Isotope-edited NOESY experiments provided crucial information on the extent of the binding interface, confirmation of intermolecular NOEs, and orientation of the peptide. The precise secondary structure in the $TolB_{22-34}$ sequence in the TolAIII-saturated state could not have been otherwise resolved in the absence of labelled peptide.

P. aeruginosa TolAIII is unusual compared with *E. coli* and *V. cholera* TolAIII molecules. Specifically, the existence of an extended C-terminus that included an alpha-helix in *P. aeruginosa* represented a crucial question mark during the determination of the TolA-TolB binding

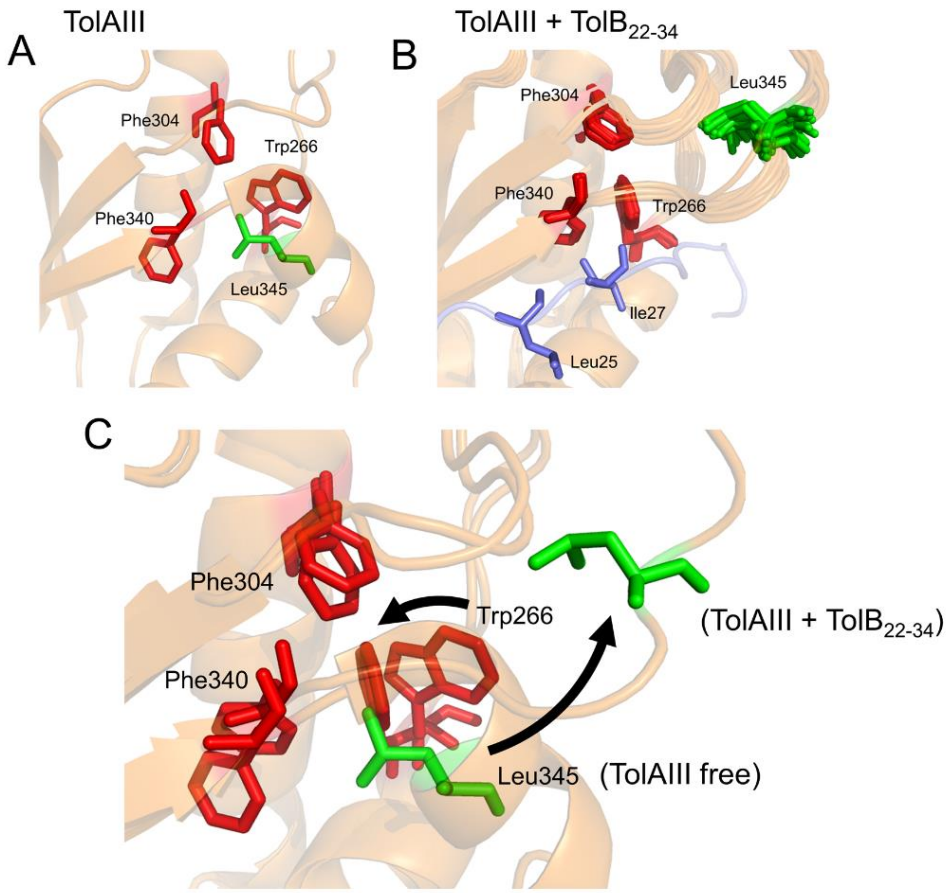


Figure 4.19 **The arrangement of aromatic sidechains in the hydrophobic core of TolAIII was found to undergo major remodelling during the formation of a complex with TolB$_{22-34}$** . (A) TolAIII in the unbound state is characterized by the intercalation of Leu345 in-between aromatic residues packing the hydrophobic core. Depicted here in cartoon form, Leu 345 sidechain is shown in green, with major aromatics indicated in red. (B) A superposition of the top 20 lowest energy structure for TolAIII in the TolB$_{22-34}$ bound state show a characteristic rearrangement of the aromatic residue sidechains in the hydrophobic core of TolAIII. The rearrangement of Trp266 is matched by a large displacement of Leu345. A structural overlay of the unbound and bound states, using the lowest energy structure is shown in (C).

mechanism. Evidence supporting a structured region at the C-terminus largely consistent with the *P. aeruginosa* TolAIII crystal structure was derived from the combined evidence of TALOS+ secondary shift predictions, heteronuclear NOE and NOE measurements. An additional piece of evidence strongly supporting the occlusion of the TolB binding site on TolAIII by an ordered C-terminal structure comes from the observed chemical shift changes in key residues before and after the saturation of TolAIII with TolB$_{22-34}$.

Aromatic rings in amino acid side chains such as tryptophan or phenylalanine typically produce magnetic anisotropic effects called ring currents. The effects of magnetic anisotropy have

been exploited for the identification of ligand binding sites before (McCoy & Wyss, 2002). Ring currents refer to an electron current circulating around the aromatic ring that strongly perturbs the chemical shift of nearby residues. Residues above the plane of the aromatic ring are typically strongly upfield shifted in an NMR experiment. As shown in Figure 4.19, the orientation of the Leu345 sidechain relative to the Trp266 sidechain is suggested to be drastically different between unbound and $TolB_{22\text{-}34}$ states. This is consistent with the observation that the δ^1H chemical shift assignments for Leu345 in the unbound state are strongly upfield shifted. Unbound Leu345 δ^1H has a value of -0.40 ppm. In the $TolB_{22\text{-}34}$ state, this returns to a value of 0.58 ppm, which is much closer to the average chemical shift expected for δ^1H of a Leucine residue in random coil (0.7-0.8 ppm). This enormous chemical shift change is very strong evidence that the Leu345 sidechain packs against the face of the Trp266 aromatic ring and experiences large ring current effects in the free state.

Similar phenomena were observed for the Hα chemical shift of Pro342. In the unbound state Pro342 is strongly deshielded, with a chemical shift value of 3.73 ppm. In the $TolB_{22\text{-}34}$ bound state the chemical shift of the Pro342 Hα is no longer strongly perturbed. The proposed mechanism for the interaction between TolAIII and $TolB_{22\text{-}34}$ suggests that the C-terminus of TolAIII is displaced by $TolB_{22\text{-}34}$ to facilitate ligand binding. Chemical shifts, specifically the chemical shift changes observed in amino acid residues that would be expected to experience ring current effects in the unbound state are more strongly consistent with random coil values in the bound state. This is fully consistent with a rearrangement of the C-terminus and strongly suggests that Leu345 does in fact pack against the face of Trp266 in the unbound state of TolAIII and occlude the TolB binding cleft.

4.4.1 Ton and TolA-TolB share a β-augmentation mechanism dependent on a rearrangement at the C-terminus

TonB forms an activated complex with liganded transporters to transduce energy from the PMF. Specifically, the interaction beteen TonB and its receptor involves a C-terminal rearrangement to accommodate the binding of an intrinsically disordered region of its receptor. This bears strong similarity to the TolA-TolB interaction. The outer membrane receptor FhuA from Escherichia coli mediates TonB-dependent import of ferrichrome (Pawelek *et al*, 2006). A comparison between the TolA mechanism and the TonB mechanism is shown in the Appendix.

5. Validation of the TolA-TolB interaction site *in vivo*

5.1 Introduction

5.1.1 The *tol* phenotype is a characteristic unstable bacterial cell wall in Gram-negative bacteria

The function of the Tol-Pal complex in Gram-negative bacteria is contingent on the presence and combined action of all *tol* proteins within the Tol-Pal gene cluster. The deletion of any of the *tolQ*, *tolR, tolA, tolB,* or *pal* genes produces a *tol* phenotype, which is a characteristic loss of membrane integrity (Godlewska *et al*, 2009). Specifically, this loss of membrane integrity includes cell replication defects, periplasmic leakage, impaired defense against toxic compounds, and the appearance of outer membrane vesicle formation (blebbing) in the presence of low-osmolarity or high-ionic strength solutions (Cascales & Bernadac, 2002; Bernadac *et al*, 1998; Webster, 1991). In *E. coli* K12*,* disruption of the Tol-Pal complex perturbs the polar localization of chemoreceptor clusters as well as causing concomitant impaired swarming motility and chemotactic defects, including biased flagellar rotation (Santos *et al*, 2014). A proposed mechanism involved Tol-Pal forming a physical barrier favouring the accumulation of chemosensory machinery at the poles and mid-cell cytokinetic machinery (Santos *et al*, 2014). In other strains, such as *P. aeruginosa,* as well as *wt E. coli* VW187, *tol* mutants are lethal (Gaspar *et al*, 2000; Dennis & Lafontaine, 1996).

Although cells presenting growth defects survive poorly in adverse conditions, in nutrient-rich LB they continue to grow. The presence of the *tol* phenotype provides a method for assessing the effect of mutations in the *tol* genes in *E. coli* K12. Mutations that produce a *tol* phenotype are inferred to disrupt Tol-Pal function. The “rescue” of this phenotype with plasmid complementation methods allows the testing of plasmid-borne genes with individual mutations designed to impair *tol-pal* function. Cells presenting with a *tol* phenotype exhibit impaired growth in the presence of SDS or other outer membrane destabilizing agents, such as EDTA, which increases the sensitivity of bacterial cells to these adverse conditions, and accelerates the death rate of starved bacteria (Russell 1969, Strange and Dark 1965). EDTA causes a non-specific increase in cellular permeability via the chelation of Mg^{+} and Ca^{2+} causing the disruption of the outer membrane.

5.2 Aims

In Gram-negative bacteria, the *tolA* and *pal* gene products are embedded in the inner and outer membranes, respectively. TolB is suggested to act as an "On-Off" switch for Tol-Pal function, as an intermediary that either binds Pal or TolA (Bonsor et al., 2009). The aim of this chapter was to functionally test structure-based mutations in the binding interface between TolA and TolB identified in chapter 4, assessed by *in vivo* growth assay. Specifically, I tested the hypothesis that mutation within a conserved TolA interaction motif in the N-terminus of TolB would abrogate binding *in vivo*, and subsequently determined the affinity of N-terminal mutants for TolA *in vitro* by spectroscopic methods.

Some Gram-negative bacteria, such as *P. aeruginosa,* possess small structural regions in the C-terminus of TolAIII, posterior to strand 3. In chapter 4 it was suggested that the unfolding of this structure was essential for the binding of TolB. Using bioinformatics, I show that Gram-negative bacteria TolAIII molecules alternatively fall into two categories, with and without C-terminal extensions, and I use an *in silico* comparative modelling approach to present an explanation for how the presence or absence of a C-terminal 'tail' might affect TolB binding.

5.3 Results

5.3.1 The TolA-TolB interaction interface is a strongly conserved binding interface in Gram-negative bacteria.

The *tol* genes are characteristically divergent in their sequences. This reflects the strong external and selective evolutionary pressure applied by the appropriation of *tol* genes by bacteriophage and bacteriocins (Sturgis, 2001). The divergence of these sequences has made the TolA-TolB binding interface impossible to identify by bioinformatic methods alone. A sequence-based alignment of TolA-TolB interaction sites in Gram-negative bacteria provided an additional layer of refinement that revealed a conserved interaction motif in TolA and TolB.

Using the knowledge of the binding interface gained from the structural and biophysical information in chapter 4, I completed a sequence-based alignment of the TolA-TolB binding interface across a representative slice of all Gram-negative bacteria possessing the Tol-Pal complex. Using the program PRED-TAT (Bagos *et al*, 2010), I predicted the Sec-signal peptide cleavage site in the N-terminus of 34 sequences representing each of the major phyla. I removed these sequences and performed an alignment of the TolB sequences weighted towards the first

40 amino acids of every species, containing the N-terminus of TolB known to be disordered in *E.coli*. In cases where the Sec-signal could not be determined, sequences were aligned with the signal peptide present. The alignment of TolB-termini with Sec-signals is shown in figure 5.1. For these genera, the position of the Sec- signal in three could not be determined. Three genera possessed N-terminal signal sequences predicted to be cleaved by the Tat pathway: Bradyrhizobia, Rhodopseudomonids and Pelagibacteria, and all correspond to the phylum alphaproteobacteria.

After removal of the Sec-signal, a ClustalOMEGA alignment of TolB N-termini reveal three strongly conserved N-terminal residues in TolB, as shown in figure 5.2. The N-terminus of TolB has been shown to undergo an order-disorder transition dependent on the binding state of the TolB β-propeller (Bonsor *et al*, 2009). The crystal structures of TolB in TolA-accessible and inaccessible states are shown, and the position of these conserved residues indicated in red. Notably, in the fully unbound state, TolB has a disordered N-terminus, of which 6 residues are not visible in the crystal structure, shown in *E. coli* by solution NMR (Bonsor *et al*, 2009). When the N-termini of TolB are aligned, the three residues strongly conserved are Proline 36, Isoleucine 27 and Isoleucine 25. Isoleucine residues 27 and 25 are most commonly substituted with Valine or Leucine residues, and form a [Hydrophobic]-X-[Hydrophobic] motif. The only genera for which the [ILV]X[ILV] motif is absent is Acidobacterium, in which the precise location of the Sec signal cleavage in TolB could not be determined, thus it is possible that the absence of this motif reflects only a poor alignment of the N-terminus to the other species.

The length of the TolB N-terminus is variable, and although cannot be precisely determined in all species without structural data, can be estimated by secondary structure prediction and conserved residues, and its position in all known crystal structures of TolB. A highly conserved Proline (Pro36) is present at the beginning of strand 1 in all Gram-negative bacteria except for three genera. The six downstream residues are strongly predicted by JPred4 prediction to possess β-strand secondary structure, hence this Proline likely represents the beginning of the structured N-domain in TolB. The position of this Proline near the end of the first strand in *E. coli* is shown in figure 5.2.

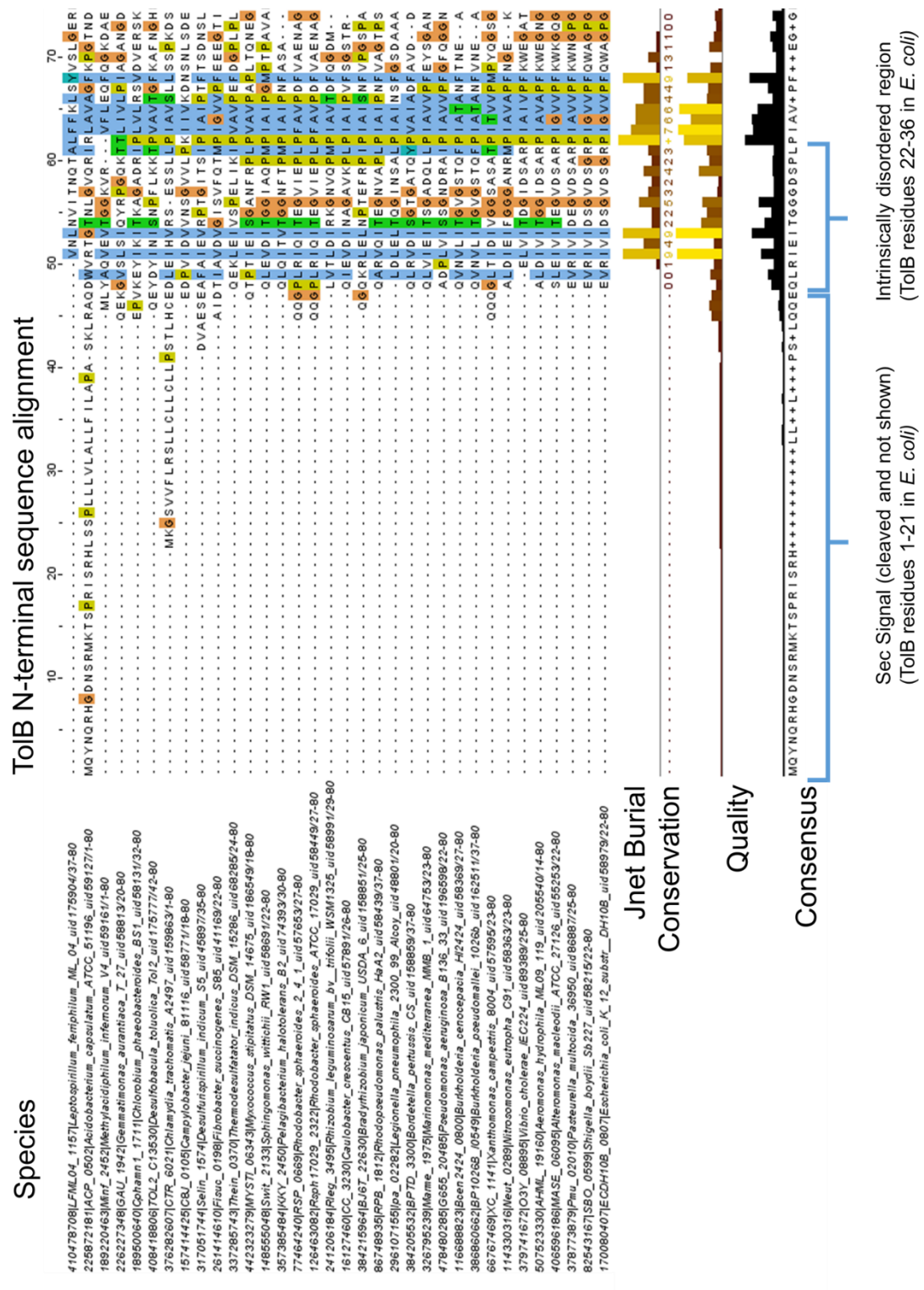


Figure 5.1 **Analysis of aligned N-termini of TolB demonstrates conserved regions preceded by a cleavable Sec/Tat signal.** A ClustalOmega alignment of TolB proteins representing 34 representative genera from across all bacteria containing the Tol-Pal cluster was aligned to show the N-terminal region conserved in TolB. Signal sequences have been removed, by predicting the cleavage site using PRED-TAT (Bagos *et al*, 2010). Two species did not possess identifiable Sec signal cleavage sites, in *A. capsulatum* and *C. trachomatis*. N-terminal weighting of the TolB alignment was achieved by using the first 40 amino acids of each species after each predicted Sec signal site (remaining residues cropped from figure). A partially conserved region is shown, consistent with the intrinsically disordered region in *E. coli* TolB, containing the TolA binding region. This alignment was used for the generation of a consensus sequence in figure 5.2

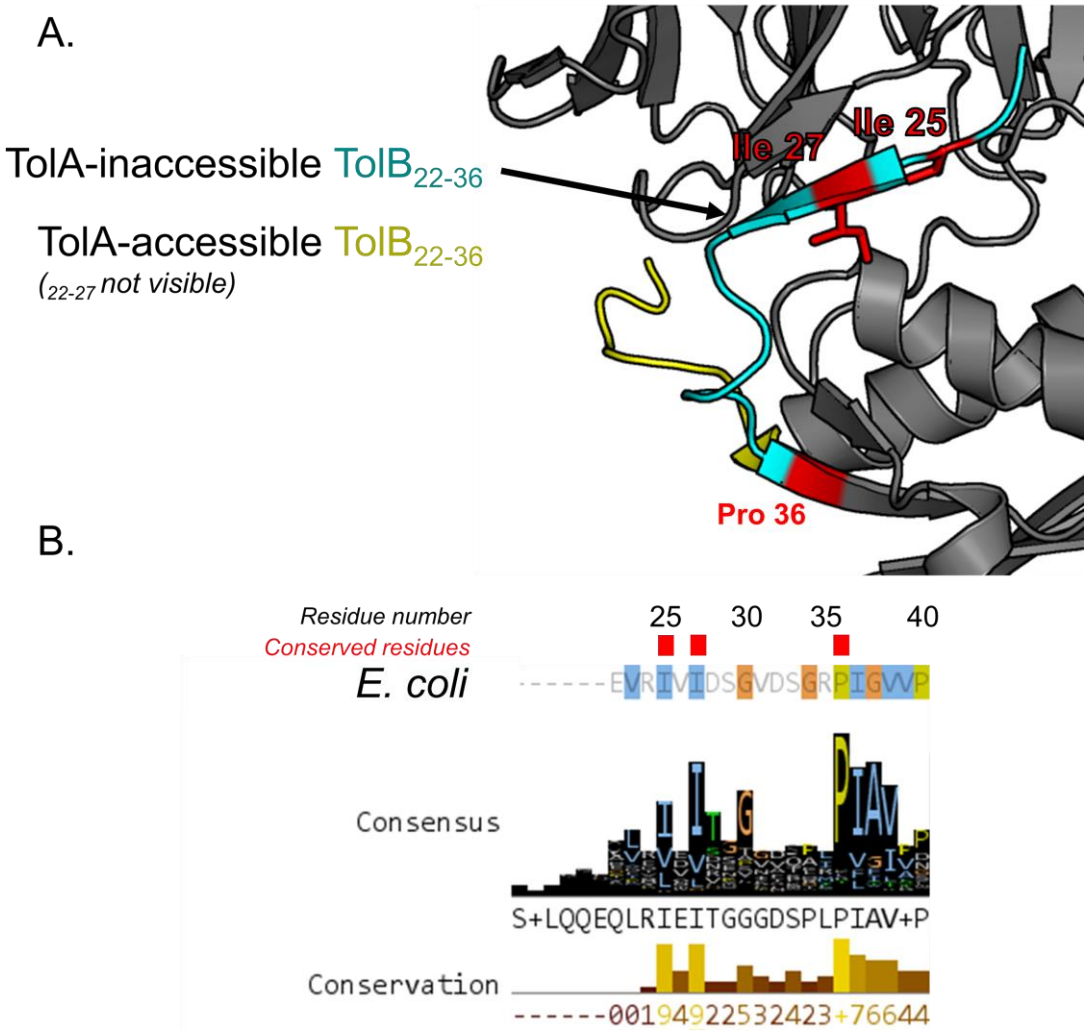


Figure 5.2 **The TolB N-terminus is an intrinsically disordered region with a conserved motif.** (A.) Depicts a cartoon representation of the Sec-cleaved N-terminal region of *E. coli* TolB in the free and Pal-bound state (PDB 1CRZ and 2HQS) (Bonsor *et al*, 2007; Abergel *et al*, 1999). In the Pal-bound state, shown here as "TolA-inaccessible", the N-terminus of TolB in cyan is sequestered away into a binding cleft between the N- and C-domains of TolB. In the free state the N-terminus is disordered in solution, in yellow, and accessible for TolA binding. (B.) Depicts the conservation of residues from 34 representative genera, and reveals that two hydrophobic residues and a downstream proline are strongly conserved. In *E. coli* these residues are isoleucines, the motif conserved across all genera is a [ILV]X[ILV] motif.

The importance of the [ILV]X[ILV] motif becomes apparent when comparing it to the binding site on TolAIII for TolB. In the *P.* aeruginosa NMR structure from chapter 4, the [ILV]X[ILV] motif is the segment of TolB responsible for strand complementation with TolAIII. Aligning the TolAIII C-terminal residues and the N-terminal TolAIII residues reveal that the corresponding residues in TolAIII that form parallel interactions with TolB are similarly conserved, as shown in figure 5.3. The conservation scores for the TolAIII and TolB binding motifs from the same genera are indicated, and correspond closely to the two pairs of residues that form the β-strand parallel complementation in the *P. aeruginosa* TolA-TolB binding mechanism.

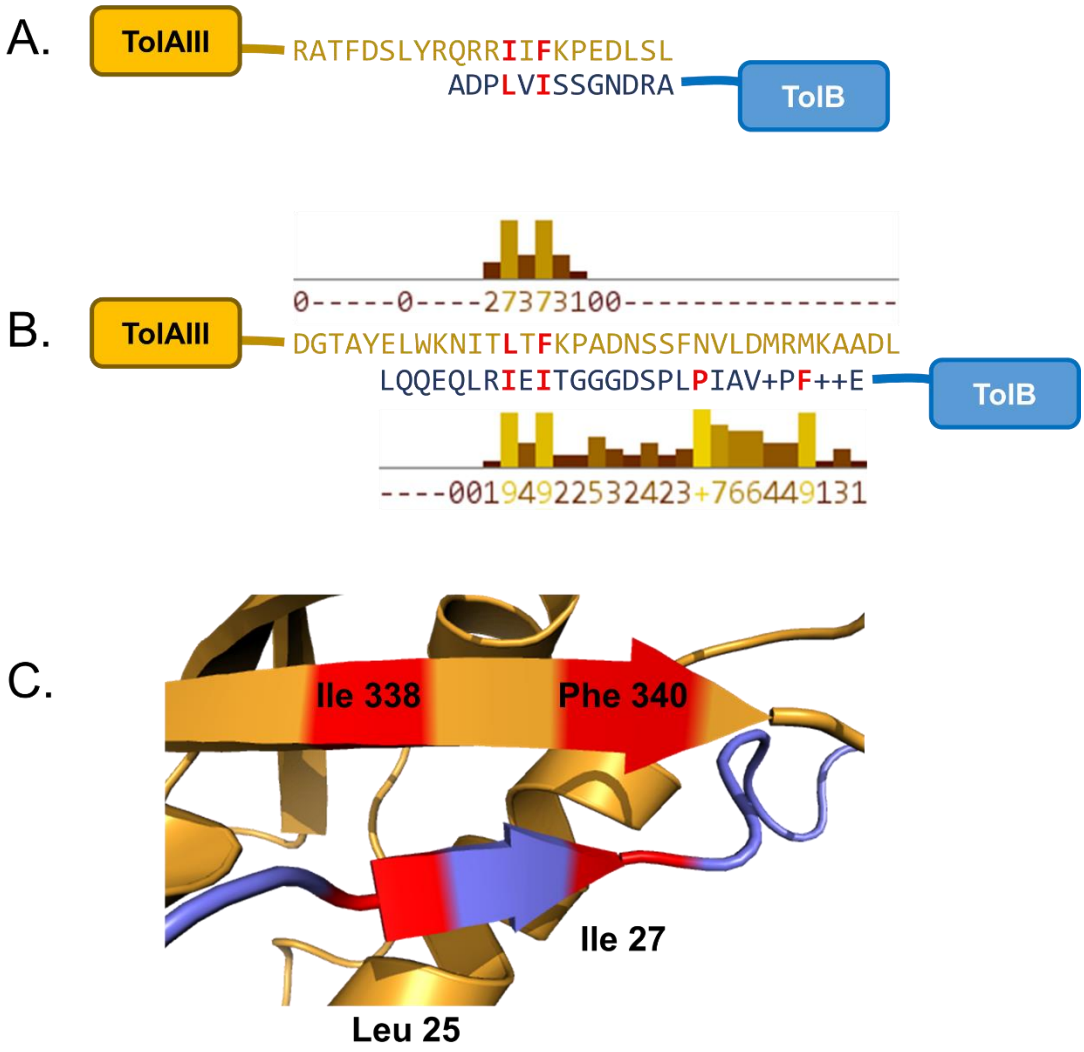


Figure 5.3 **The TolA-TolB binding interface is strongly conserved in Gram negative bacteria.** (A) A depiction of the binding interface of TolA and TolB shows the two residues in TolA and the two residues in TolB that form an aligned pair in strand complementation, as discussed in chapter 4. (B) A ClustalOmega alignment of representative genera of TolB N-termini superimposed with a ClustalOmega alignment of TolA C-termini reveals the parallel complementation mechanism has conserved residues across Gram-negative bacteria. (C) A cartoon representation of the 4 conserved residues, shown in red, that in TolA and TolB form the TolA-TolB binding interface. Conservation statistics determined by Clustal Omega and visualized using JalVIEW.

The TolB-binding motif in TolA is more divergent than the TolA-binding motif in TolB. Specifically, although the conserved residues often fit the [hydrophobic]X[hydrophobic] motif seen in the TolB [IVL]X[IVL] motif, in some genera, such as Acidobacteria or Bordetella, one of the hydrophobic residues is substituted with a Tyrosine. From secondary structure prediction however, as discussed in Chapter 3, the overall secondary structure of the TolAIII globular domain is predicted to remain largely consistent between species despite the low sequence similarity. What is clear is the presence of largely conserved pair of hydrophobic residues on TolA and TolB corresponding to the TolA-TolB binding site in chapter 4.

5.3.2 The *tol* phenotype in Δ*tolb E. coli JW5100* can be rescued by plasmid complementation

In order to disrupt the TolA-TolB binding site and observe the effects on the cell, I performed an assay for the measurement of cell defects *in vivo.* Δ*tolb* cells were grown in on solid LB-agar medium in the presence and absence of SDS. A delayed onset of growth for these cells is the standard reporter for a *tol* phenotype (Bernadac *et al*, 1998). This was then repeated in liquid LB medium, and is shown in figure 5.4. The phenotypic impact of a *tolB* deletion is clearly visible, as cells that are exposed to 0.4% w/v and 2.0% w/v SDS growth conditions display a clear reduction in cell growth for the Δ*tolB* strain. Impaired growth in Δ*tolB* is seen in the liquid growth assay as well (Figure 5.4B) in 2% w/v SDS. Δ*tolB* strain cells begin to demonstrate strongly impaired growth at SDS concentrations of 0.01-0.001% w/v SDS. Complementation of Δ*tolB* cells with pDAB17 encoding *tolB* reverses this growth defect. Induction with arabinose was tested, however as the pDAB17 plasmid exhibited leaky expression, this was not used in subsequent experiments. Cells are susceptible to SDS at concentrations similar to that of *wt* cells, and in a solid agar growth assay demonstrate improved growth over Δ*tolB.* There is a small decrease in cell growth between Δ*tolB* + *tolB* and *wt* strain cells, which can be quantified as an increase in susceptibility to SDS at concentrations greater than 0.002% w/v, which could be attributed to the inefficiency of expressing *tolB* outside of the gene cluster. TolB is lethal when overexpressed (data not shown) suggesting possibly that although *tolB* in Δ*tolB* + *tolB* uses leaky expression from the pDAB17 plasmid, the concentration of TolB in the cell must normally be tightly regulated.

5.3.3 *tolB* mutants that disrupt the [IVL]X[IVL] binding motif result in a *tol* phenotype

I designed TolB mutants to disrupt the TolA-TolB interaction, by site-directed mutagenesis of the pDAB17 plasmid. N-terminal truncations of TolB up to residue 25 and a full truncation of TolB up to residue 33 in *E. coli* have been observed to produce a *tol* phenotype and corresponding loss of binding *in vitro* (Bonsor *et al*, 2009). The aim of this experiment was to determine if a *tol* phenotype could be produced by mutagenesis of the [IVL]X[IVL] motif in TolB. Constructs that corresponded to a triple alanine mutation of Ile 25, Val 26 and Ile 27, as well as a double alanine substitution of Ile 25 and 27, and single alanine substitutions of Ile 25 and Ile 27 were produced. The result of complementation of these plasmids with Δ*tolB* cells demonstrated that in *E. coli* even a single point mutation within the N-terminus of TolB cannot be tolerated for the maintenance of Tol-Pal function. In Figure 5.5 the resistance of Δ*tolB* cells to 2% w/v SDS on

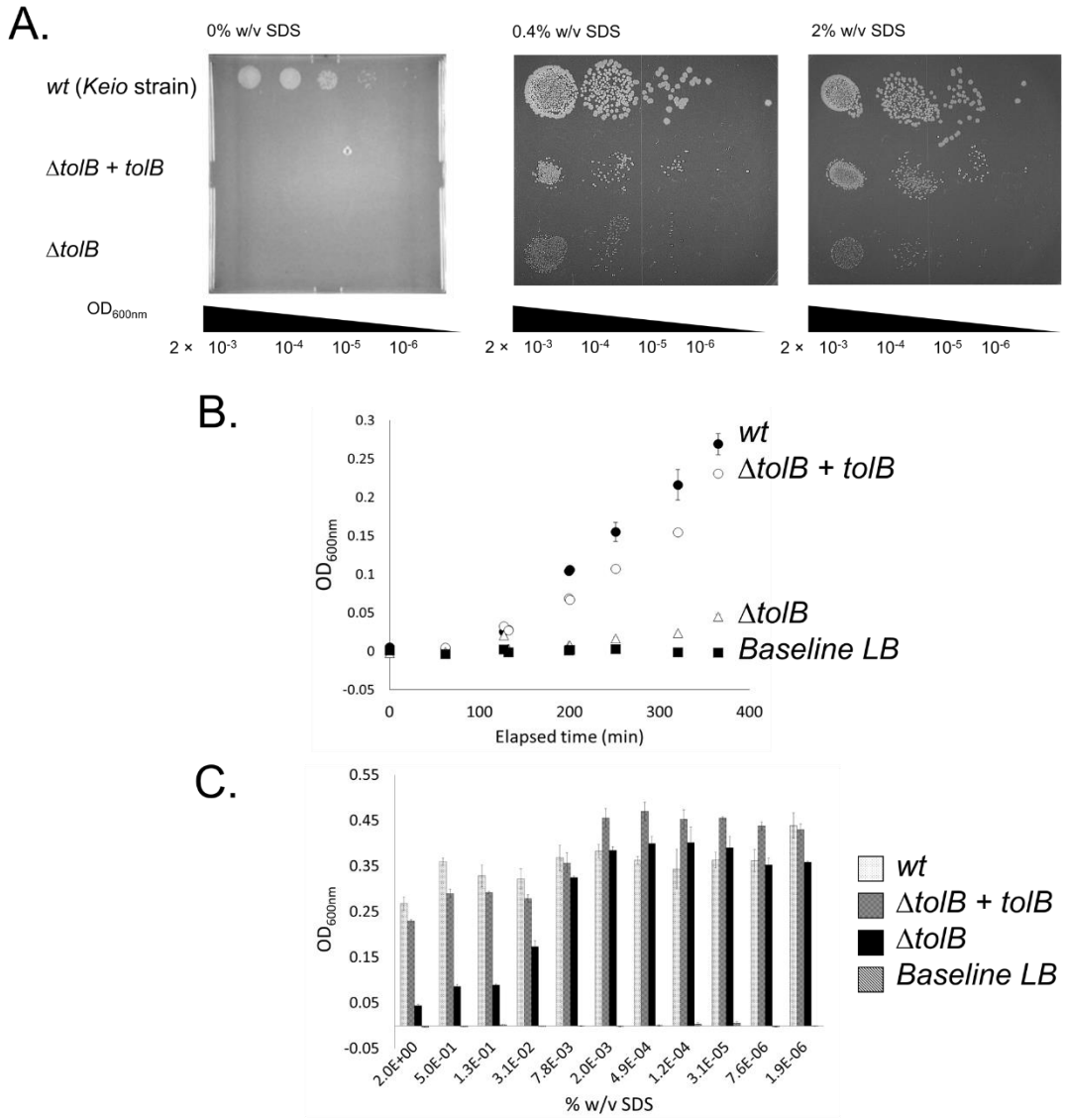


Figure 5.4 Δ***tolB* cells from the Keio collection exhibit a *tol* phenotype that is recovered using a plasmid encoding *tolB*.** The growth of Δ*tolB* cells relative to *wt* and plasmid-complemented cells was determined in liquid LB and solid LB agar at variable concentration of SDS. Each assay was conducted in triplicate. (A) The growth phenotype of TolB-deficient cells can be rescued with the pDAB17 plasmid encoding *tolB.* Varying cell concentrations were spotted on dry LB-Agar and grown overnight. (B) A Liquid growth assay confirms impaired growth in the dry LB-Agar assay. (C) The effective concentration of SDS is 0.002% w/v, as negative control Δ*tolB* cells grown 300 minutes show no significant impairment of growth for concentrations lower than this.

nutrient-rich agar media is shown in comparison to SDS-free media under 3 different growth conditions. Each growth condition was replicated 3 times, (replicates not shown). Although initial knockout of the TolB TolA interaction was attempted using a triple mutant, as seen in Figure 5.5 a single alanine mutation in the binding motif is sufficient for complete loss of Tol-Pal function.

Δ*tolB* cells were subsequently reassessed using an *in vivo* liquid culture assay for the determination of cell densities over time in 2% w/v SDS growth media. Similar to Figure 5.5, and

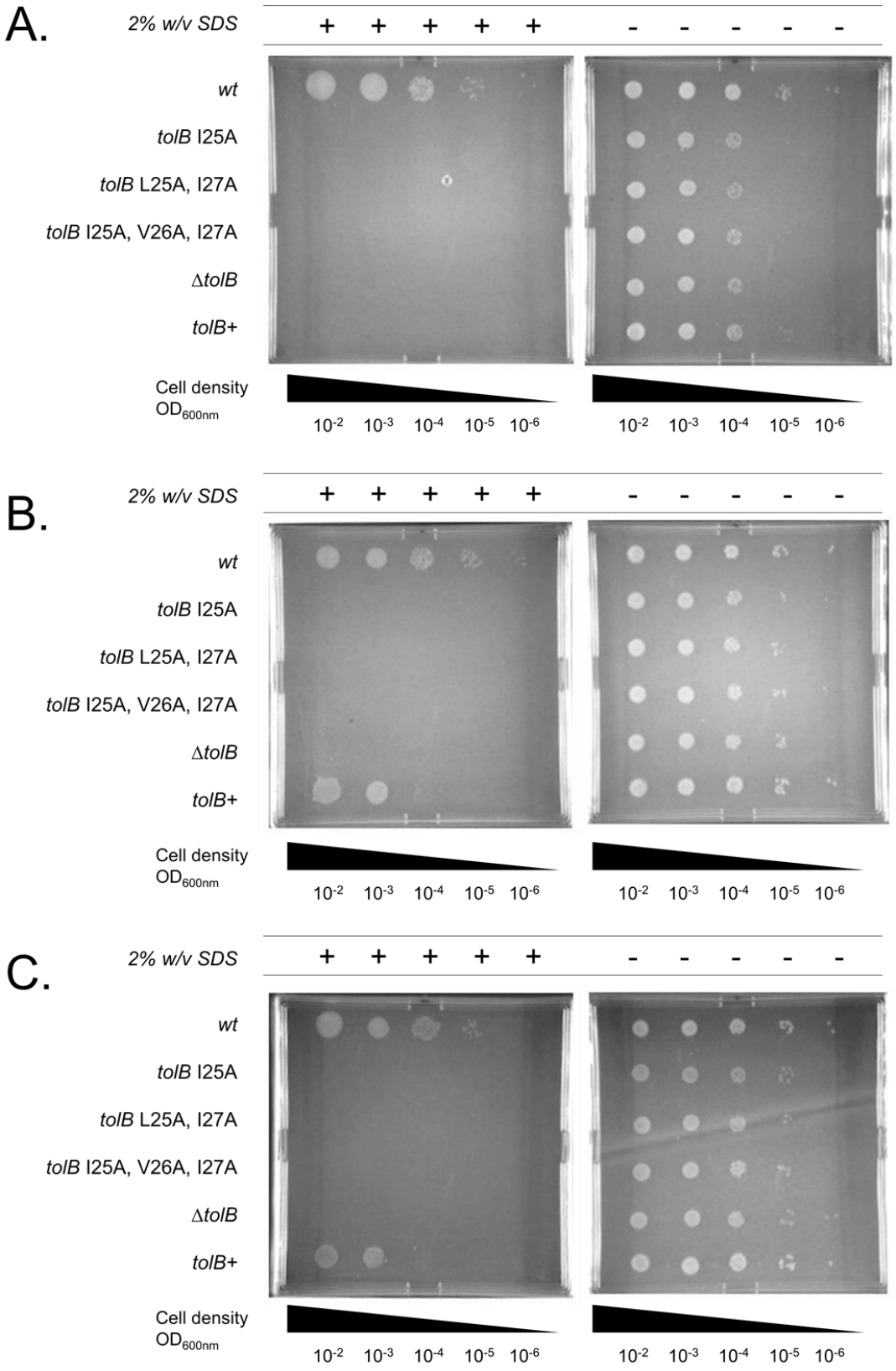


Figure 5.5 ***tolB* alanine mutations in the [IVL]X[IVL] motif result in abrogation of Tol-Pal activity.** (A), (B) and (C) show the effect of 2% SDS at varying incubation conditions. Leaky expression fromt the pDAB17 is temperature sensitive, as the 22°C overnight incubation did not recover from a *tol* phenotype. Alanine mutants targeting the entire TolA binding motif in TolB, as well as an alanine point mutant in the motif do not recover the *tol* phenotype, at any spotted cell density.

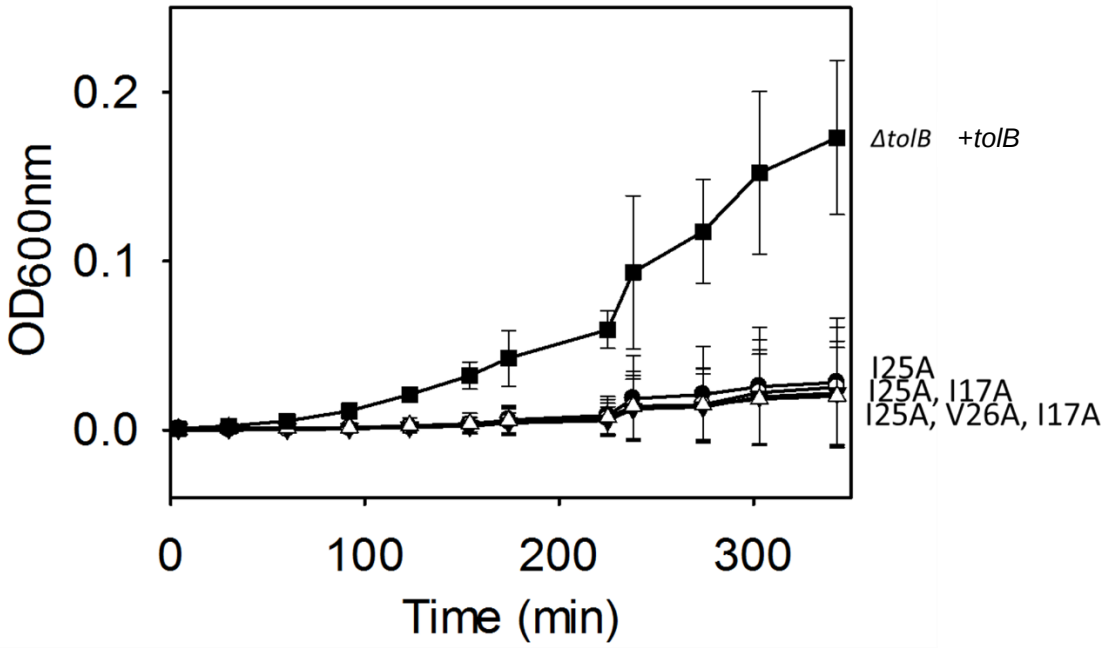


Figure 5.6 ***tolB* alanine mutations in the [IVL]X[IVL] motif result in abrogation of Tol-Pal activity via liquid growth assay.** In 2% SDS, grown over 350 minutes, at 37 °C, OD600 nm densities for *tolB* suppressor mutants relative to a complemented Δ*tolB* strain show no improvement or recovery of a *tol* phenotype.

seen in Figure 5.6, Binding-motif mutations in Δ*tolB* + *tolB* cultures subjected to SDS treatment displayed impaired growth. None of the constructs tested showed any partially impaired growth, suggesting that the presence of the [IVL]X[IVL] motif in TolB is all-or-nothing. Future experiments that involve cross-species mutations, for example I25L or I27V mutants designed to tune the affinity of the TolA binding site would clarify the evolutionary sensitivity of this motif.

5.3.4 Alanine mutants weaken TolB$_{22\text{-}34}$ binding to TolAIII

Although the mutations in TolB *in vivo* in section 5.3.3 suggested the interaction between TolA and TolB required a fully intact [IVL]X[IVL] motif, I sought to confirm this finding *in vitro* using synthetic TolB peptide with residue substitutions mimicking the observed *in vivo* alanine mutants. Using *P. aeruginosa* TolAIII and *P. aeruginosa* TolB$_{22\text{-}34}$ peptide, I performed competition experiments using fluorescence anisotropy. 1 μM TolB$_{22\text{-}34}$FITC was saturated with excess 2 mM TolAIII. This was then challenged with 0-3 mM *wt,* single and double TolB alanine mutant peptides. The overall average anisotropic signal from the TolB$_{22\text{-}34}$FITC reporter molecule changes in response to displacement from TolAIII from a competitor, indicating if a competitor peptide out-competes TolB$_{22\text{-}34}$FITC for TolAIII.

TolB alanine mutant sequences did not alter $TolB_{22\text{-}34}$FITC reporter signal at any tested concentration, suggesting these peptides faiiled to compete against the *wt* sequence TolAIII. Conversely, the *wt* peptide successfully outcompetes $TolB_{22\text{-}34}$FITC at a concentration approaching 3 mM. This suggests that in the absence of an intact [IVL]X[IVL] motif, TolB is unable to bind TolAIII, and supports the *in vivo* finding in *E. coli* that the TolB [IVL]X[IVL] motif is essential for Tol-Pal function.

5.3.5 Molecular modelling of the TolA-TolB interaction in *E. coli.*

The interaction between TolA and TolB in *P. aeruginosa* is characterized by the rearrangement of the C-terminus away from an occluding position in the TolB binding cleft. Although the [IVL]X[IVL] motif is strongly conserved in TolB, and a similar [hydrophobic]X[hydrophobic] motif is found in the complimentary parallel strand 3, the presence of a C-terminal structure is not a conserved feature. As discussed in section 5.3.1, the only major conserved region of TolAIII in the last 15 C-terminal residues is the aforementioned hydrophobic motif. The length of the C-terminal tail in TolAIII after strand 3 varies from one residue in *E. coli* to up to 18 residues in Rhizobium. Approximately half of species variants of TolAIII contain additional residues after strand 3. The remainder terminate at a similar position to *E. coli.*

As determined in Chapter 4, the presence or absence of a C-terminal extension in *P. aeruginosa* has has a greater than 2-fold effect on the binding affinity of TolAIII. Strikingly, the *E. coli* structure of TolAIII does possess a structure similar to the C-terminal 1.5-turn helix 4 in the *P. aeruginosa* crystal structure. In *E.coli* helix 2 is positioned over the binding site for g3p and must be partially displaced (Lubkowski *et al*, 1999a; Li *et al*, 2012). Based on this evidence, I used a structure-based comparative modelling strategy software package, MODELLER (Eswar *et al*, 2007), to calculate and predict a preliminary 3-D model of *E. coli* TolAIII in a TolB-bound state. I used the *P. aeruginosa* TolAIII structure as a template to which the *E. coli* sequence was aligned. A plausible model for how TolAIII and TolB molecules may interact in species that lack C-terminal extensions may therefore involve a C-terminal rearrangement mechanism similar to TolAIII, with the primary difference being a rearrangement in an internal structure absent from those molecules with a C-terminal extension.

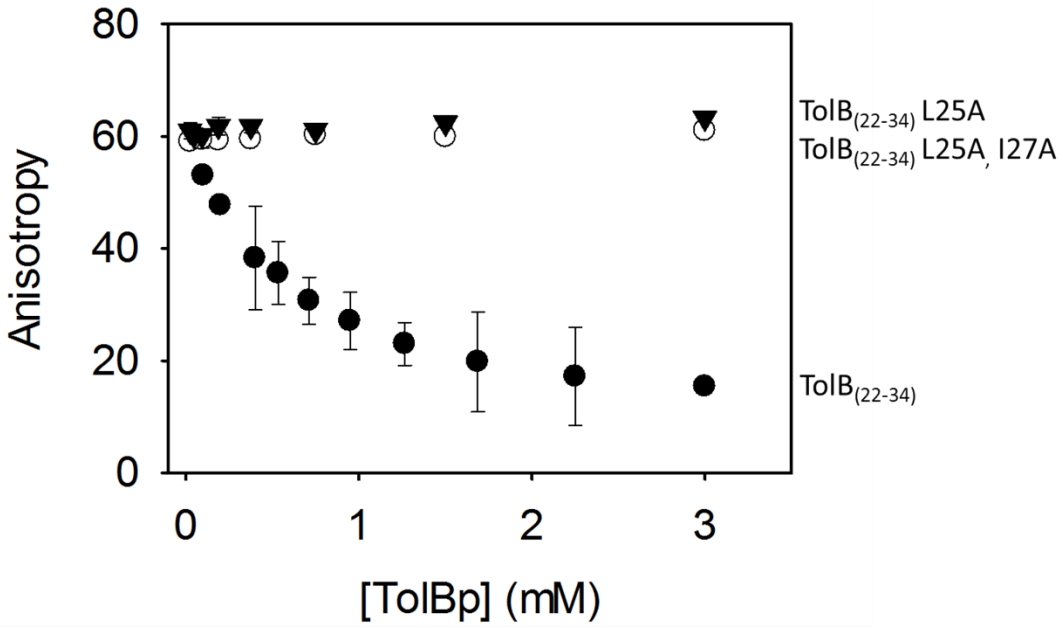


Figure 5.7 ***P. aeruginosa*** **TolB alanine mutant sequences do not out-compete a TolB$_{22-34}$FITC reporter for binding TolAIII.** TolB$_{22-34}$ peptide readily displaces the reporter at a concentration up to 3 mM, whereas at no concentration do sequences with alanine substitutions inside the [IVL]X[IVL] binding motif successfully displace the reporter. Anisotropy values determined at room temperature, n=3.

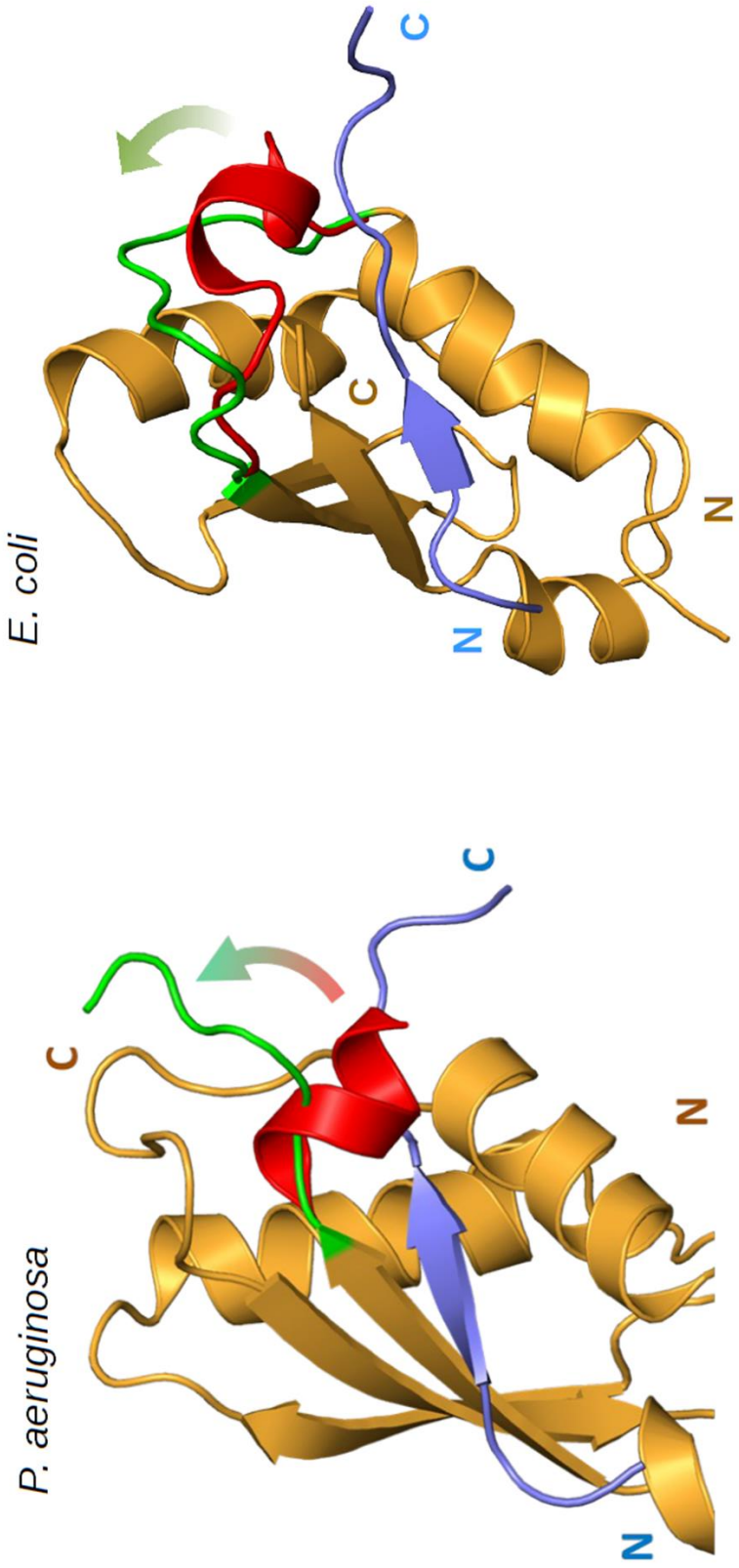


Figure 5.8 **A structural comparative model of the TolA-TolB interaction in *Escherichia coli* compared with the novel TolAIII-TolB$_{22\text{-}34}$ from *P. aeruginosa*.** By using the software MODELLER, an approximation was made of what the TolA-TolB interaction may look like in species that lack a C-terminal structure. Notably, TolAIII in *E. coli* possess a small second helix that potentially occludes the TolB-binding site in a similar manner to helix 4 in *P. aeruginosa.* In this cartoon depiction, the bound state of *E. coli* TolAIII is consistent with the remodelling of this second helix to facilitate TolB binding.

5.4 Discussion

Using a detailed structure-based bioinformatics study of the TolA-TolB interaction, I determined the identity of a conserved TolA interaction motif in the N-terminus of TolB. Using *in vivo* growth assays, this motif was removed from TolB and the effect on the cell observed. The conclusion of this study was that an intact [IVL]X[IVL] motif is vital for TolA-TolB interaction, as the loss of even a single residue in this motif resulted in the complete abrogation of binding *in vivo.*

Using bioinformatics, I further show that Gram-negative bacteria possess one of two types of TolAIII. Specifically those with and those without C-terminal extensions, an in silico comparative modelling approach suggests that even in the absence of a C-terminal extension, alternate internal structures, such as helix 2 in *E. coli* may still undergo order-disorder transitions on binding TolB. This therefore represents a new avenue of study, specifically into the conservation of the order-disorder conformational change mechanism in structures occluding the TolB binding site in TolAIII.

6. Conclusion

6.1 General discussion of major findings

6.1.1 Structural investigations of TolA preceding this study

Prior to the studies performed in this thesis, structural investigations into the mode of action of TolA protein complexes focused largely on the interactions of TolAIII with bacteriocins and bacteriophage. The first structure of TolAIII published was determined from *E. coli* in complex with the gene 3 protein (g3p) of filamentous phage (Jacek Lubkowski, Hennecke, Pluckthun, & Wlodawer, 1999). The endogenous binding partner for g3p N1 domain in filamentous phage before it recruits TolAIII is the phage N2 domain, for which a structural interaction had been reported (Lubkowski *et al*, 1998). TolAIII and phage N2 domains do not share structural motifs or homology, nor does TolAIII share homology with any other protein folds outside of TonB and the TolAIII fold proved to be a novel discovery (Lubkowski *et al*, 1999a).

In 2002 a structure of TolAIII protein in the unbound state was determined from *P. aeruginosa.* This permitted a clear comparison with the structure of *E. coli* TolAIII bound to g3p. Firstly, the structures revealed a remarkable structural homology (r.m.s.d. of 1.5 Å for 69 equivalently matched Cα atoms) that was unexpected given the relatively low sequence identity between species. Also evident was a lack of major structural changes between the bound and unbound states of TolAIII (Witty *et al*, 2002). In 2005, a solution structure of *E. coli* TolAIII was determined, which supported these observations, concluding that the free and bound form of TolAIII protein adopted essentially the same fold (Deprez *et al*, 2005).

Conversely, studies on the interaction between colicin A and TolAIII at first suggested that this interaction induced a strongly destabilizing effect on TolAIII (Deprez *et al*, 2002). This hypothesis was later shown to be inaccurate. Although colicin A binds to the opposite face of the TolAIII β-sheet via a classic β-strand complementation mechanism, this interaction did not appear to cause any large conformational change in the overall fold of TolAIII (Li *et al*, 2012). As shown in Chapter 1 of this thesis in Figure 1.6, the emerging structural hypothesis for TolAIII was that it exhibited a largely stable fold with compatible binding interfaces for both bacteriophage and bacteriocin protein binding partners. In all cases, TolAIII was predicted to interact via a β-strand complementation mechanism on either strand one or strand three of the TolAIII β-sheet. Of unusual note however, was the observation that the binding of g3p N1 to TolAIII induced large

chemical shift changes in TolAIII that were distinctly away from the binding site (Deprez *et al*, 2005). Recent studies on the interaction between *V. cholera* TolAIII and CTX bacteriophage, as well as between *E. coli* TolAIII and fd phage indicated that they share binding interfaces on TolAIII with previously described bacteriocin and phage protein interactions (Ford *et al*, 2012; Lorenz *et al*, 2011). Notwithstanding these observations, determination of the molecular features underlying a protein-protein complex between TolAIII and an endogenous binding partner had never been accomplished.

6.1.2 The TolA TolB interaction *in vitro* in *P. aeruginosa* is weak

In order to address this question, in this study we sought to provide a comprehensive understanding of the TolA-TolB structure and determine its mode of interaction. The binding of a synthetic peptide designed to mimic the N-terminus of TolB and its interaction with TolAIII, facilitated the opportunity to explore the structural biophysical details underlying dynamic TolA-TolB interactions. These interactions were determined to be remarkably weak for an important protein-protein interaction within the periplasm, especially since the loss of the TolA-TolB interaction *in vivo* had already been demonstrated to generate a classic *tol* phenotype. Initial detailed structural investigations of TolA-TolB interactions were strongly hampered by the inherent weakness underlying the interaction.

Determination of the binding affinity of the TolA-TolB complex was itself difficult, due to protein solubility limits and the concentrations required to make a detailed thermodynamic investigation using ITC techniques unfeasible. Thermophoretic binding assays, surface plasmon resonance and fluorescence shift assays were all performed yet demonstrated inadequate signal to noise ratios, given available solubility limits for TolB peptide, TolB and TolAIII. Although crystallographic techniques were attempted, these were similarly unsuccessful, again most likely due to the extremely weak interaction between TolA and TolB.

To determine biophysical measurements for the affinity and kinetics of TolA-TolB complex, I used fluorescence anisotropic techniques, involving a structurally-based positioning of the fluorescent reporter, to determine high precision binding data. This approach, in combination with NMR-based titration data, resolved a dissociation constant (K_D) of 250 μM for the weak interaction between TolA and TolB. In *E. coli* the TolA-TolB interaction is 40 μM, which mechanistically can be explained by the lack of a C-terminal sequence in TolAIII that occludes

the TolB binding site. In *P. aeruginosa* the affinity increases greater than 2-fold in the absence of a C-terminal sequence.

6.1.3 The structure of the TolA TolB complex was determined by solution NMR

I addressed the incongruities between the *E. coli* and *P. aeruginosa* TolAIII structures by employing an NMR-based investigation. Heteronuclear NOE data, in combination with residual dipolar coupling experiments and chemical shift data, strongly suggest that the ordering of the 20 N-terminal residues in the *P. aeruginosa* crystal structure are due to crystallographic artifacts. Furthermore, I demonstrated that the C-terminal helix 4 of TolAIII from *P. aeruginosa* is only partially formed compared with the TolAIII crystal structure in *P. aeruginosa*. NOE data, TALOS secondary structure prediction and residual dipolar coupling data are all consistent with the hypothesis that the C-terminus occludes the hydrophobic core of TolAIII in the unbound state, similar to observations from the TolAIII crystal structure. However, there is evidence that the last two residues of this helix are in fact disordered. The overall fold of *P. aeruginosa* was otherwise confirmed by NMR analysis, and the strong conservation of this fold across Gram-negative species with a *tol-pal* operon was determined through a comprehensive bioinformatics assessment of fully assembled genomes.

By analysis of the NMR studies, a peculiar juxtaposition was evident between the weak measured dissociation constant for TolA-TolB, and the spectral characteristics of the interaction. This involved large chemical shift changes upon binding, paired with every observable amide in slow exchange. Chemical shift assignment revealed that the binding of $TolB_{22-34}$ induced large amide chemical shift changes across the entirety of the globular TolAIII domain, and changes could not be localized to a single binding site. Although amide chemical shift changes are sensitive to small changes in the local chemical environment, and are not necessarily indicative of conformational change, by determining the full assignment of the bound and unbound states of TolAIII, a structural basis for this effect is now clearly established.

My NMR-based structural investigation of the TolAIII-$TolB_{22-34}$ complex determined for the first time the mode of action of TolA-TolB complex interactions. Whereas previous studies with bacteriophage and bacteriocin proteins involved antiparallel β-complementation mechanisms, the endogenous interaction was found to involve a parallel β-complementation mechanism. The hydrophobic core of TolAIII was found to be strongly perturbed by the binding of TolB.

Specifically, the displacement of Phe266 in TolAIII by Leu27 in TolB results in a novel arrangement of the phenylalanine aromatic ring. Concomitant aromatic ring-current effects strongly perturb the chemical shifts of nearby residues, and the displacement of this residue from the hydrophobic core. In addition, the deep penetration of TolB sidechains from the conserved [I/V/l]X[I/V/L] motif into the hydrophobic core would result in minor perturbations throughout TolAIII that are consistent with the observed chemical shift changes. Specifically for *P. aeruginosa,* a structural rearrangement and displacement of the occluding C-terminal residues of TolAIII from a binding cleft is necessary for the interaction of TolB with TolAIII. A full analysis of carbon chemical shift changes, which are more sensitive to structural perturbation, supports a hypothesis which predicts a major rearrangement in the C-terminus of TolAIII as a consequence of TolB binding.

6.1.4 The structure of the TolA TolB interaction provides insight into the underlying binding mechanism

These results also provided a testable hypothesis for the unusually slow exchange phenomena observed by NMR spectroscopic methods and that either the association rate or dissociation rate constant for the TolA-TolB interaction might be negatively perturbed by the presence of the TolAIII C-terminal sequence. Using stopped-flow kinetic methods, I demonstrated that the deletion of this C-terminal structure leads to a greater than 2-fold increase in the overall binding affinity of TolB for TolA. Furthermore, I determined that although the dissociation rate constant for this complex increases by 1.5-fold, the association rate constant experiences a concomitant 3-fold increase. This further provides an explanation for the observed disparity in binding affinities between *P. aeruginosa* and *E. coli* TolA-TolB interactions. The *E.coli* TolAIII molecule does not possess a C-terminal structure that fully occludes the TolB binding site. Removal of this structure in *P. aeruginosa* therefore results in a binding affinity within the range previously determined for *E. coli.*

6.1.5 A structure-based bioinformatics study revealed conservation of the TolA-TolB interaction across Gram-negative bacteria

A strong motivating factor for determining the structure of the TolA-TolB interaction was to provide insight into how the TolA-TolB interaction may be widely conserved in the face of evolutionary pressure from parasitisation by invading bacteriocins and bacteriophages. A structure-based alignment of the TolA-TolB binding interface revealed two paired conserved hydrophobic interaction motifs. The conservation of this binding interface, in combination with the

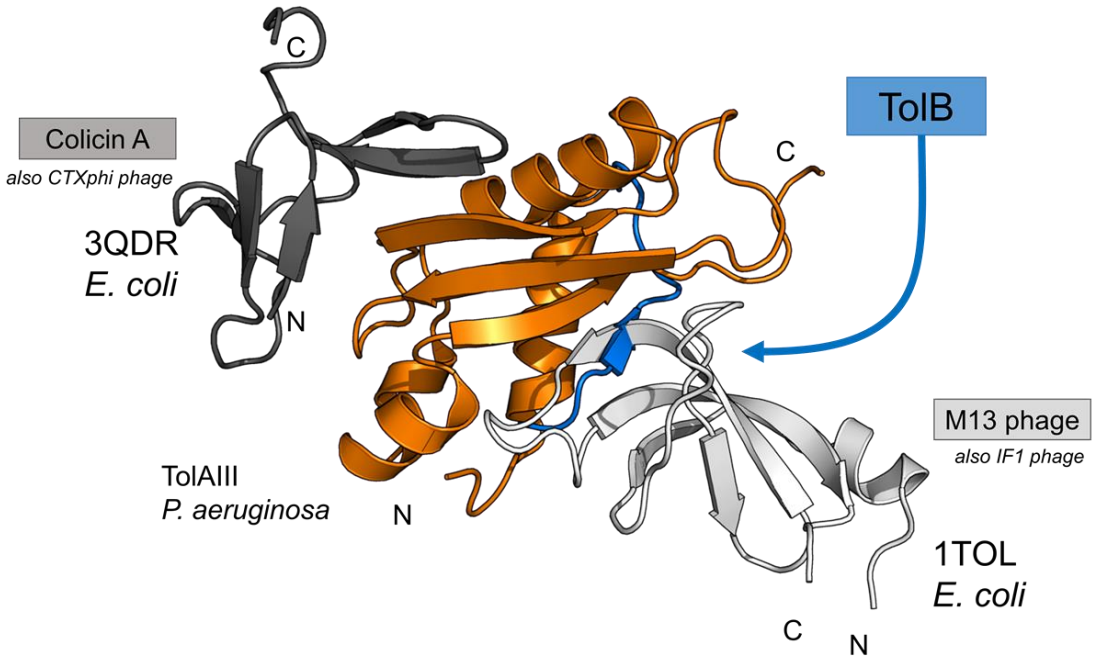


Figure 6.1 **The new state of knowledge on the TolA and TolB interaction via an intrinsically disordered region.** In this cartoon scheme the new paradigm TolAIII as a protein-protein interaction hub is depicted. This figure combines the cartoon depictions for the Colicin A interaction with *E. coli* TolA, the cartoon depiction of the M13 phage interaction with TolA, both previously used in Figure 1.6, with the novel structure of *P.aeruginosa* TolAIII in complex with $TolB_{22\text{-}34}$. In this superposition of figures, all known binding interfaces on TolAIII are shown. TolB is indicated in blue, overlapping with the M13 phage site.

strong secondary structure conservation of TolAIII predicted by JPred3 analysis, suggests that the TolA-TolB interaction is ubiquitous in Gram-negative bacteria with a conserved *tol-pal* operon.

6.1.6 The effectiveness of the TolA-TolB binding interface is abolished *in vivo* and *in vitro* by a single alanine substitution within the conserved [I/V/L]X[I/V/L] motif

In order to understand the importance of these findings *in vivo*, I studied the role of the conserved [I/V/l]X[I/V/L] motif in *E. coli* TolAIII. I determined that not only does a triple mutant knockout of the [I/V/l]X[I/V/L] motif result in complete abrogation of TolA binding, but that a single mutant is sufficient. Disruption of this motif produces a *tol* phenotype and indicates a breakdown in Tol-Pal functionality. To understand the effect of alanine substitutions in the [I/V/l]X[I/V/L] motif, I completed an *in vitro* competition binding assay which revealed that [I/V/l]X[I/V/L] mutants are binding incompetent, therefore these molecules cannot displace *wt* TolB under any of the tested conditions. The remarkable sequence divergence between TolA and TolB molecules across all bacteria, as a consequence of strong evolutionary pressures is understandable. The TolA-TolB interaction is very likely strongly conserved because both core hydrophobic residues in TolA and the essential TolB [I/V/l]X[I/V/L] motif are conserved.

6.2 Future outlook

Through the studies performed in this thesis, I have defined the structural molecular details underlying the TolA-TolB interaction in Gram-negative bacteria. I have described the biophysical characteristics that define this interaction, as well as provided a structural rationale that underlies the binding mechanism. Many avenues for future investigation remain. The most critical question that remains outstanding is to determine the precise mechanism by which TolA may transmit the proton motive force to TolB. In Chapter 3 I was able to demonstrate that the N-terminus of TolA has intrinsic disorder, and represents a barrier to signalling mechanisms that rely upon cumulative or sequential conformational rearrangements. Furthermore, no large conformational changes to the fold of TolA in the presence of TolB were observed. The structure of TolB has further been reported to not show any conformational rearrangement, outside of its N-terminus, upon binding to TolAIII (Bonsor *et al*, 2009). Additionally, the presence or absence of a TolAIII N-terminal 'cap', in this case approximated by a soluble Im9 molecule, both with and without the disordered N-terminal residues of TolAIII appears to have no effect on the binding characteristics of the TolA-TolB interaction. All this evidence strongly points to the idea that a

conformational cascade is unlikely to be the method by which the proton motive force is transferred to TolB.

Very recent studies in analogous Ton systems have suggested that proton-motive-force (PMF) linked complexes may act to spatially displace TonB in a piston-like or rotational motion (Celia et al., 2016). Although major work remains to be done to confirm this, in Chapter 4, I demonstrated how the TolA-TolB complex binding mechanism is analogous to Ton-dependent receptor interaction mechanisms. One method of testing the hypothesis that the functional mode of action by which Tol and Ton proteins transmit the PMF to their receptors would be to carry out further mutational studies to those performed in Chapter 5. I attempted to fuse TolA and TolB together via their C and N-termini respectively, to observe the *in vivo* effect of this mutant in a Δ*tolAtolB* strain, although at present the cloning of this mutant remained technically unsuccessful. This line of investigation may provide an avenue to further test the spatial nature and importance of the TolA-TolB interaction.

More challenging technically, but an equally viable research avenue, would be to test the effect of mutations on the recruitment of TolB and TolA to the septation apparatus during cell division. TolB has previously been reported to be recruited to the septal ring (Gerding *et al*, 2007a). This was attempted in the present study, and TolB-GFP was expressed and observed by fluorescence microscopy, however no clear recruitment to the septation site could be observed. If this system were optimised and recruitment was in future replicated, mutations to TolAIII could be assessed for their effect on TolB recruitment. A TolAIII mutant with a null alanine-substituted TolB binding site would be expected to produce a *tol* phenotype and septal recruitment of TolB might be inhibited. This approach would allow for tests involving the elongation or truncation of TolAIII and observation of the effect on TolB recruitment.

During the course of this work, I suggest that the evolutionary resilience of the Tol-Pal system to external pressures is contingent on preservation of the TolA-TolB binding interface. 'Like to like' mutants of this binding interaction that mimic alternate species would reveal the tunable character of the [I/V/L]X[I/V/L] interface. Specifically, if full complementation of TolA can be achieved by interface with TolB molecules from different species presenting the correct binding motif. As this interaction is critical *in vivo,* a clear understanding of the tunability of the TolA-TolB

interface would provide insight into design of peptide-based therapeutics that would interrupt this interaction.

Another feature of this study demonstrated unusual protein-protein binding characteristics, wherein a kinetically slow dissociation of a ligand resulted in NMR spectra with amides presenting slow exchange dynamics for an atypically weak complex. The poor solubility exhibited by full length TolB and the weak affinity within the components of the TolAIII-TolB complex, meant that a detailed description of the full length TolB-TolAIII complex was outside the scope of this study. Further analysis of the full length complex remains an interesting suggestion for future experiments. ^{2}H $^{13}C/^{15}N$ enrichment of TolA and TolB could be used with TROSY experiments, or methyl TROSY techniques. The key *in vitro* experiment to explore would be a competition experiment involving heterolabelled TolAIII, full length unlabelled TolB and full length unlabelled Pal. This would allow a direct test of the hypothesis that Pal acts as an on/off switch for the Tol-Pal complex, as proposed by Bonsor et al. (Bonsor *et al*, 2009). Specifically, with an increase in concentration of Pal, a shift in the bound to unbound state of TolAIII would indicate direct competition. Alternatively, if TolB may simultaneously bind TolAIII and Pal, you would expect the appearance of a third characteristic set of chemical shifts indicating a novel third species during the titration. Due to the size of this complex (>50 kDa), this kind of experiment may only be accessible to Methyl TROSY techniques.

When taken together, the studies presented in this thesis on the Tol-Pal complex from *P. aeruginosa* have significantly advanced our understanding of key protein-protein interactions within the periplasm of Gram-negative bacteria. They provide a sound basis for further investigation “a priori”, as human kind must address the molecular mechanisms underlying bacteria function in order to elucidate new strategies to combat bacterial infection.

7. Appendix

7.1 List of primers used for the amplification of DNA.

Name	Primer Sequence
TBgAVIfp	CGCGGAGAATTCGAAGTCCGCGCAGTGATCGACAGCGGTGTAGATTCC
TBgAVIrp	GGAATCTACACCGCTGTCGATCACTGCGCGGACTTCGAATTCTCCGCG
TBgAAAfp	GGAGAATTCGAAGTCCGCGCAGCAGCAGACAGCGGTGTAGATTCCGG
TBgAAArp	CCGGAATCTACACCGCTGTCTGCTGCTGCGCGGACTTCGAATTCTCC
ABDMf	GGAGATCACCAGCGGGTCGGAGACTCAAATCCTCCGGTTTAAAGATGATG
ABDMr	GGAGGATTTGAGTCTCCGACCCGCTGGTGATCTCC
TABCf	CGCGCGAATTCTTGAAGCAACAGTTCGAACGCT
TABCr	CGCGCGAAGCTTTCAGTTCAGGTAAGGGGACCAGG
TAPSfp	GCCGCATCATCTTTAAATAAGAGGATTTGAGTCTGTAAG
TAPSrp	CTTACAGACTCAAATCCTCTTATTTAAAGATGATGCGGC
TAGSfp	CGCATCATCTTTAAACCGTAAGATTTGAGTCTGTAAGC
TAGSrp	GCTTACAGACTCAAATCTTACGGTTTAAAGATGATGCG
TAgnfp	CATCATCATCCATGGTGTCAAAGGCAACCGAACAAAACG
TAgnrp	ATGATGATGTCTAGATTACGGTTTGAAGTCCAATGGCGCG
TBAVI2fp	CTGCATGCTGAAGTCCGCGCAGTGATCGACAGCGGTGTAGATTCC
TBAVI2rp	GGAATCTACACCGCTGTCGATCACTGCGCGGACTTCAGCATGCAG
TBAVAfp	GCATGCTGAAGTCCGCGCAGTGGCAGACAGCGGTGTAG
TBAVArp	CTACACCGCTGTCTGCCACTGCGCGGACTTCAGCATGC
TBIVAfp	GCATGCTGAAGTCCGCATTGTGGCAGACAGCGGTGTAG
TBIVArp	CTACACCGCTGTCTGCCACAATGCGGACTTCAGCATGC
TBAVIfp	GCATGCTGAAGTCCGCGCAGTGATCGACAGCGGTGTAG
TBAVIrp	CTACACCGCTGTCGATCACTGCGCGGACTTCAGCATGC
TBIVIfp	GCATGCTGAAGTCCGCATTGTGATCGACAGCGGTGTAG
TBIVIrp	CTACACCGCTGTCGATCACAATGCGGACTTCAGCATGC
TBIVI22AAAfp	ATGCTGAAGTCCGCGCAGCAGCAGACAGCGGTGTA
TBIVI22AAArp	TACACCGCTGTCTGCTGCTGCGCGGACTTCAGCAT
TABfp	GCAGCTCCATGGCAAAGGCAACCGAACAAAACG
TArp	GATTCATCTAGATTACGGTTTGAAGTCCAATGGCG
TABrp	GATTCATCTAGATTATCACAGATACGGCGACCAGG
TBRKfp	AAGCTTTGGATGGCCGACCCGCTGGTGATCTCCAGCGGTAACGACCGTGCCTAA
TBRKrp	TGGATGGCCGACCCGCTGGTGATCTCCAGCGGTAACGACCGTGCCTAAGGATCC

7.2 List of peptides used in this work.

Peptide	Notation	Notes	Purity
EVRIVIDSGVDS	$TolB_{22-33}$	Synthetic	>95%
EVRIVIDSGVDSG-NH_2	$TolB_{22-34}$ [AMID]	Synthetic	>95%
ADPLVISSGNDRA	$TolB_{22-34}$	Synthetic	>95%
ADPLVISSGNDRAK-FITC	$TolB_{22-34}$ [FITC]	Synthetic	>98%
ADPLVISSGNDRA-NH_2	$TolB_{22-34}$ [AMID]	Synthetic	>95%
ADPLVISSGNDRA-NH_2	$TolB_{22-34}$ [RE]	Recombinant expression (TrpLE)	>95%
ADPLVISSGNDRA-NH_2	$^{13}C^{15}N$ $TolB_{22-34}$ [RE]	Recombinant expression (TrpLE) ^{13}C ^{15}N	>80%
ADPLVISSGNDREDYE	$TolB_{22-33}$EDYE	Synthetic	>95%
ADPLVISSGNDRAEDYE	$TolB_{22-34}$EDYE	Synthetic	>95%
ADPLVISSGNDRKYRK	$TolB_{22-33}$KYRK	Synthetic	>95%
ADPLVISSGNDRAKYRK	$TolB_{22-34}$KYRK	Synthetic	>95%
ADPLVISSGNDRYNGR	$TolB_{22-33}$YNGR	Synthetic	>95%
ADPAVISSGNDRA-NH_2	$TolB_{22-34}$L25A	Synthetic	>95%
ADPAVASSGNDRA-NH_2	$TolB_{22-34}$L25A,I27A	Synthetic	>95%
ADPAAASSGNDRA-NH_2	$TolB_{22-34}$L25A,V26A,I27A	Synthetic	>95%

7.3 List of proteins recombinantly expressed or used for *in vivo* studies in this work

pDAB17 – *E. coli* secTolB residues 1-430;
MKQALRVAFGFLILWASVLHAEVRIVIDSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADL
RNSGKFNPLDRARLPQQPGSAQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVDTGGA
PGTVLAQNSYKVNKQWLRYAGHTASDEVFEKLTGIKGAFRTRIAYVVQTNGGQFPYELRVSD
YDGYNQFVVHRSPQPLMSPAWSPDGSKLAYVTFESGRSALVIQTLANGAVRQVASFPRHNGA
PAFSPDGSKLAFALSKTGSLNLYVMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAG
RPQVYKVNINGGAPQRITWEGSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGVQVLS
STFLDETPSLAPNGTMVIYSSSQGMGSVLNLVSTDGRFKARLPATDGQVKFPAWSPYL

pDAB18 – *E. coli* TolB residues 22-430;
MEVRIVIDSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADLRNSGKFNPLDRARLPQQPGS
AQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVDTGGAPGTVLAQNSYKVNKQWLRYA
GHTASDEVFEKLTGIKGAFRTRIAYVVQTNGGQFPYELRVSDYDGYNQFVVHRSPQPLMSPA
WSPDGSKLAYVTFESGRSALVIQTLANGAVRQVASFPRHNGAPAFSPDGSKLAFALSKTGSL
NLYVMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAGRPQVYKVNINGGAPQRITWE
GSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGVQVLSSTFLDETPSLAPNGTMVIYS
SSQGMGSVLNLVSTDGRFKARLPATDGQVKFPAWSPYL

pREN28 – *P.aeruginosa* Im9-FXa-TolAIII residues 224-347;
MKHHHHHHHNMELKHSISDYTEAEFLQLVTTICNADTSSEEELVKLVTHFEEMTEHPSGSDL
IYYPKEGDDDSPSGIVNTVKQWRAANGKSGFKQIEGRHMRALAELLSDTTERQQALADEV
GSEVTGSLDDLIVNLVSQQWRRPPSARNGMSVEVLIEMLPDGTITNASVSRSSGDKPFDS
SAVAAVRNVGRIPEMQQLPRATFDSLYRQRRIIFKPEDLSL

pPH01 – *P.aeruginosa* TrpLE-CNBr site-TolB residues 22-34;
MHHHHHHHHHKAIFVLKGSLDRDLDSRIELELRTDHKELSEHLLLVDLARNDLARIATPGSR
YVADLTKVDRYSYVLHLVSRVVGELRHDLDALHAYRAALNLGTLSGAPKVRAKLWMADPLVI
SSGNDRA

pPH02 – *P.aeruginosa* Im9-FXa-TolAIII residues 224-346;
MKHHHHHHHNMELKHSISDYTEAEFLQLVTTICNADTSSEEELVKLVTHFEEMTEHPSGSDL
IYYPKEGDDDSPSGIVNTVKQWRAANGKSGFKQIEGRHMRALAELLSDTTERQQALADEVGS
EVTGSLDDLIVNLVSQQWRRPPSARNGMSVEVLIEMLPDGTITNASVSRSSGDKPFDSSAVA
AVRNVGRIPEMQQLPRATFDSLYRQRRIIFKPEDLS

pPH03 – *P.aeruginosa* Im9-FXa-TolAIII residues 224-341;
MKHHHHHHHNMELKHSISDYTEAEFLQLVTTICNADTSSEEELVKLVTHFEEMTEHPSGSDL
IYYPKEGDDDSPSGIVNTVKQWRAANGKSGFKQIEGRHMRALAELLSDTTERQQALADEVGS
EVTGSLDDLIVNLVSQQWRRPPSARNGMSVEVLIEMLPDGTITNASVSRSSGDKPFDSSAVA
AVRNVGRIPEMQQLPRATFDSLYRQRRIIFK

pPH04 – *P.aeruginosa* Im9-FXa-TolAIII residues 224-342;
MKHHHHHHHNMELKHSISDYTEAEFLQLVTTICNADTSSEEELVKLVTHFEEMTEHPSGSDL
IYYPKEGDDDSPSGIVNTVKQWRAANGKSGFKQIEGRHMRALAELLSDTTERQQALADEVGS
EVTGSLDDLIVNLVSQQWRRPPSARNGMSVEVLIEMLPDGTITNASVSRSSGDKPFDSSAVA
AVRNVGRIPEMQQLPRATFDSLYRQRRIIFKP

pPH05 – *E. coli* secTolB residues 1-430 AVI mutant;
MKQALRVAFGFLILWASVLHAEVRAVIDSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADL
RNSGKFNPLDRARLPQQPGSAQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVDTGGA
PGTVLAQNSYKVNKQWLRYAGHTASDEVFEKLTGIKGAFRTRIAYVVQTNGGQFPYELRVSD
YDGYNQFVVHRSPQPLMSPAWSPDGSKLAYVTFESGRSALVIQTLANGAVRQVASFPRHNGA
PAFSPDGSKLAFALSKTGSLNLYVMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAG

RPQVYKVNINGGAPQRITWEGSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGVQVLS
STFLDETPSLAPNGTMVIYSSSQGMGSVLNLVSTDGRFKARLPATDGQVKFPAWSPYL

pPH06 – *E. coli* secTolB residues 1-430 AVA mutant;
MKQALRVAFGFLILWASVLHAEVRAVADSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADL
RNSGKFNPLDRARLPQQPGSAQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVDTGGA
PGTVLAQNSYKVNKQWLRYAGHTASDEVFEKLTGIKGAFRTRIAYVVQTNGGQFPYELRVSD
YDGYNQFVVHRSPQPLMSPAWSPDGSKLAYVTFESGRSALVIQTLANGAVRQVASFPRHNGA
PAFSPDGSKLAFALSKTGSLNLYVMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAG
RPQVYKVNINGGAPQRITWEGSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGVQVLS
STFLDETPSLAPNGTMVIYSSSQGMGSVLNLVSTDGRFKARLPATDGQVKFPAWSPYL

pPH07 – *E. coli* secTolB residues 1-430 IVA mutant;
MKQALRVAFGFLILWASVLHAEVRIVADSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADL
RNSGKFNPLDRARLPQQPGSAQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVDTGGA
PGTVLAQNSYKVNKQWLRYAGHTASDEVFEKLTGIKGAFRTRIAYVVQTNGGQFPYELRVSD
YDGYNQFVVHRSPQPLMSPAWSPDGSKLAYVTFESGRSALVIQTLANGAVRQVASFPRHNGA
PAFSPDGSKLAFALSKTGSLNLYVMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAG
RPQVYKVNINGGAPQRITWEGSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGVQVLS
STFLDETPSLAPNGTMVIYSSSQGMGSVLNLVSTDGRFKARLPATDGQVKFPAWSPYL

pPH08 – *E. coli* secTolB residues 1-430 AAA mutant;
MKQALRVAFGFLILWASVLHAEVRAAADSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADL
RNSGKFNPLDRARLPQQPGSAQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVDTGGA
PGTVLAQNSYKVNKQWLRYAGHTASDEVFEKLTGIKGAFRTRIAYVVQTNGGQFPYELRVSD
YDGYNQFVVHRSPQPLMSPAWSPDGSKLAYVTFESGRSALVIQTLANGAVRQVASFPRHNGA
PAFSPDGSKLAFALSKTGSLNLYVMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAG
RPQVYKVNINGGAPQRITWEGSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGVQVLS
STFLDETPSLAPNGTMVIYSSSQGMGSVLNLVSTDGRFKARLPATDGQVKFPAWSPYL

7.4 Acquisition parameters for triple resonance experiments used for protein backbone assignment

Experiment	Spectrometer	Sweep Width			Number of Complex Points			Spectrometer Frequency (MHz)	Centre Reference Chemical Shift (ppm)			Number of Scans
		^{1}H	^{15}N	^{13}C [^{1}H]	^{1}H	^{15}N	^{13}C [^{1}H]		^{1}H	^{15}N	^{13}C [^{1}H]	
HNCO	Bruker 500	12.46	27.50	15.00	512	32	50	500.0123	4.8	117.75	172.50	8
HNCA	Bruker 500	12.46	27.50	30.00	512	32	50	500.0123	4.8	117.75	52	16
CBCANH	Bruker 500	12.46	27.50	59.35	512	32	49	500.0123	4.8	117.75	41.5	20
CBCA(CO)NH	Bruker 500	12.46	27.50	59.35	512	32	49	500.0123	4.8	117.75	41.5	8
HcC(CO)NH	Bruker 500	12.46	27.50	10.64	512	32	89	500.0123	4.8	117.75	4.8	16
HCc(CO)NH	Bruker 500	12.46	27.50	66.0	512	32	64	500.0123	4.8	117.75	40	16
HBHA(CO)NH	Bruker 500	12.46	27.50	9.52	512	28	49	500.0123	4.8	117.75	4.8	8
^{1}H,^{15}N-HSQC	Bruker 500	12.46	27.50	-	1024	128	-	500.0123	4.8	117.75	4.8	4
^{15}N NOESY edited, filtered	Bruker 500	12.46	27.50	-	512	128	1	500.0123	4.8	117.75	4.8	160
^{13}C NOESY edited, filtered	Bruker 500	12.46	-	80	512	1	128	500.0123	4.8	4.8	40	160
^{1}H,^{13}C-HSQC	Bruker 500	12.46	-	70.00	1024	-	128	500.0123	4.8	4.8	40	8
ZZ-exchange spectroscopy	Bruker 500	12.46	27.5	-	1024	128	-	500.0123	4.8	117.75	-	16

7.5 Acquisition parameters for ^{15}N-edited NMR experiments.

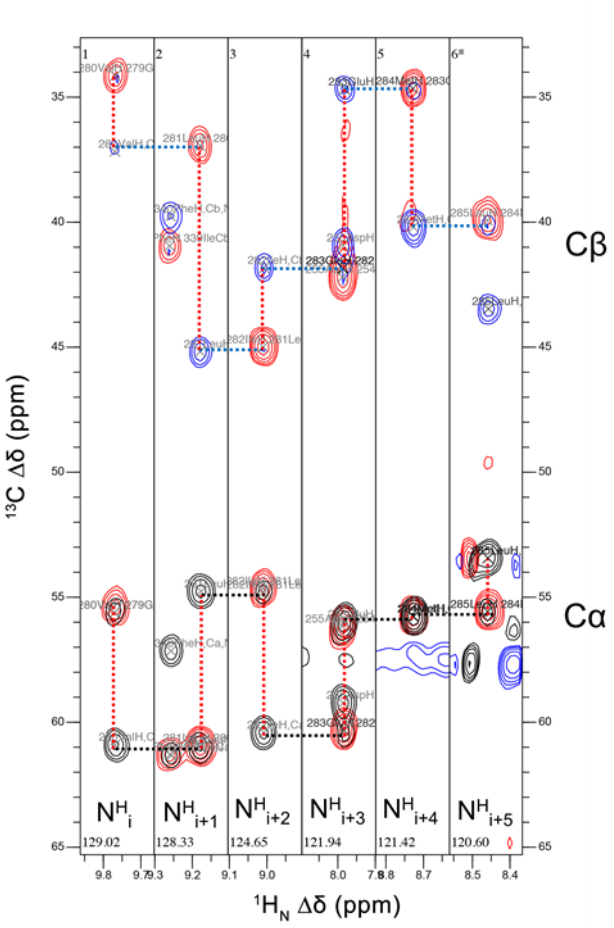
Experiment	Sweep Width			Number of Complex Points			Spectrometer Frequency (MHz)	Centre of Resonance Chemical Shift (ppm)			Number of Scans	Mixing Time (ms)
	^{1}H (Hz)	^{15}N (Hz)	^{1}H (Hz)	^{1}H	^{15}N	^{1}H		^{1}H	^{15}N	^{1}H		
^{15}N-TOCSY-HSQC	9345	2500	8333	512	32	128	749.95	4.8	117.75	4.8	8	50
^{15}N-NOESY-HSQC	9345	2500	8333	512	32	128	749.95	4.8	117.75	4.8	8	100
Heteronuclear NOE	9345	2500	-	1024	-	128	749.95	4.8	117.75	4.8	64	
1H,15N-HSQC (RDC:ISO)	9345	2500	-	1024	128	-	749.95	4.8	117.75	4.8	96	-
1H,15N-HSQC (RDC:ANISO)	9345	2500	-	1024	128	-	749.95	4.8	117.75	4.8	96	-

Based on the content of the alignment medium, ISO and ANISO refer to isotropic or anisotropic samples

7.6 Acquisition parameters for ^{13}C-edited experiment

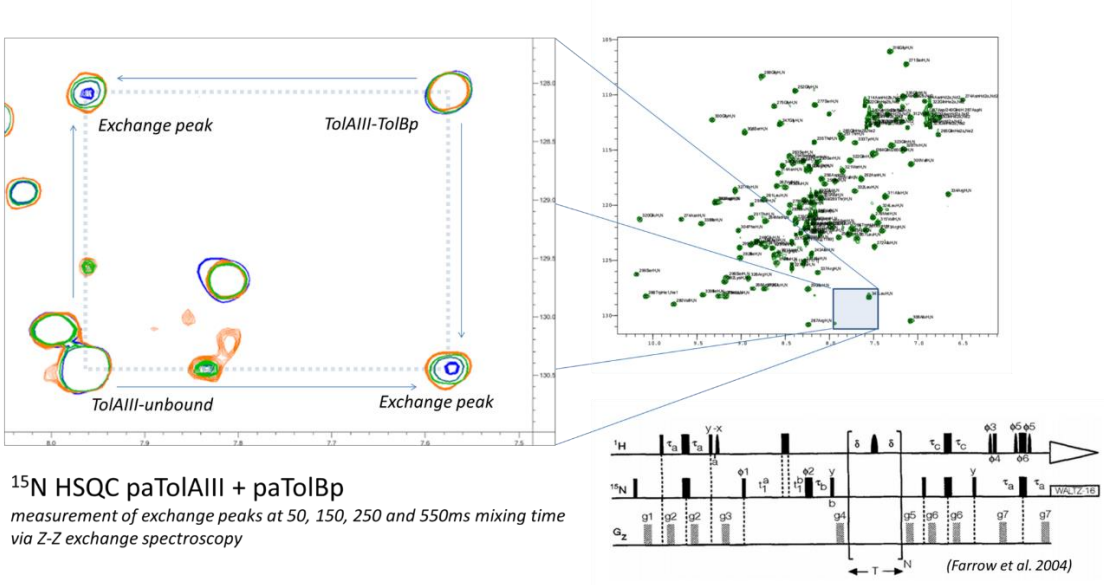
Experiment	Sweep Width			Number of Complex Points			Spectrometer Frequency (MHz)	Centre Reference Chemical Shift (ppm)			Number of Scans	Mixing Time (ms)
	^{1}H	^{13}C	^{1}H	^{1}H	^{13}C	^{1}H		^{1}H	^{13}C	^{1}H		
950 OMEGA	9345	11364	7874	512	64	128	749.95	4.8	40	4.8	8	15
950 OMEGA	11904	14706	10000	512	64	128	949.79	4.8	40	4.8	8	100
950 OMEGA	11904	14706		1024	128	-	949.79	4.8	40	4.8	16	-

7.7 Sample slice for the assignment strategy for TolAIII using CBCACNH/CBCACONH for *de novo* assignment of backbone residues



Sample slices from 13C CBCACONH and CBCANH spectra used for standard triple resonance backbone assignment for {15N, 13C}-TolAIII saturated with $TolB_{22-34}$.

7.8 ZZ-exchange spectroscopy

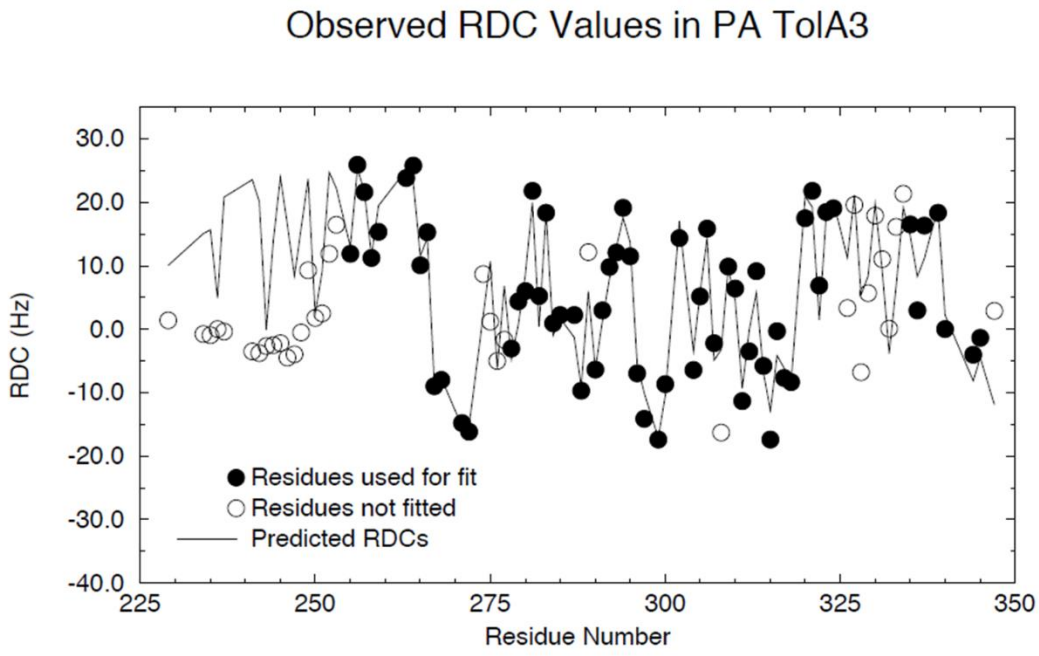


^{15}N HSQC paTolAIII + paTolBp

measurement of exchange peaks at 50, 150, 250 and 550ms mixing time via Z-Z exchange spectroscopy

As a slowly exchanging complex, the interaction between TolAIII and TolB is accessible to Z-Z exchange spectroscopy. An example segment of the 15N HSQC spectra collected has been shown to demonstrate the effect of mixing time on the intensity of exchange cross peaks. Mixing times depicted as follows: Orange=550ms, Light green= 250ms, Dark Green= 150ms, Blue=50ms. Pulse sequence used as described in (Farrow et al. 2004), using preloaded sequences on the Bruker 500 with cryoprobe in the TopSpin program. Spectra collected over the course of 4 days. Since ‘forward’ and ‘reverse’ crosspeaks between the bound and bound state occur in ZZ exchange spectroscopy, it is a useful tool for assigning spectra in slow exchange.

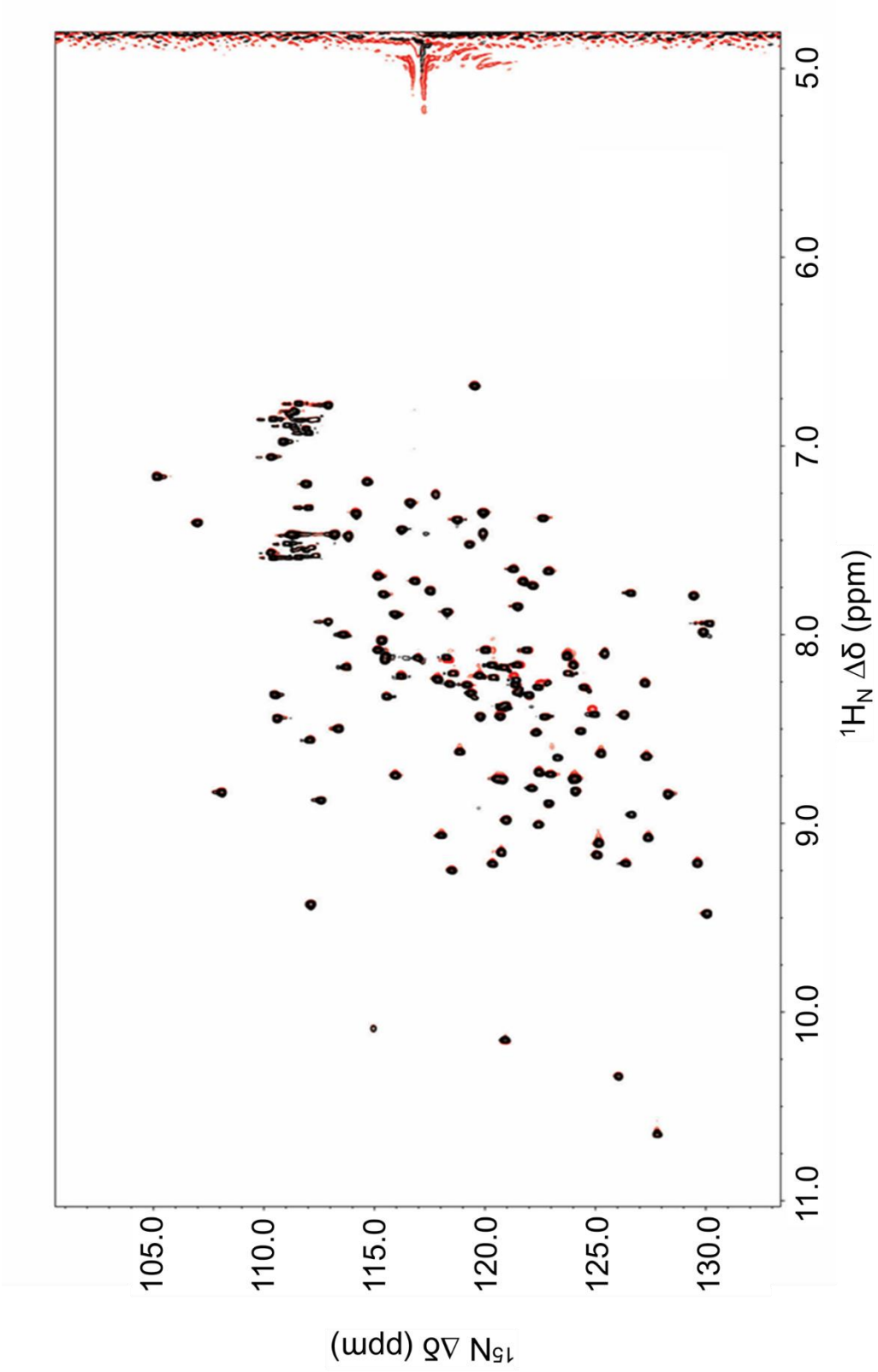
7.9 RDC analysis is largely consistent with the crystal structure for *P. aeruginosa* TolAIII



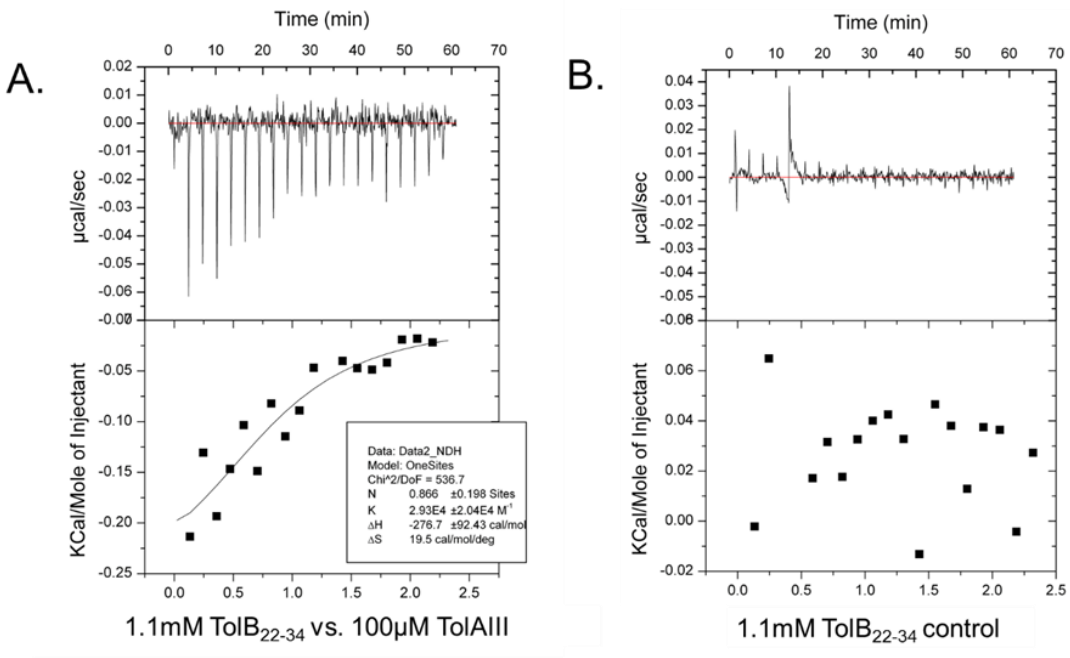
A fit of the observed RDC's with the RDCs predicted by the ILR0 crystal structure shows good agreement across most of the folded globular domain. The low Q factor of 0.19 is an indicator of quality of the fit (Tjandra, 2013).This suggests that the unbound state of TolAIII is largely consistent with the crystal structure of TolAIII. The poor agreement of the RDCs before residue 255 confirm that the extended helix found in the X-ray structure is not present in solution. Fitting parameters in the unbound state were:

Da=13.785, R=0.22, q=0.19 theta=98.59, phi=43.81, psi=124.71

7.10 An overlay of two HSQC spectra for TolAIII within a range of 6.4-7.5 shows no pH-induced major chemical shift changes.

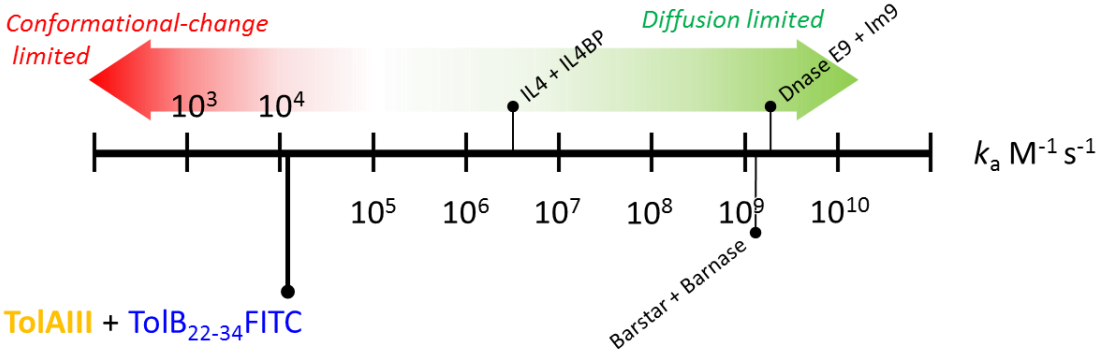


7.11 The TolAIII-TolB$_{22-34}$ interaction is exothermic by ITC



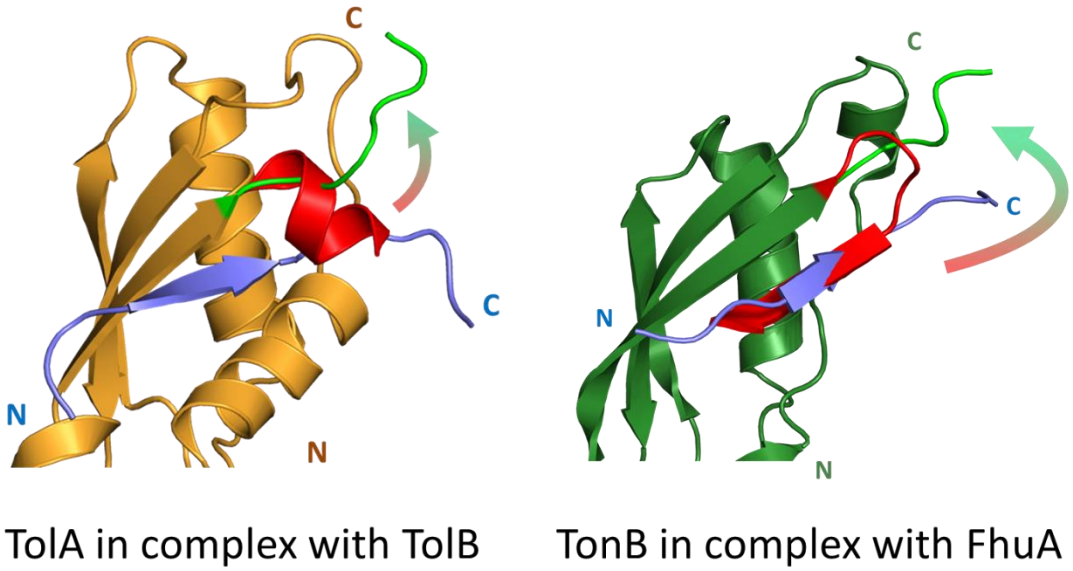
Signal to noise for a sample of 100 µM TolB$_{22-34}$ titrated against 1.1 mM TolAIII was too small to resolve a binding affinity for the TolA-TolB interaction by ITC (see methods Chapter 2). The heats of association however were sufficiently large to demonstrate that the TolAIII-TolB$_{22-34}$ interaction is an exothermic reaction.

7.12 The TolA-TolB interaction is in a range consistent with conformational change limited interactions.



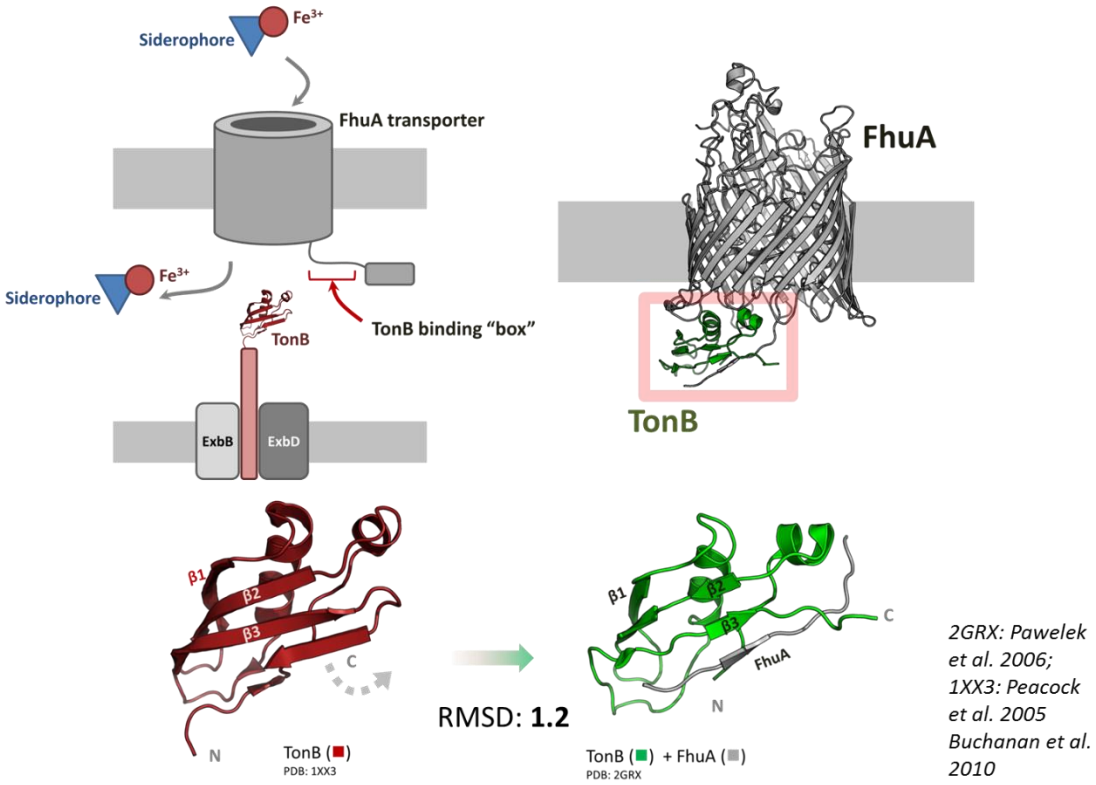
The TolA-TolB interaction has a comparatively slow rate of association, placing it in the range of interactions that are conformational-change limited. Reaction rate scale adapted from (Schreiber *et al*, 2009).

7.13 The TonB import pathway involves a rearrangement at the C-terminus of TonB, in a mechanism similar to TolA-TolB



TolA in complex with TolB TonB in complex with FhuA

TonB forms a complex with an outer membrane receptor to facilitate the entry of a siderophore. This interaction involves the displacement of a C-temrinal structure in TonB, specifically the 4th beta-strand to allow for the association of its endogenous receptor, in this case FhuA, through parallel beta-strand complementation.



2GRX: Pawelek et al. 2006; 1XX3: Peacock et al. 2005 Buchanan et al. 2010

References

Abergel C, Bouveret E, Claverie JM, Brown K, Rigal A, Lazdunski C & Bénédetti H (1999) Structure of the Escherichia coli TolB protein determined by MAD methods at 1.95 A resolution. *Structure* **7:** 1291–300

Abergel C, Walburger A, Chenivesse S & Lazdunski C (2001) Crystallization and preliminary crystallographic study of the peptidoglycan-associated lipoprotein from Escherichia coli. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57:** 317–319

Adams H, Zeder-Lutz G, Schalk I, Pattus F & Celia H (2006) Interaction of TonB with the outer membrane receptor FpvA of Pseudomonas aeruginosa. *J. Bacteriol.* **188:** 5752–5761

Angelini A, Cendron L, Goncalves S, Zanotti G & Terradot L (2008) Structural and enzymatic characterization of HP0496, a YbgC thioesterase from Helicobacter pylori. *Proteins* **72:** 1212–21

Anwari K, Poggio S, Perry A, Gatsos X, Ramarathinam SH, Williamson NA, Noinaj N, Buchanan S, Gabriel K, Purcell AW, Jacobs-Wagner C & Lithgow T (2010) A modular BAM complex in the outer membrane of the alpha-proteobacterium Caulobacter crescentus. *PLoS One* **5:** e8619

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL & Mori H (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2:** 2006.0008

Bagos PG, Nikolaou EP, Liakopoulos TD & Tsirigos KD (2010) Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics* **26:** 2811–2817

Baron S, Salton M & Kim K (1996) 'Structure'. In *Baron's Medical Microbiology (4th Ed.)* Univ. of Texas Medical Branch

Beckwith J (2013) The Sec-dependent pathway. *Res. Microbiol.* **164:** 497–504

Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D & Holden HM (1998) The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from Pseudomonas sp. Strain CBS-3. *J. Biol. Chem.* **273:** 33572–9

Berjanskii M V. & Wishart DS (2008) Application of the random coil index to studying protein flexibility. *J. Biomol. NMR* **40:** 31–48

Berks BC, Palmer T & Sargent F (2005) Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Curr. Opin. Microbiol.* **8:** 174–181

Berks BC, Sargent F & Palmer T (2000) The Tat protein export pathway. *Mol. Microbiol.* **35:** 260–274

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res.* **28:** 235–242

Bernadac A, Gavioli M, Lazzaroni J-C, Raina S & Lloubès R (1998) Escherichia coli tol-pal mutants form outer membrane vesicles. *J. Bacteriol.* **180:** 4872–4878

Bernard CS, Sadasivam M, Shiomi D & Margolin W (2007a) An altered FtsA can compensate for the loss of essential cell division protein FtsN in Escherichia coli. *Mol. Microbiol.* **64:** 1289–1305

Bernstein A, Rolfe B & Onodera K (1972) Pleiotropic properties and genetic organization of the tolA, B locus of Escherichia coli K-12. *J. Bacteriol.* **112:** 74–83

Bonsor DA, Grishkovskaya I, Dodson EJ & Kleanthous C (2007) Molecular mimicry enables competitive recruitment by a natively disordered protein. *J. Am. Chem. Soc.* **129:** 4800–4807

Bonsor DA, Hecht O, Vankemmelbeke M, Sharma A, Krachler AM, Housden NG, Lilly KJ, James R, Moore GR & Kleanthous C (2009) Allosteric beta-propeller signalling in TolB and its manipulation by translocating colicins. *EMBO J.* **28:** 2846–57

Bourdineaud J, Howard S & Lazdunski C (1989) Localization and assembly into the Escherichia coli envelope of a protein required for entry of colicin A. *J. Bacteriol.* **171:** 2458–2465

Bouveret E, Bénédetti H, Rigal A, Loret E & Lazdunski C (1999) In vitro characterization of peptidoglycan-associated lipoprotein (PAL)–peptidoglycan and PAL–TolB interactions. *J. Bacteriol.* **181:** 6306–6311

Bouveret E, Derouiche R, Rigal A, Lloubès R, Lazdunski C & Bénédetti H (1995) Peptidoglycan-associated lipoprotein-TolB interaction. A possible key to explaining the formation of contact sites between the inner and outer membranes of Escherichia coli. *J. Biol. Chem.* **270:** 11071–7

Braun V & Herrmann C (1993) Evolutionary relationship of uptake systems for biopolymers in Escherichia coli: cross-complementation between the TonB-ExbB-ExbD and the TolA-TolQ-TolR proteins. *Mol. Microbiol.* **8:** 261–268

Browning DF, Wells TJ, França FLS, Morris FC, Sevastsyanovich YR, Bryant JA, Johnson MD, Lund PA, Cunningham AF, Hobman JL, May RC, Webber MA & Henderson IR (2013) Laboratory adapted Escherichia coli K-12 becomes a pathogen of Caenorhabditis elegans upon restoration of O antigen biosynthesis. *Mol. Microbiol.* **87:** 939–950

Cannon GC, Bradburne CE, Aldrich HC, Baker SH, Heinhorst S & Shively JM (2001) Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra. *Appl. Environ. Microbiol.* **67:** 5351–5361

Carr S, Penfold CN, Bamford V, James R & Hemmings a M (2000) The structure of TolB, an essential component of the tol-dependent translocation system, and its protein-protein interaction with the translocation domain of colicin E9. *Structure* **8:** 57–66

Carugo O & Argos P (1997) Protein-protein crystal-packing contacts. *Protein Sci.* **6:** 2261–3

Cascales E & Bernadac A (2002) Pal lipoprotein of Escherichia coli plays a major role in outer membrane integrity. *J. Bacteriol.* **184:** 754–759

Cascales E, Buchanan SK, Duché D, Kleanthous C, Lloubès R, Postle K, Riley M, Slatin S & Cavard D (2007) Colicin biology. *Microbiol. Mol. Biol. Rev.* **71:** 158–229

Cascales E, Gavioli M, Sturgis JN & Lloubès R (2000) Proton motive force drives the interaction of the inner membrane TolA and outer membrane pal proteins in Escherichia coli. *Mol. Microbiol.* **38:** 904–15

Cascales E & Lloubès R (2003) Deletion analyses of the peptidoglycan-associated lipoprotein Pal reveals three independent binding sequences including a TolA box. *Mol. Microbiol.* **51:** 873–885

Cascales E, Lloubès R & Sturgis JN (2001) The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. *Mol. Microbiol.* **42:** 795–807

Cassels E (2012) Characterisation of the E.coli and Pseudomonas aeruginosa TolA-TolB interaction. Thesis. Available at: http://etheses.whiterose.ac.uk/3890/

Cheung MS, Maguire ML, Stevens TJ & Broadhurst RW (2010) DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J. Magn. Reson.* **202:** 223–233

Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science*. **311**(5765):1283-7. (Erratum in: *Science.* 2006 **312**(5774):697.)

Claridge JK & Schnell JR (2012) Bacterial production and solution NMR studies of a viral membrane ion channel. *Methods Mol. Biol.* **831:** 165–179

Clavel T, Germon P, Vianney A, Portalier R, Lazzaroni JC & Lyon I (1998) TolB protein of Escherichia coli K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal , Lpp and OmpA. **29:** 359–367

Clavel T, Lazzaroni JC, Vianney A & Portalier R (1996) K-12 is controlled by the RcsC sensor protein involved in capsule synthesis. *Mol. Microbiol.* **19:** 19–25

Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M & Waksman G (2015) Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat. Rev. Microbiol.* **13:** 343–359

Coyette J & Van Der Ende A (2008) Peptidoglycan: The bacterial Achilles heel. *FEMS Microbiol. Rev.* **32:** 147–148

Daley D, Rapp M, Granseth E, Melen K, Drew D & von Heijne G (2005) Global topology analysis of the Escherichia coli inner membrane proteome. *Sci. …* **308:** 1321–1323

Dartigalongue C, Nikaido H & Raina S (2000) Protein folding in the periplasm in the absence of primary oxidant DsbA: modulation of redox potential in periplasmic space via OmpL porin. *EMBO J.* **19:** 5980–8

Dasgupta S, Iyer GH, Bryant SH, Lawrence CE & Bell JA (1997) Extent and nature of contacts between protein molecules in crystal lattices and between subunits of protein oligomers. *Proteins Struct. Funct. Genet.* **28:** 494–514

Datsenko K a & Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97:** 6640–6645

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293

DeLano WL (2002) The PyMOL Molecular Graphics System. *Schrödinger LLC* **Version 1.:** http://www.pymol.org

Dennis JJ, Lafontaine ER & Sokol P a (1996) Identification and characterization of the tolQRA genes of Pseudomonas aeruginosa. *J. Bacteriol.* **178:** 7059–7068

Deprez C, Blanchard L, Guerlesquin F, Gavioli M, Simorre J-P, Lazdunski C, Marion D & Lloubès R (2002) Macromolecular import into Escherichia coli: the TolA C-terminal domain changes conformation when interacting with the colicin A toxin. *Biochemistry* **41:** 2589–98

Deprez C, Lloubès R, Gavioli M, Marion D, Guerlesquin F & Blanchard L (2005) Solution structure of the E.coli TolA C-terminal domain reveals conformational changes upon binding to the phage g3p N-terminal domain. *J. Mol. Biol.* **346:** 1047–57

Derouiche R, Bénédetti H, Lazzaroni JC, Lazdunski C & Lloubès R (1995) Protein complex within Escherichia coli inner membrane. TolA N-terminal domain interacts with TolQ and TolR proteins. *J. Biol. Chem.* **270:** 11078–84

Derouiche R, Lloubès R, Sasso S, Bouteille H, Oughideni R, Lazdunski C & Loret E (1999) Circular dichroism and molecular modeling of the E. coli TolA periplasmic domains. *Biospectroscopy* **5:** 189–98

Driessen AJM, Manting EH & van der Does C (2001) The structural basis of protein targeting and translocation in bacteria. *Nat. Struct. Biol.* **8:** 492–498

Dubuisson JF, Vianney A, Claude J & Lazzaroni JC (2002) Mutational Analysis of the TolA C-Terminal Domain of Escherichia coli and Genetic Evidence for an Interaction between TolA and TolB. *J Bacteriol.* **184**(16):4620-5

Eckler SA, Kuehn R & Gautam M (2005) Deletion of N-terminal rapsyn domains disrupts clustering and has dominant negative effects on clustering of full-length rapsyn. *Neuroscience* **131:** 661–670

Ehrmann, M., & American Society for Microbiology (2007). The periplasm. Washington, D.C: *ASM Press.*

Eswar N, Webb B, Marti-Renom M a, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U & Sali A (2007) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **Chapter 2:** Unit 2.9

Faraldo-Gómez JD & Sansom MSP (2003) Acquisition of siderophores in gram-negative bacteria. *Nat. Rev. Mol. Cell Biol.* **4:** 105–116

Farrow N, Muhandiram R, Singer U, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD & Kay LE (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33:** 5984–6003

Ford CG, Kolappan S, Phan HTH, Waldor MK, Winther-Larsen HC & Craig L (2012) Crystal structures of a CTXphi pIII domain unbound and in complex with a Vibrio cholerae TolA domain reveal novel interaction interfaces. *J. Biol. Chem.* **287:** 36258–72

Gaspar J, Thomas J, Marolda CL & Valvano M (2000) Surface expression of O-specific lipopolysaccharide in Escherichia coli requires the function of the TolA protein. *Mol. Microbiol.* **38:** 262–75

Gennity JM, Inouye M (1991) The protein sequence responsible for lipoprotein membrane localization in Escherichia coli exhibits remarkable specificity. *J Biol Chem.* **266**(25):16458-64.

Gerding MA, Ogata Y, Pecora ND, Niki H & De Boer PAJ (2007b) 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. *Mol. Microbiol.* **63:** 1008–1025

Germon P, Ray M, Vianney A, Lazzaroni C, Ray L & Lazzaroni JC (2001) Energy-Dependent Conformational Change in the TolA Protein of Escherichia coli Involves Its N-Terminal Domain , TolQ , and TolR. *J Bacteriol*. **183**(14):4110-4

Gill SC & von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182:** 319–26

Glauert AM & Thornley MJ (1969) The topography of the bacterial cell wall. *Annu. Rev. Microbiol.* **23:** 159–98

Godlewska R, Wiśniewska K, Pietras Z & Jagusztyn-Krynicka EK (2009) Peptidoglycan-associated lipoprotein (Pal) of Gram-negative bacteria: function, structure, role in pathogenesis and potential application in immunoprophylaxis. *FEMS Microbiol. Lett.* **298:** 1–11

Gokce I, Raggett EM, Hong Q, Virden R, Cooper A & Lakey JH (2000) The TolA-recognition site of colicin N. ITC, SPR and stopped-flow fluorescence define a crucial 27-residue segment. *J. Mol. Biol.* **304:** 621–32

Goley ED, Yeh Y-C, Hong S-H, Fero MJ, Abeliuk E, McAdams HH & Shapiro L (2011) Assembly of the Caulobacter cell division machine. *Mol. Microbiol.* **80:** 1680–98

Graham L, Beveridge TJ & Nanninga N (1991a) Periplasmic space and the concept of the periplasm. *Trends Biochem Sci.* **9:** 328–329

Graham LL & Beveridge TJ (1991) Periplasmic space and the concept of the periplasm. *Trends Biochem. Sci.* **16:** 328–329

Graham LL, Harris R, Villiger W & Beveridge TJ (1991b) Freeze-substitution of gram-negative eubacteria: general cell morphology and envelope profiles. *J. Bacteriol.* **173:** 1623–33

Gratia A (1925) Sur un rémarquable example d´ antagonisme entre deux souches de colibacille. *Biochimie* **84:** 1040–1041

Gray AN, Egan AJF, van't Veer IL, Verheul J, Colavin A, Koumoutsi A, Biboy J, Altelaar MAF, Damen MJ, Huang KC, Simorre JP, Breukink E, den Blaauwen T, Typas A, Gross CA & Vollmer W (2015) Coordination of peptidoglycan synthesis and outer membrane constriction during Escherichia coli cell division. *Elife* **4:** 1–29

Gully D & Bouveret E (2006) A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in Escherichia coli. *Proteomics* **6:** 282–293

Habeck M, Rieping W, Linge JP & Nilges M (2004) NOE assignment with ARIA 2.0: the nuts and bolts. *Methods Mol. Biol.* **278:** 379–402

Hagan CL, Silhavy TJ & Kahne D (2011) β-Barrel Membrane Protein Assembly by the Bam Complex. *Annu. Rev. Biochem.* **80:** 189–210

Han L, Zheng J, Wang Y, Yang X, Liu Y, Sun C, Cao B, Zhou H, Ni D, Lou J, Zhao Y & Huang Y (2016) Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23:** 192–196

Hecht O, Ridley H, Lakey JH & Moore GR (2009) A common interaction for the entry of colicin N and filamentous phage into Escherichia coli. *J. Mol. Biol.* **388:** 880–93

Hecht O, Zhang Y, Li C, Penfold CN, James R & Moore GR (2010) Characterisation of the interaction of colicin A with its co-receptor TolA. *FEBS Lett.* **584:** 2249–52

Higgs PI, Larsen R & Postle K (2002) Quantification of known components of the Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA. *Mol. Microbiol.* **44:** 271–81

Hizukuri Y, Morton JF, Yakushi T, Kojima S & Homma M (2009) The peptidoglycan-binding (PGB) domain of the Escherichia coli pal protein can also function as the PGB domain in E. coli flagellar motor protein MotB. *J. Biochem.* **146:** 219–29

Housden NG, Hopper JTS, Lukoyanova N, Rodriguez-Larrea D, Wojdyla J., Klein A, Kaminska R, Bayley H, Saibil HR, Robinson C V. & Kleanthous C (2013) Intrinsically Disordered Protein Threads Through the Bacterial Outer-Membrane Porin OmpF. *Science (80).* **340:**

Iancu C V, Morris DM, Dou Z, Heinhorst S, Cannon GC & Jensen GJ (2010) Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. *J. Mol. Biol.* **396:** 105–17

Ikura M & Bax A (1992) Isotope-filtered 2D NMR of a protein-peptide complex: study of a skeletal muscle myosin light chain kinase fragment bound to calmodulin. *J. Am. Chem. Soc.* **114:** 2433–2440

Jakob RP, Geitner A-J, Weininger U, Balbach J, Dobbek H & Schmid FX (2012) Structural and energetic basis of infection by the filamentous bacteriophage IKe. *Mol. Microbiol.* **84:** 1124–38

Jeske O, Schüler M, Schumann P, Schneider A, Boedeker C, Jogler M, Bollschweiler D, Rohde M, Mayer C, Engelhardt H, Spring S & Jogler C (2015) Planctomycetes do possess a peptidoglycan cell wall. *Nat. Commun.* **6:** 7116

Kampfenkel K & Braun V (1993) Membrane topologies of the TolQ and TolR proteins of Escherichia coli: inactivation of TolQ by a missense mutation in the proposed first transmembrane segment. *J. Bacteriol.* **175:** 4485–91

Karlsson M, Hannavy K & Higgins CF (1993) A sequence-specific function for the N-terminal signal-like sequence of the TonB protein. *Mol. Microbiol.* **8:** 379–88

Kay LE (2011) Solution NMR spectroscopy of supra-molecular systems, why bother? A methyl-TROSY view. *J. Magn. Reson.* **210:** 159–170

Kerkhoven R, van Enckevort FHJ, Boekhorst J, Molenaar D & Siezen RJ (2004) Visualization for genomics: The Microbial Genome Viewer. *Bioinformatics* **20:** 1812–1814

Kim KH, Aulakh S & Paetzel M (2012) The bacterial outer membrane β-barrel assembly machinery. *Protein Sci.* **21:** 751–768

Kleanthous C (2010) Swimming against the tide: progress and challenges in our understanding of colicin translocation. *Nat. Rev. Microbiol.* **8:** 843–8

Koch L (1998) The biophysics of the gram-negative periplasmic space. *Crit. Rev. Microbiol.* **24:** 23–59

Koebnik R (1995) Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins. *Mol. Microbiol.* **16:** 1269–70

Koronakis V, Sharff A, Koronakis E, Luisi B & Hughes C (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405:** 914–9

Krachler AM (2009) Thesis: Functional and Structural Characterization of the Tol System-Associated Protein YbgF.

Krachler AM, Sharma A, Cauldwell A, Papadakos G & Kleanthous C (2010a) TolA modulates the oligomeric status of YbgF in the bacterial periplasm. *J. Mol. Biol.* **403:** 270–85

Krachler AM, Sharma A & Kleanthous C (2010b) Self-association of TPR domains: Lessons learned from a designed, consensus-based TPR oligomer. *Proteins Struct. Funct. Bioinforma.* **78:** 2131–2143

Lafontaine ER & Sokol PA (1998) Effects of Iron and Temperature on Expression of the Pseudomonas aeruginosa tolQRA Genes : Role of the Ferric Uptake Regulator. *J Bacteriol.* **180:**(11):2836-41

Laskowski RA (2009) PDBsum: new things. *Nucleic Acids Res.* **37:** D355–9

Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML & Thornton JM (1997) PDBsum: A Web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.* **22:** 488–490

Laskowski RA & Swindells MB (2011) LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51:** 2778–2786

Lazdunski CJ, Bouveret E, Rigal A, Journet L, Lloubès R & Bénédetti H (1998) Colicin import into Escherichia coli cells. *J. Bacteriol.* **180:** 4993–5002

Lazzaroni JC, Vianney A, Popot JL, Bénédetti H, Samatey F, Lazdunski C, Portalier R & Géli V (1995) Transmembrane alpha-helix interactions are required for the functional assembly of the Escherichia coli Tol complex. *J. Mol. Biol.* **246:** 1–7

Lee W, Revington MJ, Arrowsmith C & Kay LE (1994) A pulsed field gradient isotope-filtered 3D 13C HMQC-NOESY experiment for extracting intermolecular NOE contacts in molecular complexes. *FEBS Lett.* **350:** 87–90

Levengood SK, Beyer WF & Webster RE (1991) TolA: a membrane protein involved in colicin uptake contains an extended helical region. *Proc. Natl. Acad. Sci. U. S. A.* **88:** 5939–43

van der Ley P & Steeghs L (2003) Lessons from an LPS-deficient Neisseria meningitidis mutant. *J. Endotoxin Res.* **9:** 124–8

Li C, Zhang Y, Vankemmelbeke M, Hecht O, Aleanizy FS, Macdonald C, Moore GR, James R & Penfold CN (2012) Structural evidence that colicin A protein binds to a novel binding site of TolA protein in Escherichia coli periplasm. *J. Biol. Chem.* **287:** 19048–57

Liao M, Cao E, Julius D & Cheng Y (2013) Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**

Linge JP, Habeck M, Rieping W & Nilges M (2003) ARIA: Automated NOE assignment and NMR structure calculation. *Bioinformatics* **19:** 315–316

Llamas MA, Ramos JL & José J (2000) Mutations in each of the tol genes of Pseudomonas putida reveal that they are critical for maintenance of outer membrane stability. *J Bacteriol.* **182**:4764-72.

Llamas MA, Rodríguez-herva JJ, Robert EW, Bitter W, Tommassen J, Juan L, Hancock REW & Ramos JL (2003) Role of Pseudomonas putida tol-oprL gene products in uptake of solutes through the cytoplasmic membrane. *J Bacteriol.* **185**(16):4707-16.

Lloubès R, Cascales E, Walburger a, Bouveret E, Lazdunski C, Bernadac a & Journet L (2001) The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity? *Res. Microbiol.* **152:** 523–9

Lloubès R, Goemaere E, Zhang X, Cascales E & Duché D (2012) Energetics of colicin import revealed by genetic cross-complementation between the Tol and Ton systems. *Biochem. Soc. Trans.* **40:** 1480–5

Loftus SR, Walker D, Maté MJ, Bonsor D a, James R, Moore GR & Kleanthous C (2006) Competitive recruitment of the periplasmic translocation portal TolB by a natively disordered domain of colicin E9. *Proc. Natl. Acad. Sci. U. S. A.* **103:** 12353–8 Lorenz SH, Jakob RP, Weininger U, Balbach J, Dobbek H & Schmid FX (2011) The filamentous phages fd and IF1 use different mechanisms to infect Escherichia coli. *J. Mol. Biol.* **405:** 989–1003 Available at: http://www.ncbi.nlm.nih.gov/pubmed/21110981 [Accessed January 28, 2013]

Lubkowski J, Hennecke F, Plueckthun A & Wlodawer A (1999a) Filamentous phage infection: Crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA. *Structure* **7:** 711–722

Lubkowski J, Hennecke F, Plückthun A & Wlodawer A (1998) The structural basis of phage display elucidated by the crystal structure of the N-terminal domains of g3p. *Nat. Struct. Biol.* **5:** 140–7

Lubkowski J, Hennecke F, Plückthun A & Wlodawer A (1999b) Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA. *Structure* **7:** 711–22

Lutkenhaus J, Pichoff S & Du S (2012) Bacterial cytokinesis: From Z ring to divisome. *Cytoskeleton* **69:** 778–790

Macnab RM (2003) How bacteria assemble flagella. *Annu. Rev. Microbiol.* **57:** 77–100

Madigan M (2012) Brock Biology of Microorganisms, 13th edn. *Int. Microbiol.***:** 550–551

Maltseva N, Kim Y, Osipiuk J, Anderson W. & Joachimiak A Crystal structure of peptidoglycan-associated outer membrane lipoprotein from Yersinia pestis CO92. *To be Publ.*

McCoy MA & Wyss DF (2002) Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J. Am. Chem. Soc.* **124:** 11758–11763

Meselson M & Yuan R (1968) DNA Restriction Enzyme from E Coli. *Nature* **217:** 1110

Meury J & Devilliers G (1999) Impairment of cell division in tolA mutants of Escherichia coli at low and high medium osmolarities. *Biol. cell* **91:** 67–75

Mielke SP & Krishnan V V. (2009) Characterization of protein secondary structure from NMR chemical shifts. *Prog. Nucl. Magn. Reson. Spectrosc.* **54:** 141–165

Mizuno T (1979) A novel peptidoglycan-associated lipoprotein found in the cell envelope of Pseudomonas aeruginosa and Escherichia coli. *J. Biochem.* **86:** 991–1000

Muhandiram DR & Kay LE (1994) Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *J. Magn. Reson. Ser. B* **103:** 203–216

Muller MM & Webster RE (1997) Characterization of the tol-pal and cyd region of Escherichia coli K- 12: Transcript analysis and identification of two new proteins encoded by the cyd operon.

J. Bacteriol. **179:** 2077–2080

Murakami S, Nakashima R, Yamashita E & Yamaguchi A (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* **419:** 587–93

Narita S-I, Matsuyama S-I & Tokuda H (2004) Lipoprotein trafficking in Escherichia coli. *Arch. Microbiol.* **182:** 1–6

Neidhardt FC & Curtiss R (1996) Escherichia coli and Salmonella : cellular and molecular biology. Washington, D.C: ASM Press

Neudecker P, Lundstrom P & Kay LE (2009) Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. *Biophys. J.* **96:** 2045–2054

Nietlispach D, Mott H & Stott K (2004) Structure determination of protein complexes by NMR. *METHODS …* **278:** 255–88

Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67:** 593–656

Nikaido H & Vaara M (1985) Molecular basis of bacterial outer membrane permeability. *Microbiol. Rev.* **49:** 1–32

Noinaj N, Guillier M, Barnard TJ & Buchanan SK (2010) TonB-dependent transporters: regulation, structure, and function. *Annu. Rev. Microbiol.* **64:** 43–60

Ostash B & Walker S (2010) Moenomycin family antibiotics: chemical synthesis, biosynthesis, biological activity. *Nat. Prod. Rep.* **27:** 1594–1617

Ottiger M, Delaglio F & Bax A (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131:** 373–378

Overmars L, Kerkhoven R, Siezen RJ & Francke C (2013) MGcV: the microbial genomic context viewer for comparative genome analysis. *BMC Genomics* **14:** 209

Papadakos G, Housden NG, Lilly KJ, Kaminska R & Kleanthous C (2012) Kinetic basis for the competitive recruitment of TolB by the intrinsically disordered translocation domain of colicin E9. *J. Mol. Biol.* **418**(5):269-80

Park E, Ménétret J-F, Gumbart JC, Ludtke SJ, Li W, Whynot A, Rapoport T a. & Akey CW (2013) Structure of the SecY channel during initiation of protein translocation. *Nature* **506**

Parry BR, Surovtsev I V., Cabeen MT, O'Hern CS, Dufresne ER & Jacobs-Wagner C (2014) The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. *Cell* **156:** 183–194

Parsons LM, Grishaev A & Bax A (2008) The periplasmic domain of TolR from Haemophilus influenzae forms a dimer with a large hydrophobic groove: NMR solution structure and comparison to SAXS data. *Biochemistry* **47:** 3131–42

Parsons LM, Lin F & Orban J (2006) Peptidoglycan recognition by Pal, an outer membrane lipoprotein. *Biochemistry* **45:** 2122–8

Paterson GK, Northen H, Cone DB, Willers C, Peters SE & Maskell DJ (2009) Deletion of tolA in Salmonella Typhimurium generates an attenuated strain with vaccine potential. *Microbiology* **155:** 220–8

Pawelek PD, Croteau N, Ng-Thow-Hing C, Khursigara CM, Moiseeva N, Allaire M & Coulton JW (2006) Structure of TonB in complex with FhuA, E. coli outer membrane receptor. *Science* **312:** 1399–402

Pellecchia M, Montgomery DL, Stevens SY, Vander Kooi CW, Feng HP, Gierasch LM & Zuiderweg ER (2000) Structural insights into substrate binding by the molecular chaperone DnaK. *Nat. Struct. Biol.* **7:** 298–303

Postle K & Kadner RJ (2003) Touch and go: Tying TonB to transport. *Mol. Microbiol.* **49:** 869–882

Pugsley AP & Schwartz M (1985) Export and secretion of proteins by bacteria. *FEMS Microbiol. Lett.* **32:** 3–38

Raetz CR (1978) Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli. *Microbiol Rev.* **42**(3):614-59.

Raetz CR (1996) Escherichia coli and Salmonella. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* pp 1035–63.

Raetz CR & Dowhan W (1990) Biosynthesis and function of phospholipids in Escherichia coli. *J. Biol. Chem.* **265:** 1235123–8

Raetz CRH & Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71:** 635–700

Ramarao MK, Bianchetta MJ, Lanken J & Cohen JB (2001) Role of Rapsyn Tetratricopeptide Repeat and Coiled-coil Domains in Self-association and Nicotinic Acetylcholine Receptor Clustering. *J. Biol. Chem.* **276:** 7475–7483

Ray M, Germon P, Vianney A, Lazzaroni JC, Ray L & Portalier R (2000) Identification by genetic suppression of Escherichia coli TolB residues important for TolB-Pal interaction. *J Bacteriol*. **182**(3):821-4.

Remaut H & Waksman G (2006) Protein-protein interaction through beta-strand addition. *Trends Biochem. Sci.* **31:** 436–44

Rinaldi PL (1983) Heteronuclear 2D-Noe Spectroscopy. *J. Am. Chem. Soc.* **105:** 5167–5168

Robertson IM, Spyracopoulos L & Sykes BD (2009) Biophysics and the Challenges of Emerging Threats. Puglisi JD (ed) Dordrecht: Springer Netherlands

Rosenzweig R & Kay LE (2014) Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR. *Annu. Rev. Biochem.* **83:** 291–315

Sankaran K & Wu HC (1994) Lipid Modification of Bacterial Prolipoprotein. *J. Biol. Chem.* **269:** 19701–19706

Santos TMA, Lin TY, Rajendran M, Anderson SM & Weibel DB (2014) Polar localization of Escherichia coli chemoreceptors requires an intact Tol-Pal complex. *Mol. Microbiol.* **92:** 985-1004

Schalk IJ, Yue WW & Buchanan SK (2004) Recognition of iron-free siderophores by TonB-dependent iron transporters. *Mol. Microbiol.* **54:** 14–22

Schreiber G, Haran G & Zhou H-X (2009) Fundamental aspects of protein− protein association kinetics. *Chem. Rev.* **109:** 839–60

Shen Y, Delaglio F, Cornilescu G & Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44:** 213–223

Silhavy TJ, Kahne D & Walker S (2010) The Bacterial Cell Envelope. *Cold Spring Harb. Perspect. Biol.* **2:** a000414

Sklar JG, Wu T, Kahne D & Silhavy TJ (2007) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev.* **21:** 2473–2484

Stratmann D, Boelens R & Bonvin AMJJ (2011) Quantitative use of chemical shifts for the modeling of protein complexes. *Proteins* **79:** 2662–70

Studier FW & Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189:** 113–130

Sturgis JN (2001) Organisation and evolution of the tol-pal gene cluster. *J. Mol. Microbiol.*

Biotechnol. **3:** 113–22

Sun TP & Webster RE (1987) Nucleotide sequence of a gene cluster involved in entry of E colicins and single-stranded DNA of infecting filamentous bacteriophages into Escherichia coli. *J. Bacteriol.* **169:** 2667–74

van Teeseling MCF, Mesman RJ, Kuru E, Espaillat A, Cava F, Brun Y V., VanNieuwenhze MS, Kartal B & van Niftrik L (2015) Anammox Planctomycetes have a peptidoglycan cell wall. *Nat. Commun.* **6:** 6878

Tibor a, Weynants V, Denoel P, Lichtfouse B, De Bolle X, Saman E, Limet JN & Letesson JJ (1994) Molecular cloning, nucleotide sequence, and occurrence of a 16.5-kilodalton outer membrane protein of Brucella abortus with similarity to pal lipoproteins. *Infect. Immun.* **62:** 3633–9

Typas A & Sourjik V (2015) Bacterial protein networks: properties and functions. *Nat Rev Micro* **13:** 559–572

Vianney A, Muller MM, Clavel T, Lazzaroni JC, Portalier R, Webster E, Vianney A, Muller MM, Clavel T, Lazzaroni JC, Portalier R & Webster RE (1996) Characterization of the tol-pal region of Escherichia coli K-12: translational control of tolR expression by TolQ and identification of a new open reading frame downstream of pal encoding a periplasmic protein. *J Bacteriol.* **178**(14):4031-8

Vinés E & Marolda C (2005) Defective O-antigen polymerization in tolA and pal mutants of Escherichia coli in response to extracytoplasmic stress. *J. Bacteriol.* **187:** 3359–3368

Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J & Laue ED (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Genet.* **59:** 687–696

de Vries SJ, van Dijk M & Bonvin AMJJ (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **5:** 883–897

Walburger A, Lazdunski C & Corda Y (2002) The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB. *Mol. Microbiol.* **44:** 695–708

Wang Y, Wang R, Jin F, Liu Y, Yu J, Fu X & Chang Z (2016) A supercomplex spanning the inner and outer membranes mediates the biogenesis of ??-barrel outer membrane proteins in bacteria. *J. Biol. Chem.* **291:** 16720–16729

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF & Jones DT (2004) Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **337:** 635–645

Webster RE (1991) The tol gene products and the import of macromolecules into Escherichia coli. *Mol. Microbiol.* **5:** 1005–1011

Whitehead L, Long BM, Price GD & Badger MR (2014) Comparing the in Vivo Function of -Carboxysomes and -Carboxysomes in Two Model Cyanobacteria. *Plant Physiol.* **165:** 398–411

Wielink J Van & Duine J (1990) How big is the periplasmic space? *Trends Biochem. Sci.***15:**136-7

Wilks JC & Slonczewski JL (2007) pH of the cytoplasm and periplasm of Escherichia coli: Rapid measurement by green fluorescent protein fluorimetry. *J. Bacteriol.* **189:** 5601–5607

Wishart DS & Sykes BD (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR* **4:** 171–80

Witty M, Sanz C, Shah A, Grossmann JG, Mizuguchi K, Perham RN & Luisi B (2002) Structure of the periplasmic domain of Pseudomonas aeruginosa TolA: evidence for an evolutionary

relationship with the TonB transporter protein. *EMBO J.* **21:** 4207–18

Wojdyla JA, Cutts E, Kaminska R, Papadakos G, Hopper JTS, Stansfeld PJ, Staunton D, Robinson C V. & Kleanthous C (2015) Structure and function of the Escherichia coli Tol-Pal stator protein TolR. *J. Biol. Chem.* **290:**

Yokota N, Kuroda T, Matsuyama S & Tokuda H (1999) Characterization of the LolA-LolB system as the general lipoprotein localization mechanism of Escherichia coli. *J. Biol. Chem.* **274:** 30995–9

Zhang G, Meredith TC & Kahne D (2013) On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr. Opin. Microbiol.* **16:** 779–785

Zhang XY-Z, Goemaere EL, Seddiki N, Célia H, Gavioli M, Cascales E & Lloubes R (2011) Mapping the interactions between Escherichia coli TolQ transmembrane segments. *J. Biol. Chem.* **286:** 11756–64

Zhang Y, Li C, Vankemmelbeke MN, Bardelang P, Paoli M, Penfold CN & James R (2010) The crystal structure of the TolB box of colicin A in complex with TolB reveals important differences in the recruitment of the common TolB translocation portal used by group A colicins. *Mol. Microbiol.* **75:** 623–36

Zhuang Z, Song F, Martin BM & Dunaway-Mariano D (2002) The YbgC protein encoded by the ybgC gene of the tol-pal gene cluster of Haemophilus influenzae catalyzes acyl-coenzyme A thioester hydrolysis. *FEBS Lett.* **516:** 161–3

Zwahlen C, Legault P, Vincent SJF, Greenblatt J, Konrat R & Kay LE (1997) Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/ boxB RNA Complex. *J. Am. Chem. Soc.* **119:** 6711–6721

Zwaig R de & Luria S (1969) New class of conditional colicin-tolerant mutants. *J. Bacteriol.* **99:** 78–84