Characterisation of the *E.coli* and *Pseudomonas aeruginosa* TolA-TolB interaction

Eoin Cassels

Doctor of Philosophy

University of York
Department of Biology

December 2012

Abstract

The Tol-Pal complex of Gram-negative bacteria is a highly conserved family of interacting proteins that span the periplasm, from inner to outer membrane. Despite decades of work on this protein complex, only the structure of a small part of the Tol complex in E.coli (TolB, Pal, TolB-Pal complex, TolA domain 3) as well as part of TolA from Pseudomonas aeruginosa have been resolved, and the native function of the Tol-Pal system remains elusive. A key interaction of the Tol-Pal system is between TolA and TolB. These two proteins bridge the periplasm, linking the inner membrane complex of TolQ, TolR and TolA with the outer membrane through the outer membrane bound Pal and periplasmic TolB. The structure of the TolA-TolB complex is not known, and although the TolA binding epitope of TolB has been localised to the intrinsically disordered N-terminus of TolB, the binding site on TolA is also unknown (Bonsor et al. 2009). The aim of this work is to address a number of questions regarding a fundamental part of the Tol-Pal system: the interaction between TolA and TolB. This thesis reports that not only is the short 22 residue N-terminus of TolB important for its interaction with TolA domain 3, but that it is the sole site of interaction between E.coli TolA and TolB. This was shown by engineering the E.coli TolB N-terminus onto another protein to create a novel interaction with E.coli TolA. In addition, a synthetic peptide of the N-terminus of TolB binding TolA can recapitulate and serve as a model for the native interaction. This work also reports that the TolA-TolB interaction is conserved between Pseudomonas aeruginosa TolA and TolB, and that the N-terminus is also important for this interaction, the first work to suggest this. Finally, through use of nuclear magnetic resonance spectroscopy, residues perturbed on Pseudomonas aeruginosa TolA domain 3 through the binding of a synthetic peptide representing Pseudomonas aeruginosa TolB have been mapped onto the TolA protein to determine the potential binding site of TolB on TolA, something which until now has been unknown.

Contents

List of figures

List of tables

Acknowledgements

There are so many people I wish to thank that I hardly know where to begin. Firstly I would like to thank my supervisor, Colin Kleanthous, for allowing me the opportunity of undertaking my PhD in his lab and for all the support, be it scientific, financial, or otherwise during my time at York. I thank you for your understanding, particularly when I was unwell. I would also like to thank my training committee, Gavin Thomas and Gideon Davies for all their support and advice during this project. From the Kleanthous lab I would like to thank all the members, both current and past. In particular I thank Nick Housden for all his scientific input and advice, and not to mention the rigid timekeeping for coffee breaks. Without your assistance in all matters biophysical then this work would read very differently. On the NMR side of things, I would like to thank Karthik Rajasekar for all his assistance in designing my experiments and assigning my proteins. I thank Greg Papadakos, Justyna Wodjyla, and Renata Kaminska for all their assistance and input over the years, it is greatly appreciated. I also thank two former members of the lab, Daniel Bonsor and Anne-Marie Krachler, who welcomed me into the lab and gave me a basis to start my work. My thanks go to Andrew Brentnall as well for all his assistance, guidance and patience when I first started using NMR. My thanks to the Technology Facility staff, in particular Andrew Leech, Berni Strongitharm, David Ashford and Adam Dowle for assistance with biophysical techniques and mass spectrometry. Finally, and perhaps most importantly, I wish to thank my partner in crime, Eliška. Without you, none of this would be possible. For all those late nights, grumpy moods and missed holidays I am sorry and so lucky that you were there for me. You have supported me emotionally, financially and even physically (!) throughout my time as a PhD student. This work is dedicated to you.

This work was funded by the BBSRC.

Declaration

I declare that the work presented within this thesis, submitted by me for the degree of Doctor of Philosophy, is my own original work, except where due reference is given to other authors, and has not been previously submitted at this or any other university.

1. Introduction

1.1 Gram-negative bacteria

One of the key ways of differentiating Gram-negative from Gram-positive bacteria is by the organisation of their membranes. Gram-negative bacteria have 3 distinct features; an inner membrane surrounding the cytosol, an outer membrane separating the cell from the extra-cellular medium, and a space between the two which is called the periplasm. Both the inner and outer membranes are lipid bilayers, although their specific composition and behaviour of their components differs. The inner membrane is comprised of a single symmetrical phospholipid bilayer (for review see Raetz 1978). In contrast, the outer membrane is more rigid, comprising of an asymmetric phospholipid bilayer on the periplasmic leaflet, and lipopolysaccharide (LPS) on the extracellular side. The outer membrane protein components are limited in their diffusion in comparison with the inner membrane (Spector et al. 2010). Although the periplasm acts as a buffering space between the two membranes, it is an organelle with a unique composition in it's own right. The periplasm is a gel-like environment with it's own set of soluble proteins, as well as a unique structure of peptidoglycan within it. The peptidoglycan adds structural strength and rigidity to the cell, helping the cell resist extracellular stresses, as well as internal turgor pressure (For review see Coyette et al. 2008). Gram-positive bacteria are contrasted with Gram-negatives wherein they do not have an outer membrane; rather they have a lipid inner membrane that is surrounded by a thick wall of peptidoglycan (Zuber et al. 2006).

1.2 The outer membrane of Gram-negative bacteria

The outer membrane is an asymmetric bilayer, the extracellular leaflet of which is composed of lipopolysaccharide (LPS); a core structural component. In *E.coli*, there are approximately 10^6 LPS molecules per cell. LPS covers approximately 75% of the cell surface and accounts for nearly 30% of the gross weight of the outer membrane (Meredith et al. 2006). In addition to providing a structural role in the outer membrane, LPS is also required in Gram-negative bacteria to promote pathogenicity (Raetz 1996).

LPS is both a highly complex and variable molecule, grouped into 3 domains; a lipid A group, which acts as a hydrophobic anchor to secure the molecule in the bacterial outer membrane; a core oligosaccharide group that both links the lipid anchor to the O-antigen, and also acts to aid the outer membrane as a barrier to antibiotics (Raetz 1996). Finally the O-antigen which stretches into the extracellular medium is a polymer of saccharide groups that is variable, depending on the strain of bacteria. It should be noted the K-12 *E.coli*, the standard laboratory strain, does not synthesise or present on the surface O-antigen, nor do many of the non-pathogenic *E.coli* strains. Strains expressing O-antigen are categorised as “smooth”, whilst those without are classed as “rough” (Raetz 1996; Meredith et al. 2006). LPS can coordinate divalent metal ions (such as Mg^{2+}) to counter the natural repulsion between LPS molecules due to the negative charge held by the molecule, which in turn adds rigidity to the outer membrane (Snyder et al. 1999).

The outer membrane serves to protect the cell from the harsh external environment, however, in doing so prevents the cell from accessing essential macromolecules (Ruiz et al. 2006). Thus the cell has a variety of porins and specialised protein transporters to allow transfer of these essential nutrients into the cell (Ruiz et al. 2006). Two classes of transmembrane transporters

and pores are found in Gram-negative bacteria. The first class are active protein transporters such as BtuB (a 22-stranded β-barrel protein which is coupled to the TonB transport system) which transports vitamin B_{12}. The second class are outer membrane protein pores (termed Porins) that allow passive diffusion across the membrane (Spector et al. 2010). An example of porins is OmpF (a trimeric 16-stranded β-barrel protein) which allows passive diffusion of small polar molecules up to 700 Da, such as water and glucose) (Cowan et al. 1995). The BtuB transporter is relatively uncommon on the outer membrane, in the region of 300 BtuB molecules per cell. Conversely, the more generalised OmpF porin has over 1000 times more copies are present in the outer membrane, in the order of 10^5 copies per cell (Cowan et al. 1995; Spector et al. 2010).

1.3 The periplasm of Gram-negative bacteria

The periplasm is a semi-fluidic space between inner and outer membrane, which also includes a layer of peptidoglycan anchored to the outer membrane through Braun's lipoprotein (Lpp) (Vollmer et al. 2008a). The periplasm can be seen as a buffering organelle between the extracellular medium and the carefully managed cytoplasm. Proteins that are located in the periplasm are expressed and translated in the cytoplasm with an encoded signal sequence, which tags them for transport across the inner membrane and into the periplasmic space, translocated by a protein apparatus such as the Sec or Tat systems (Pugsley 1993, Zalucki et al. 2011). Numerous proteins involved in monitoring stress conditions are located in the periplasm, as well as proteins involved in nutrient transport (such as the TonB system for vitamin B_{12} transport) (Braun et al. 1993) and those that act in a protecting manner, such as proteases (Vollmer et al. 2008b). Additionally, as the periplasm contains peptidoglycan, many of the proteins involved in peptidoglycan synthesis and turn-over are present in the periplasm. Peptidoglycan consists of covalently linked murein peptide

subunits. These subunits are made up of β-1,4 N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) with alanine pentapeptide attached to the muramic acid (for review see Vollmer et al. 2008b). It is estimated that the periplasm spans between 100 and 150 Å, depending on the environmental conditions that the bacterial cell experiences (Collins et al. 2007).

Peptidoglycan precursors are synthesised in the cytoplasm from disaccharide peptide monomer subunits by the MurA-B-C-D-E-F protein pathway, to yield UDP-N-Acetylmuramic acid (UDP-MurNAc) attached to pentapeptide (Fiuza et al. 2008). This subunit is then transferred to the MraY inner membrane translocase which adds the lipid I component. This is then passed onto another inner membrane protein, MurG which adds UDP-N-Acetylglucosamine onto the lipid I PG intermediate, creating lipid II. This intermediate is then flipped into the periplasm by the inner membrane FtsW, whereupon a glycotransferase polymerises a nascent peptidoglycan chain onto the lipid II. Following this, a series of transpeptidases trim and crosslink the peptidoglycan molecule to be part of the PG layer network (Fiuza et al. 2008; Typas et al. 2012).

1.4 Inner membrane

The inner membrane is a phospholipid bilayer that surrounds the cytoplasm, and divides it from the periplasm. It is typically composed of 70-80% phosphatidylethanolamine, 15-20% phosphatidylglycerol and approximately 5% diphosphatidylglycerol (Raetz 1978). By surrounding the cytoplasm, the inner membrane helps to maintain the delicate set of conditions in the cytoplasm required for cell viability. Unlike limited diffusion of the outer membrane, wherein outer membrane proteins are can diffuse only within a limited region, proteins within the inner membrane appear to be freely diffusing (Spector et al. 2010). It is estimated that up to 20% of all translated

proteins in *E.coli* are inner membrane proteins, with a diverse selection of structures and functions, although the majority of proteins seem to have alpha-helical structure (Luirink et al. 2005). Proteins of the inner membrane consist of 2 major groups; lipoproteins, which are proteins modified with a lipid on their N-terminus and inserted into the outer leaflet of the inner membrane through a signal sequence (Daley et al. 2005) and inner membrane proteins, which are proteins inserted and anchored to the inner membrane via transmembrane helices (Narita et al. 2004). Movement of macromolecules across the inner membrane from the periplasm to the cytoplasm (and vice versa) is dependant on active ATP driven transport (Pugsley 1993). Generation of ATP at the inner membrane for use in ATP-dependant transport is by an inner membrane ATP-synthase. The ATP-synthase is in turn dependent on the proton motive force, a proton gradient across the inner membrane (Capaldi et al. 2002).

1.5 The *E.coli* Tol-Pal complex

The Tol-Pal complex is a family of interacting proteins located in both the inner and outer membranes of most Gram-negative bacteria, as well as spanning the periplasmic space. Tol proteins are encoded on the same operon, in the following order: YbgC-TolQ-TolR-TolA-TolB-Pal-YbgF (figure 1.1). The tol operon is located at 17 minutes on the +strand of the *E.coli* chromosome (Webster 1991). Tol can be further divided into 2 subcomplexes, Tol's Q, R and A comprising of one group, TolB and Pal the other. Although encoded on the same operon, *tolb-pal-ybgf* also has an internal promoter. The *tol-pal* operon promoters are regulated is several ways, through either consituitive expression, iron-regulation (Muller et al 1997), or by RcsC, a protein of RcsBC, a regulatory system involved in the regulation of the *cps* (capsular polysaccharide) genes. The *cps* genes encode biosynthesis machinery of cholanic acid, the major component of the capsule, which is induced in response to cell envelope stresses (Clavel et al.

1996). The stoichiometry of the Tol-Pal complex is poorly defined, although it is known that each gene is expressed at different levels. Pal is one of the most abundant outer membrane associated proteins, with estimates of 8000-40000 copies depending on cell morphology (Sturgis et al. 2001). TolA and TolR abundance have been measured as about 600 (Levengood et al. 1991) and 2500 (Müller et al. 1993) copies per cell, respectively.

tolb-pal-ybgf are mainly expressed independently of the other genes on the *tol-pal* operon through use of the internal promoter. The upsteam (P_1) promotor is constituitive, whereas the internal promoter (P_B) is iron regulated. It is estimated that 70% of *tolb-pal-ybgf* expression comes from the internal promoter (Muller et al. 1997). In addition, the expression of the other *tol-pal* genes (*ybgc-tolq-tolr-tola*) is tightly regulated; for example *tolr* is only expressed following the successful translation of *tolq*, allowing the production of these proteins to be co-ordinated, which is relevant as they form a complex together (Muller et al. 1997).

Figure 1.1 The *E.coli* Tol operon

1.6 *E.coli* TolA (eTolA)

TolA is an inner membrane protein arranged in 3 domains, anchored by a single transmembrane domain (domain 1). The N-terminal domain 1 is very short and consists of a short cytoplasmic stretch of 13 residues, followed by a 21 residue transmembrane helix (Levengood et al. 1991). Domains 2 and 3 are periplasmic, separated by a glycine rich region. The first domain after the membrane anchor has been predicted to be a triple stranded coiled-coil,

based on analytic ultracentrifugation, circular dichoism (Derouiche et al. 1999) and solution X-ray scattering data (Witty et al. 2002). This allows the protein to span the periplasm and interact with outer membrane components. The second periplasmic domain is globular, located at the C-terminus (Figure 1.2), and is essential for TolA function (Click et al. 1997). The C-terminal domain is also important in eTolA's native interaction with eTolB, as well as being parasitized by both colicins and bacteriophages to facilitate their entry into the cell (described later) (Bonsor et al. 2008).

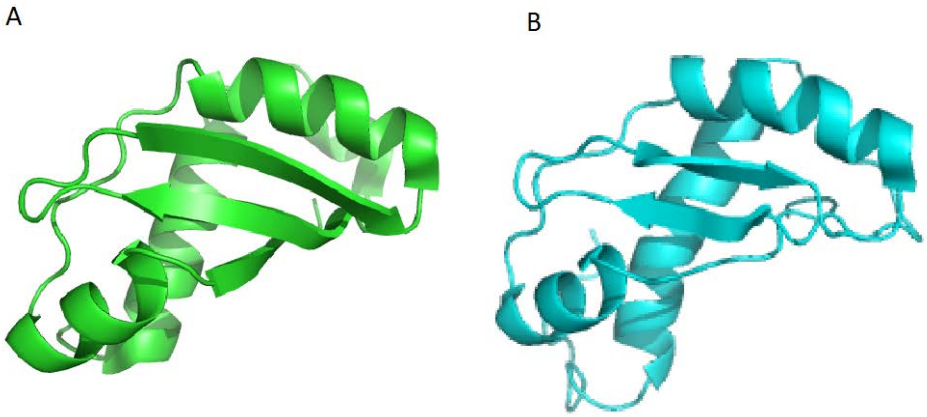


Figure 1.2 Cartoon representation of *E.coli* TolA domain 3 structure as obtained from X-ray crystallography and nuclear magnetic resonance spectroscopy experiments. (A) *E.coli* TolA3 (residues 302-421) crystal structure, PBD ID: 3QDP (Li et al. 2012). (B) *E.coli* TolA3 (325-421) solution NMR structure, PDB ID: 1S62 (Deprez et al. 2005).

1.7 *E.coli* TolQ and TolR

TolR is a 142 residue protein that like TolA is anchored to the inner membrane via a single TMD, but unlike TolA only has a single periplasmic c-terminal domain (Kampfenkel et al. 1993). *E.coli* TolR has been reported to have homology with both the inner membrane flagella motor protein MotB and shares not only homology with the TonB system component ExbD, but ExbD is also capable of limited complementation of TolR when *tolr* gene has

been deleted (Braun et al. 1993). TolR assembles with TolQ and TolA to form an inner membrane complex that may act as a molecular motor, or some form of inner membrane pore, and as such may be energised by the proton motive force (Cascales et al. 2001; Lloubes et al. 2001). TolQ is a membrane protein of 230 residues. It that spans the inner membrane three times, and has a large N-terminal region residing in the periplasm, and a large 90 residue cytoplasmic domain. eTolQ has been reported to have homology with both the inner membrane flagella motor protein MotA and the TonB system component ExbB (Braun et al. 1993). Work by Zhang et al. suggests that TolR may rotate within the membrane, and that TolQ may act as either a pore or stator, much like the association of the flagellal motor proteins MotA and MotB (Zhang et al. 2009). The stoichiometry of the *E.coli* TolR-TolQ-TolA complex is poorly defined, but is thought to be in the region of 1 TolA to 2 TolR's to 4-6 TolQ's (Guihard et al. 1994; Cascales et al. 2001).

1.8 *E.coli* TolB

E.coli TolB is a 409 residue periplasmic protein that associates with the outer membrane via Pal. TolB is a 44 kDa protein consisting of 2 domains (Bonsor et al. 2008). When expressed by the cell, *E.coli* TolB has an additional 22 residues on the extreme N-terminus that direct it for export into the periplasm. These 22 extra residues are then cleaved by signal peptidase, following export to periplasm in a SecYEG dependent manner (Zalucki et al. 2011). The larger of the 2 domains has a beta-propeller motif consisting of 6 blades. It is within this beta-propeller domain that *E.coli* Pal interacts with TolB (Cascales et al. 2007). Additionally, it has a smaller N-terminal domain, the key feature of which is the intrinsically disordered N-terminus (figure 1.3). This disordered N-terminus has been found to a site of interaction between *E.coli* TolB and TolA (Bonsor et al. 2009), something that will be further discussed in subsequent chapters.

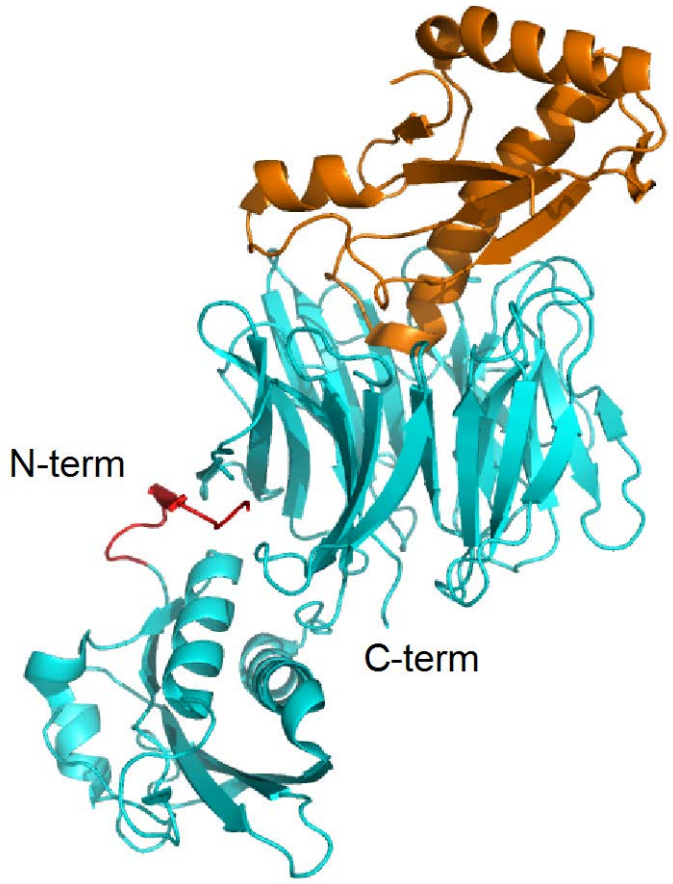


Figure 1.3 Cartoon representation of structure of *E.coli* TolB bound to *E.coli* Pal. *E.coli* TolB (cyan, residues 22-430) structure as determined by protein crystallography. N-terminal strand involved in the interaction between TolA and TolB (red, residues 22-34) is in ordered (bound back) conformation when TolB is bound to Pal (orange, residues 65-173). (PDB ID: 2W8B) (Bonsor et al. 2009).

1.9 *E.coli* Pal (Peptidoglycan-associated lipoprotein)

Pal is a 13kDa outer membrane anchored lipoprotein that is believed to be important for the Tol system's native function (Cascales et al. 2007). As it's name suggests, Pal associates with the peptidoglycan layer in the periplasm of *E.coli* cells, and as it is anchored to the outer membrane of the cell with a

lipoyl tether which is connected to the protein by a 40 residue disordered linker, and thus may act to anchor the PG layer to the outer membrane (Cascales et al. 2004). Pal's fold is that of an OmpA-like domain, which means that Pal can interact with the peptidoglycan in a non-covalent fashion (figure 1.4). Pal has been reported to interact with both TolA and TolB, as well as peptidoglycan (Cascales et al. 2007), although the interaction between TolA and Pal remains controversial (Bonsor et al. 2009). In work by Bonsor et al. it has been suggested that *E.coli* Pal may act as an off switch for the *E.coli* TolA-TolB interaction, as when *E.coli* TolB is bound to Pal it is unable to interact with TolA (Bonsor et al. 2009).

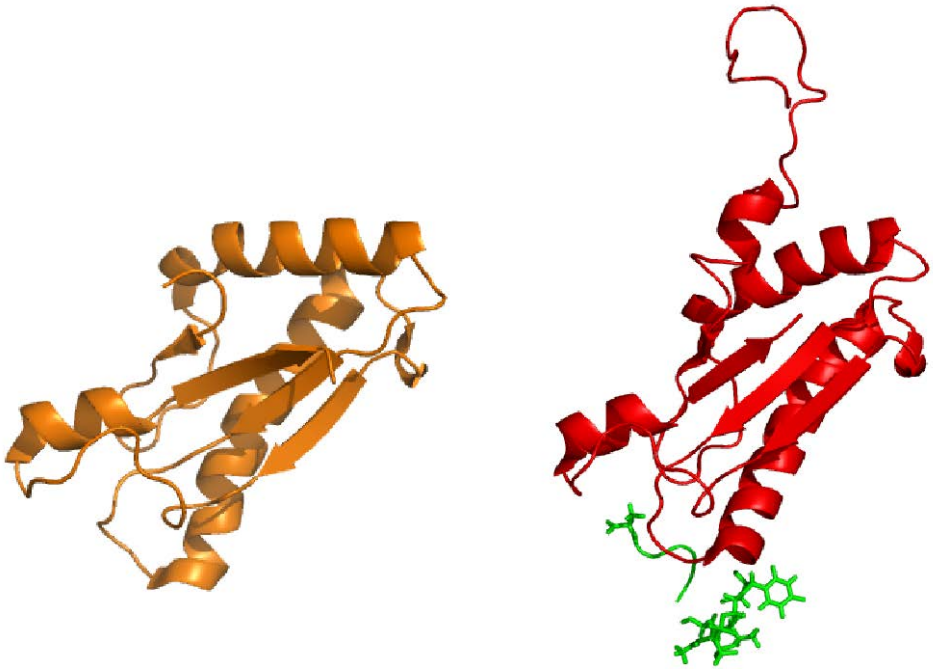


Figure 1.4 Cartoon representation of structure of *E.coli* Pal and Haemophilus influenza Pal bound to peptidoglycan precursor. *E.coli* Pal structure (orange), as obtained by protein crystallography, residues 65-173. PDB ID: 2W8B (Bonsor et al. 2009). *Haemophilus influenza* Pal (red), residues 20-153 bound to UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-meso-2,6-diaminopimeloyl-D-alanyl-D-alanine (shown as sticks). PBD ID: 2AIZ (Parsons et al. 2006).

1.10 *E.coli* YbgF

YbgF is a periplasmic protein of the Tol family with a tetratrico peptide repeat (TPR)-like structural motif of unknown function (Krachler et al. 2010). The TPR structural motif has been found to mediate protein–protein interactions and the assembly of multiprotein complexes. The motif consists of between 3 and 16 tandem repeats of 34 amino acids residues, the consensus sequence of which is defined by a pattern of small and large amino acids, although there is little specific conservation of residues. These repeats can be widely dispersed throughout the protein. Proteins with TPR motifs have been found throughout Prokaryotes and Eukaryotes, and are involved in a wide range of biological processes, including cell cycle regulation, transcriptional regulation and protein transport (For reviews see Andrea et al. 2003, Schapire et al. 2006). In isolation in the periplasm, YbgF forms a homotrimeric complex (figure 1.5), and upon binding TolA, this trimer dissociates, and single a YbgF subunit forms a heterodimeric complex with TolA. YbgF has been shown to interact with the domain 2 of TolA both *in vivo* and *in vitro*, however, the function of this interaction is not known. (Krachler et al. 2010).

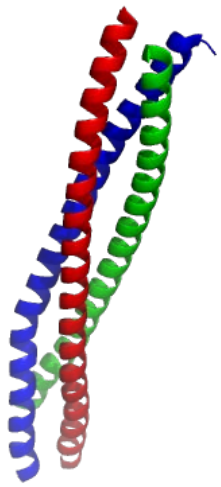


Figure 1.5 *E.coli* YbgF is a homotrimer. Cartoon representation of three *E.coli* YbgF subunits (residues 35-109) in trimeric state. PDB ID: 2XDJ (Krachler et al. 2010).

1.11 *E.coli* YbgC

YbgC is a cytoplasmically located thioesterase that is believed to be involved in phospholipid metabolism. It has been reported to have acyl-CoA thioesterase activity on malonyl-CoA (demonstrated *in vitro*), catalysing the hydrolysis of the thioester bond. It has also been reported to interact with a number of other proteins involved in phospholipid metabolism at the inner membrane, however, no interaction with other Tol family proteins has been reported (Gully et al. 2006, Krachler 2010).

1.12 The *E.coli* Tol-Pal complex

Tol-Pal is organised to span across the inner membrane, through periplasm, to the outer membrane (figure 1.6). Residing in the inner membrane, TolQ, TolR and TolA form the inner membrane complex. TolB and YbgF reside within the periplasm and TolB is able to bind the 3rd (C-terminal) domain of TolA. YbgF is reported to bind the 2nd (long triple helical) domain of TolA (Krachler et al. 2010). Although disputed (Bonsor et al. 2009), TolA's 3rd domain is reported to interact with Pal (Cascales et al. 2000), which although periplasmic, is also anchored in the inner leaflet of the outer membrane. Pal is also reported to bind TolB (Bonsor et al. 2009). Finally, the YbgC component of the Tol operon appears to be located in the cytoplasm (Gully et al. 2006), and thus may suggest a function that allows Tol to communicate from the cytoplasm via YbgC, through the inner membrane with TolQRA, across the periplasm to the outer membrane via TolA, TolB and Pal, however, to date there is no evidence to support this. No function for the Tol-Pal complex has been confirmed.

Work by Goemaere et al. has indicated that the transmembrane domains of TolQ and TolR may form a network of ionisable and hydophillic groups that promote the transit of protons through a pore channel contained within the

2nd and 3rd TMD's of TolQ (Goemaere et al. 2007). It has also been suggested that TolQ/R may regulate TolA's interaction with other proteins. While the role of pmf in Tol function is unresolved it is known to influence the interactions of TolA with TolB and Pal (Goemaere et al. 2007). There is some debate in the literature regarding both the localisation and organisation of Tol, as well as the interactions between the respective protein components. Work by Henry et al. (Henry et al 2004) suggest that the N-terminal domain of TolA is localised within the peptidoglycan layer, whereas Lazzaroni et al. (Lazzaroni et al. 2002) contend that both TolB and Pal is located with the PG layer. Due to the length of TolA domains and the report that TolA and Pal interact *in vivo* it is likely that TolA can span the periplasm, potentially interacting simultaneously at both the inner and outer membranes (Cascales et al. 2000).

Although the specific function of the Tol complex or the individual Tol proteins is not known, speculated Tol functions can be categorised into 3 main areas. Tol has been shown to be involved in the maintenance and stabilisation of the outer membrane of Gram negative bacteria and is involved in the import of both colicins and bacteriophages from the extracellular environment into the bacterial cell (for review see Cascales et al. 2007). Tol and related homologues can also be broadly categorised by their function; they are energy transducing proteins that can couple electrochemical gradients of the inner membrane to support energy dependant processes in the periplasm and outer membrane (Postle et al. 2007).

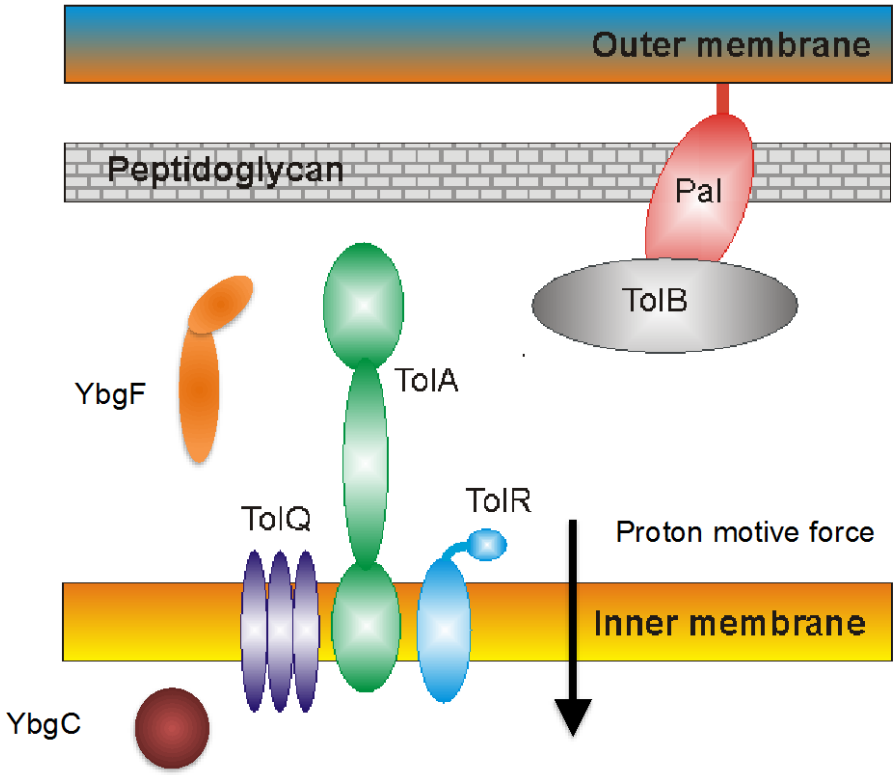


Figure 1.6 Schematic diagram of organisation of the *E.coli* Tol complex (not to scale). Pal is associated with peptidoglycan and forms a complex with TolB. YbgF is a homotrimer and resides in the periplasm. TolQ, TolR and TolA form a complex within the inner membrane and YbgC is located in the cytoplasm.

A *tol* phenotype describes the phenotype that arises when any of the *tol* genes (with the exception of *ybgf* or *ybgc*) are deleted. Cells display a similar phenotype of membrane instability. Specifically, cells are hypersensitive to drugs normally excluded by the outer membrane; hypersensitive to detergents such as SDS; they release of outer membrane vesicles (OMV's) in excess of normal release activity and release of proteins normally found within the periplasm, all of which indicates a "leaky" outer membrane (Deprez et al. 2002). In addition, *tol* minus *E.coli* cells have problems with cell division, specifically they cannot separate into mother and daughter cells, and thus form long chains of cells (Bernadac et al. 1998) (Cascales et al. 2002) (Gerding et al. 2007).

Released outer membrane vesicles have been reported to contain lipopolysaccharide, phospholipids from the outer membrane and periplasmic proteins. In addition, Henry et al. found that upon deletion of Tol complex, release of outer membrane vesicles was vastly increased, leading to the possibility for development of a novel technique for harvesting bacterially expressed, but non-secreted proteins of interest (Henry et al. 2004).

It has been suggested by some that the Tol system has a role in maintaining the structure of the bacterial cell membranes through an architectural function; forming a bridge network from the inner membrane, across the periplasm and onto the outer membrane (Henry et al. 2004; Cascales et al. 2007). The outer membrane maintenance function of Tol-Pal requires the proton motive force (pmf). (Lloubes et al. 2001; Cascales et al. 2007; Goemaere et al. 2007). Other potential functions have been inferred from a variety of evidence, including a role in cell division. GFP-Tol protein fusions locate to the cell division site, and as previously stated, *tol* mutant cells have cell division problems, such as the formation of long “chains” of cells. As Tol proteins may form some kind of transenvelope bridge from outer to inner membrane, it has been proposed that Tol proteins may act to pull the outer and inner membranes together during cell division, as part of the invagination process (Gerding et al. 2007) although this model relies on *E.coli* TolA and Pal forming a complex, something that has not been found *in vitro* (Bonsor et al. 2009). In addition, it has been proposed that Tol-Pal has a role in the assembly and localisation of lipopolysaccharide (Gaspar et al. 2000) and localisation and assembly of outer membrane proteins, in particular porins (Cascales et al. 2007).

In particular, lipopolysaccharide (LPS), a core component of the outer membrane, has been reported to be at lower levels in the outer leaflet of the outer membrane in *tol-pal* mutant strains of *E.coli* (Gaspar et al. 2000;

Lloubes et al. 2001; Vines et al. 2005; Gerding et al. 2007). It has also been reported that in *tolA* mutants, LPS biosynthesis is not reduced, rather a reduction was seen in the levels of LPS in the outer leaflet of the outer membrane, likely as a result of the LPS post-synthesis processing and transportation (Gaspar et al. 2000).

1.13 The *E.coli* Ton system

Due to the rarity of the essential molecules and elements such as vitamin B_{12} and iron in the extracellular medium, bacteria have evolved a high affinity transporter system to transfer them into the cell. Vitamin B_{12} is translocated across the outer membrane by BtuB, and receptors such as FhuA and FecA transport iron complexes. The outer membrane transporters, such as BtuB, consist of two domains; a 22-stranded beta-barrel, and a globular "plug" domain which is located on the cytoplasmic side of the protein and prevents loss of periplasmic contents (Andrews et al. 2003). In order for ligands to be internalised, energy is required, however there is no ATP in the periplasm. Thus, by use of proton motive force by the Ton system, the cell is able to transduce energy to the outer membrane to facilitate transport (Karlsson et al. 1993). The Ton system consists of three proteins; TonB, ExbB and ExbD. TonB has a single transmembrane domain, which forms an inner membrane complex with ExbB and ExbD, and has a long periplasmic domain (Karlsson et al. 1993) and a globular C-terminal domain, of which there are numerous solved structures (Weiner 2005). ExbB has 3 membrane spanning domains, and has both a large cytoplasmic loop, and a small periplasmic domain. ExbB is related to TolQ in that it shares 62% sequence homology and considerable structural similarity (Braun et al. 1993). ExbD has a single transmembrane domain, and a C-terminal periplasmic domain and shares 66% homology with TolR (Braun et al. 1993). The structure of the C-terminal domain as been solved by NMR (Garcia-Herrero et al. 2007).

The C-terminal domain of TonB stretches across the periplasm, and comes into contact a conserved TonB box epitope on the N-terminal “plug” domain of the receptor (Schauer et al. 2008). How exactly the energy of the pmf is transduced through TonB to import of Vitamin B_{12} or iron sideophores is not known, although it has been suggested that as ExbB and ExbD of Ton are related to MotA and MotB of the flagellal rotor motor, that some form of rotation in TonB occurs, pushing the plug out of the way and allowing import (Chang et al. 2001).

1.14 *E.coli* Tol complex homologues

The Tol-Pal and Ton protein families share some sequence identity and Tol QRA are structurally similar to TonB/ExbB/ExbD, and it is believed that TolA and TonB may share a common evolutionary ancestor (Witty et al. 2002). Despite this, mutations in *tonB/exbD/exbB* do not result in outer membrane instability, indicating a different specific function. Although TolA and TonB only share approximately 30% sequence identity (despite similar folds) (Witty et al. 2002), ExbB and ExbD share considerable sequence homology and have very similar transmembrane domains as TolQ and TolR. Both ExbB and ExbD have been shown to be capable of partially cross-complementing TolQ and TolR (Braun et al. 1993). It should be noted that despite their topological similarities TonB and TolA have not been found to be capable of cross complimenting one another (Braun et al. 1993). TolQ and TolR are also homologous with flagella motility proteins MotA and MotB. MotA and MotB are part of the flagella motor complex, and drive rotation of the flagella in a proton motive force dependant manner. TolQ and TolR have been shown to be influenced by pmf, indicating a potentially similar function to the MotA/B (Cascales et al. 2001). Both the Tol and Ton protein complexes can be subverted by different groups of bacteriophages and colicins to facilitate their entry into the bacterial cell (Bouveret et al. 2002; Lazzaroni et al. 2002; Cascales et al. 2007; Goemaere et al. 2007; Postle et al. 2007).

In addition to similarities with the Ton and Mot systems, the use by Tol of order-disorder signalling is not unique. It is estimated that up to 35% of prokaryotic proteins contain some significant degree of disorder (Tompa 2012). In addition to the intrinsically disordered N-terminus of colicin E9, used for binding and translocation of the colicin into the cell (Housden et al 2005, Bonsor et al. 2009), systems such as the DegS protease undergo order-disorder transition to go from an inactive to active state (Wilken et al. 2004). It has also been documented that the binding of a protein intrinsically disordered domains of another can occur at binding sites distal of one another. In a recent study it was shown that in the regulation of the bacterial phd/doc toxin-antitoxin operon involves the toxin protein Doc as co- or derepressor for PhD. A monomer of Doc binding Phd dimers in two unrelated and distal binding sites causes the intrinsically disordered C-terminal domain of Phd to structure its N-terminal DNA-binding domain illustrating allosteric coupling between disordered domains (Garcia-Pino et al. 2010).

1.15 Tol proteins of other Gram-negative bacteria

As previously stated, the Tol family of proteins are highly conserved throughout most Gram-negative bacteria, and although most work to date published on the Tol proteins have been also entirely focused on the *E.coli* Tol family. Tol proteins have been confirmed in 31 Gram-negative genera, with the proteins in this genera ranging from 10-100% sequence identity with *E.coli* (Deatherage et al. 2009).

The Tol operon of *Pseudomonas aeruginosa* is arranged in an identical order to that of *E.coli.* However, as little work has characterised *Pseudomonas aeruginosa* Tol proteins, some annotation is lacking of specific genes. Like *E.coli* Tol which is arranged in the order of YbgC-TolQ-TolR-TolA-TolB-Pal-YbgF, *Psuedomonas aeruginosa* Tol operon is arranged

as pa0968/orf1-tolq-tolr-tola-tolb-oprl-pa0974/orf2 (Dennis et al. 1996). PA0968 is of unknown function, however has homology with *E.coli* YbgC. OprL is a homologue of *E.coli* Pal (Lim et al. 1997), and PA0974 is also of unknown function, however is predicted to be a TPR repeat protein, and has homology with *E.coli* YbgF (Winsor et al. 2011).

The only structure available for any *Pseudomonas aeruginosa* Tol protein is a crystal structure published by Witty et al. in 2002 of *Pseudomonas aeruginosa* TolA consisting of domain 3 (C-terminal globular domain) with a short region of domain 2 (helical). When compared with *E.coli* TolA, both proteins share a near identical fold, although they only share approximately 20% sequence identity (Witty et al. 2002). No further information is known on either the interactions of *Pseudomonas aeruginosa* TolA, or it's function. Little is known regarding *Pseudomonas aeruginosa* TolB. Based on homology with *E.coli* TolB, psTolB is likely to be organised into 2 domains, including a C-terminal beta-propeller protein. psTolB shares approximately 45% sequence identity with eTolB. No further information is available for either the interactions of psTolB, it's function, or structure (Winsor et al 2011).

Little is known regarding other *Pseudomonas aeruginosa* Tol proteins including Pal/OprL, TolQ, TolR, YbgF/PA0974 or YbgC/PA0968 as no structural or function information on these proteins has been reported.

1.16 Bacteriocins and colicins

Colicins are protein based antimicrobial agents that are released by *E.coli* under stress conditions in order to kill competing organisms (For review see Cascales et al. 2007). Although colicin only refers to proteins released from *E.coli*, many other gram negative species are capable of releasing cytotoxic proteins or peptides, classified under the umbrella term "bacteriocin".

Colicins confer a competitive advantage to the producing culture from a plasmid pCol (Hardy et al. 1973). *E.coli* cells that contain a pCol plasmid are termed colicinogenic, and these plasmids are divided into 2 groups; type I and type II. Type I are small 6-10 kb plasmids, of which there are approximately 20 copies per cell, and mainly encode group A colicins. Type II are much larger plasmids, of approximately 40kb, and are usually present only as a single copy in bacteria. Type II plasmids mainly encode group B colicins (Cascales et al. 2007). Colicins are composed of 3 domains (figure 1.7); N-terminal translocation domain (T-domain), Receptor binding domain (R-domain) and C-terminal cytotoxic domain (may be either nuclease or pore forming domain which causes cell death). An additional protein, co-translated by the *E.coli* cell that produces the colicin, is called the Immunity protein, which binds with very high affinity to the cytotoxic domain to prevent suicide in the producing cell (Kleanthous & Walker 2001). Colicins can utilise a variety of outer membrane receptors to allow their entry into the target cell, including BtuB. OmpF, OmpA and Cir (Braun et al. 2002), and facilitate their entry across the periplasm from outer to inner membrane through either the Tol (for group A colicins) or Ton (for group B colicins) protein families. As mentioned above, Tol proteins are so named due to the Δ*tol* phenotype, whereby, in addition to other phenotypes (described below), cells were observed to be “tolerant to colicin” (Cascales et al. 2007). Although colicins are specific to their outer membrane receptors (for example, colicin E3 utilises BtuB for initial contact with cell through R-domain, and OmpF to translocate across the membrane), it is possible for these domains to be swapped between colicins to make them dependent on different receptors. Colicin Ia utilises Cir (Colicin I Receptor) as both it’s receptor and translocator. When the R-domain of colicin Ia is replaced with that of colicin E3 R-domain, the subsequent hybrid is dependant on both BtuB and Cir (Jakes et al. 2010).

Colicins kill off competing cells through several modes of action; some colicins are nucleases, whereby upon entry into the cytoplasm they destroy the target cells nucleic acids, both DNA and RNA (Cascales et al. 2007). Other colicins, including colicin A, colicin Ia, colicin E1, colicin N and colicin B form pores in the target cell's inner membranes, depolarising the inner membrane and destroying ionic gradients, such as the proton motive force that are vital for cell survival (Tilley et al. 2006). Colicins are grouped according to their method of entry into the cell. All group A colicins subvert the Tol-Pal system to facilitate entry. The second class, termed Group B colicins act in the same cytotoxic manner as group A colicins, although they parasitise the Ton-Exb family of proteins to facilitate translocation (Cascales et al. 2007).

In addition to a differing import system, group A and B colicins utilise different methods of export from the host cell. Group A colicins are released by autolysis through a co-translated lysis protein, causing death of the host cell. Group B colicins do not produce lysis proteins and thus are not released in this way (Toba et al. 1986). Lysis proteins are small lipoproteins of 27 to 35 amino acids (Wu et al. 1996), and although each colicin has it's own respective lysis protein, they are highly conserved between each colicin subtype. The lysis protein causes lysis through modifications of the structure of the cell envelope; activation of OmpLA (outer membrane phospholipase A) and ultimately death of the producing cell. This may well be an altruistic event, by which the host cell dies for the good of the colony. The cells become lysed through activation of OmpLA which promotes the formation of lysophospholipids (Pugsley & Schwartz. 1984). Lysophospholipids are detergents and permeabilise the outer membrane and, subsequently, the inner membrane of the cells. However, it should be noted that both colicin A release and lysis occur in the absence of an active OmpLA (Cascales et al. 2007).

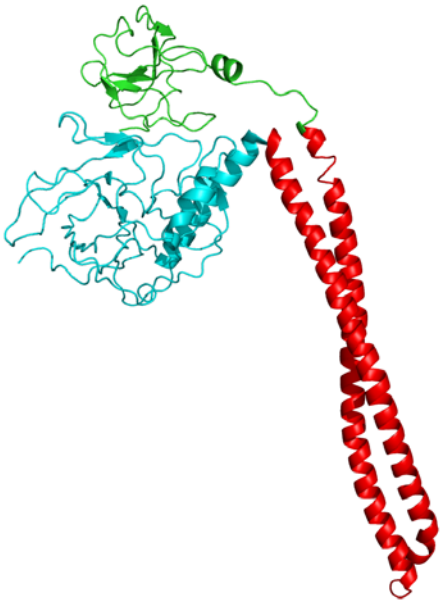


Figure 1.7 Three dimensional representation of colicin E3 secondary structure and domain organisation. Red helices represent Receptor binding domain, green presents cytotoxic RNase domain, cyan represents translocation domain (T-domain). N-terminus of T-domain is disordered, and divided into 3 distinct regions, 2 OmpF binding sites and a TolB binding box. PDB ID: 1JCH (Soelaiman et al. 2001). Figure draw with pymol 0.99a (Delano Scientific).

1.17 Cells with impaired Tol-Pal system are tolerant to colicin's

The Tol proteins are so named after *in vivo* observations that when *tol* genes (with the exception of *pal*) were deleted, bacterial cells became tolerant to colicin. Both colicins and bacteriophages subvert and parasitise the Tol system and outer membrane porins to facilitate their entry into the cell. Other proteins formerly had the name Tol, and have since been renamed, but all share a common link (like the Tol complex itself) that they confer tolerance to colicin when knocked out (Cascales et al. 2007). However, *Δtol* cells, if

placed in environmentally stressful conditions, such as osmotic shock, where their outer membrane instability is exacerbated, upon being challenged with colicin, were still killed. This indicated that the Tol proteins must have a role solely in allowing the translocation of colicin's across the outer membrane, and that once in the periplasm do not have a role in translocating the colicin through the periplasm and across the inner membrane (Tilby et al. 1978).

1.18 Pyocins

Pyocins are a type of bacteriocin produced by over 90% of *Pseudomonas* strains of bacteria during environmental stress/competition conditions. Type R and F pyocins are similar in structure and function to bacteriophages, whereas type S are evolutionarily closer to colicins. It has been reported that pyocin S1, S2 and AP41 may share a common ancestor with colicin E2, including a C-terminal cytotoxic DNase domain. Like colicins pyocins have pore forming, DNase or RNase killing activity. It has been reported that pyocin AP41 is dependent on the *Pseudomonas aeruginosa* Tol protein family for translocation, and deletion of Tol proteins yields the "tolerant to colicin" phenotype (Michel-Briand et al. 2002).

1.19 Colicin E9 (ColE9)

Colicin E9 is a member of the group A family of colicin's. When produced by the cell, ColE9 co-translated with its immunity protein called Im9. Im9 binds the cytotoxic domain of the colicin to prevent suicide in the cell that produced the colicin. Im9 dissociates from the colicin upon binding to target cell. ColE9 enters the cell (figure 1.8) by firstly binding on the outer membrane to BtuB with its R-domain. The translocation domain of ColE9 then recruits an OmpF trimer by threading it's N-terminal intrinsically unstructured N-terminus through one of the 3 OmpF pores, binding OmpF via one of 2 OmpF binding sites (OBS1, OBS2) (Housden et al. 2005). This disordered N-terminus then

recruits and binds TolB via a TolB binding epitope (TBE), which in turn promotes the interaction of TolA and TolB by driving the N-terminus of TolB into a disordered state, making the epitope available to bind TolA (Bonsor et al. 2009). It is this process of binding BtuB and OmpF on the outer membrane, and recruiting both TolB and TolA in the periplasm that somehow drives the uptake of the colicin into the periplasm (Housden et al. 2005). There are differences in the interactions of colicin E9, colicin N and colicin A with the Tol system, despite all 3 colicin's being dependant on Tol. Colicin N interacts solely with the 3rd domain of TolA, and does not interact with TolB at all. In addition, colicin N appears to bind TolA domain 3 in a similar binding interface as the bacteriophage g3p. Colicin A conversely interacts with both TolA domain 3 and TolB separately and does not require TolB to interact with TolA, however it requires both proteins in order to translocate across the outer membrane (Hecht et al. 2010). It should be noted that some colicins appear to mimic the interaction of Pal with TolB in order to recruit TolB to an outer membrane pore and thus facilitate the formation of a translocon, and ultimately allow import of the colicin cytotoxic domain (Bonsor et al. 2007; Cascales et al. 2007; Bonsor et al. 2009).

Although the specific details of how the colicin translocates across the outer membrane is not known, some detail is known about how colicin E9 translocates across the inner membrane. Whereas once pore forming colicin's reach the inner membrane their journey is over, as they create a pore in the membrane to de-couple the inner membrane, nuclease colicins such as E9 must also translocate across the inner membrane. It is thought that following some form of proteolytic processing in the periplasm, the colicin may be translocated through the inner membrane AAA+ ATPase (part of inner membrane secretion system) and FtsH protease via a direct interaction with the inner membrane (Mosbahi et al. 2002; Kleanthous 2010).

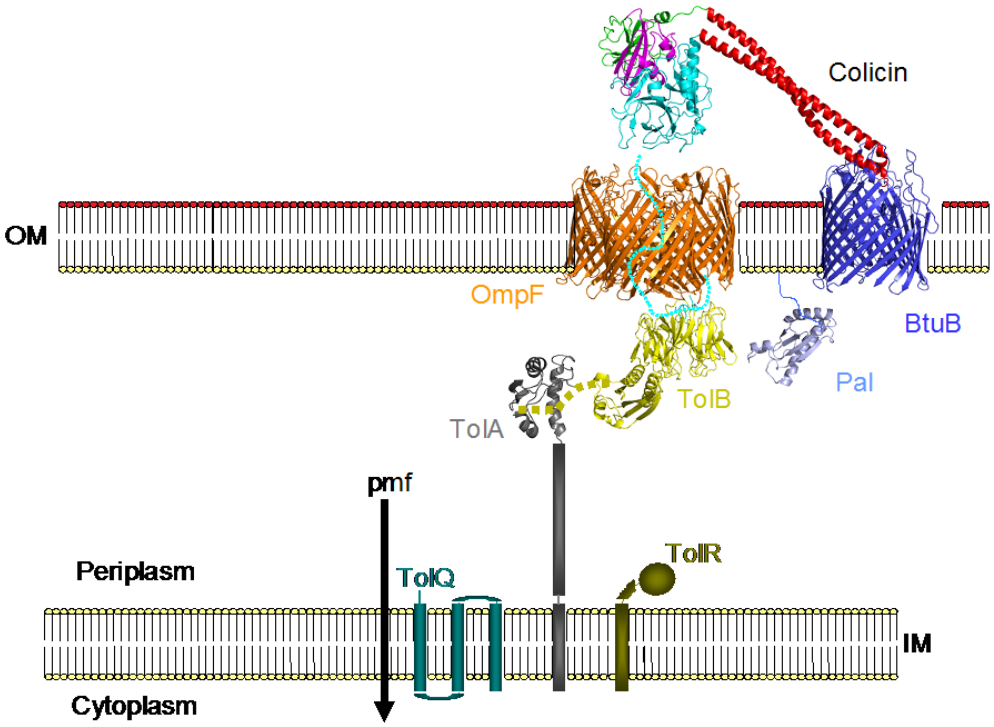


Figure 1.8 Assembly of colicin E3 bound to outer membrane porins and TolB, prior to outer membrane translocation. There are 3 stages for group A colicin entry into the cell; firstly the colicin's receptor binding domain (shown in red) binds a specific receptor, such as BtuB in the case of group A nuclease colicins. Following this the colicin's N-terminal translocation domain (shown in green) translocates through the outer membrane porin, OmpF. Finally the C-terminal cytotoxic domain (shown in cyan), is somehow translocated across the porin into the periplasm. It is unknown if the N-terminal translocation domain remains bound at the outer membrane, or translocates across with the C-terminal domain. Tol proteins are also subverted in a similar way by bacteriophages to facilitate their entry into the cell (Cascales et al. 2007). Figure used and modified with the permission of Dr Nicholas Housden, University of York.

1.20 Aims

Although the Tol system of Gram-negative bacteria has been studied for several decades, a number of key questions have yet to be answered. While it is known that the N-terminus of *E.coli* TolB is involved in it's interaction with the C-terminal domain of *E.coli* TolA, it is unclear if this is the sole site of interaction, as well as the location of the binding site on TolA is unknown. It is also unknown if the interaction between TolA and TolB is conserved throughout Gram-negative bacteria. Through a variety of biophysical techniques this work will seek to confirm the importance of the *E.coli* TolB N-terminus, and determine if it is indeed the sole site of interaction with *E.coli* TolA. This work will also determine if the TolA-TolB interaction is present in *Pseudomonas aeruginosa,* and if this interaction is dependent on the TolB N-terminus. Finally, it is the intention of this work to determine the binding site of TolB on TolA, the structure of which, to date, is unknown.

2. Materials and methods

2.1 Molecular Biology

2.1.1 Media and bacterial strains

E. coli cultures were either grown in M9 minimal media, lysogeny broth (LB) or plated on LB-agar. LB contained 10 g tryptone, 5 g yeast extract and 10 g NaCl per 1000 ml of medium. M9 media contained 4.8 mM Na_2HPO_4, 2.2 mM KH_2PO_4, 0.7 mM NaCl, 0.2 mM Na_2SO_4, 1.9 mM NH_4Cl, 2 mM $MgCl_2$, 0.1 mM $CaCl_2$ and 22 mM D-glucose. When plated, 1.5% (w/v) final concentration of agar was added to media. Antibiotic selection was performed by use of ampicillin (amp) (100 µg/ml), carbenicillin (car) (100 µg/ml), chloramphenicol (chl) (34 µg/ml) and/or kanamycin (kan) (50 µg/ml) where indicated (Melford). The bacterial strains used in this work are listed in table 2.1

Table 2.1 Bacterial Strains

DH5α (F⁻ φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(rk^-, mk^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-)	Competent cells used for transformation of plasmids for propagation and after ligation.
BL21(DE3) (F^-, *omp*T, *hsd*S_B (r_B^-, m_B^-), *dcm*, *gal*, λ(DE3)).	Competent cells used for induced gene expression from recombinant plasmids under the regulation of T7 promoter. Gene expression controlled

	by isopropyl-β-D-thiogalactopyranoside (IPTG).
Rosetta 2 (BL21 (DE3)) F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE2 (Chl^R)	Competent cells used for expression of genes containing rare codons (E.g. *Pseudomonas aeruginosa* proteins) as contained pRARE2 plasmid, which encodes rare tRNA's for AGA, AGG, AUA, CUA, GGA, CCC, and CGG. Expression of recombinant plasmid under regulation of T7 promoter, controlled by IPTG.

2.1.2 Vectors and plasmid isolation

pET11c, pET15b and pET21d were purchased from Novagen/Merck (Darmstadt). pMA-T was purchased as part of synthetic gene production from Geneart (Regensburg). Plasmids were prepared and purified using QIAprep Spin Miniprep or QIAgen Plasmid Midi kits (Qiagen). The plasmids used and generated in this work are listed and described in appendix table 7.1.

2.1.3 Preparation of competent cells

Competent cells were prepared from 35 ml *E. coli* cultures that had reached an OD_{600nm} of ~0.3-0.6. Cells were harvested by centrifugation (5000g, 15mins 4 °C) washed in 20 ml of ice-cold 20 mM Tris-HCl, 50 mM $CaCl_2$, pH 8.0 and left on ice for 1 hour. Cells were harvested again by centrifugation and resuspended in 2 ml of ice-cold 20 mM Tris-HCl, 50 mM $CaCl_2$, 20% w/v glycerol pH 8.0. Aliquots of 200 µl were made and stored at -80 °C.

2.1.4 Transformation of competent cells

Aliquot of prepared competent cells were thawed on ice before addition of 100 ng of plasmid DNA. Cells were incubated on ice for 30 minutes, heat shocked at 42 °C for 45-90 seconds and incubated for a further 5 minutes on ice. Cells were then incubated at 37 °C for 1 hour after the addition of 800 µl of LB. Cells were plated onto LB-agar (containing appropriate antibiotic(s)) and incubated overnight at 37 °C.

2.1.5 Genomic DNA

Genomic DNA from *Pseudomonas aeruginosa* (PA-01 strain) was kindly provided by Prof. Ben Luisi, University of Cambridge. Genomic DNA from *Xanthomonas campesteris* was kindly gifted by Dr. Max Dow, University College Cork. Genomic DNA from *E.coli* (K12 strain) was kindly gifted by Dr. Anne-Marie Krachler, formerly University of York.

2.1.6 DNA primers

DNA primers were synthesised by Eurofins MWG Operon. See appendix table 7.1 for full list.

2.1.7 Polymerase chain reaction (PCR)

Amplification of the genes and gene fragments of interest was performed by the polymerase chain reaction (PCR) using genomic DNA as a template. Primers were designed to contain restriction sites flanking the region of interest. PCR was performed with an Eppendorf Mastercycler Personal Thermal Cycler in a total volume of 50 µl. Typical PCR reaction mixture and cycling parameters are listed in table 2.2 and 2.3 respectively.

Polymerase used was either Pfu Turbo (Stratagene) or Pfu Ultra II (Aligent) and dNTP's purchased from Invitrogen.

Component	**Volume/µl**
Autoclaved Milli-Q water	40
10× *Pfu Turbo/Ultra II* Polymerase Reaction Buffer	5
dNTPs (10 mM stock)	1
Plasmid template (50 ng/µl)	1
Forward primer (125 ng/µl)	1
Reverse primer (125 ng/µl)	1
Pfu Turbo/Ultra II Polymerase (2.5 U/µl)	1
Total volume	**50**

Table 2.2 Typical PCR mixture

No. of cycles.	**Temperature**	**Duration**
1	95°C	2 mins
30	95°C	30 secs
	Primer T_M - 5°C (typically 55°C)	30 secs
	72°C	x* min per Kb
1	72°C	10 mins

***** Extension time of 2 min/kb for *Pfu Turbo*, 1min/kb for *Ultra II*

Table 2.3 Typical PCR cycling parameters

2.1.8 Custom gene synthesis

Recombinant genes for colicin E9-TolB fusion proteins (plasmids pEC7 and pEC8) and *Pseudomonas aeruginosa* TolB (plasmid pEC9) were synthetically made by Geneart AG (Invitrogen), Regensburg, Germany.

2.1.9 DNA Restriction digests

Restriction digests of plasmids and PCR products were performed using enzymes and buffers purchased from New England Biosciences (NEB). Using appropriate buffer and BSA if required, reactions were performed in total volume of either 20 µL or 50 µL. Typically for a 50 µl digestion 1-2 µg DNA was digested with 10-20 units of each restriction enzyme with 0.1mg/ml BSA for 2 hours at 37 °C. Linearised vectors and digested PCR products were purified by Agarose gel electrophoresis followed by gel extraction using the QIAquick Gel Extraction Kit (Qiagen).

2.1.10 DNA Ligations

Ligation of double stranded DNA was performed using T4 DNA ligase (NEB). An excess of PCR product was mixed with linear vector and incubated for 1-2 hours at room temperature before transformation in competent DH5α cells. Typical ligation mix is listed in table 2.4.

Component	**Volume/µL**
Linear vector DNA (100ng/µL)	3
Insert DNA (3x excess) (100ng/µL)	9
10x T4 ligation buffer	2
Autoclaved MilliQ water	5.5
T4 Ligase (50U/µL)	0.5
Total volume	**20**

Table 2.4 Typical DNA ligation

2.1.11 Agarose Gel-electrophoresis

Electrophoresis grade Agarose (Invitrogen) was dissolved in Tris-Borate-EDTA-buffer (TBE) to make up appropriate percentage gel, typically 2% for small PCR products (100-800bp) and 0.5% for plasmids. 5 µl of SYBR Safe (10000× concentrate) (Invitrogen) was added per 50 ml of gel.

Gels were run in TBE at a constant voltage of 80 V until sufficient resolution had been achieved.

2.1.12 Whole plasmid mutagenesis

Whole plasmid mutagenesis was used to introduce either point mutations or in-frame deletions into existing plasmids. PCR reactions were assembled as described in table 2.5. and run according to the program in table 2.6. Mutagenesis primers were designed to contain the mismatch in the centre of the sequence and to have a melting temperature greater than 60 °C. Melting temperature was calculated using the following formula:

$T_m = 81.5 + (0.41 \times \%GC) – (675/N)$

Where %GC is the percentage of guanine and cytosine and N the number of bases in the primer sequence.

10 µl of PCR product were analysed on 0.7% (w/v) agarose gel to check for amplification of the plasmid relative to the negative control. For plasmids that were successfully amplified, 1 µl of DpnI (20U/µl, NEB) was added to the reactions and incubated at 37 °C for 2 hours. 5 µl of DpnI treated PCR product was transformed into competent DH5α cells.

Component	**Volume/µl**
Autoclaved Milli-Q water	40
10× *Pfu Turbo/Ultra II* Polymerase Reaction Buffer	5
dNTPs (10 mM stock)	1
Plasmid template (50 ng/µl)	1
Forward primer (125 ng/µl)	1
Reverse primer (125 ng/µl)	1
Pfu Turbo/Ultra II Polymerase (2.5 U/µl)	1
Total volume	**50**

Table 2.5 A typical PCR reaction mixture

No. of cycles.	Temperature	Duration
1	95°C	30 secs
16	95°C	30 secs
	Primer T_M - 5°C (typically 55°C)	30 secs
	67.5°C	x* min per kb of plasmid
1	72°C	10 mins

***** Extension time of 2 min/kb for *Pfu Turbo*, 1min/kb for *Ultra II*

Table 2.6. PCR cycling parameters for whole plasmid mutagenesis.

2.2 Sub-cloning

2.2.1 Cloning of tagless *Pseudomonas aeruginosa* TolA domain 3 (pEC1)

The region encoding TolA domain 3 (identified by homology with *E.coli* TolA domain 3, residues 226-347) was PCR amplified from *Pseudomonas aeruginosa* genomic DNA (PA-01 strain) using appropriate primers (see appendix table 7.2 for primers). The PCR amplified fragment was then purified by gel electrophoresis/extraction, digested with BamHI and NdeI, and subsequently ligated into pre-digested pET11c, to make pEC1.

2.2.2 Cloning of tagless *Xanthomonas campesteris* TolA domain 3 (pEC3)

The region encoding TolA domain 3 (based on homology with *E.coli and Pseudomonas aeruginosa,* residues 224 - 345) was PCR amplified from

Xanthomonas campesteris genomic DNA (8004 strain) using appropriate primers (see appendix table 7.2 for primers). The PCR amplified fragment was then purified by gel electrophoresis/extraction, digested with BamHI and NdeI, and subsequently ligated into pre-digested pET11c, to make pEC3.

2.2.3 Cloning of non-cleavable his-tagged *Pseudomonas aeruginosa* TolA domain 3 (pEC4)

The region encoding TolA domain 3 (residues 226-347) was PCR amplified from *Pseudomonas aeruginosa* genomic DNA (PA-01 strain) using appropriate primers (see appendix table 7.2 for primers). The PCR amplified fragment was then purified by gel electrophoresis/extraction, digested with BamHI and NdeI, and subsequently ligated into pre-digested pET21d (to introduce C-terminal 6xHis tag), to make pEC4.

2.2.4 Cloning of colicin E9-*E.coli* TolB fusion protein (TolA replacing TolB binding epitope, pEC7)

DNA fragment of the region encoding the N-terminus of colicin E9 (residues 1-83, with the addition of KpnI site at the 5' end and XhoI site at 3' end, encoding for OmpF binding site 1), TolA binding site from disordered N-terminus of TolB (EVRIVIDSGVDS) in place of TolB binding site, followed by OmpF binding site 2 was synthesised by Geneart AG. This fragment was subsequently excised from Geneart delivery plasmid (pMAT-EC7) with KpnI and XhoI and ligated into pre-digested pCS4 (mutated through whole plasmid mutagenesis to introduce KpnI site). This ligation yielded pEC7, which encoded for full length colicin E9 (with respective Immunity protein, Im9), with TolA binding epitope from TolB in place of TolB binding epitope. In all other respects, this colicin E9 mutant was as wild type.

2.2.5 Cloning of colicin E9-*E.coli* TolB fusion protein (N-terminal TolA binding epitope, pEC8)

DNA fragment of the region encoding the N-terminus of colicin E9 (residues 1-83, with the addition of KpnI site at the 5' end and XhoI site at 3' end), to encode for TolA binding site from disordered N-terminus of TolB (EVRIVIDSGVDS) in place of OmpF binding site 1, TolB binding site with 3 key residues mutated to abolish TolB binding (D35A, S37A, W39A), followed by OmpF binding site 2 was synthesised by Geneart AG. The DNA fragment was subsequently excised from Geneart delivery plasmid (pMAT-EC8) with KpnI and XhoI and ligated into pre-digested pCS4 (mutated through whole plasmid mutagenesis to introduce KpnI site). This ligation yielded pEC8, which encoded for full length colicin E9 (with respective Immunity protein, Im9), with TolA binding epitope from TolB in place of TolB binding epitope. In all other respects, this colicin E9 mutant was as wild type.

2.2.6 Cloning of colicin E9-*E.coli* TolB fusion protein (N-terminal TolA binding epitope, TolB binding epitope present, pEC12)

A mutant based on pEC8, where the TolB binding epitope was mutated back to wild type by whole plasmid mutagenesis to reintroduce TolB binding to this protein. In all other respects plasmid is as pEC8.

2.2.7 Cloning of his-tagged *Pseudomonas aeruginosa* TolB (pEC14)

The region corresponding to *Pseudomonas aeruginosa* TolB (residues 22 – 432, with the addition of BamHI site at the 5' end and XhoI site at 3' end) was synthesised by Geneart AG. The DNA fragment was subsequently excised from Geneart delivery plasmid (pMAT-psTolB) with BamHI and XhoI and ligated into pre-digested pET21d. This ligation

introduced codons for a C-terminal 6xHis tag onto the TolB gene, to make pEC14.

2.2.8 Cloning of *Pseudomonas aeruginosa* Peptidoglycan associated lipoprotein (pEC15)

The region corresponding to periplasmic soluble domain of Pal (residues 60-168, mature protein predicted by bioinformatic analysis, see appendix section 7.2 for details) was PCR amplified from *Pseudomonas aeruginosa* genomic DNA (PA-01 strain) using appropriate primers. The PCR amplified fragment was purified by gel electrophoresis and subsequently digested with NcoI and XhoI, and subsequently T4 ligated into pET21d to make pEC15.

2.2.9 DNA sequencing

Fidelity of all DNA constructs were verified by T7 forward and reverse primer sequencing, performed by Beckman Coulter (Takely, UK).

2.3 Protein Purification

2.3.1 Expression and purification of tagless *E.coli* TolA domain 3 (pAK108)

The plasmid pAK108 (supplied by Anne-Marie Krachler, Bonsor et al. 2009) was transformed into BL21(DE3) pLysS competent cells for expression and purification of TolA (domain 3, residues 293-421). A 1:1000 dilution of a 50 ml overnight from a single colony was used to inoculate (typically) 6 flasks (0.8 L per flask, giving a total of 4.8 L). Cultures were grown at 37 °C on a Innova 2300 platform shaker (New Brunswick Scientific) at 120 rpm until an OD_{600nm} of ~ 0.6 was reached. Gene expression was

induced with IPTG (Melford) (Figure 2.1) at a final concentration of 1 mM. Cells continued to grow on the orbital shaker at 37 °C for a further 4 hours before harvesting (SLC-6000 rotor, 4500 rpm, 4 °C for 12 minutes). Cell pellets were resuspended in 40 ml (total) 50 mM Tris-HCl pH 8.0, PMSF (1 mM), DNase-I (40 µg/ml) magnesium chloride (5 mM) and lysozyme (60 mg) were added. Cells were sonicated on ice using a S-4000 sonicator with ½ inch stud horn titanium probe (Misonix) with 60 x 3 seconds pulses at 70 W (a 7 second "off" spacing between pulses). The cell debris was pelleted (19776 rotor, 10000 rpm, 4 °C for 30 minutes) and supernatant decanted. The supernatant was then subjected to two ammonium sulphate ($(NH_4)_2 SO_4$) fractionations at 4°C; a 40 % saturation was achieved by the slow addition of $(NH_4)_2 SO_4$ (22.6 g/100ml of sample) over 10 mins, whilst being gently stirred. This was left to equilibrate for 1 hour at 4 °C before centrifugation to remove precipitated proteins (19776 rotor, 10000 rpm, 4 °C for 30 minutes). The supernatant was decanted and a second $(NH_4)_2 SO_4$ cut was performed (70 % saturation) by the addition of 18.7 g/100 ml of supernatant. This was left for a further hour at 4 °C before centrifugation (as before). The supernatant was removed and precipitated proteins were re-dissolved in ~20 ml of 50 mM Tris-HCl pH 8.0 and subsequently dialysed in the same buffer overnight at 4 °C to remove excess ammonium sulphate.

The protein solution was loaded onto a 10 ml pre-equilibrated (50 mM Tris-HCl, pH 8.0) DE52 anion exchange column (Whatman Resin). The unbound flow-through was collected before the elution of bound proteins with application of a 50 mM Tris-HCl pH 8.0, 1 M NaCl, gradient. Flow-through fractions were pooled and dialysed against 50 mM Tris-HCl, 250 mM NaCl, pH 7.5 overnight at 4 °C for size exclusion chromatography.

TolA was further purified by gel filtration using a Hiload 26/60 Superdex 75 column (GE Healthcare). Ten ml of protein were injected onto the pre-equilibrated S75 column (50 mM Tris-HCl, 250 mM NaCl, pH 7.5) and eluted at a flow rate of 3 ml/min. Protein elution was monitored by absorbance at 280 nm. All peaks were analysed by SDS PAGE to verify presence of

protein. Peak fractions were pooled and dialysed against 50 mM Sodium Acetate, pH 5 overnight at 4 °C.

Following dialysis, sample was loaded onto a pre-equilibrated 5/50 Mono S cation exchange column (GE Healthcare), and once unbound material had been collected in the flow-through, a NaCl gradient (1M) was applied over 10 column volumes. Fractions were analysed by SDS PAGE, peak fractions pooled, and once protein concentration had been estimated by absorbtion spectrophotometry, samples were divided into 1ml aliquots and stored at -20°C.

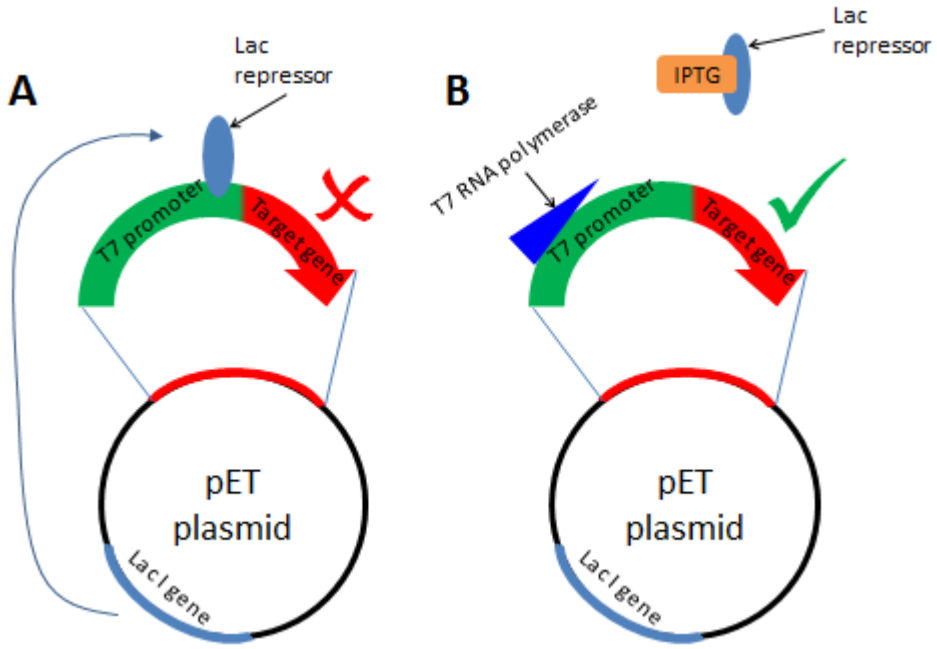


Figure 2.1 IPTG induction for gene expression Under normal conditions (A), the Lac repressor protein is translated from the constituitively expressed LacI gene. This Lac repressor protein binds to the T7 promotor region of the target gene preventing expression of the target gene. Upon addition of IPTG (B), the IPTG molecule binds to the Lac repressor protein, causing it to dissociate from the promotor region, allowing the T7 RNA polymerase to bind the promotor and translate the mRNA of the target gene (Bell et al. 2000).

2.3.2 Expression and purification of tagless, cytoplasmic *E.coli* TolB (pDAB18)

pDAB18 was transformed into BL21(DE3) pLysS competent cells for expression and purification of *E.coli* TolB (residues 22-430, with the addition of single N-terminal Methionine residue, Clone supplied by Dr Daniel Bonsor). A 1:1000 dilution of a 50 ml overnight from a single colony was used to inoculate (typically) 6 flasks (0.8 L per flask, giving a total of 4.8 L). Cultures were grown at 37 °C on a Innova 2300 platform shaker (New Brunswick Scientific) at 120 rpm until an OD_{600nm} of ~ 0.6 was reached. Protein expression was induced with IPTG (Melford) at a final concentration of 1 mM. Cells continued to grow on the orbital shaker at 37 °C for a further 4 hours before harvesting (SLC-6000 rotor, 4500 rpm, 4 °C for 12 minutes). Cell pellets were resuspended in 40 ml of 50 mM Tris-HCl pH 8.0 and PMSF (1mM), DNase-I (40 µg/ml) magnesium chloride (5 mM) and lysozyme (60 mg) were added. Cells were sonicated on ice using a S-4000 sonicator with ½ inch stud horn titanium probe (Misonix) with 60x3 seconds pulses at 70 W (a 7 second “off” spacing between pulses). The cell debris was spun out (19776 rotor, 10000rpm, 4 °C for 30 minutes) and supernatant decanted. The supernatant was then subjected to an ammonium sulphate ($(NH_4)_2 SO_4$) fractionation at 4°C; a 40 % saturation was achieved by the slow addition of $(NH_4)_2 SO_4$ (22.6 g/100ml of sample) over 10 mins, whilst being gently stirred. This was left to equilibrate for 1 hour at 4 °C before centrifugation to remove precipitated proteins (19776 rotor, 10000 rpm, 4 °C for 30 minutes). The supernatant was removed and precipitated proteins were re-dissolved in ~20 ml of 50 mM Tris-HCl pH 8.0 and subsequently dialysed in the same buffer overnight at 4 °C to remove excess ammonium sulphate.

The protein solution was loaded onto a 10 ml pre-equilibrated (50m M Tris-HCl, pH 8.0) DE52 anion exchange column (Whatman Resin). The unbound flow-through was collected before elution of bound proteins with application of 50 mM Tris-HCl, 1 M NaCl, pH 8.0 gradient. Flow-through was

pooled and dialysed against 50 mM Tris-HCl, 250 mM NaCl, pH 7.5 overnight at 4 °C for size exclusion chromatography.

TolB was further purified by gel filtration using a Hiload 26/60 Superdex 75 column (GE Healthcare). 10 ml of protein was injected onto the pre-equilibrated S75 column (50 mM Tris-HCl, 250 mM NaCl, pH 7.5) and was eluted at a flow rate of 3 ml/min. Protein elution was monitored by absorbance at 280 nm. All peaks were analysed by SDS PAGE to verify presence of protein. Peak fractions were pooled, and once protein concentration had been estimated by absorbtion spectrophotometry, samples were divided into 1ml aliquots and stored at -20°C.

2.3.3 Expression and purification of tagless *Pseudomonas aeruginosa* TolA domain 3 (pEC1)

pEC1 was transformed into BL21 (DE3) competent cells for expression and purification of tagless *Pseudomonas aeruginosa* TolA3. A 1:1000 dilution of a 50 ml overnight from a single colony was used to inoculate (typically) 6 flasks (0.8L per flask, giving a total of 4.8L). Cultures were grown at 37 °C on Innova 2300 platform shaker (New Brunswick Scientific) at 120 rpm until an OD_{600nm} of ~ 0.6 was reached. Gene expression was induced with IPTG (Melford) at a final concentration of 1 mM. Cells continued to grow on the orbital shaker at 37 °C for a further 4 hours before harvesting (SLC-6000 rotor, 4500 rpm, 4 °C for 12 minutes). Cell pellets were resuspended in ~30 ml 50 mM Tris-HCl pH 8.0. PMSF (1mM), DNase-I (40 µg/ml) magnesium chloride (5 mM) and lysozyme (60 mg) were added. Cells were sonicated on ice using a S-4000 sonicator with ½ inch stud horn titanium probe (Misonix) with 60x3 seconds pulses at 70 W (a 7 second “off” spacing between pulses). The cell debris was spun out (19776 rotor, 10000 rpm, 4 °C for 30 minutes) and supernatant decanted.

The supernatant was subsequently loaded onto a 10 ml pre-equilibrated (50m M Tris-HCl, pH 8.0) DE52 anion exchange column (Whatman Resin).

The unbound flow-through was collected before elution of bound proteins with application of 50 mM Tris-HCl, 1 M NaCl, pH 8.0 gradient. Flow-through was pooled and dialysed against 50 mM Tris-HCl pH 9, overnight at 4 °C. This sample was then loaded onto a pre-equilibrated MonoQ (strong anion exchange) column (GE Healthcare). Once unbound material had eluted, and 280 nm absorbance had returned to baseline, a 0-150mM NaCl gradient was applied over 30 column volumes. Eluted fractions were analysed on 16% SDS-PAGE gels. Peak fractions were pooled, and dialysed against gel filtration buffer (50 mM Tris-HCl, 250 mM NaCl, pH 7.5) overnight at 4 °C.

This sample was further purified by gel filtration using a Hiload 26/60 Superdex 75 column (GE Healthcare). 10 ml of protein was injected onto the pre-equilibrated S75 column and was eluted at a flow rate of 3 ml/min. Protein elution was monitored by absorbance at 280 nm. All peaks were analysed by SDS PAGE. Peak fractions were pooled and dialysed against 50 mM Tris pH 9 overnight at 4 °C for second MonoQ purification step. This protein sample was then loaded onto a pre-equilibrated MonoQ (strong anion exchange) column. Once unbound material had eluted, and 280nm absorbance had returned to baseline, a 0-150 mM NaCl gradient was applied over 30 column volumes. Eluted fractions were analysed on 16% SDS-PAGE gels. Fractions containing pure protein of interest were pooled and stored at -20 °C until further use.

2.3.4 Expression and purification of non-cleavable his-tagged *Pseudomonas aeruginosa* TolA domain 3 (pEC4)

pEC4 was transformed into BL21 (DE3) competent cells for expression and purification of his-tagged *Pseudomonas aeruginosa* TolA3. Growth and harvesting was as pEC1 (2.3.3) with the exception that pelleted cells were resuspended in 30 ml 1x binding buffer (40 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM imidazole) prior to sonication.

The cell debris was spun out (19776 rotor, 10000rpm, 4 °C for 30 minutes) and supernatant decanted. The supernatant was subsequently loaded onto a 10 ml column containing HisBind resin (Novagen) charged with 3 column volumes (c.v.) of 50 mM $NiSO_4$ and equilibrated with 3 c.v. of binding buffer. The column was washed with binding buffer at a flow rate of 1.5 ml/min until the absorbance of the eluate returned to background levels (monitored by 280 nm absorbance). Bound protein was then eluted using a linear gradient of 5-500 mM imidazole over 10 c.v. Fractions were analyzed by SDS-PAGE and those containing the protein of interest were pooled and dialyzed against gel-filtration buffer (50 mM Tris-HCl pH 7.5, 250 mM NaCl) overnight at 4 °C.

Protein was further purified by gel filtration using a Hiload 26/60 Superdex 75 column (GE Healthcare). 10 ml of protein was injected onto the pre-equilibrated S75 column (50 mM Tris-HCl, 250 mM NaCl, pH 7.5) and was eluted at a flow rate of 3 ml/min. Protein elution was monitored by absorbance at 280 nm. All peaks were analysed by SDS PAGE and peak fractions were pooled and stored at -20 °C until further use.

2.3.5 Expression and purification of colicin E9-*E.coli* TolB fusion protein (pEC7)

pEC7 was transformed into BL21 (DE3) cells and subsequently expressed and purified as for pEC4, with the following modification; instead of eluting bound protein from HisBind resin with imidazole elution buffer, a step elution (100%) with 6 M Guanidinium chloride (GnHCl) was applied. Colicin E9 mutants are co-translated with their high affinity binding partner Im9, which contains a 6xHis-tag. This complex is translated together, and the colicin will remain bound to the immunity protein when it itself is bound to the Nickel-NTA resin. 6 M GnHCl is applied to unfold the colicin, so that it is

no longer bound to the immunity protein (which remains bound to the HisBind resin) and collected in eluted fractions. The eluted fractions of interest were extensively dialysed against gel filtration buffer (50 mM Tris-HCl pH 7.5, 250 mM NaCl, 3x buffer changes) prior to gel filtration on pre-equilibrated Hiload 26/60 Superdex 200 column (GE Healthcare). Subsequent steps are as pEC4.

2.3.6 Expression and purification of colicin E9-*E.coli* TolB fusion proteins (pEC8 and pEC12)

Expression and purification of pEC8 and pEC12 was identical to that of pEC7.

2.3.7 Expression and purification of his-tagged *Pseudomonas aeruginosa* TolB (pEC14).

Expression and purification of pEC14 was identical to that of pEC4.

2.3.9 Expression and purification of N-terminal thrombin cleavable his-tagged *Pseudomonas aeruginosa* TolA domain 3 (pEC16)

pEC16 was transformed into BL21 (DE3) competent cells for expression and purification of thrombin cleavable his-tagged *Pseudomonas aeruginosa* TolA3. Growth and harvesting was as pEC1 (2.3.3) with the exception that pelleted cells were resuspended in 30 ml 1x binding buffer (40 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM imidazole) prior to sonication.

The cell debris was spun out (19776 rotor, 10000rpm, 4 °C for 30 minutes) and supernatant decanted. The supernatant was subsequently loaded onto a 5 ml HisTrap HP column (GE Healthcare) charged with 3 column volumes (c.v.) of 50 mM $NiSO_4$ and equilibrated with 3 c.v. of binding

buffer. The column was washed with binding buffer at a flow rate of 1.5 ml/min until the absorbance of the eluate returned to background levels (monitored by 280nm absorbance). Bound protein was then eluted using a linear gradient of 5-500 mM imidazole over 10 c.v. Fractions were analyzed by SDS-PAGE and those containing the protein of interest were pooled and buffer exchanged with PD10 desalting column (GE Healthcare) into 50 mM Tris-HCl pH 7.5, 250 mM NaCl, 2.5mM $CaCl_2$. Sample was then cleaved with thrombin for 16 hours at 4°C. Following cleavage, sample was again loaded onto pre-equilbrated 5 ml HisTrap HP column to remove the tag. Cleaved protein collected in flowthrough.

Protein was further purified by gel filtration using a Hiload 26/60 Superdex 75 column (GE Healthcare). 10 ml of protein was injected onto the pre-equilibrated S75 column (50 mM Tris-HCl, 250 mM NaCl, pH 7.5) and was eluted at a flow rate of 3 ml/min. Protein elution was monitored by absorbance at 280 nm. All peaks were analysed by SDS PAGE. Peak fractions were pooled and stored at -20 °C until further use.

2.3.10 Other proteins

Wild type colicin E9 was kindly supplied by Dr Nick Housden, University of York. Periplasmically translocated and processed *E.coli* TolB and Δ34 mutant of eTolB (lacking 12 amino acids of disordered N-terminus used to bind eTolA) were supplied by Dr Daniel Bonsor (University of Maryland). eTolA23 (*E.coli* TolA domains 2 and 3 with N-terminal Cysteine mutation) was kindly supplied by Dr Anne Marie Krachler (University of Texas, Southwestern Medical Center at Dallas).

2.3.11 TolA binding epitope synthetic peptides

Synthetic TolA binding epitope peptides were custom synthesised by Activeotec (Cambridge, UK) or Pepceuticals (Enderby, UK) in 50 mg or 100

mg batches at >95% purity. *E.coli* TolA binding peptide was of sequence EVRIVIDSGVDSWKKK and *Pseudomonas aeruginosa* TolA binding peptide was of the sequence ADPLVISSGNDRWKKK.

2.3.12 SDS polyacrylamide gel-electrophoresis

A stock of 30 % acrylamide and 0.8 % bis-acrylamide (Protogel, National Diagnostics) was diluted into 375 mM Tris-HCl pH 8.8 and 0.1 % sodium dodecyl sulphate (SDS) to give final acrylamide concentrations ranging from 10 – 20 % as required for the running gel. Ammonium persulfate and N, N, N', N'-tetramethyl ethylenediamine (TEMED) were added to a final concentration of 0.08 % w/v and 0.13 % v/v, respectively to initiate gel polymerisation. Stacking gels were prepared similarly but contained 5 % acrylamide and 250 mM Tris- HCl pH 6.8. Samples were mixed with loading buffer to give a final concentration of 50 mM Tris-HCl pH 6.8, 2.5 % β-mercaptoethanol, 2 % SDS, 0.1 % bromophenol blue and 10 % glycerol. Samples were usually boiled for 5 minutes prior to gel loading.

Unstained protein molecular weight marker with a range of 14.4 to 116 kDa (Fermentas) was loaded on gels alongside samples to enable approximate molecular weight determination of protein samples. Samples were separated by gel electrophoresis in 25 mM Tris, 193 mM glycine, 10 % SDS at a constant current of 30 mA per gel until sufficient separation had been achieved (usually approximately 30 mins).

Proteins were visualized by staining with 0.2 % Coomassie brilliant blue R250 (Pierce) in 10 % v/v acetic acid and 50 % v/v ethanol and destained in the same solution lacking Coomassie blue dye.

2.3.13 Protein estimation

Protein concentrations were measured by the absorbance at 280 nm using a Biophotometer (Eppendorf) and the Beer-Lambeth law. Theoretical

molar absorbance coefficients (ε_{280nm}) for each protein were calculated according to the following formula;

$$\varepsilon_{280nm} = (N_{cys} \times \varepsilon_{280nm\ of\ cys}) + (N_{tyr} \times \varepsilon_{280nm\ of\ tyr}) + (N_{trp} \times \varepsilon_{280nm\ of\ trp})$$

where N is the number of amino acids (cys, cystine; tyr, tyrosine; trp, tryptophan) and ε_{280nm} is the molar absorption coefficient for the amino acids (Cys, Tyr, Trp; 125 $M^{-1}cm^{-1}$, 1490 $M^{-1}cm^{-1}$ and 5500 $M^{-1}cm^{-1}$ respectively) (Pace et al. 1995).

A list of molar absorbance coefficients for the proteins used in this work are presented in table 2.7.

Protein	Theoretical molar absorbance co-efficient (M^{-1} cm^{-1})
eTolA3 (pAK108)	5960
eTolA3 (domain 3 only)	6085
eTolB	57870
ColE9 (WT)	57040
ePal	11920
psTolA	6990
psTolB	41830

Table 2.7. List of theoretical molar absorbance coefficients.

2.3.14 Electrospray ionisation mass spectrometry

To verify the molecular mass of purified proteins, sample protein (typically 50-100 µM) was dialysed in dH_2O overnight and then diluted 1:10 in 50% acetonitrile, 50% dH_2O and 0.1% trifluoroacetic acid prior to analysis on either Waters LCT Premier XE or ABI-QStar Tandem Mass spectrometer, both connected to electrospray ion source by Berni Strongitharm, Andrew Leech, Adam Dowle (University of York Technology Facility) or Renata Kaminski (Kleanthous Lab).

2.3.15 Formaldehyde cross-linking of purified proteins

Purified proteins were dialyzed overnight against 10 mM sodium phosphate pH 6.3-8.5 at 4 □C. Individual proteins and complexes (final concentration of 10 µM each) were incubated at 37 □C for 15 minutes before addition of formaldehyde (final concentration 1% w/v) to achieve a final volume of 10 µl. Proteins were incubated for a further 15 minutes at 37 □C before the addition of 5 µl of 100 mM Tris-HCl, pH 6.8 to quench the cross-linking reaction. Samples were mixed with 4 x SDS loading buffer, incubated for 15 minutes at 37 □C and run on 13% either Coomassie blue staining or Western blotting.

2.3.16 DSP ((Dithiobis(succinimidylpropionate), Lomants reagent) crosslinking of purified proteins

Purified proteins were dialyzed overnight against 10 mM sodium phosphate pH 7.0 at 4 □C and diluted to final concentrations of 0.5-100 µM. 80 mg/ml stock of DSP dissolved in DMSO was added to reaction mixture to give a 10-fold molar excess (unless stated otherwise). The reaction was left to proceed at 22 □C for 30 minutes and subsequently quenched by adding a 100-fold molar excess of Tris-HCl pH 8.0. Samples were analyzed by non-reducing SDS-PAGE to maintain cross link or reduced by adding 20 mM DTT prior to gel loading to abolish crosslink.

2.3.17 Cross-linked protein mass spectrometry

Purified *E.coli* TolA3 (eTolA3) was dialysed overnight against 10 mM Sodium phosphate, pH 6.3 – 8.5) at 4 □C and diluted to final concentrations of 0.5-100 µM. Synthetic *E.coli* TolA binding peptide (eTABp) stock was made up to appropriate concentration in identical buffer conditions. eTolA3 and eTABp were then crosslinked in the presence of either formaldehyde or DSP (described in section 2.3.15-16). Following quenching of crosslinking

reaction, samples were separated with SDS-PAGE and desired bands excised from the gel. Samples were then delivered to Proteomics Lab (Technology Facility) where in gel trypsin digest was performed by Adam Dowle. Samples were then diluted into 50% acetronitrile, 50% dH_2O and analysed using MALDI-TOF (ABI-QStar instrument with electrospray ionisation) mass spectrometry performed by Dr David Ashford. Trypsinised fragments from eTolA3 alone and eTolA3 crosslinked with eTABp were then compared and analysed to determine if firstly any additional mass from the peptide crosslinking to the protein could be detected, and secondly if any trypsin fragments were lost due to protected trypsin digestion site as a result of peptide crosslinking to peptide.

2.3.18 Western blotting

Antibodies

All antibodies were supplied by Dr Daniel Bonsor (Bonsor 2009). Briefly, polyclonal antibodies raised against *E.coli* TolA, TolB, Pal and ColE9 were produced in rabbits (Eurogentec). Antibodies for TolB, ColE9-Im9 and tagless Pal were produced against the purified proteins. Anti-rabbit IgG peroxidase conjugate (Sigma) was used as secondary antibody for Western blots.

Antibody	Dilution
ecTolB	1:100000
ecTolA3	1:1000
ecTolA2-3	1:12500
ecPal	1:1000
Rabbit IgG	1:2000

Table 2.8 Typical dilutions of antibodies used for Western Blots

Blotting procedure

Following protein separation by SDS-PAGE, samples were transferred onto nitrocellulose membranes via semi-dry electroblotting. Polyacrylamide gels were first incubated in transfer buffer (25 mM Tris, 150 mM glycine, pH 8.3 in 20 % v/v methanol) for 1 minute and then sandwiched between three layers of Quickdraw blotting paper (Sigma), a sheet of Hybond-ECL nitrocellulose membrane (GE Healthcare) and another three layers of blotting paper, all equilibrated in transfer buffer. The transfer was carried out using a V10-SDB semi-dry blotter (Fisher) under a constant current of 1 mA/cm^2 for 45 minutes.

Blotted nitrocellulose membranes were washed for 3×5 min with 1×TBS-T (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.05 % v/v Tween 20) and incubated with blocking buffer (8 % w/v Marvel milk powder in 1×TBS-T) overnight. The blot was washed for 3 × 5 min with 1×TBS-T before a one hour incubation with 0.25 ml/cm^2 membrane of primary antibody diluted into 1×TBS-T containing 4 % w/v milk powder. The membrane was washed 3 × 5 min with 1×TBS-T again, followed by incubation with the secondary antibody diluted 1:1000 into 1×TBS-T + 4 % w/v milk powder.

Following another three wash steps, the membrane was incubated with ECL western blotting analysis system (GE Heathcare) for one minute and immediately exposed to a Kodak BioMax light film (Sigma) for 30 seconds to 15 minutes inside a Hypercassette (GRI). The film was developed using a Compact X4 Film Processor (Xograph Healthcare Ltd.).

2.3.19 Circular Dichroism (CD)

Protein samples were dialysed into buffer (typically 10 mM Sodium Phosphate, pH 6.5-9) overnight at 4 °C at concentrations ranging from 5-50 μM. Spectra were collected on a Jasco J810 CD spectrophotometer with

peltier temperature control attachment using 0.1 or 1 mm pathlength quartz cuvettes. Scans were performed from 260-190 nm at a temperature of 20 °C with a scanning rate of 100 nm/min and a data pitch of 1 nm. 5-10 scans were averaged to give final spectra. For thermal denaturation experiments, scans were performed in 5 °C increments with an equilibration phase of one minute at each temperature and a heating rate of 0.5 °C/min (unless stated otherwise) over a temperature range of 20 - 90°C. Samples were subsequently cooled to 20 °C, allowed to equilibrate for 10 mins, and spectra was collected again. For all proteins, efficiency of refolding was greater than 95%.

To calculate molar ellipticity (θ, deg cm^2 $dmol^{-1}$), the following formula was applied;

$$\theta = CDsignal[mdeg] / (L \times c \times n)$$

(where L = pathlength (mm), c = protein concentration (M) and n = number of residues)

2.3.20 Isothermal Titration Calorimetry (ITC)

Purified protein samples were dialysed against appropriate buffers overnight at 4 °C, and any precipitates removed by centrifugation (10000g, 10mins). Concentrations of the samples were measured (typically 100-5000 μM for syringe sample, 30-500 μM for cell sample), diluted in appropriate dialysis buffer and degassed for 10 minutes using a Thermovac degassing unit (when using VP-ITC) (Microcal/GE Healthcare). Samples were loaded into the cell and syringe of either VP-ITC or ITC-200 microcalorimeter (Microcal/GE Healthcare). When using VP-ITC, a typical full titration consisted of 35 injections (1x2 μl, 34x8 μl) measured at 20 °C with an interval of 270 seconds between injections, and a stirring speed of 307 rpm (unless otherwise stated). When using ITC-200, a typical full titration

consisted of 20 injections (1x0.4 µl, 19x2 µl) measured at 20 °C with an interval of 240 seconds between injections, and a stirring speed of 1000 rpm (unless otherwise stated). Heats of dilutions were measured by injecting syringe samples into buffer under identical titration conditions and subtracted from each data set. Data was analysed using Origin 8.0 software, and fitted to single site binding model. Parameters obtained from experiment were stoichiometry, K_d, ΔH, ΔS of probed interaction. See appendix section 7.6 for details of ITC and single site binding model equation.

2.3.21 Surface Plasmon Resonance (SPR)

To measure binding affinities by SPR between immobilised *E.coli* TolA3 and binding partners, a Biacore T100/200 instrument (GE Healthcare) was employed. Protein to be immobilised was an *E.coli* TolA domain 2 and 3 (eTolA23) mutant, with the addition of an extreme N-terminal Cysteine residue, encoded by pAK123 construct. For preparation of SPR chip and protein immobilisation, distilled H_2O (dH_2O) was used as the running solvent.

To immobilise the protein a C1 SPR chip was activated by two 120s injections of 0.1M glycine pH 12 with 0.3 % triton X-100 at a flow rate of 30 µl/min. All consecutive steps were conducted at a flow rate of 10 µl/min. The chip surface was first modified by injecting a 1:1 mixture of freshly thawed NHS:EDC (stock of 0.4 M EDC and 0.1 M NHS in dH2O) for 420 seconds. This was followed by a 420 second injection of 0.1 M ethylenediamine in 0.1 M sodium borate pH 8.5 and an injection of 50 mM N-[γ-maleimidobutyryloxy]sulfosuccinimide ester (sulfo-GMBS) in 0.1 M sodium borate pH 8.5 for 240 seconds. eTolA23 at a concentration of 8-10 µM in 10 mM sodium phosphate pH 7.0 was subsequently injected into the sample channel only for 420 seconds, resulting in a typical immobilization of 600 RU. Next, reactive groups remaining on both sample and reference channel were blocked by injecting 50 mM cysteine in 1 M NaCl, 0.1 M sodium phosphate

pH 7.0 for 240 seconds. After changing the running solvent to binding buffer (50 mM Hepes pH 7.5, 50 mM NaCl, 0.05 % P20), binding experiments were carried out with a flow rate of 10 µl/min at 25 °C. Prior to titration experiments to ensure the quality of chip, a 100 µM test sample of eTolB was passed over the immobilised eTolA23, ensuring a good response curve was obtained. Following a 6 M GnHCl unfolding/regeneration step, to remove eTolB from the eTolA23, 300 Response units (RU) of stable eTolA23 were left attached to the chip, ready for binding experiments. Subsequent samples at concentrations ranging from 1-150 µM for protein binding partners, and 10-1000 µM for TolA binding peptides were injected for 300 seconds, followed by an equilibration phase (buffer injection for 300 seconds). The surface was regenerated after each binding experiment by injecting 6 M guanidine chloride in 30 mM Tris-HCl pH 7.5, 5 mM Imidazole, 200 mM NaCl for 60 seconds, followed by an equilibration period of 600 seconds prior to the next binding experiment. Binding affinities were derived by equilibrium analysis (assuming an equilibrium 4 seconds prior to end of the injection) of blank-subtracted sensograms using the Biacore evaluation software. See appendix section 7.5 for further details of surface plasmon resonance immoblisation.

2.3.22 Protein crystallisation

Numerous attempts to crystallise eTolA3, eTolA3-eTABp complex, psTolA3, psTolA3-psTABp complex were made with protein concentrations ranging from 1-150 mg/ml (with peptides in 1-10x molar excess), set up in MRC-Wilden crystallisation plates using Hydra and Mosquito robots against commercially available screens including Peg-Ion 1&2, Hampton 1&2, Morpheus, PACT and Index, using the sitting drop method.

2.3.23 Analytical ultracentrifugation (AUC)

Analytical ultracentrifugation experiments were performed by Dr Andrew Leech, University of York using Beckman Optima XL/I analytic ultracentrifuge equipped with Beckman 12 mm path length double sector charcoal filled Epon centrepieces and sapphire windows in an AN-60Ti rotor (3 cells plus counterbalance). Approximately 420 µl reference buffer and 416 µl protein sample were loaded into each cell, following which absorbance scans were performed at 3000 rpm to verify loading concentrations and samples were uniformly distributed. Cells were then removed, agitated, and replaced. Samples were then spun at 50000 rpm for 10 hours at 20 °C with sample scans collected every 180 seconds at 302 nm until either sedimentation was complete, or plateau region had disappeared. Data was analysed, fitted and transformed using SEDFIT software (Schuck 2000). See appendix section 7.7 for details of AUC fitting.

2.3.24 Nuclear Magnetic Resonance (NMR)

^{15}N eTolA3 and psTolA3 protein expression and purification

Proteins to be used in Nuclear Magnetic Resonance experiments were grown in supplemented M9 minimal media for labelling. Briefly, plasmid was transformed into BL21 (DE3) cells, from which a single colony was used to inoculate a 50ml culture of M9 minimal media, supplemented with 20% (w/v) glucose and 8.5% (w/v) Yeast Nitrogen Base (YNB) without amino acids or ammonium sulfate (Sigma). A 1 in 20 dilution was used to inoculate 2x750ml of M9 (as above, with the exception of ammonium chloride was ^{15}N enriched (CK Gas Products). Cells were grown at 37 °C on Innova 2300 platform shaker (New Brunswick Scientific) at 120 rpm until an OD_{600nm} of ~ 0.6 was reached. Gene expression was induced with IPTG (Melford) at a final concentration of 1 mM. Cells continued to grow on the orbital shaker at

37 °C for a further 16 hours before harvesting (SLC-6000 rotor, 4500 rpm, 4 °C for 12 minutes). All subsequent purification procedures were as tagless eTolA3 (pAK108 construct) or His-tagged psTolA3 (pEC4 construct).

^{15}N ^{13}C psTolA3 protein expression and purification

Double labelled protein was purified as ^{15}N psTolA3 described above, with the exception of glucose used in 750ml cultures was ^{13}C enriched (CK Gas Products).

NMR Data acquisition and analysis

Following purification and concentration, labelled proteins were dialysed against either 50mM Potassium phosphate, 50mM NaCl, pH 7.5 (in case of eTolA3) or 20mM Sodium phosphate, pH7.5 (in case of psTolA3) overnight at 4 °C. NMR samples were made by mixing 540 µM of protein with 60 µl D_2O. 0.05% (w/v) Sodium Azide was added as a preservative. Final protein concentrations ranged from 200 µM for titration experiments and 600-1000 µM for 3D experiments for assignment. Samples were loaded into Norell 600MHz tuned tubes (Sigma-Aldrich) and 2D-HSQC experiments recorded for titration experiments using singly labelled protein, and CBCANH, CBCACONH, HNCO, HNCACO, Trosy-HSQC and NOESY experiments (see appendix section 7.8 for details of NMR experiment types) recorded with double labelled protein on a 700MHz Avance II Spectrometer (Bruker) with triple resonance probe. The probe was operating at ^{1}H frequency of 700.13 MHz, ^{13}C frequency of 176.05 MHz and ^{15}N frequencies of 70.93 MHz, using pulse sequences supplied and modified by Bruker Topspin 3.0 software. All spectra were collected at 20°C. To verify that spectra had been successfully collected, data were initially processed with Bruker Topspin 3.0 software. Following this, spectra were processed and

phased using NMRDraw (Delaglio et al. 1995), converted to Azara format using NMRPipe (Delaglio et al. 1995) and analysed with CCPN Analysis 2.1.5 (Vranken et al. 2005).

2.3.25 *In vivo* colicin cell killing assay

Bacterial strains (JM83, unless otherwise stated) were grown to an OD_{600} of 0.6, diluted 1:20 into 0.75 % (w/v) top-agar at 42 °C and the suspension was spread on top of pre-poured LB-agar plates. Once top-agar had set, serial dilutions of wild type colicin E9 or colicin E9 mutants were spotted onto the plates. After 16 hours of incubation at 37 °C, zones of clearance in the bacterial lawn were indicative of cell death.

3. The N-terminus of *E.coli* TolB is the sole determinant for *E.coli* TolA binding

3.1 Introduction

In work reported by Bonsor et al (2009), it was shown primarily via isothermal titration calorimetry (ITC) experiments that not only was the N-terminal strand of *E.coli* TolB (eTolB) a site of interaction with *E.coli* TolA3 (eTolA3), but that *E.coli* Pal (ePal) did not interact with *E.coli* TolA3, as had been previously reported (Cascales et al. 2002). As these two findings conflicted with one another, further investigation of the interactions between the *E.coli* TolA, *E.coli* TolB and *E.coli* Pal proteins was addressed in the following work. In addition, the work of Bonsor et al. (2009) showed that the interaction of TolB with TolA is dependent on the instrinsically unstructured N-terminus of TolB being in the disordered conformation (Bonsor et al. 2009). This work also showed that binding of the T-domain of colicin E9 (a group A colicin) to TolB caused this instrinsically unstructured N-terminus to enter into its disordered conformation and thus promoted it's interaction with TolA (Bonsor et al 2009). To further investigate this phenomenon colicin E9-TolB fusion proteins were designed to ascertain if TolB could be bypassed and a direct interaction between colicin E9 and *E.coli* TolA could be engineered and function, both *in vivo* and *in vitro*.

3.1.1 *E.coli* TolA

E.coli TolA is a 43 kDa, 421 residue, 3 domain protein that spans the periplasm of the cell. It consists of domain 1 (residues 1-48), a short transmembrane domain that anchors the protein into the inner membrane of the cell, domain 2 (residues 48-310), which is a long helical domain arranged into a triple helix (Cascales et al. 2007), and domain 3 (residues 311-421), a folded globular domain (Lubkowski et al. 1999). Previous work published has indicated that it is domain 3 of TolA that interacts with *E.coli* TolB (Bonsor et

al. 2009), however, the binding site on TolA is currently unknown. *E.coli* TolA has been reported to interact with *E.coli* Pal in a proton motive force dependent manner (Cascales et al. 2000; Cascales et al. 2002), however this finding was at odds with other work that had shown that there no interaction *in vitro* between *E.coli* TolA and Pal (Bonsor et al. 2009). A crystal structure of TolA domain 3 from *Pseudomonas aeruginosa* (figure 3.1B) was the first structure to be published for a TolA protein on its own (Witty et al. 2002). In addition, a solution NMR structure for *E.coli* TolA domain 3 has been published (Deprez et al. 2005) (figure 3.1C), as well as more recently a crystal structure (Li et al. 2012) (figure 3.1A). As can be seen in figure 3.1D, *Pseudomonas aeruginosa* TolA3 and *E.coli* TolA3 both have a similar fold, despite sharing only 20% sequence identity (Witty et al. 2002).

3.1.3 *E.coli* TolB

E.coli TolB is a 44 kDa, 430 residue protein consisting of 2 domains. The larger of the 2 domains has a beta-propeller motif consisting of 6 blades. It is within this beta-propeller domain that both *E.coli* Pal and the translocation domain of colicin E9 interact with TolB. Additionally, it has a smaller N-terminal domain, the key feature of which is the natively unstructured N-terminus (Bonsor et al. 2009). When produced by the cell, *E.coli* TolB has 22 additional residues on the extreme N-terminus that direct it for Sec-dependant export into the periplasm. These 22 extra residues are then cleaved by signal peptidase, leaving the mature protein (Isnard et al. 1994). The first 12 residues (EVRIVIDSGVDS) on the N-terminus of the mature protein have been found to be important for the interaction that occurs between TolB and TolA (Bonsor et al. 2009). Without these 12 residues on the N-terminus, no binding was detected *in vitro* (ITC and formaldehyde cross-linking) between TolB and TolA (Bonsor et al. 2009). Work on subsequent pages refers to a mutant of TolB that lacks the N-terminal 12

residues (in addition to the 22 residues removed by signal peptidase) is termed “Δ34 TolB mutant”. When TolB is in isolation, these N-terminal residues are presumed to be in dynamic equilibrium between an ordered and disordered state (from evidence obtained from NMR experiments) (figure 3.2, b). It is in the disordered state that TolA is predicted to interact with TolB (figure 3.2, c). When Pal binds TolB however it causes a conformational change in TolB that causes its N-terminus to become ordered and bound back to the body of the TolB protein. When in this ordered conformation, the N-terminus is unavailable to bind TolA (figure 3.2, a), and thus Pal binding can be seen as an “off” switch for the TolA-TolB interaction. Conversely, when the translocation domain of colicin E9 binds TolB (in the same site as Pal), the opposite occurs, wherein the binding of colicin E9 promotes disorder in the TolB N-terminus and therefore promotes the interaction with TolA (Bonsor et al. 2009).

3.1.4 *E.coli* Pal

Pal is a small (173 residue, 19 kDa) protein that is normally attached to the outer leaflet of the outer membrane via a lipoyl tether (Lazzaroni et al. 1992). The name Pal stands for Peptidoglycan Associated Lipoprotein, and as it’s name suggests, Pal has been found to associate with the peptidoglycan layer of Gram-negative bacteria (Mizuno 1979). When Pal is bound to the beta-propeller domain of TolB, it causes a conformational change in TolB, which acts as an allosteric switch to mediate the conformation of the N-terminus of TolB from that of dynamic equilibrium to ordered state (figure 3.2, a) (Bonsor et al. 2009). Pal has also been reported to interact with *E.coli* TolA *in vivo* in a proton-motive force dependent manner, found by *in vivo* formaldehyde cross-linking and immunoprecipitation experiments. This work found that when cells were treated with CCCP (abolishing pmf) then the TolA-Pal interaction was lost (Cascales et al. 2000). However, this interaction has not been verified *in vitro* (Bonsor et al. 2009).

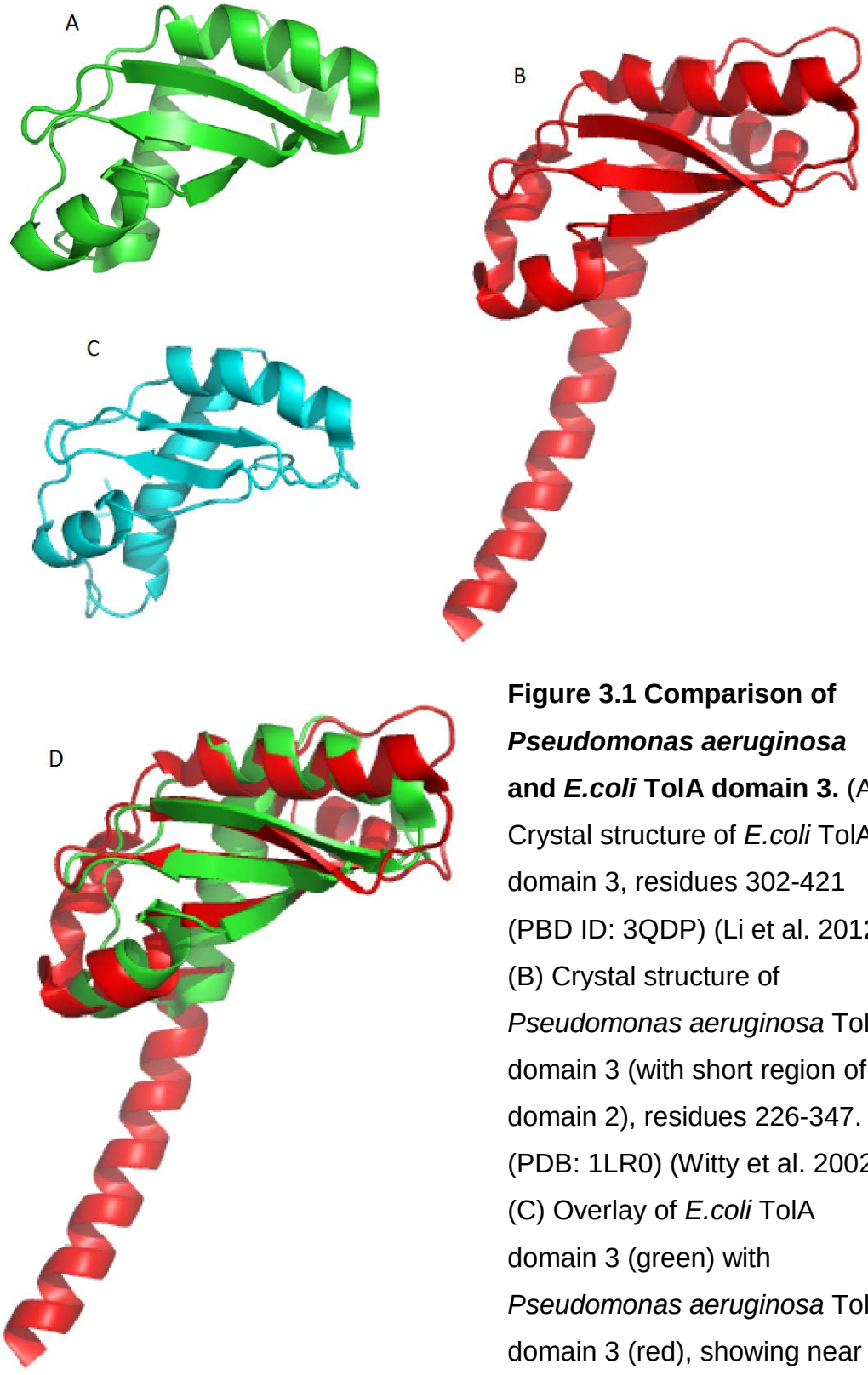


Figure 3.1 Comparison of *Pseudomonas aeruginosa* and *E.coli* TolA domain 3. (A) Crystal structure of *E.coli* TolA domain 3, residues 302-421 (PBD ID: 3QDP) (Li et al. 2012). (B) Crystal structure of *Pseudomonas aeruginosa* TolA domain 3 (with short region of domain 2), residues 226-347. (PDB: 1LR0) (Witty et al. 2002). (C) Overlay of *E.coli* TolA domain 3 (green) with *Pseudomonas aeruginosa* TolA domain 3 (red), showing near identical fold.

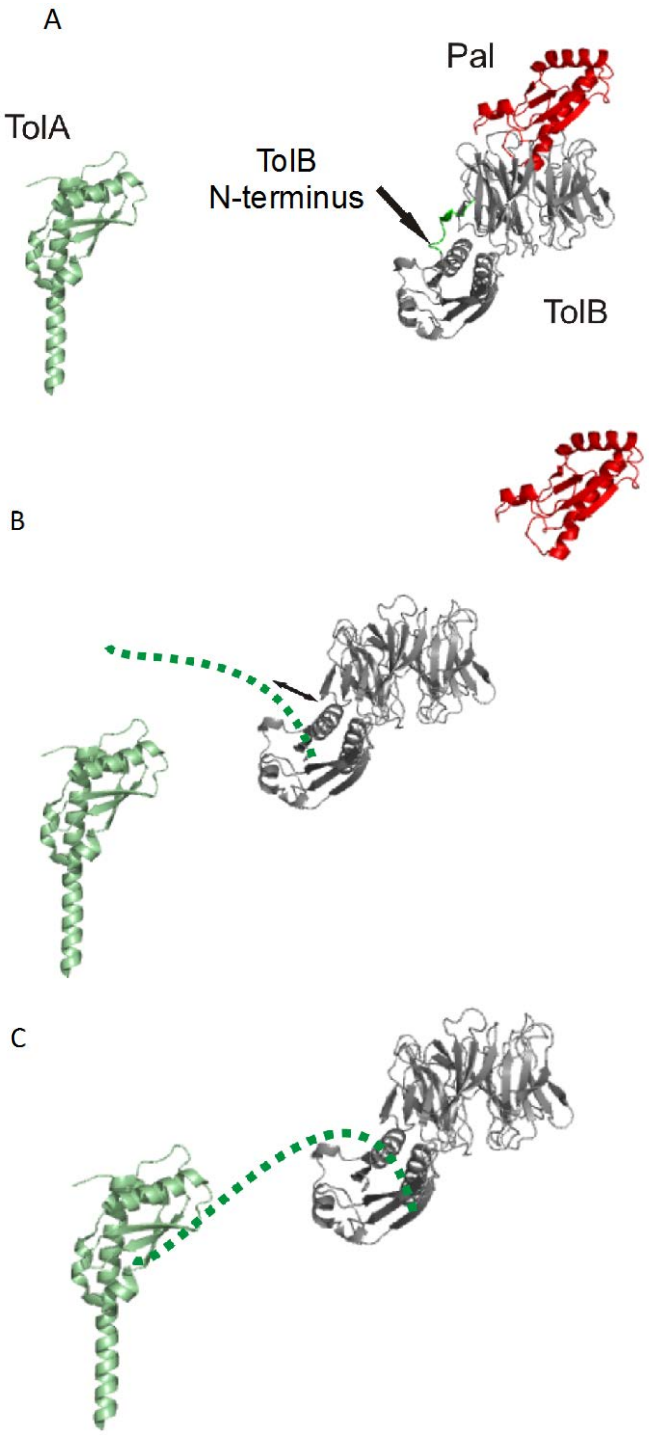


Figure 3.2 *E.coli* TolA-TolB-Pal interaction network. (a) When TolB is in complex with Pal, it's intrinsically disordered N-terminus is in ordered conformation, unavailable for TolA. (b) When not bound to Pal, the N-terminus is in dynamic equilibrium between ordered and disordered state. (c) In disordered conformation, TolB's N-terminus can bind TolA.

3.1.5 Colicin E9

Colicin E9 is a Tol dependent colicin that kills *E.coli* cells via a cytotoxic DNase action. It is arranged in 3 domains, the intrinsically disordered N-terminal translocation domain, the receptor domain that binds to BtuB and finally the C-terminal cytotoxic DNase domain. The N-terminal translocation domain is subdivided into 3 binding sites; OmpF binding site 1 (OBS1), *E.coli* TolB binding site and OmpF binding site 2 (OBS2). To enter into the cell, and therefore cause cell death, colicin E9 first binds BtuB through it's receptor domain, secondly threads it's intrinsically disordered translocation domain through OmpF (Housden et al. 2005) (Housden et al. 2010), to bind with TolB through TolB's beta-propeller domain. This binding event promotes the N-terminus of TolB into a disordered conformation, driving it's interaction with TolA (Bonsor et al. 2009). By contacting TolA in the periplasm, the translocation of colicin E9 across the outer membrane, periplasm and ultimately to the inner membrane is somehow driven, allowing the cytotoxic DNase domain to enter into the cell (Housden et al. 2005). The specific role that the Tol proteins have in this event is currently not known.

3.1.7 The *E.coli* TolA-TolB-Pal interaction

As reported by Bonsor et al (2009), the eTolA3 interaction with eTolB is dependent on the N-terminus of eTolB in the disordered conformation when challenged via ITC. Mutants that lack this N-terminus (of sequence ^{23}EVRIVIDSGVDS^{34}) do not show any heats of binding when titrated against with eTolA3. As no structure is currently available for any form of TolA-TolB, the binding site of TolB on TolA is not known, nor is it known if the N-terminus is the sole site of interaction between TolA and TolB.

In addition, when a preformed complex of eTolB and ePal is titrated against with eTolA3, no heats are detected (Bonsor et al. 2009). Furthermore, when

the crystal structures of eTolB alone (PDB ID: 1C5K) (Carr et al. 2000), and eTolB in complex with ePal (PDB ID: 2W8B) (Bonsor et al. 2009) are compared, in the presence of ePal, the N-terminal strand of eTolB is in an ordered conformation, bound back to eTolB. This indicated that ePal controlled the conformation of the disordered N-terminus of eTolB. When the N-terminus is in an ordered state (bound back conformation) it is prevented from binding with eTolA3. When eTolB is not bound to ePal, the N-terminus of eTolB is presumed to be in dynamic equilibrium between ordered and disordered state, as NMR data showed that peaks corresponding to the N-terminus were in slow exchange (i.e. double the number of peaks found in spectrum than expected, indicating that the N-terminal peaks were in 2 populations). When the N-terminal residues were deleted, these peaks were lost from the NMR spectrum. Additionally, although previous work had suggested that an interaction between eTolA3 and ePal occurred (Cascales et al. 2000), work by Bonsor et al. disagreed with these findings. It was the aim of this work to independently investigate these findings through an alternative biophysical technique (Surface Plasmon Resonance).

3.1.8 Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry can be used to measure the thermodynamic properties of a protein-protein interaction. ITC determines the binding equilibrium of the interaction by measuring the amount of heat released or absorbed on association of a ligand with it`s binding partner. From an ITC experiment, the binding constant (K_a), stoichiometry (n), enthalpy of binding (ΔH_b) and entropy of binding (ΔS_b) can be determined.

An ITC instrument comprises of 2 identical chambers or cells (one sample cell, one reference cell) made from a material with a very high thermal efficiency (such as gold), both of which are surrounded by a thermal jacket, which is regulated by either a circulating water bath or, more commonly, a

Peltier device. A highly sensitive thermocouple monitors the temperature of the 2 cells, and regulates the thermal jacket to ensure that both cells maintain an identical temperature (figure 3.3).

To perform an ITC experiment, a buffered protein sample is placed into the sample cell and allowed to equilibrate in terms of temperature. In the reference cell, either a sample of identical buffer without the protein, or water is placed. A constant power is applied to the reference cell, creating a baseline signal. This controls the sample heater, which reacts to maintain the sample cell at the same temperature as the reference. The directly observed data for an ITC experiment is the time dependent power level applied to the cells to maintain an identical temperature. When a sample is injected (titrated) into the sample cell, and binds with the protein in the sample cell, heat is either released or absorbed from the environment, depending on whether the interaction is exothermic or endothermic, respectively. If the interaction (reaction) is exothermic, the temperature of the sample cell will rise, causing the sample cell heater to be deactivated (or in the case of a Peltier, the device being activated to remove heat). If the interaction is endothermic, the opposite will occur, i.e. the temperature in the sample cell will drop, causing activation of the sample cell heater to return the temperature to that of the reference cell. The amount of power over time that is required to do either of these tasks is what is directly measured.

Over the course of an ITC experiment, the amount of heat that is released or absorbed by the binding event (injection) is directly proportional to the fraction of ligand bound to the sample cell protein. During the initial experimental titrations, most of the ligand injected into the sample cell will be immediately bound to the sample protein, causing a large change in temperature. As more titrant is injected into the sample cell, less sample protein will be available for binding, and thus the amount of heat released or

absorbed will decrease, until saturation is reached, wherein there is no more free sample protein available for binding with the titrant (Pierce et al. 1999).

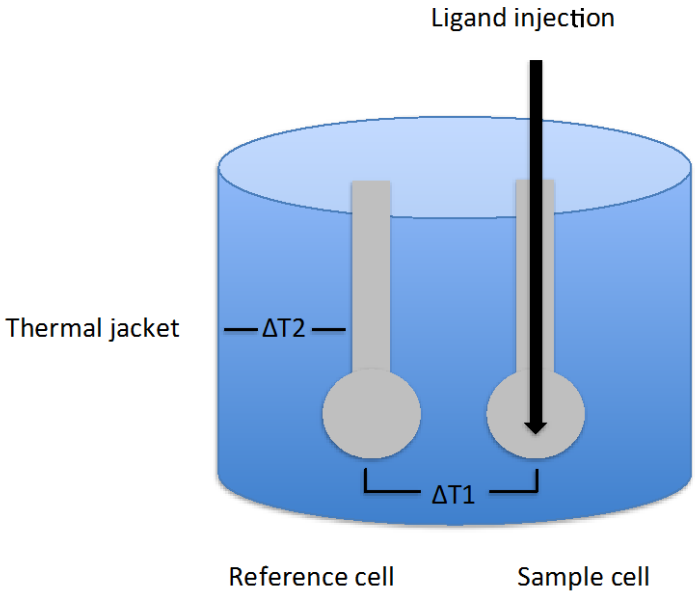


Figure 3.3 Representation of typical Isothermal Titration Calorimetry instrument. ΔT1 compares temperature of reference cell with sample cell during ligand injection. ΔT2 compares reference cell with thermal jacket, i.e. baseline to ensure that any heat released or absorbed during titration is as a result of interaction in sample cell and not changes in the external environment (Microcal 2004).

The amount of heat released or absorbed by the addition of a titrated ligand into the sample cell can be represented by the following, assuming a single site of binding on the target protein:

$$Q = V_0 \Delta H_b [M]_t K_a [L] / (1 + K_a [L])$$

Where Q is heat, V_0 is the sample cell volume, ΔH_b is the enthalpy per mole of the titrated ligand, $[M]_t$ is the total sample protein concentration including bound and free fractions, K_a is the binding constant and [L] is the free titrated ligand concentration (Microcal 2004).

From the heat trace (energy released or absorbed) of the titration (reference cell vs sample cell differential power [µcal / s] against time [s]), it is possible to integrate the area of the peaks (representing the total heat released for a given injection) and plot them against the molar ratio of ligand and protein, creating the binding isotherm (Figure 3.4). From this isotherm, it is possible to calculate the stoichiometry of the interaction (at ligand-protein ratio of 1), as well as the affinity ($1/K_d$) and enthalpy (and thus entropy) of binding (K_a is obtained from equation fit, ΔS is obtained from Gibbs equation [$\Delta G = \Delta H - T\Delta S$] using the measured ΔH and ΔG obtained from affinity of interaction [$\Delta G = RTIn(K_d)$] (Pierce et al. 1999). For further details of ITC single site binding model, see appendix section 7.6.

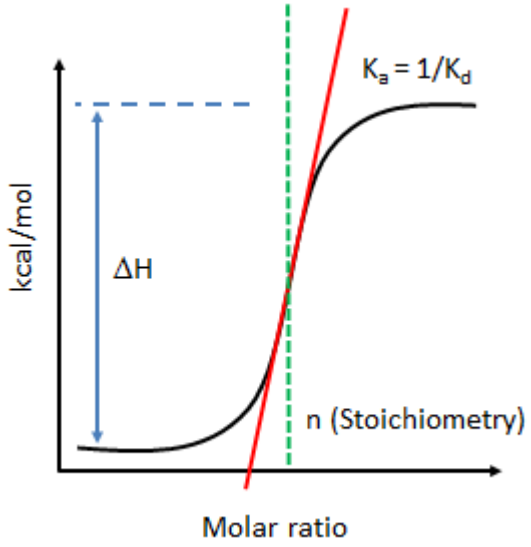


Figure 3.4 Isothermal titration calorimetry binding isotherm

The heat trace peaks are intergrated in order to give the binding isotherm. From this isotherm, the ΔH can be calculated (blue) from the kcal/mol, the stoichiometry (green) from the mid-point of the isotherm, and the $K_{affinity}$ from the slope of the line (red). From the K_a the dissociation constant can be calculated ($K_a = 1/K_d$) (Microcal 2004).

3.1.9 Surface Plasmon Resonance (SPR)

Surface plasmon resonance is a technique that can be used to obtain both the equilibrium dissociation constant for a given interaction, as well as potentially association/dissociation kinetics. SPR occurs when light is produced from a high-refractive index medium, such as a prism, toward an interface containing a low-refractive index material, such as protein sample solution. When a thin film of gold is placed at the interface of the two media, an evanescent wave of photons created by the rear illumination of the high-refractive index media couple with the electrons of the gold film, termed plasmons. This coupling can only occur when resonance of the photons occurs, which is dependent on the polarisation of the light against the gold film, the angle of incidence of the light, and the wavelength of the light. Coupling of the photons and plasmons results in an exchange of momentum, causing a decrease in intensity of the reflected light at the interface. Therefore, at a specific angle, termed the SPR angle (at a specific wavelength), minimal reflectivity occurs. Thus, the SPR signal can be monitored by the incidence angle or wavelength, and correlated to any resultant change in reflectivity (monitored by an optical detector).
In a given system, if a binding event occurs (i.e. a substrate binds to the gold film), a change will occur in the interface by the gold film (the change in refraction is sensitive to up to 300 nm at the gold film surface), and thus, cause a change in resonance, which will in turn cause a change in the intensity of the reflected light. This change in reflected light will cause a change in the SPR signal, which is monitored by an optical sensor, and transformed into a real-time event, measured in response units (RU) (Rusling et al. 2010).

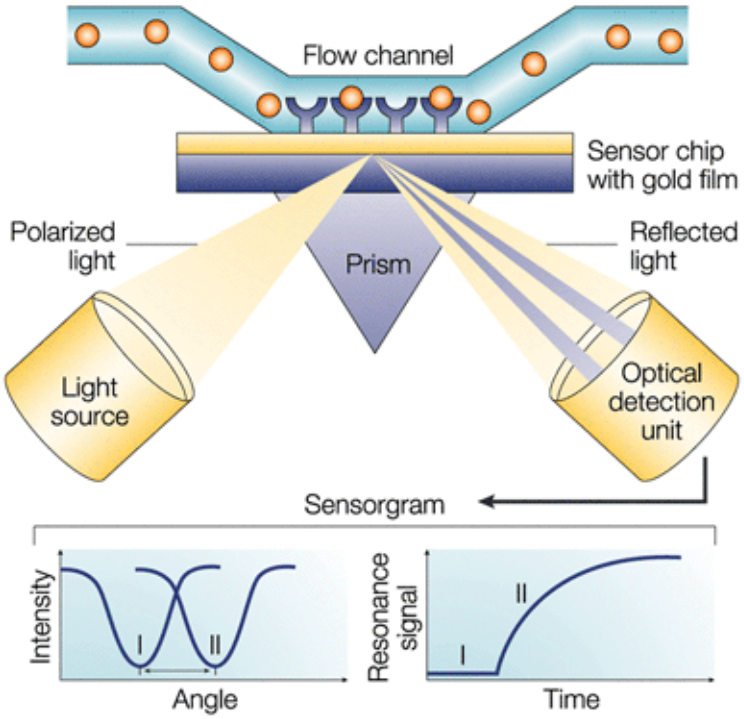


Figure 3.5 Surface plasmon resonance Surface plasmon resonance (SPR) detects changes in the refractive index in the immediate vicinity of the surface layer of a sensor chip, which causes a change in the SPR angle. This angle shifts (from I to II) when biomolecules bind to the surface and change the mass of the surface layer. This change in resonant angle is monitored in real time as a plot of resonance signal/response units (proportional to mass change) as a function of time. Figure reproduced from (Cooper 2002).

3.1.10 Aims

It is the aim of this work in this chapter to test several hypotheses; (1) to confirm that the TolB N-terminus (residues 23-34) is required for it's interaction with TolA; (2) as reported by Bonsor et al in 2009, but in conflict with early work (Cascales et al. 2000), that TolA domain 3 does not interact with Pal; and finally (3) that binding of Pal to TolB mediates the interaction

between TolA and TolB. These questions will be addressed biophysically using Surface Plasmon Resonance.

In addition, once the 12 residue N-terminus of TolB has been investigated in it's native interaction with TolA, this work will investigate (4) that when this short 12 residue stretch is added to another protein (colicin E9) creating a colicin E9-*E.coli* TolB fusion protein, that this short sequence is capable of not only driving a novel interaction between colicin E9 and TolA *in vitro*, but that TolA can complement TolB to drive the *in vivo* translocation of colicin E9 fusion proteins (that cannot interact with TolB) into *E.coli* cells and thus kill them.

3.2 Results

3.2.1 Probing the interactions of *E.coli* TolB, TolA and Pal with surface plasmon resonance

In order to further investigate the interaction network of *E.coli* TolA, TolB and Pal with SPR, a series of experiments were devised. Firstly genes of interest were overexpressed and proteins purified and subsequently characterised in terms of their secondary structure (for quality control purposes due to the lack of a biochemical assay for activity of any of the Tol proteins). To investigate the *in vitro* interactions of *E.coli* TolA, TolB and Pal, *E.coli* TolA protein consisting of domains 2 and 3 (with addition of cysteine residue at the extreme N-terminus of domain 2) was immobilised on a C1 SPR chip via thiol-coupling and titrated against with both *E.coli* TolB and Pal. This work was to confirm the following; that the N-terminus of TolB is required for TolA binding, that *E.coli* Pal does not interact with *E.coli* TolA (as reported by Bonsor et al in 2009, but disputed by Cascales et al. in 2000) and finally, to verify that the binding of *E.coli* Pal to *E.coli* TolB act as an allosteric switch for the *E.coli* TolA-TolB interaction. When *E.coli* Pal is bound to *E.coli* TolB, it causes a conformational change driving the N-terminus (presumed to be in a state of dynamic equilibrium between ordered and disordered conformation when TolB is in isolation) into it's ordered (bound back onto *E.coli* TolB) conformation, and thus is unavailable for binding with *E.coli* TolA (Bonsor et al. 2009).

3.2.1.1 Purification of *E.coli* TolA domain 3 (pAK108 construct)

E.coli TolA domain 3, residues 293-421 was purified from BL21 (DE3) cells transformed with pAK108, grown in 4.8 L of LB media. Protein was purified as described in section 2.3.1, and steps of purification are shown in figure 3.6. A typical protein yield of 20 mg/L of culture was obtained. Mass

spectrometry of data for protein size is reported in table 3.1. Protein sequence reported in appendix section 7.3.

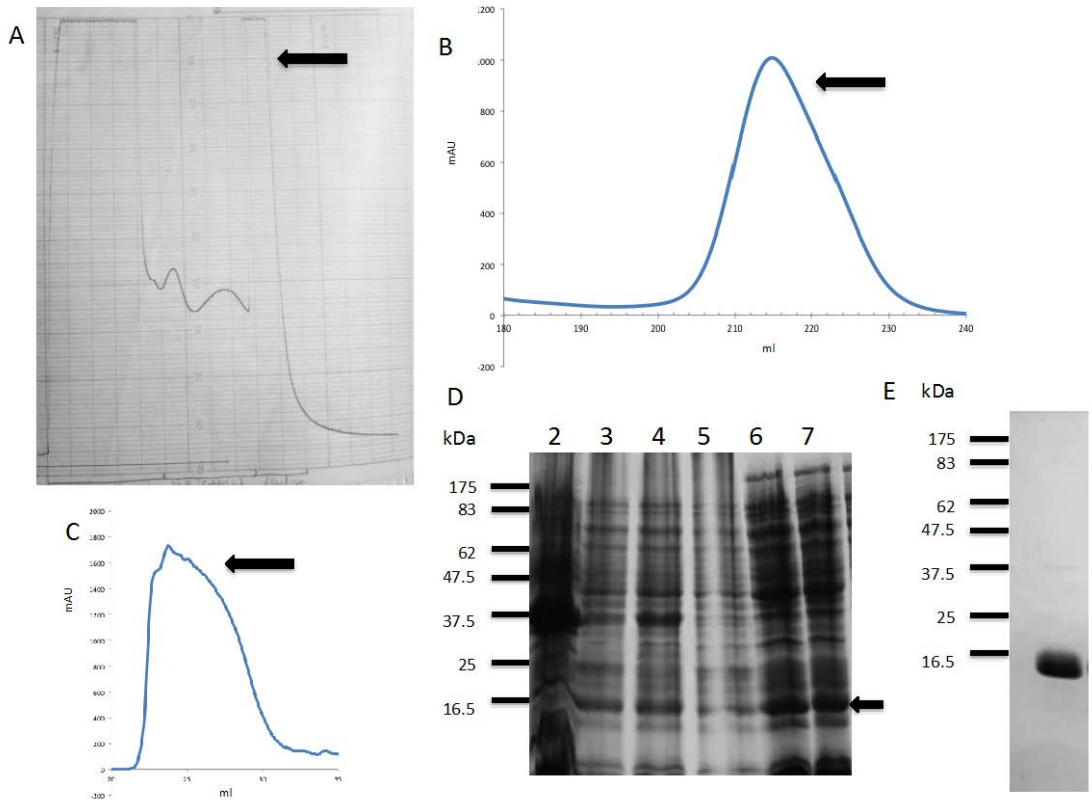


Figure 3.6 Purification of tagless eTolA3 (pAK108 construct) (A) 280 nm absorbance profile from DE52 weak-anion exchange chromatography. Protein (indicated by arrow) eluted with 1M NaCl step gradient. (B) Gel-filtration of eTolA3 on Superdex 75 26/60 column. eTolA3 eluted between 200 and 230 ml. (C) Strong cation exchange (MonoS) chromatography purification step, as monitored at 280 nm. 1M NaCl elution gradient applied. (D) 16% SDS-PAGE gel of purification steps (arrow indicates expected size of eTolA3); D2: 40% Ammonium sulphate precipitated fraction, D3: 40% Ammonium sulphate soluble fraction, D4: 70% Ammonium sulphate precipitated fraction, D5: 70% Ammonium sulphate soluble fraction, D6/7: sample loaded onto DE52 column. (E) 16% SDS-PAGE gel of purified eTolA3 following MonoS purification step.

3.2.1.2 Purification of *E.coli* TolB (pDAB18 construct)

E.coli TolB (residues 22-430) was purified from BL21 (DE3) cells transformed with pDAB18, grown in 4.8L of LB media. Protein was purified as described in chapter 2, section 2.3.2, and steps of purification are shown in figure 3.7. A typical protein yield of 40 mg/L of culture was obtained. Mass spectrometry of data for protein size is reported in table 3.1. Protein sequence reported in appendix section 7.3

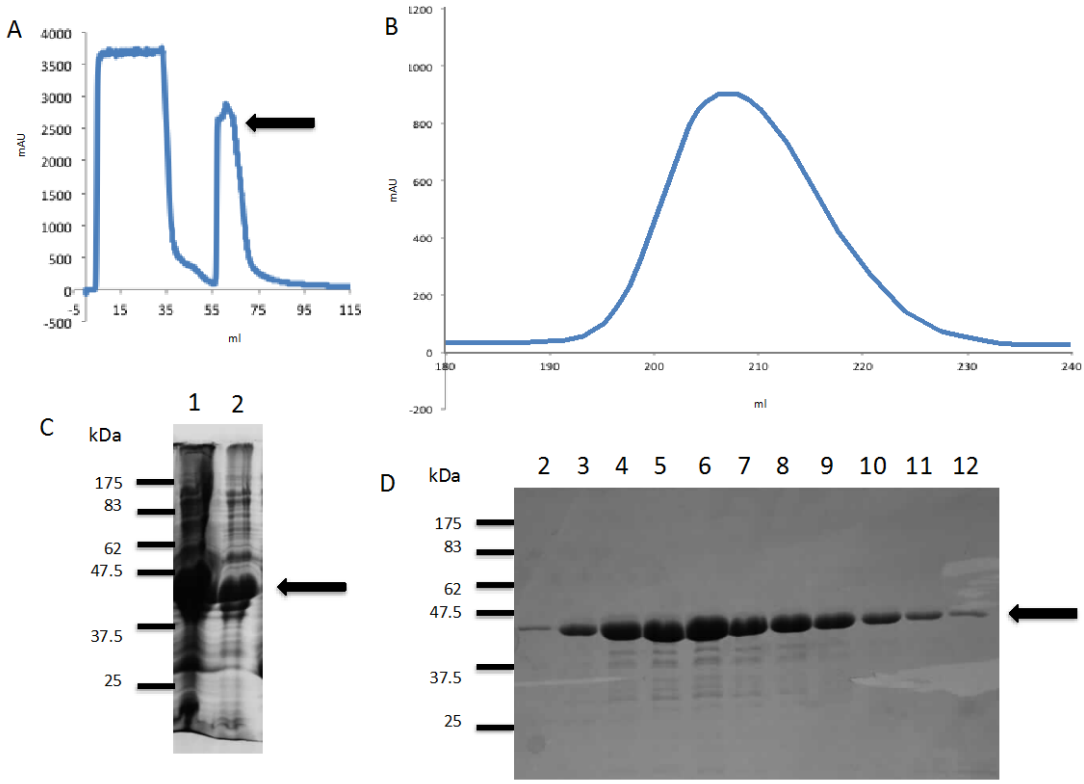


Figure 3.7 Purification of tagless eTolB (pDAB18 construct) (A) 280 nm absorbance profile from DE52 weak-anion exchange chromatography. Protein (indicated by arrow) eluted with 1M NaCl step gradient. (B) Gel-filtration of eTolB on Superdex 75 26/60 column. eTolB eluted between 190 and 220 ml. (C) Sample after Am_2SO_4 cut prior to loading on DE52 column (1) and sample (2) following DE52 weak anion exchange chromatography. (D) 13% SDS-PAGE gel of purified eTolB following S75 gel filtration purification step. Arrow indicates purified protein.

3.2.1.3 Mass spectrometry of purified proteins to verify fidelity of mass.

To verify fidelity of purified proteins, electrospray mass spectrometry was performed as described in section 2.3.14. Expected and observed values indicated in table 3.1.

Protein	Expected Mass	Observed Mass
eTolB (pDAB18)	43733 Da	43735 (±4) Da
eTolA3 (pAK108)	13000 Da	12997 (±7) Da

Table 3.1 Expected vs observed masses of purified proteins

3.2.1.4 Characterisation of secondary structure of purified *E.coli* TolA3 and *E.coli* TolB proteins by circular dichroism spectroscopy.

As there is no functional assay to determine if purified Tol proteins are active, biophysical characterisation of proteins is required, to ensure that they are folded. Far UV (190-260 nm) circular dichroism spectra were collected for eTolA3 (figure 3.8) and eTolB (figure 3.9) proteins. Data was corrected for protein concentration to give molar ellipticity, and subsequently deconvoluted with CDNN 2.1 (Bohm et al. 1992). CDNN analysis estimated that both purified eTolA3 and eTolB had the expected secondary structure in comparison to previously reported purifications (Bonsor 2009; Krachler 2009), indicating that they were correctly folded. Additionally, proteins were also characterised by thermal denaturation to determine both their stability of fold, and efficiency at refolding. Proteins were incubated from 20-90 °C for eTolA3 and 20-70 °C for eTolB with 5 °C spacings before cycle repeated to verify efficiency of refolding. Melting temperature for eTolA3 was 55 °C and for eTolB was 50 °C. More than 95% of both proteins refolded following thermal denaturation (calculated based on total signal from CD spectra).

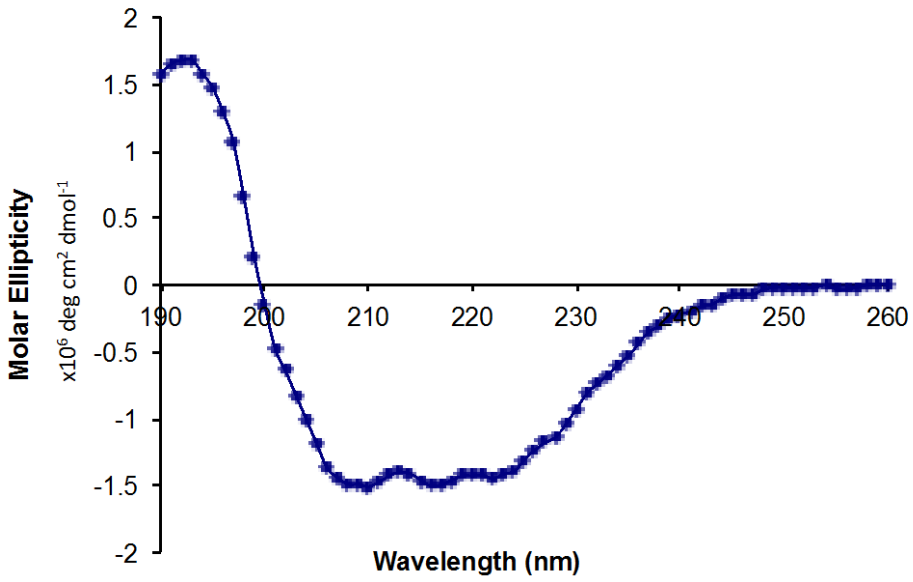


Figure 3.8 Far UV Circular Dichroism spectrum of *E.coli* TolA3 (pAK108 construct). CD spectrum recorded for 50 µM eTolA3 in 10 mM Sodium phosphate, pH7 at 20 °C (0.1 mm pathlength) indicates high alpha helical content (as expected from NMR solution structure; PDB 1S62).

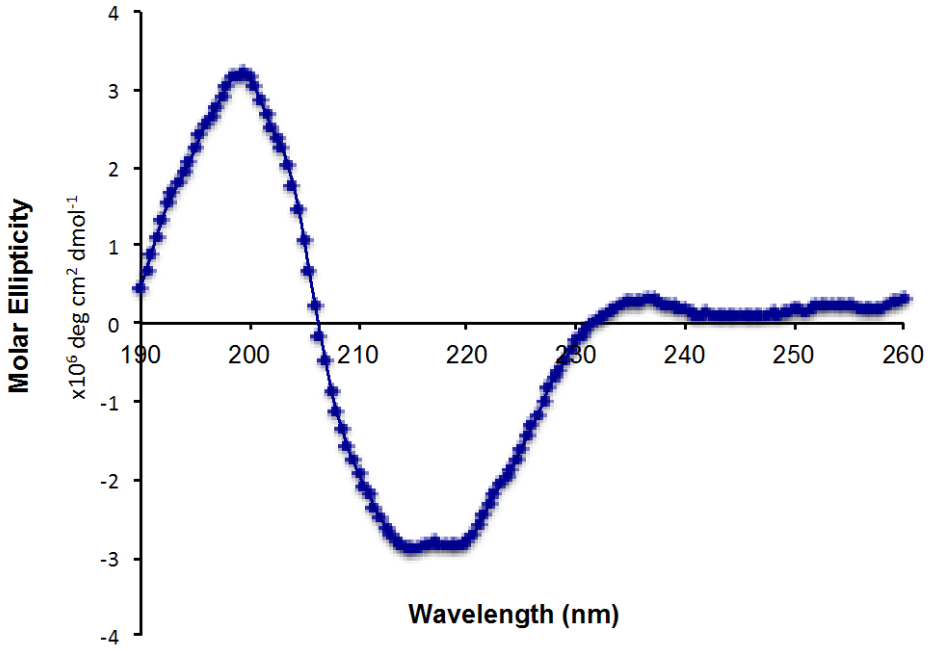


Figure 3.9 Far UV Circular Dichroism spectrum of *E.coli* TolB (pDAB18 construct). CD spectrum recorded for 30 µM eTolB in 10 mM Sodium phosphate, pH7 at 20 °C (0.1 mm pathlength).

3.2.2 Investigating the *E.coli* TolA, TolB, Pal interactions by Surface Plasmon Resonance

To investigate the *E.coli* TolA3-TolB-Pal interaction (figure 3.10) with surface plasmon resonance, *E.coli* TolA mutant consisting of domains 2 (long helical domain) and 3 (C-terminal globular domain involved in interaction with eTolB) with additional N-terminal Cysteine residue (at end of long helical domain 2, residues 74-421, pAK123 construct, supplied by Dr Anne-Marie Krachler) was immoblised on 1 of 2 flow channels on a C1 SPR chip (see section 3.2.2.1 for details on immobilisation of eTolA23 to C1 chip, and appendix section 7.5 for details of chemistry used for protein immobilisation on SPR chip). This immobilised TolA was then challenged with a number of potential binding partners; free eTolB (section 3.2.2.2), free ePal (section 3.2.2.3) and a pre-formed complex of eTolB-ePal (section 3.2.2.4). A binding event between the immobilised eTolA23 and the challenging protein would cause a change in response units as detected by Biacore T-100 biosensor from which affinities of binding could be calculated. The pAK123 construct that consists of a cysteine on the end of the long helical domain 2 is advantageous for this work as it ensures that the domain 3 of TolA (distal to the cysteine) is distant from the chip surface, and thus available for eTolB binding.

In addition to eTolA23 immobilisation on flow channel 1, a second channel was prepared in an identical manner as eTolA23 channel. However, instead of protein, the second flow channel was blocked with free cysteine. This ensured that any change in response measured by SPR was due to genuine protein binding and not non-specific interactions with C1 chip surface. This is particularly relevant as previous SPR work investigating the binding interactions of immobilised eTolB had used a CM5 (dextran coated) chip (Hands et al. 2005). This approach was not used in this work as anecdotal data collected when challenging eTolB to bind eTolA3 immobilised on CM5

chip indicated that eTolB binds with equal affinity to both reference and protein channel (Dr Anne-Marie Krachler, personal communication). Therefore, a C1 chip was chosen for this work, as it only presents single carboxyl groups on its surface for immobilisation chemistry.

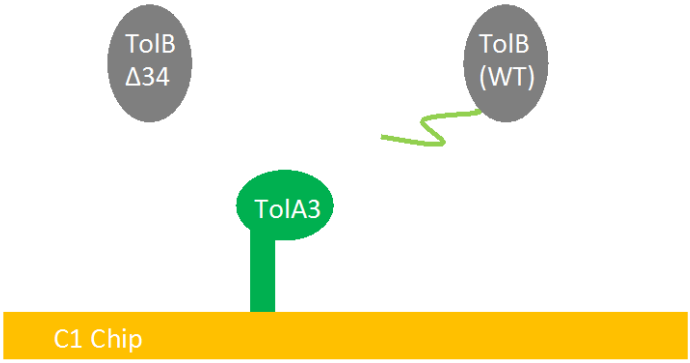


Figure 3.10 Schematic diagram of *E.coli* TolA/TolB SPR experimental setup. TolA (dark green) is immobilized onto C1 SPR chip and challenged with wild type TolB and Δ34 TolB mutant (that lacks the disordered N-terminus, shown here in light green).

Once immobilised on C1 chip, eTolA23 was challenged with increasing concentrations of wild type eTolB and Δ34 eTolB (lacking N-terminal strand) as well as ePal, and eTolB in pre-made complex with ePal. Work previously published reported an affinity of ~40 μM for wild type eTolB – eTolA3 interaction via ITC. No binding was detected for either Δ34 eTolB, ePal, or ePal-eTolB complex (Bonsor et al. 2009).

3.2.2.1 *E.coli* TolA was successfully immobilised on a C1 SPR chip.

To challenge *E.coli* TolA domain binding via SPR, eTolA23 (pAK123) was immobilised onto a C1 chip. Briefly, C1 SPR was prepared and activated to thiol-couple eTolA23 to the chip (see appendix section 7.5 for details). Subsequently, any non-coupled (but active) groups were blocked with free

cysteine, and the immobilised eTolA23's binding efficiency was challenged with a test injection of eTolB (figure 3.11). The reference channel was also prepared by an identical protocol, with the exception that no protein was applied, and all activated groups on reference channel were blocked with free cysteine. See section 2.3.21 for full details of SPR chip immobilisation method. Typically 300 Response Units's (RU) of eTolA23 was immoblised on C1 chip.

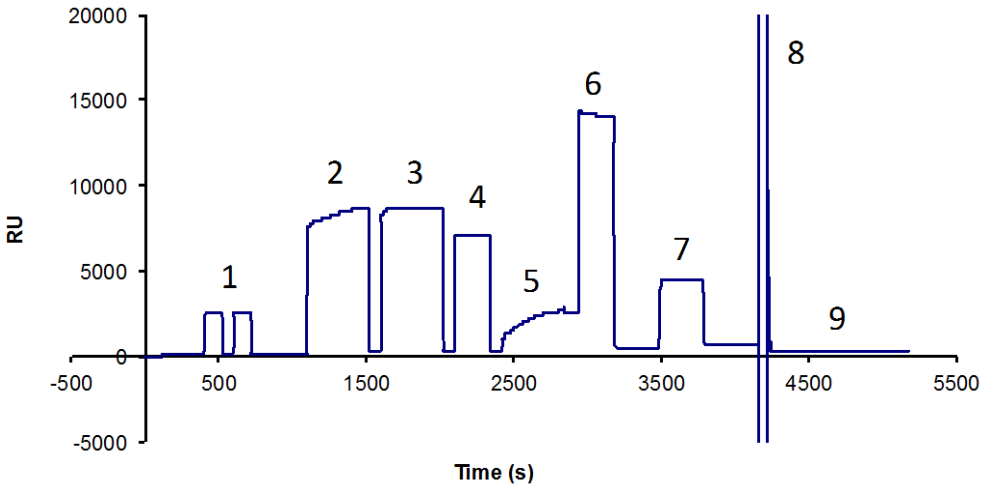


Figure 3.11 Immobilisation of eTolA23 on C1 SPR chip. (1) Application of 0.1M glycine pH 12 with 0.3 % triton X-100. (2) Application of EDC:NHS. (3) Application of Ethylenediamine. (4) Application of Sulfo-GMBS. (5) Application of eTolA23 (~2000 RU's). (6) Cysteine block step. (7) 100 µM Test eTolB injection. (8) 6M GnHCl unfolding/regeneration. (9) Final (stable) immobilised eTolA23 (~300 RU's).

3.2.2.2 The dependence of the *E.coli* TolA/TolB interaction on the N-terminus of TolB is confirmed by Surface Plasmon Resonance.

Having successfully immobilised eTolA23 onto C1 chip, and the integrity of the chip verified via test injections of eTolB, immobilised eTolA23 was challenged with both wild type eTolB, and Δ34 eTolB mutant, which lacks the N-terminal 12 amino acids (figure 3.10) under equilibrium conditions.

Increasing concentrations of appropriate eTolB from 1-150 μM were titrated against the immobilised eTolA23 (as well as simultaneous titrations against reference channel), monitoring for change in response units of chip (an increase in response units is indicative of a protein complex forming between eTolA23 and titrated protein). Following the binding/unbinding event, any tightly bound protein complex that formed was dissociated by unfolding the proteins with 6 M GnHCl. Once unbound, titrated protein had been washed away, immobilised protein was refolded in running buffer (50 mM Hepes, 50 mM NaCl, pH 7.5, 0.02% P20 detergent). Periodic test injections of 100 μM eTolB were also titrated against chip in order to verify the condition of eTolA23, to ensure that 6 M GnHCl step did not cause permanent denaturation of protein. Average refolding efficiency of eTolA23 over each titration data set (full titration of 1-150 μM) was greater than 95%. Results for both wild type and mutant eTolB titrations are shown in figures 3.12 and 3.13.

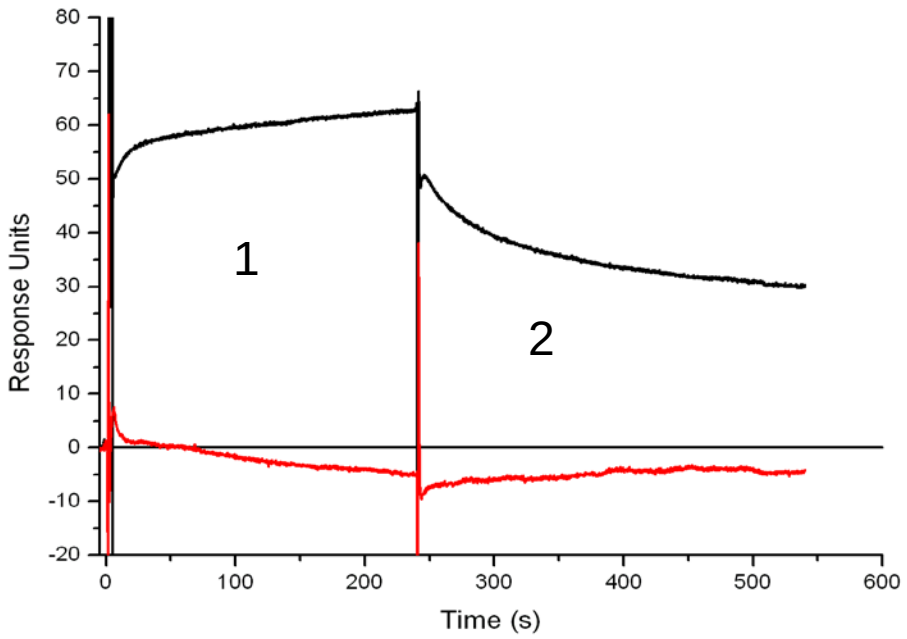


Figure 3.12 SPR sensorgrams of eTolA23 association/dissociation with wild type eTolB and Δ34 eTolB mutant. (1) Association phase, (2) Dissociation phase. Black sensorgram indicates change in response units for association and dissociation of wild type eTolB with immobilised eTolA23.

Shape of sensorgram is indicative of fast on, slow off binding of a weak protein complex with 1:1 (Langmuir) binding (Schuck 1997). Red sensorgram indicates no binding between eTolA23 and Δ34 eTolB mutant as it lacks characteristic association/dissociation phases present in black sensorgram. Both eTolB variants at 150 μM in 50 mM Hepes, 50 mM NaCl, pH 7.5, 0.02% P20 detergent, flow rate 10 μl/min, temperature: 20° C.

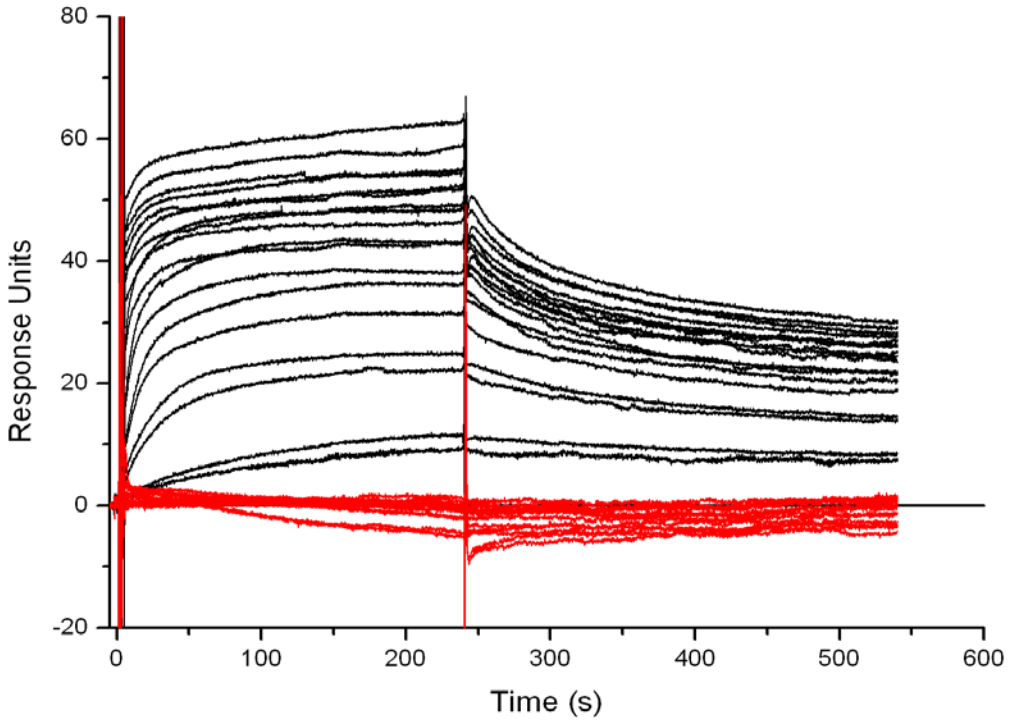


Figure 3.13 SPR titration sensorgrams of eTolA23 association and dissociation with various concentrations of wild type eTolB and Δ34 eTolB mutant. Black sensorgrams indicate change in response units for association and dissociation of wild type eTolB with immobilised eTolA23 with increasing eTolB concentration, from 1-150 μM. Red sensorgram for eTolA23 and Δ34 eTolB indicates no detectable binding compared to wild type eTolB (black sensorgram) in identical conditions. Conditions: 50 mM Hepes, 50 mM NaCl, pH 7.5, 0.02% P20 detergent, flow rate 10 μl/min, temperature: 20°C. Protein injected for 300 s, followed by dissociation for 600 s. Chip surface regenerated with 6 M Guanidine-HCl for 60s following each titration.

As seen in figure 3.12, black trace of 150 μM eTolB titrated against eTolA3 shows characteristic association/dissocation shape displayed when protein binding occurs and is detected by SPR (Schuck 1997). In addition, as shown in figure 3.13, with increasing concentration of eTolB (from 0-150 μM), there is a progressive increase in response units. Saturation is reached by 150 μM. Conversely, at equivalent concentrations of Δ34 eTolB mutant (red trace), no increase in response units is correlated to increase in eTolB concentration.

This data confirm ITC experiments that the N-terminus of eTolB is important for eTolB to interact with eTolA. This titration was then fitted using single site binding model (Biacore) to calculate dissociation constant (figure 3.14), which was estimated as 10 μM (± 2 μM, calculated from 6 replicate experiments). This was in agreement with dissociation constant for eTolA3-eTolB as estimated from ITC data was 40 μM (Bonsor et al. 2009). These SPR data confirmed not only the importance of the N-terminus of eTolB, but also that the eTolA-eTolB is a weak interaction *in vitro.* It should be noted that during dissociation phase, sensorgram does not return to zero RU's, and as such it is not possible to analyse this data in terms of kinetics of eTolA23-eTolB interaction. Increase or decrease of flow rate did not improve quality of sensorgrams.

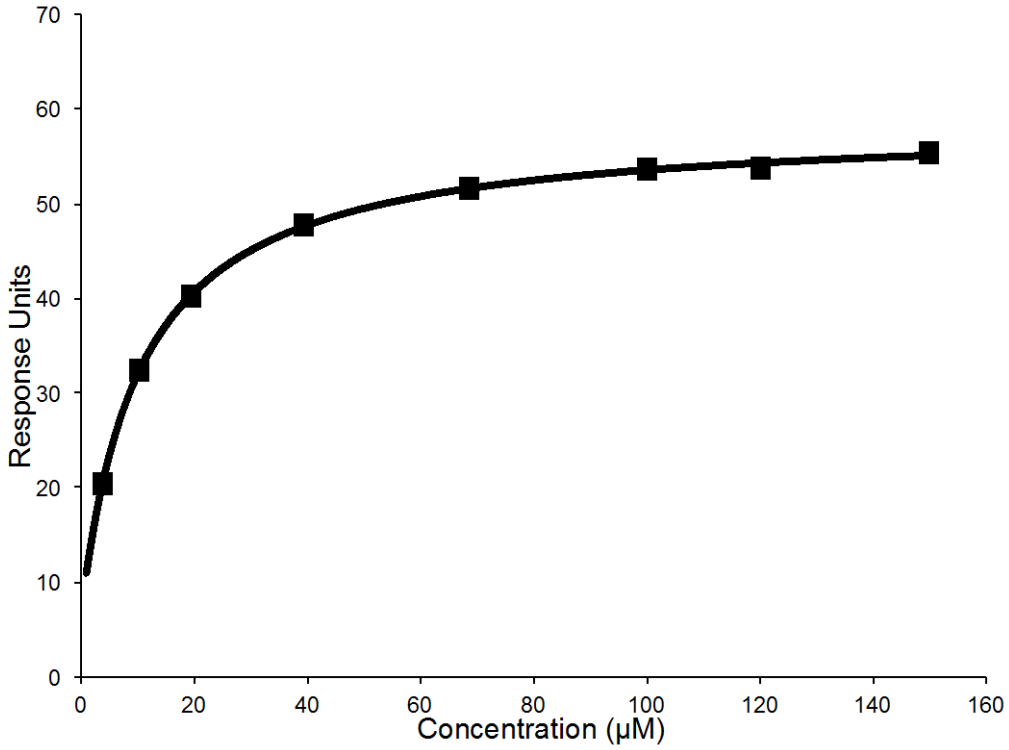


Figure 3.14 SPR titration data of eTolA23 association with various concentrations of wild type eTolB at maximal response. Maximal response is measured for each eTolB (wild type) concentration and plotted against appropriate concentration. Saturation of the TolA23 appears to be achieved, as higher TolB concentration response units plateau. When fitted single site binding model (Biacore) data yields calculated Kd of 10 μM ± 2 μM.

3.2.2.3 No binding is detected *in vitro* between *E.coli* TolA and *E.coli* Pal by surface plasmon resonance.

Having confirmed the importance of N-terminus of eTolB in it's interaction with eTolA, the next step was to investigate the ability of ePal to bind eTolA *in vitro*. Although is was previously reported that eTolA interacted with ePal (Cascales et al. 2000), this finding has been disputed as no complex was found between these two proteins by formaldehyde crosslinking, ITC or NMR (Bonsor et al. 2009). Using the same experimental setup as that investigating the interactions between eTolA and eTolB, immobilised

eTolA23 was titrated against with ePal (from 0-150 μM, figure 3.14). As can be seen in figure 3.15, no increase in response units was detected when titrated against with ePal, indicating no complex between eTolA23 and ePal was formed. Validity of eTolA23 binding was verified during experimental timeline with periodic 100 μM eTolB test injections onto chip to confirm the ability of immobilised eTolA23 to bind eTolB (data not shown).

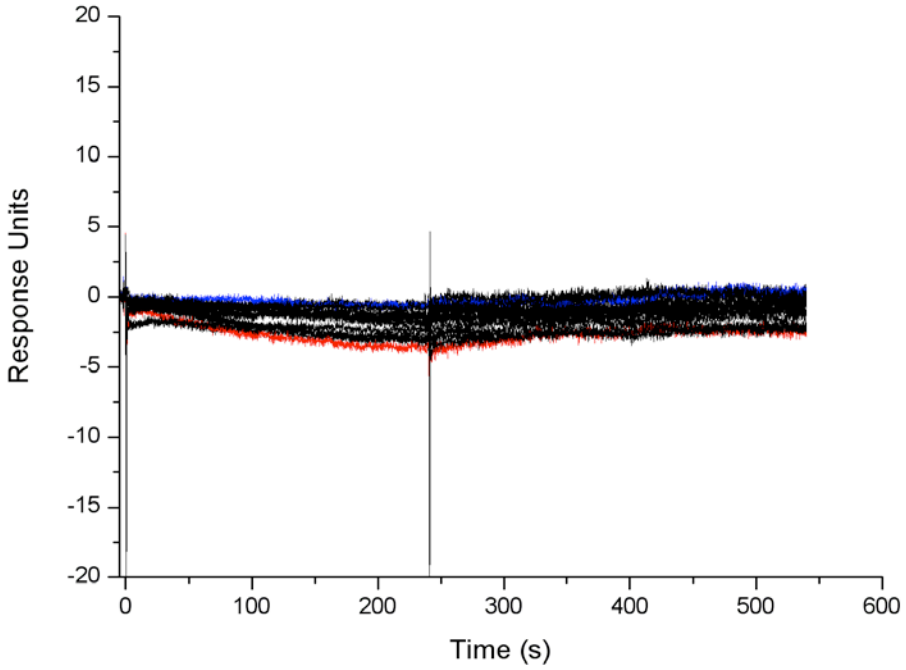


Figure 3.15 *E.coli* TolA23 immobilised on C1 SPR chip does not interact with *E.coli* Pal. Black sensorgrams indicate ePal at concentrations from 1-150 μM. Blue sensorgram indicates 1 μM titration, red indicates 150 μM titration. Conditions: 50 mM Hepes, 50 mM NaCl, pH 7.5, 0.02% P20 detergent, flow rate 10 μl/min, temperature: 20 °C. Protein injected for 300 s, followed by dissociation for 600 s. Chip surface regenerated with 6 M Guanidine-HCl for 60s following each titration.

3.2.2.4 *E.coli* Pal influences the interaction of *E.coli* TolB with *E.coli* TolA

Work published by Bonsor et al. in 2009 suggested that the binding of *E.coli* Pal to *E.coli* TolB acted as an allosteric switch by driving the N-terminus of TolB from a disordered to an ordered state, and thus mediating it's interaction with TolA3. SPR titration experiments against immobilised eTolA23 were conducted, investigating the effect of ePal on the eTolA-eTolB interaction. Using the same experimental setup as that investigating the interactions between eTolA and eTolB, immobilised eTolA23 was titrated against with a pre-made complex of eTolB-ePal (from 1-150 μM of eTolB, with 1.1:1 ePal:eTolB molar ratio to ensure no free eTolB was available in solution). As can be seen in figure 3.16, no increase in response units was detected when titrated against with ePal, indicating no complex between eTolA23 and eTolB-ePal complex was formed. Validity of eTolA23 binding was verified during experimental timeline with periodic 100 μM eTolB test injections onto chip to confirm the ability of immobilised eTolA23 to bind eTolB (data not shown).

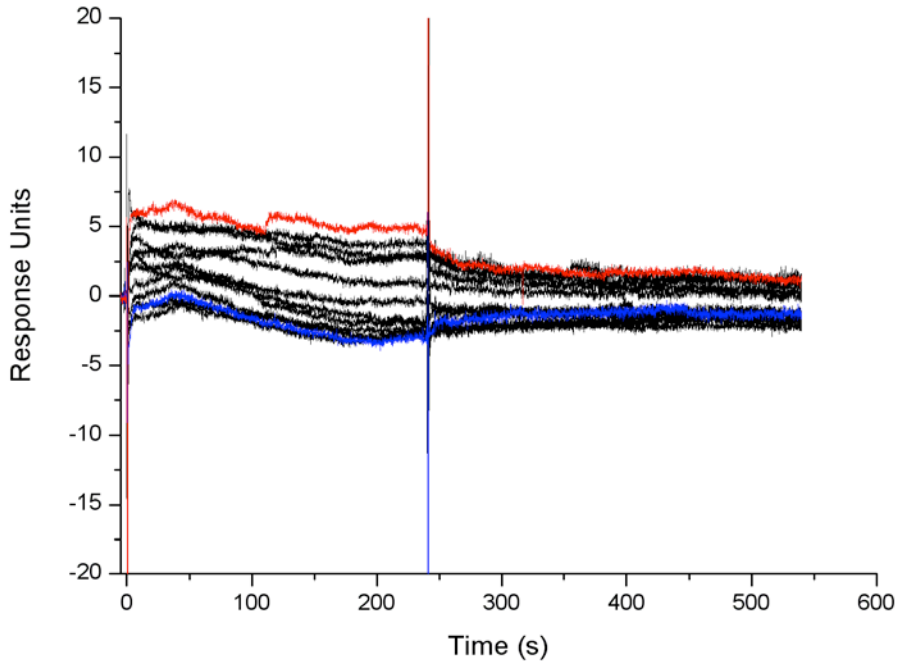


Figure 3.16 *E.coli* TolA23 immobilised on C1 SPR chip does not interact with *E.coli* TolB in pre-made complex with Pal. Black sensorgrams indicate no binding between eTolA23 and eTolB-ePal complex at concentrations from 1-150 μM. Blue sensorgram indicates 1 μM titration, red indicates 150 μM titration. Conditions: 50mM Hepes, 50mM NaCl, pH 7.5, 0.02% P20 detergent, flow rate 10μl/min, temperature: 20 °C. Protein injected for 300s, followed by dissociation for 600s. Chip surface regenerated with 6M Guanidine-HCl for 60s following each titration.

3.3 Complementing *E.coli* TolB with *E.coli* TolA to drive colicin E9 uptake.

As shown in previous sections (3.2.2.2), as well as in the work of Bonsor et al (2009), the N-terminus of eTolB is required for its interaction with eTolA and that it must be in the disordered state to be available for binding. In addition, colicin E9 requires TolB to translocate across the outer membrane and into the cell. This is achieved through the binding of the E9 translocation domain (T-domain) in the beta-propellor domain of TolB, in a similar way to

Pal. However, whereas Pal binding TolB prevents TolA binding, T-domain binding to TolB promotes disorder in the N-terminus, thus promoting TolA binding (Bonsor et al. 2009). Some other colicins however, do not require TolB for translocation. colicin N is known to be solely dependent on eTolA to drive it's uptake and ultimately it's ability to kill invaded cells (Hecht et al. 2009). Colicin A, is dependent on both eTolA and eTolB (Gokce et al. 2000) and has separate eTolA and eTolB binding epitopes. Although both colicin N and A interact with TolA3 in a TolB-independent manner, they do not share a common binding site. The colicins interact with TolA3 on opposite faces of the protein. The colicin N binding face is shared with bacteriophage g3p (figure 3.17) (Hecht et al. 2010; Li et al. 2012).

The work reported in this section tested a series of hypotheses. Firstly I set out to determine if the sequence from TolB's N-terminus known to be required for the interaction with eTolA could drive an *in vitro* interaction with proteins known to not directly interact with eTolA, such as colicin E9. The second objective was to determine if this sequence was required to be extremely N-terminal, as it is on *E.coli* TolB, or if location was not relevant. The final aim was to determine that providing an interaction was found *in vitro,* if colicin E9-TolB fusion proteins were capable of killing cells *in vivo*. This would prove that by bypassing TolB and interacting with TolA directly, TolA is capable of directly and independently driving the translocation of the colicin into the cell.

Three cloning strategies (figure 3.18) were devised; the first clone would replace the eTolB binding epitope of colicin E9 with the sequence known to interact with eTolA3 from the N-terminus of eTolB (sequence: EVRIVIDSGVDS); the second construct would replace the OmpF binding site 1 with the sequence known to interact with eTolA3 from the N-terminus of eTolB, and the eTolB binding epitope would be disrupted; the third construct would replace the OmpF binding site 1 with the sequence known to

interact with eTolA3 from the N-terminus of eTolB, as well as maintaining the eTolB binding epitope. These genes would then be expressed and proteins purified. To determine if the eTolA3 binding epitope for the N-terminus of eTolB is sufficient create a novel interaction between eTolA3 and the ColE9-eTolB fusion proteins, the eTolA3 and eTolB *in vitro* binding was probed via ITC, and secondly, these fusion proteins were challenged with *in vivo* cell killing assays, to determine if the interaction with eTolA3 is sufficient to drive entry of the colicin E9 into the cell, thus killing the cell.

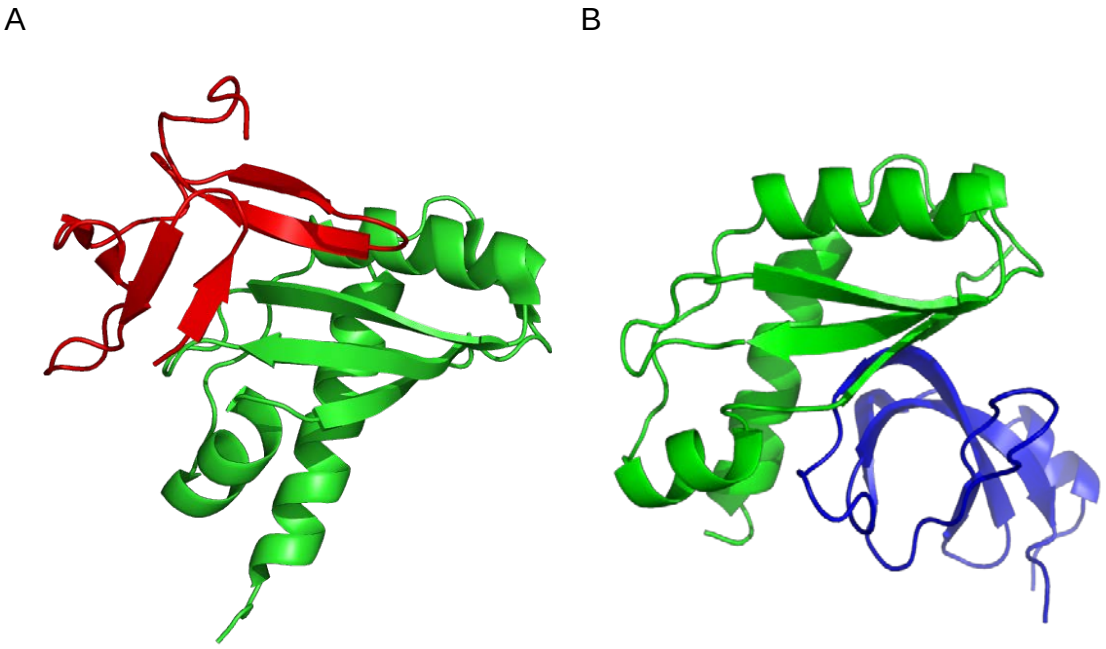


Figure 3.17 Colicin A and bacteriophage g3p bind on opposite sides of *E.coli* TolA domain 3. (A) Crystal structure of colicin A (residues 53-107, red) in complex with domain 3 (residues 302-421, green) of *E.coli* TolA. PBD ID: 3QDR (Li et al. 2012). (B) Crystal structure of bacteriophage g3p (1-86, blue), in complex with domain 3 (residues 295-421) of *E.coli* TolA. PDB ID: 1TOL (Lubkowski et al. 1999).

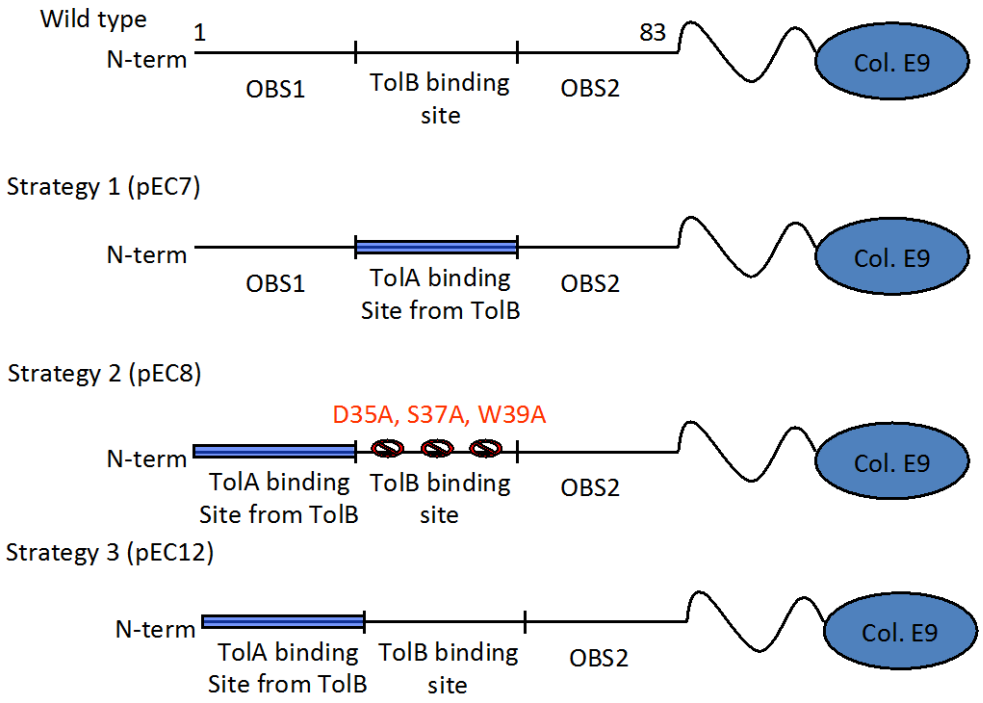


Figure 3.18 Summary of colicin E9-*E.coli* TolB fusion protein constructs. Wild type is as encoded by pCS4 construct. Residues 1-83 contain OmpF binding site 1 (OBS1), TolB binding site, and OmpF binding site 2 (OBS2). EC7 fusion (pEC7 construct) consists of *E.coli* TolA binding epitope from disordered N-terminus of *E.coli* TolB replacing the colicin E9 TolB binding epitope. EC8 fusion (pEC8 construct) consists of *E.coli* TolA binding epitope from disordered N-terminus of *E.coli* TolB replacing the OBS1, and TolB binding site being knocked out by mutating 3 key residues; D35A S37A, W39A (Garinot-Schneider et al. 1997). EC12 fusion (pEC12) consists of *E.coli* TolA binding epitope from disordered N-terminus of *E.coli* TolB replacing the OBS1, with TolB binding epitope being maintained.

3.3.1 Purification of colicin E9-TolB fusion proteins

Colicin E9-TolB mutants (pEC7, pEC8, pEC12 constructs) were purified as described in section 2.3.5-2.3.6. Briefly, BL21 (DE3) cells were transformed with relevant plasmid and grown up in 4.8L LB media for purification.

Example purification (EC12 fusion protein) is displayed in figure 3.19. Other purified fusion proteins shown in figure 3.20. Typical yield for each protein was 50 mg/ml. Purified protein mass was confirmed by electrospray ionisation mass spectrometry (ABI Qstar instrument).

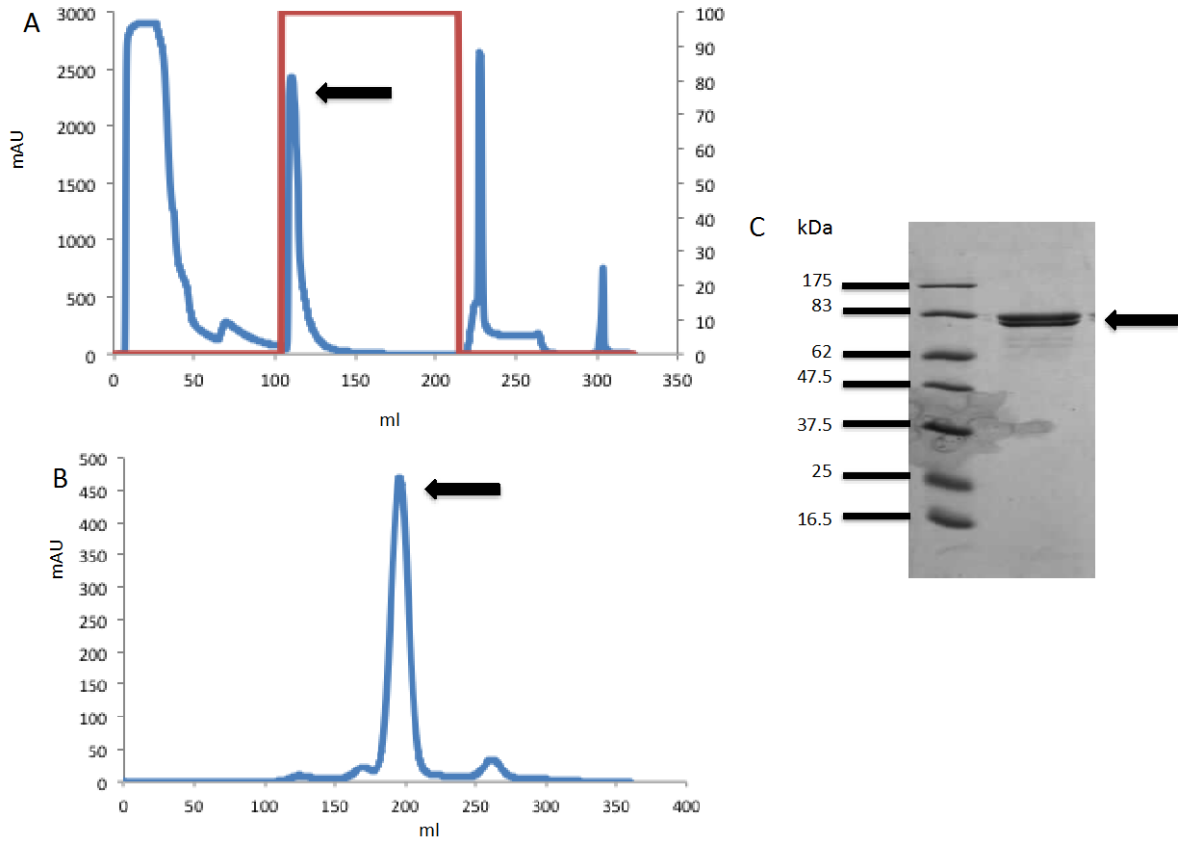


Figure 3.19 Purification of colicin E9-TolB fusion EC12 (pEC12 construct) (A) 280nm Absorbance profile from Ni-NTA affinity chromatography. Protein (indicated by arrow) eluted with 6 M GnHCl step elution (red). (B) Gel-filtration of EC12 fusion on Superdex 200 26/60 column. EC12 eluted between 175 and 200 ml. (C) 13% SDS-PAGE gel of pure pooled colicin E9 mutant EC12 following S200 gel filtration.

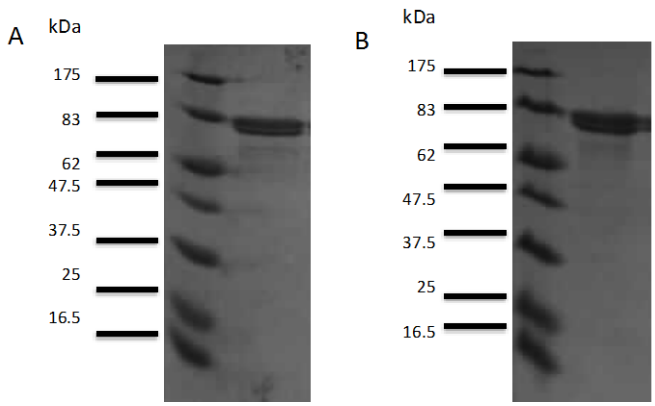


Figure 3.20 Purity of colicin E9-TolB fusion proteins. (A) Fusion protein EC7 (pEC7 construct: OBS1-TolA site-OBS2). (B) Fusion protein EC8 (pEC8 construct: (TolA site-*ΔTolB*-OBS2). Analysed on 13% SDS-PAGE gels stained with coomassie blue.

3.3.2 The *E.coli* TolA binding epitope of *E.coli* TolB is sufficient for the *in vitro* interaction between colicin E9-TolB fusion proteins and *E.coli* TolA domain 3.

To determine if ColE9-TolB fusion proteins were capable of interacting with eTolA3, and if the location of the binding epitope had any impact on the binding affinity, ITC titration data were collected for each fusion titrated with eTolA3 and eTolB (figure 3.21, table 3.2), in the absence of calcium ions (Loftus 2006).

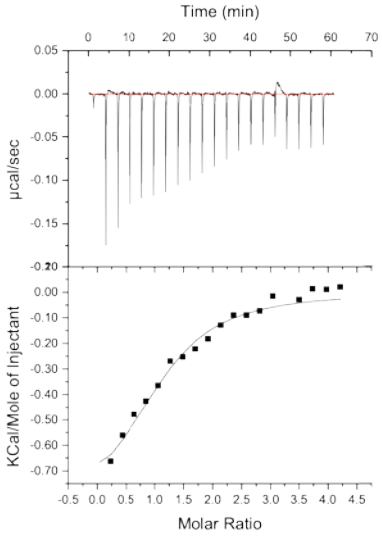
Low heats of binding were detected, and a K_d of 14 (±4) μM, 39 (±9) μM and 37 (±7) μM for fusions EC7, EC8 and EC12 when titrated against with eTolA3 were calculated (figure 3.21 A, B, C, respectively). These are comparable numbers as obtained for native eTolA3-eTolB ITC data (43 ± 2 μM, Figure 3.20E and (Bonsor et al. 2009) and SPR data reported above (10 μM) (section 3.2.2.2). As binding isotherms (figures 3.21 A, B, C, E) are not of the typical sigmoidal shape (figure 3.4), there is potential for error in the ΔH and stoichiometry estimations. It is interesting to note that the location of the TolA binding epitope appears to be relatively unimportant, given that

both orderings of the disordered T-domain yielded similar thermodynamic data. The fusion proteins had differing thermodynamics to that of the native TolB interaction, most likely due to the fusion already being in disordered confirmation, whereas the TolB N-terminus is likely in a state of dynamic equilibrium between ordered and disordered state. Additionally, when EC12 fusion was titrated against eTolB (figure 3.21, D), the K_d was calculated as 450 nM, which is again similar to the published values (~1 μM in the absence of Ca^{2+} ions) for wild type eTolB with colicin E9 as measured by ITC (Loftus 2006).

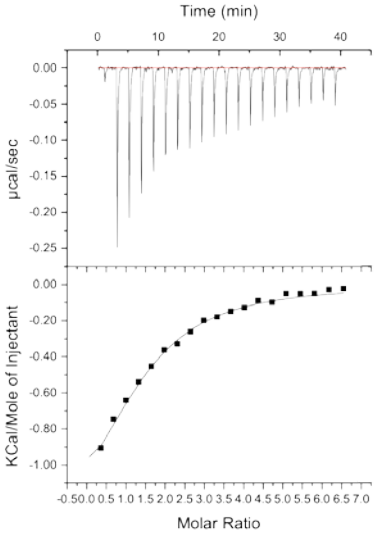
Components titrated against eTolA3	**ΔH (kcal mol^{-1})**	**ΔS (cal K^{-1} mol^{-1})**	N	**K_d (μM)**
(A) EC7 fusion protein (OBS1-TolA site-OBS2)	- 0.88 (± 0.01)	+ 19.3 (±0.14)	1.12 (±0.09)	14 (±4)
(B) EC8 fusion protein (TolA site-*ΔTolB*-OBS2)	- 1.57 (±0.02)	+ 14.9 (± 0.08)	1.40 (±0.08)	39 (±9)
(C) EC12 fusion protein (TolA site-WT TolB site-OBS2)	- 1.42 (± 0.03)	+ 15.5 (± 0.12)	1.39 (± 0.2)	37 (±7)
(D) Control: eTolB titrated against EC12	- 7.13 (± 0.01)	+ 5.1 (± 0.05)	1.21 (±0.01)	0.45 (±0.01)
(E) eTolB (Bonsor 2009)	+ 2.92 (±0.02)	+ 30.0 (±0.1)	1.00 (±0.03)	43 (±2)

Table 3.2 Comparision of eTolB-ColE9 fusion/wild type eTolB protein affinities for eTolA3.

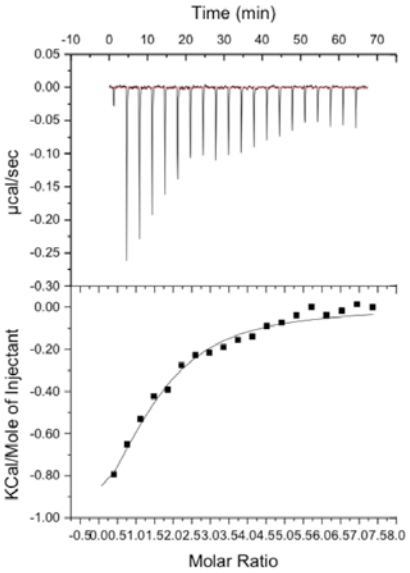
A


Time (min)
µcal/sec
KCal/Mole of Injectant
Molar Ratio

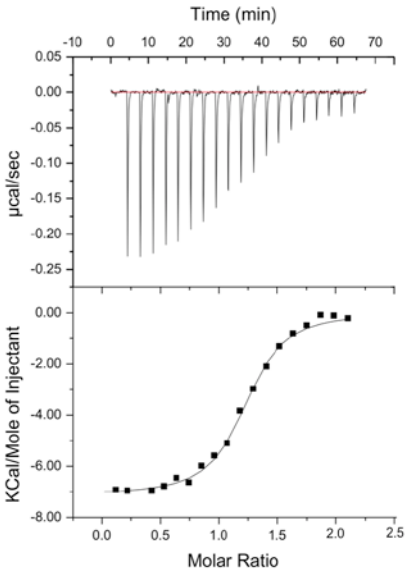
B


Time (min)
µcal/sec
KCal/Mole of Injectant
Molar Ratio

C


Time (min)
µcal/sec
KCal/Mole of Injectant
Molar Ratio

D


Time (min)
µcal/sec
KCal/Mole of Injectant
Molar Ratio

E

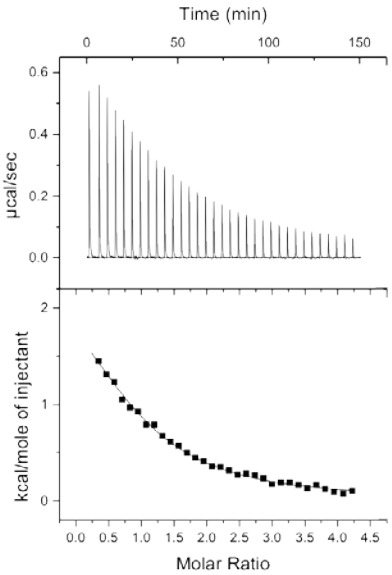


Figure 3.21 ITC data for *E.coli* TolA domain interaction with colicin E9-TolB fusion proteins. (A) eTolA3 (800 µM) titrated in 19 2 µl into colicin E9-TolB fusion protein EC7 (40 µM). K_d determined as 14 µM. (Peak integrations adjusted for noise, NDH +0.4). (B) eTolA3 (800 µM) titrated in into colicin E9-TolB fusion protein EC8 (40 µM).) K_d determined as 39 µM. (Peak integrations adjusted for noise, NDH +0.2). (C) eTolA3 (800 µM) titrated in into colicin E9-TolB fusion protein EC12 (40 µM).) K_d determined as 37 µM. (Peak integrations adjusted for noise, NDH +0.2). (D) eTolB (200 µM) titrated into colicin E9-TolB fusion protein EC12 (20 µM). K_d determined as 450 nM. (Peak integrations adjusted for noise, NDH +0.2). (E) 1.3 mM eTolA3 titrated in 34 x 8 µl injections into 60 µM of wild type eTolB. K_d determined as 43 µM. Titrations conducted at 20 °C with 240 s spacing (270 s spacing for (E)) between each injection of eTolA3 (Bonsor 2009). Proteins dialysed in 50 mM Hepes, 50 mM NaCl, pH 7.5 prior to experiment. Data fitted to single site binding model using Origin 8.0 (Microcal/GE Healthcare).

3.3.3. Colicin E9-TolB fusion proteins cannot kill *E.coli* cells without the presence of the *E.coli* TolB binding epitope.

To determine if the TolA binding epitope from TolB is sufficient to for the entry of colicin E9 into *E.coli* cells, cell-killing assays were performed (see section 2.3.25 for details). JM83 cells were challenged with various concentrations of both wild type colicin E9 and colicin E9-TolB fusion proteins. If cells died (creating a zone of clearance) this would mean that colicin E9-TolB fusions interacted with eTolA *in vivo* thereby bypassing periplasmic TolB and driving translocation of the colicin into the cell.

As shown in figure 3.22, fusion proteins EC7 and EC8 did not kill cells, even at very high concentrations (160 μM). EC12 fusion protein does kill, however, killing is severely attenuated. This attenuated killing is due to the OmpF binding site being replaced with TolA binding epitope. This attenuated phenotype has been previously reported by Housden et al (Housden et al. 2005), whereby colicin E9 lacking the first OmpF binding site could not bind OmpF at the outer membrane as efficiently as wild type, and therefore had reduced cell killing ability (although killing was not abolished completely). As neither EC7 or EC8 kills cells, whereas EC12 does, this indicates that colicin E9 remains dependent on eTolB for translocation, and ultimately cell killing, and that eTolA is not sufficient to drive translocation.

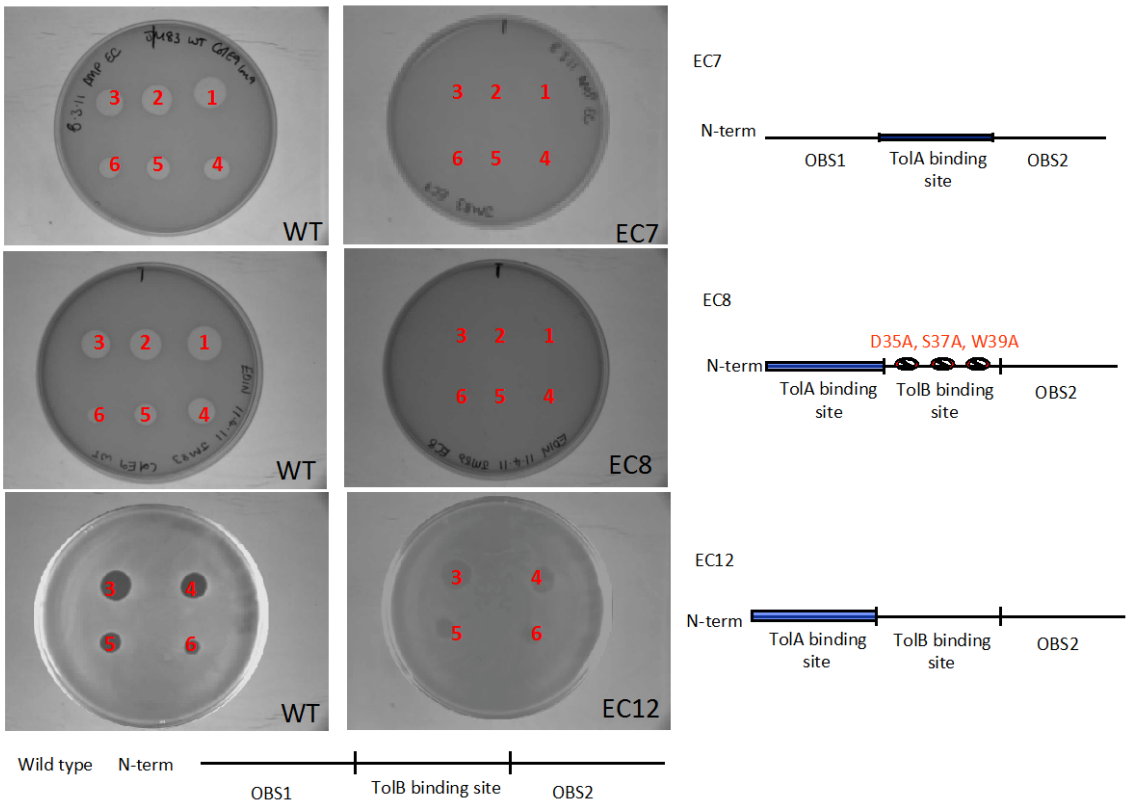


Figure 3.22 *In vivo* cell killing assay testing ability of colicin E9-TolB fusion proteins ability to kill JM83 cells. 2 µl wild type colicin E9 (WT) or colicin E9-TolB fusion proteins were spotted onto JM83 top agar plates in the following concentrations; 1) 160 µM 2) 30 µM 3) 6 µM 4) 1.2 µM 5) 250nM 6) 50nM. Fusion proteins EC7 and EC8 do not kill, even at high concentrations (160 µM). Fusion protein EC12 does have partial killing activity at concentrations down to 50 nM.

3.4 Discussion

3.4.1 The *E.coli* TolA-TolB-Pal interaction

Previously published work (Bonsor et al. 2009) had suggested that not only did *E.coli* TolB interact with TolA domain 3 via it's intrinsically disorder N-terminus, but that the binding of *E.coli* Pal to TolB mediated the equilibrium

position for the N-terminus. NMR data suggested that when bound *E.coli* TolB was bound to Pal the N-terminus moved from a dynamic position between disordered and ordered state, to a state of order (bound back position) that prevented TolA domain 3 from binding. This work also suggested that when the translocation domain of colicin E9 bound TolB (in a common binding site to that of Pal), the N-terminus moved into a state of disorder, and thus promoted the interaction between TolB and TolA. Finally, contradictory to work published regarding a potential interaction between *E.coli* TolA and Pal (Cascales et al. 2000), no evidence of an interaction between these two proteins was found in work published by Bonsor et al. To independently validate the dependence of the *E.coli* TolA-TolB interaction's dependence on the N-terminus, as well as determining if Pal interacts with eTolA3 in a manner not detectable through ITC and the effect of Pal on TolB's ability to bind TolA3, SPR experiments were performed.

Having successfully immobilised a mutant of *E.coli* TolA23 on a C1 SPR chip, this TolA was challenged with 3 potential binding partners; TolB alone, Pal alone and TolB in complex with Pal. In agreement with work published by Bonsor et al., TolB was found to bind TolA, and a K_d of ~10 μM was calculated. This value is similar to the value determined by ITC (~40 μM) (Bonsor et al. 2009). In addition, and again in agreement to the work of Bonsor et al., not only was no binding detected by SPR for either TolA titrated against with Pal, and also TolB in complex with Pal. Given that these data that show that TolB cannot bind TolA when in complex with Pal again confirms the role that Pal plays controlling the ablity of TolB to bind TolA (Bonsor et al. 2009).

This work, and the work of Bonsor et al., disagrees with work by Cascales et al. (Cascales et al. 2000 and Cascales et al. 2002) suggesting an interaction between TolA and Pal. In addition, yeast 2 hybrid experiments published by Walburger et al. (Walburger et al. 2002) did not find any interaction between

TolA and Pal. However, as the work by Cascales et al. also detected a complex between TolA and Pal *in vivo* by coimmunoprecipitation of crosslinked TolA and Pal. It is therefore possible that the *E.coli* TolA-Pal interaction is dependent on the proton motive force, which is not present in the *in vitro* work presented in this thesis. Cascales et al. suggest that the TolA-Pal interaction is indeed PMF dependent, and that it can be abolished with carbonyl cyanide m-chlorophenylhydrazone (CCCP). Alternatively, as the coimmunoprecipitation was performed *in vivo* with chemical crosslinking, it is possible that a terniary complex between TolA-TolB and Pal was captured. However, in their work published in 2000, Cascales et al. also state that they can detect a TolA-Pal complex *in vitro*, something that both the work by Walburger et al., Bonsor et al. and this work, do not. Given the work previously published in this area (chemical crosslinking, ITC and NMR), as well as the results from this work obtained through SPR, and given the limitations of the experimental techniques of work by Cascales et al., it must therefore be concluded that there is no interaction between *E.coli* TolA and Pal, *in vitro*.

3.4.2 Creating novel *E.coli* TolA domain 3 interactions with colicin E9-TolB fusion proteins

Having confirmed that the *E.coli* TolA-TolB interaction is dependent on the N-terminus of TolB, I next set out to investigate whether or not this short 12 residue sequence known to bind *E.coli* TolA domain 3 was sufficient to generate a novel interaction with TolA domain 3 to a protein that it does not normally bind (in this case, colicin E9). The aim of this work was two-fold; firstly to determine if *in vitro* the *E.coli* TolB N-terminal sequence could drive a *de novo* interaction between proteins that would not normally interact. Secondly, it was to deterimine if these proteins (*E.coli* TolA and colicin E9) could interact, and if this interaction was sufficient to bypass *E.coli* TolB and drive translocation of colicin E9 into the cell and cause cell killing, *in vivo*.

Using *E.coli* TolA as the sole translocation mechanism for colicin entry into the cell has been reported for colicin N, which does not bind *E.coli* TolB (Gokce et al. 2000). Bacteriophage g3p entry protein has also been reported to be solely dependent on TolA for entry into the cell, and thus ultimately causing cell death. In addition, colicin A is also dependent on TolA for cell killing, however it is also dependent on TolB, although the interaction between Col A and TolA is not dependent on TolB, unlike colicin E9, which requires an interaction with TolB to drive an interaction with TolA to allow translocation into the cell, and cause cell death (Hecht et al. 2010).

Through *in vitro* ITC experiments it was confirmed that the novel colicin E9-TolB fusion proteins could interact with *E.coli* TolA3, and that the interactions were similar to that of wild type eTolB-eTolA3 (40 μM), with K_d calculated as between 14-39 μM. In addition, not only does it appear as though the 12 residues in isolation from the N-terminus of eTolB are sufficient to create an interaction with eTolA3 *in vitro,* but it is also not necessary for the eTolA3 binding epitope to be N-terminal, as it is in eTolB. Whether or not the eTolA3 binding epitope must be within a disordered region of a protein (as is the case of eTolB and these colicin E9-eTolB fusions) or if it could be presented as part of a structured domain is unknown.

The *in vivo* cell killing results however suggested that colicin E9 requires the presence of TolB to translocate across the outer membrane, periplasm and inner membrane to kill cells through its DNase activity. Fusion proteins EC7 and EC8 which do not contain the TolB binding epitope, could not kill cells *in* vivo, despite being able to bind *E.coli* TolA3 *in vitro.* Fusion protein EC12 could interact with both *E.coli* TolA3 and TolB*,* and this was the only fusion protein that could kill cells. Thus it would seem as though *E.coli* TolA3 is insufficient to drive colicin E9 translocation. Although fusion protein EC12 did kill cells, its killing was reduced due to the replacement of the OmpF binding site 1 with TolA binding site. It has been previously reported that when either

of the OmpF binding sites are replaced, colicin E9 can still kill cells, but at lower efficiency (due to the reduced efficiency at which colicin E9 can translocate across the outer membrane) (Housden et al. 2005).

If fusions EC7 and EC8 which lack the ability to bind TolB, but retain the ability to bind TolA3 cannot translocate across the membrane, it may be a case of polypeptide length, whereby the T-domain of ColE9-TolB fusions is not long enough to reach TolA at the inner membrane, and therefore cannot translocate. As estimates put the width of the periplasmic space at between 100-150 Å (Collins et al. 2007), the disordered N-terminus of the colicin E9 translocation domain may be of insufficient length, or it may not be able to penetrate the peptidoglycan layer to reach TolA at the inner membrane. However, TolA is reported to interact with Pal and major outer membrane porins, and is also the sole Tol protein required for the translocation of some other colicins (colicin N and A are reported to interact directly with TolA domain 3, and TolA drives their entry into the cell) (Cascales et al. 2007). Therefore it is unlikely that ColE9-TolB fusions are unable to reach TolA, unless colicin's N and A have a specific method of penetrating the PG layer, something that the colicin E9 fusions are not capable of, especially as the T-domain of colicin N is of similar total length (90 residues) (Hecht et al. 2010) as the fusion protein T-domains.

Alternatively, it is possible that the colicin E9 is capable of binding TolA *in vivo*, however, the interaction is insufficient to drive translocation. However, this would seem to be unlikely, as the *in vitro* affinity for TolA determined by ITC was similar for fusion proteins as for the wild type TolA-TolB. In addition, translocation of wild type colicin E9 across the outer membrane is dependent on both TolA and TolB. Thus, as fusion protein-TolA is of similar affinity to TolB-TolA, this should be sufficient to drive the translocation of the colicin across the outer membrane.

3.4.3 Summary

The work reported in this chapter has confirmed the importance of the N-terminus of *E.coli* TolB for its interaction with *E.coli* TolA domain 3, and, as previously reported (Bonsor et al. 2009), no interaction occurs between *E.coli* Pal and TolA domain 3. In addition, by SPR it has been shown that the binding of *E.coli* Pal to TolB controls the ability of TolB to bind TolA. This is likely through mediation of order/disorder equilibrium of eTolB's N-terminus that allows TolA3 to bind when in the disordered (Pal not bound) state, and prevents TolA3 binding when in the ordered (Pal bound state) (Bonsor et al. 2009). This work has also shown that the 12 residues of the N-terminus of eTolB known to interact with eTolA3 are sufficient, when engineered onto a protein that does not usually interact with eTolA3, can create a novel interaction with an engineered colicin E9 fusion protein, *in vitro.* However, despite this novel interaction being of similar affinity to that of wild type eTolB-eTolA3, it is not sufficient to drive the entry of a novel colicin E9 fusion into the cell in the absence of TolB.

4. The Gram-negative TolA-TolB complex

4.1 Introduction

In the previous chapter it was shown that the 12 residues of the N-terminus of *E.coli* TolB are the sole determining factor in it's interaction with *E.coli* TolA domain 3 and that when engineered onto a protein that does not normally interact with *E.coli* TolA3, it is sufficient to bind TolA3 with similar affinity to the native interaction. The aim of the work in this chapter was two-fold; (1) to determine if a synthetic peptide, corresponding to the sequence of the N-terminus of *E.coli* TolB could, in isolation, interact with *E.coli* TolA domain 3, and (2), if the TolA3-TolB interaction of *Pseudomonas aeruginosa* was conserved. The work in this work aimed to characterise the *Pseudomonas aeruginosa* TolA and TolB interaction, and determine if, like their *E.coli* counterparts, the interaction is dependent on the 12 residues at the extreme N-terminus of the TolB. Due to the failure previous attempts with protein crystallography to determine the *E.coli* TolA-TolB complex it was hoped that either a complex of the synthetic peptide with *E.coli* TolA3 or the *Pseudomonas aeruginosa* proteins would be more amenable to complex formation, and thus not only determine the site of TolB's interaction on TolA, but to finally obtain the structure of the TolA-TolB complex, something which has eluded researchers for several decades.

4.2 The Gram-negative Tol complex

The Tol family of proteins are highly conserved throughout most Gram-negative bacteria (Sturgis 2001). The *tol-oprL* genes (OprL is the name for Pal in *Pseudomonas aeruginosa*) are organised into three operons, orf1-tolQRA, tolB and oprL-orf2, the upstream of which is constituitive and the other two of which are under iron regulation. *Pseudomonas aeruginosa*

represents the only bacterium where the *tol-oprL* operons are known to be regulated by the ferric uptake regulator (Fur). In work characterising the *orf1-tolqra* operon and investigated the function of Orf1 it was found that it was not possible to create *tolq* and *tola* knockout strains, and thus they are likely to be essential genes (*tol* genes are not essential in *E.coli*), A viable *orf1* knockout strain could also be created suggesting a non-essential function. This mutant exhibited altered cell and colony morphology (Duan et al. 2000, Wei et al. 2009), something that has also been seen in *E.coli* cells with *ybgc* knocked out (Krachler 2010).

Structural information for some Gram-negative Tol proteins have been reported, including the crystal structure of *Pseudomonas aeruginosa* TolA domain 3 (Witty et al. 2002). When compared to the crystal structure reported for *E.coli* TolA domain 3 in complex with bacteriophage g3p (Lubkowski et al. 1999), the two structures share a near identical fold (figure 4.1), with a root mean square fit of 1.5 Å over 69 equivalent atoms. Similarly, the recently published crystal structure of *E.coli* TolA domain 3 appears to share a similar fold, with a root mean square fit of 1.7 Å and 2.1 Å, for the crystal structure of TolA3-g3p and *Pseudomonas aeruginosa* TolA3, respectively (Li et al. 2012). This similarity in fold between *E.coli* and *Pseudomonas aeruginosa* TolA3 is despite the fact that they share only 20% sequence identity (Witty et al. 2002). When comparing the TolB sequences from *E.coli* and *Pseudomonas aeruginosa,* these two proteins share approximately 44% sequence identity. When comparing over 100 TolA and TolB proteins from a variety of Gram-negatives, most TolB's share approximately 45% sequence identity between one another, whereas most TolA proteins share approximately 25% sequence identity. To the best of our knowledge, this is the first study of the *Pseudomonas aeruginosa* TolA-TolB system. The *tol* genes in *Pseudomonas aeruginosa* were first identified in 1996 (Dennis et al. 1996), and in the same work it was reported that *Pseudomonas aeruginosa* Tol proteins were functionally unable to

complement *E. coli tol* mutants, although *Pseudomonas aeruginosa* TolQ was able to complement the iron-limited growth of an *E. coli* exbB mutant.

Despite a published crystal structure little else is known about the interaction network of *Pseudomonas aeruginosa* TolA, or if it interacts with *Pseudomonas aeruginosa* TolB. Given that Tol proteins are conserved throughout most Gram-negative bacteria (Sturgis 2001), we could therefore hypothesize that the interaction between TolA and TolB is conserved, and based on homology between Gram-negative TolB's, that the N-terminus of TolB is the site of interaction. Given the lack of success in determining the *E.coli* TolA-TolB complex structure, it was hoped that other Gram-negative bacterial Tol proteins would be more amenable to structural determination of the complex.

Little is known about *Pseudomonas aeruginosa* TolB as no biophysical or structural data have been published for this protein. It is unknown if, like it's *E.coli* homologue, it interacts with either *Pseudomonas aeruginosa* TolA or Pal, or if it is parasitised by *Pseudomonas aeruginosa* bacteriocins (Pyocins) to facilite their translocation, although this is likely as the introduction of *tolqra* genes in the tol-like mutant PAO 1652 (which does not display the *tol* phenotype of membrane instability, only that it makes the *Pseudomonas aeruginosa* cells tolerant to pyocins) restored pyocin AR41 killing (Dennis et al. 1996). Based on sequence identity with *E.coli* TolB, the *Pseudomonas aeruginosa* TolB is predicted to be arranged in a similar domain organisation; i.e. 2 domains including a C-terminal beta-propeller domain, and a second smaller domain that encodes an N-terminal signal peptide (at the same residue number as *E.coli* TolB, between alanine 21 and 22) that is cleaved to yield mature protein (Duan et al. 2000). In addition, when comparing the N-termini of *Pseudomonas aeruginosa* TolB to that of *E.coli,* it appears to follow a similar consensus pattern of hydrophobic residues (figure 4.2).

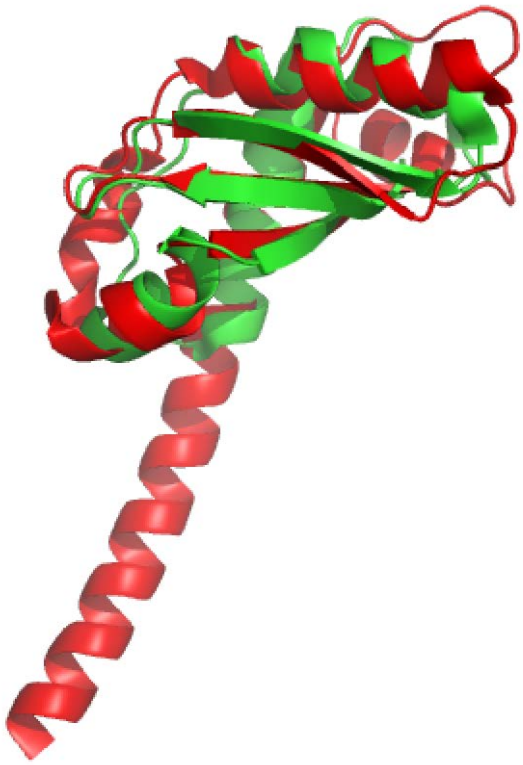


Figure 4.1 Comparison of *Pseudomonas aeruginosa* and *E.coli* TolA domain 3's. Overlay of *E.coli* TolA domain 3 (green), residues 302-421 (PBD ID: 3QDP) (Li et al. 2012) with *Pseudomonas aeruginosa* TolA domain 3 (red), residues 226-347 (PDB: 1LR0) (Witty et al. 2002).

E.Coli	EVRIVIDSGVDS
Pseudomonas aeruginosa	ADPLVISSGNDR

Figure 4.2 Alignment of N-termini of mature TolB from *E.coli* and *Pseudomonas aeruginosa*. Red represents conserved residues.

4.3 Aims

The aim of the work in this chapter is to investigate the 12 residue N-terminal sequence from *E.coli* TolB as an isolated synthetic peptide to determine if it is capable of recaptitulating the interaction between *E.coli* TolA domain 3 and TolB. In addition, this work aims to determine if, as previously hypothesised, that the TolA-TolB interaction is conserved in Gram-negative bacteria. To achieve these aims two approaches were adopted; a synthetic peptide was designed consisting of the *E.coli* TolA binding box, and through a series of *in vitro* crosslinking and ITC experiments, the peptide's interaction with *E.coli* TolA domain 3 was characterised. Additionally, as previous attempts to determine the *E.coli* TolA-TolB complex had failed, attempts were made through crystallography and nuclear magnetic resonance spectroscopy to observe both *E.coli* TolA domain 3 alone and a complex of *E.coli* TolA binding peptide with TolA domain 3.

Secondly, to investigate the TolA-TolB interaction found in other Gram-negative bacteria, 2 organisms were chosen to target the Tol system; *Pseudomonas aeruginosa* and *Xanthomonas campestris.* Using bioinformatics analysis, predictions were made in the domain organisation of the TolA proteins and constructs were designed to encode both domain 3, and a small section of domain 2, identical to *E.coli* TolA constructs used in previous work (Bonsor 2009, Krachler 2010). *Pseudomonas aeruginosa* genes were expressed and proteins purifed (*Xanthomonas campestris* clones failed to express and thus was abandoned, and all work focused on *Pseudomonas aeruginosa* proteins). Having characterised the TolA proteins, *Pseudomonas aeruginosa* TolB constructs were generated. Once purified *in vitro* experiments using chemical crosslinking and ITC were performed to characterise any interaction between *Pseudomonas aeruginosa* TolA domain 3 and TolB. In parallel, a synthetic *Pseudomonas aeruginosa* TolA binding peptide was designed and used to probe the interaction between

Pseudomonas aeruginosa TolA and TolB. Finally, given the potentially conserved nature of the TolA-TolB interaction in Gram-negative bacteria, the specificity of the TolA-TolB interaction was probed to determine if any non-cognate interactions occurred between *E.coli* and *Pseudomonas aeruginosa* Tol proteins.

4.2 Results

4.2.1 Studies of the *Pseudomonas aeruginosa* TolA-TolB interaction

4.2.1.1 Purification of tagless *Pseudomonas aeruginosa* TolA domain 3 (pEC1 construct) and his-tagged *Pseudomonas aeruginosa* TolA domain 3 (pEC4 construct)

Pseudomonas aeruginosa TolA (psTolA3) domain 3 (residues 226-347) was purified from BL21 (DE3) cells transformed with pEC1 or pEC4 and grown in 4.8L of LB media. Proteins were purified as described in section 2.3.3/2.3.4, and steps of purification are shown in figure 4.2 (A and B) and figure 4.3. A typical protein yield of 1 mg/L of culture was obtained for pEC1 construct, and 2 mg/L for pEC4 construct. Electrospray mass spectrometry data indicated that proteins were of the expected size (table 4.1)

4.2.1.2 Purification of C-terminal his-tagged *Pseudomonas aeruginosa* TolB (pEC14 construct)

His-tagged *Pseudomonas aeruginosa* TolB (psTolB, residues 22-432) was purified from BL21 (DE3) cells transformed with pEC14, grown in 4.8L of LB media. Protein was purified as described in section 2.3.7, and steps of purification are shown in figure 4.4. A typical protein yield of 1.5 mg/L of culture was obtained. Electrospray mass spectrometry indicated that protein was of the expected size (table 4.1)

Unfortunately psTolB is an unstable protein. As part of the purification procedure, on two occasions the entire protein preparation precipitated and was lost. The limit of solubility for this protein is approximately 50 μM. Even at concentrations below 50 μM the protein readily becomes insoluble.

Numerous buffer screens were tried to improve solubility, including the Optimum Solubility Screen (Jancarik et al. 2004), with little success. The buffer that the protein was most stable in is sodium phosphate, in a pH range of 7.5-9. This lack of solubilty has caused experimental problems, as will be addressed below.

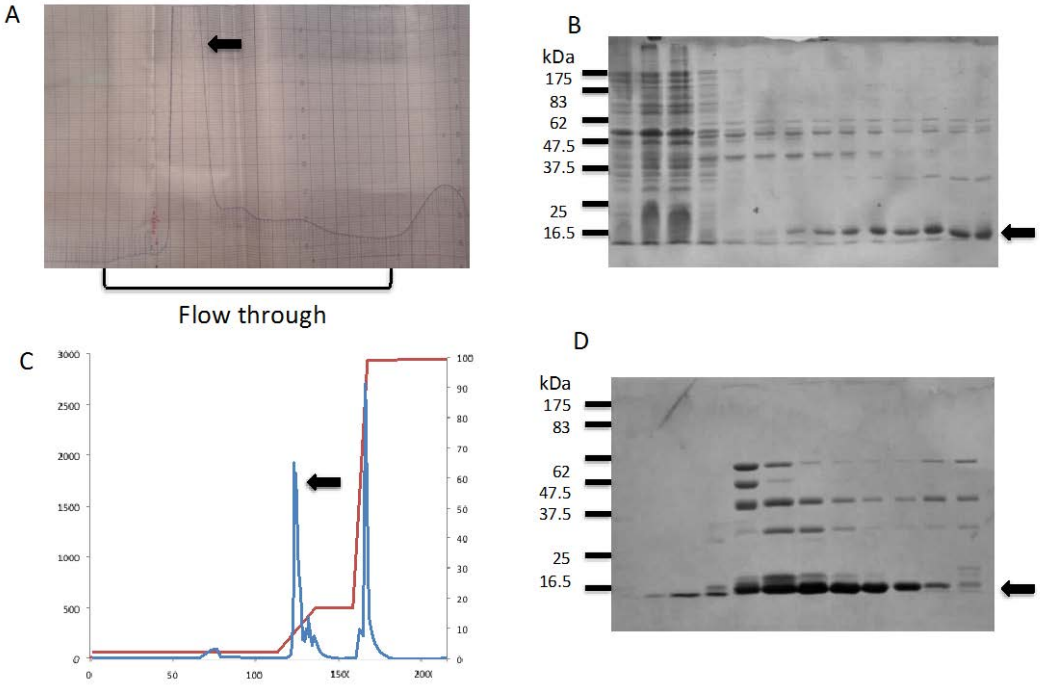


Figure 4.2A Purification of tagless psTolA3 (pEC1 construct) (A) 280nm absorbance profile from DE52 weak-anion exchange chromatography. Protein (indicated by arrow) eluted as part of non-binding flow through. (B) 16% SDS-PAGE gel to verify presence of psTolA3 (indicated by arrow). (C) Strong anion exchange (MonoQ) chromatography purification step, as monitored at 280nm. 1M NaCl elution gradient applied (D) 16% SDS-PAGE gel to verify presence of psTolA3 (indicated by arrow).

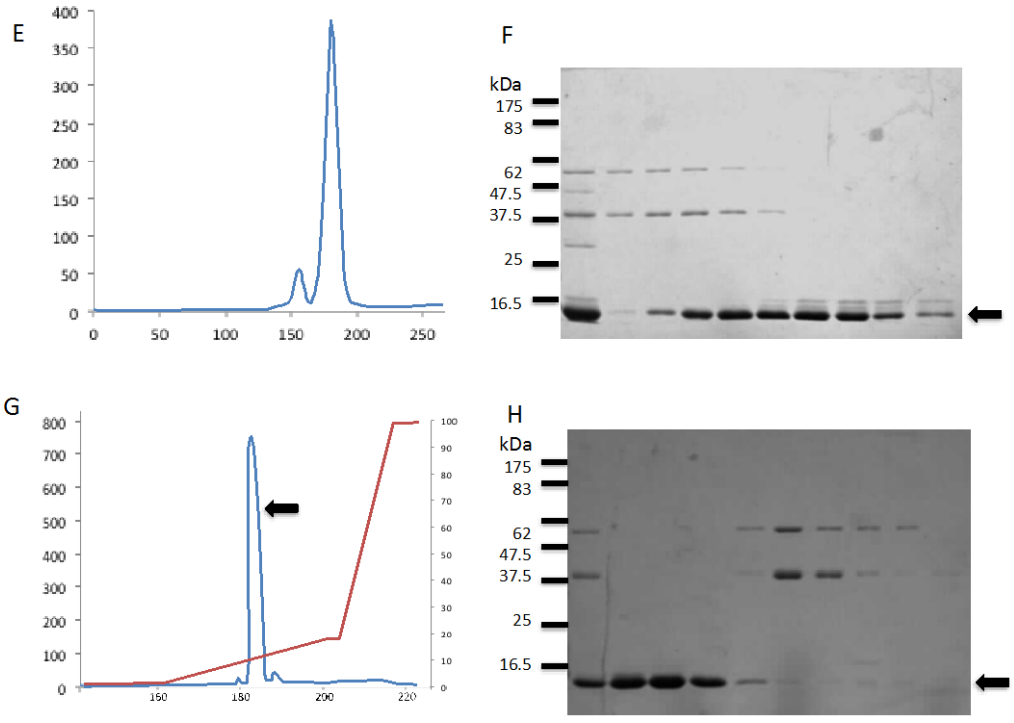


Figure 4.2B Purification of tagless psTolA3 (pEC1 construct) (E) Gel-filtration of psTolA3 on Superdex 75 26/60 column. psTolA3 eluted between 170 and 200ml. (F) 16% SDS-PAGE gel to verify presence of psTolA3 (indicated by arrow) (G) Second strong anion exchange (MonoQ) chromatography purification step, as monitored at 280nm. 1M NaCl elution gradient applied. (H) 16% SDS-PAGE gel to verify presence of pure psTolA3 (indicated by arrow).

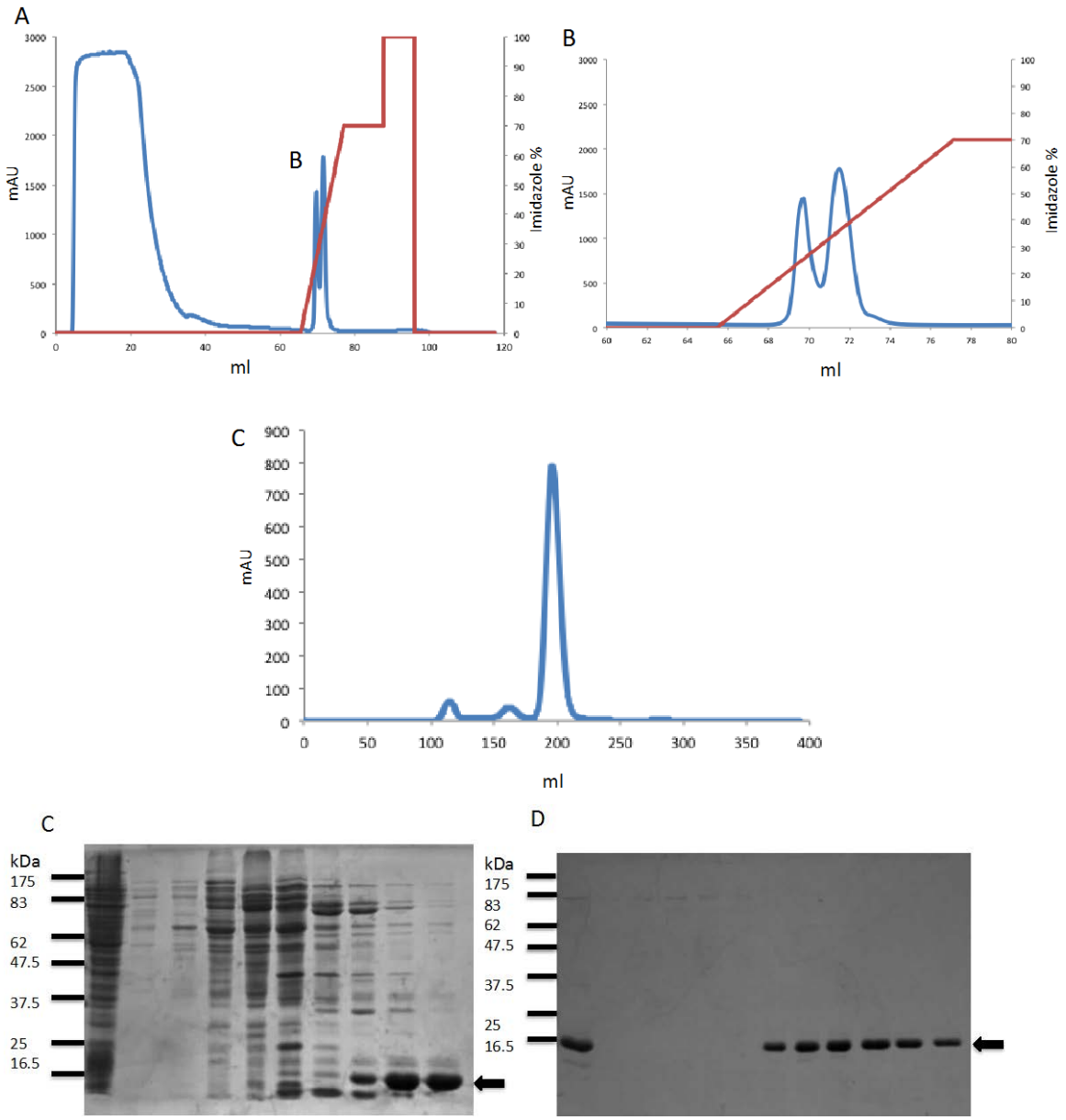


Figure 4.3 Purification of C-terminal his-tagged psTolA3 (pEC4 construct) (A) 280nm Absorbance profile from Ni-NTA affinity chromatography. (B) denotes expanded view of elution peak. Protein (indicated by arrow) eluted with 0-500mM Imidazole elution over 10 column volumes. (C) Gel-filtration histogram of eTolA3 on Superdex 75 26/60 column. psTolA eluted between 180 and 210 ml. (C) 16% SDS-PAGE gel to verify presence of psTolA3 in imidazole elution fraction (indicated by arrow). (D) 16% SDS-PAGE gel to verify presence of pure psTolA3 (indicated by arrow) after gel filtration.

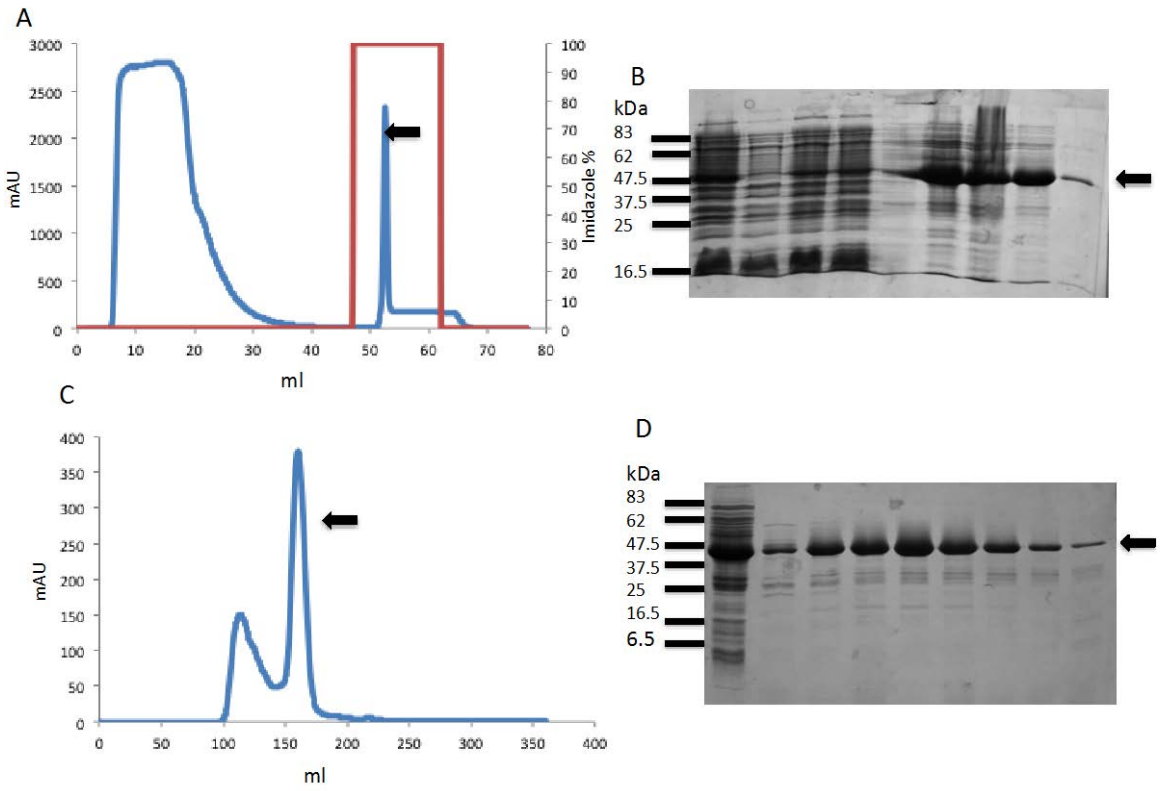


Figure 4.4 Purification of his-tagged psTolB (pEC14 construct) (A) 280nm Absorbance profile from Ni-NTA affinity chromatography. Protein (indicated by arrow) eluted with 500mM Imidazole step elution. (B) 13% SDS-PAGE gel to verify presence of psTolB (indicated by arrow). (C) Gel-filtration of eTolA3 on Superdex 75 26/60 column. psTolB eluted between 150 and 175 ml. (C) 13% SDS-PAGE gel to verify presence of pure psTolB (indicated by arrow).

Protein	Expected mass	Observed mass
psTolA3 (tagless, pEC1)	13430 Da	13433 (±9) Da
psTolA3 (C-terminal 6x his-tag, pEC4)	14495 Da	14492 (±5) Da
psTolB (C-terminal 6x his-tag, pEC14)	46459 Da	46454 (±8) Da

Table 4.1 Expected vs observed masses of purified proteins

4.2.1.4 *Pseudomonas aeruginosa* TolA domain 3 and TolB proteins are folded.

To ensure that not only were purified psTolA3 and psTolB proteins folded, but they also had similar secondary structure to their *E.coli* homologues, far UV (190-260 nm) circular dichroism spectra were collected for psTolA3 (figure 4.5) and psTolB (figure 4.6) proteins, to be compared with CD spectra of respective *E.coli* proteins.

CD spectra were subsequently deconvoluted with CDNN 2.1 (Applied Photophysics). CDNN analysis estimated that psTolA3 had 26% helical structure, 22% antiparallel, 10% parallel, 15% beta-turn and 29% random coil, compared to eTolA had 37% helical structure, 4% antiparallel, 9% parallel, 15% beta-turn and 38% random coil. psTolB has 21% helical structure, 29% antiparallel, 10% parallel, 15% beta-turn, and 33% random coil, compared with eTolB which was estimated to contain 12% helical structure, 29% antiparallel, 10% parallel, 15% beta-turn and 37% random coil. As these statistics are similar, it can be assumed that both psTolA3 and eTolA3 as well as psTolB and eTolB have a similar secondary structure.

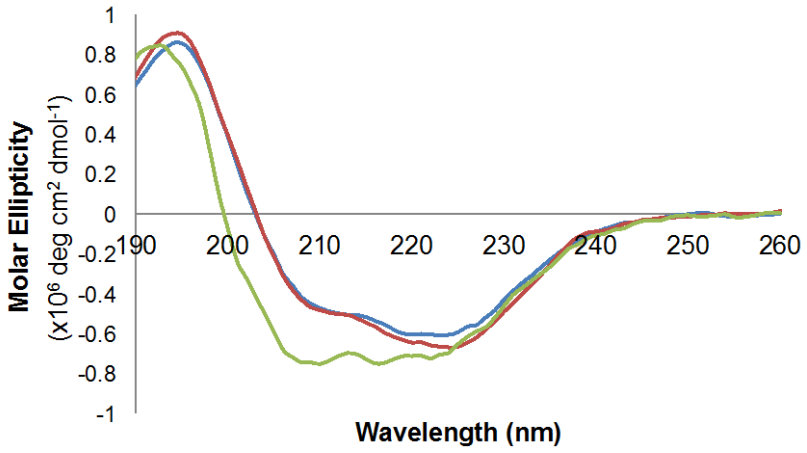


Figure 4.5 Far UV circular dichroism spectra comparing *Pseudomonas aeruginosa* TolA3 variants (tagless and his-tagged) with *E.coli* TolA3. CD spectrum collected in 10mM Sodium phosphate, pH7 at 20 °C. psTolA3 (pEC1 construct) is represented in blue, C-terminal his-tagged psTolA3 (pEC4 construct) is represented in red and eTolA3 (pAK108) is represented in green. All proteins at 50 µM final concentration.

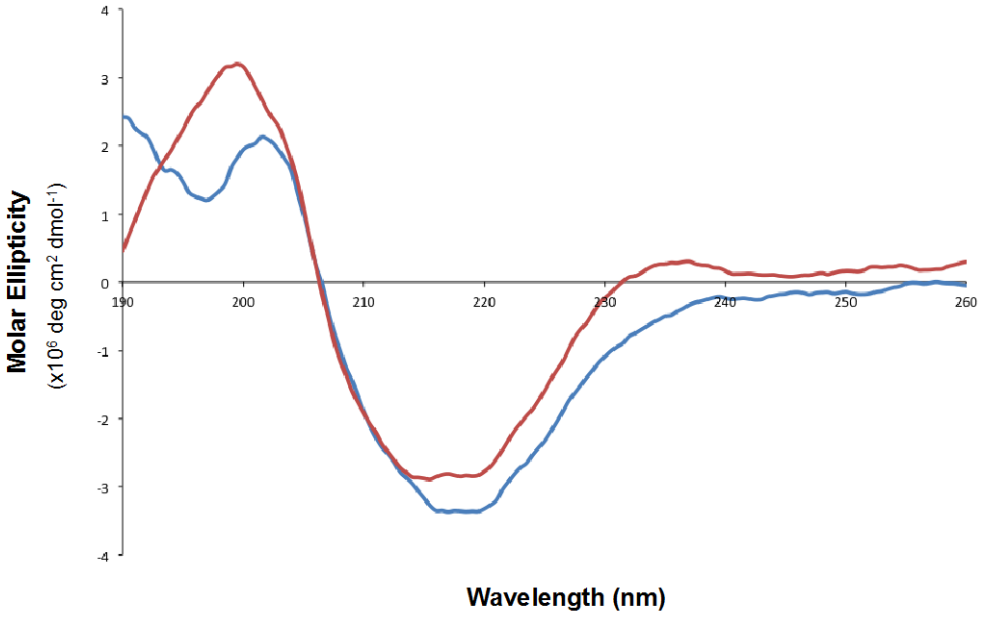


Figure 4.6 Far-UV Circular dichroism spectra comparing *Pseudomonas aeruginosa* TolB with *E.coli* TolB. CD spectrum recorded in 10mM Sodium phosphate, pH7 at 20 °C. psTolA3 is represented in blue and eTolB is represented in red. All proteins at 50 µM final concentration.

4.2.1.5 Pseudomonas aeruginosa TolA domain 3 and Pseudomonas aeruginosa TolB interact in vitro

To ascertain whether or not psTolA3 and psTolB interacted *in vitro*, a simple formaldehyde crosslinking experiment was performed (perfomed as described in section 2.3.15). psTolA3 and psTolB were incubated both together and in isolation (along with *E.coli* TolA3-TolB positive controls, data not shown), and once any complexes had been crosslinked with formaldehyde (see appendix section 7.9 for details on crosslinking reaction) and excess cross-linker quenched, the samples were loaded onto 13% SDS-PAGE gels for analysis (figure 4.7).

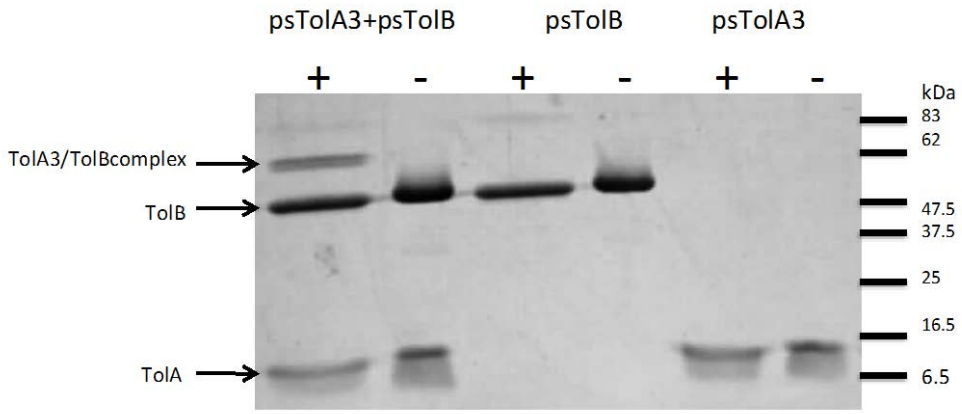


Figure 4.7 Formaldehyde crosslinking experiment showing interaction of psTolA3 and psTolB *in vitro*. Each protein at 10 μM (final) was incubated in presence of formaldehyde crosslinker (presence of crosslinker denoted by "+" symbol, "-" indicates control sample with no crosslinker), as described in section 2.3.15. Following quenching of the reaction, samples were run out on 13% SDS-PAGE gels, which were subsequently stained with Coomassie blue. Crosslinking performed at pH8.

As shown in figure 4.7, there is a band of approximately 60 kDa that corresponds to a complex between psTolA3 and psTolB, which is not

present in the samples corresponding to individual proteins alone. The band corresponding to psTolA3-psTolB was excised from the gel and analysed with MALDI mass spectrometry (University of York Technology Facility) to confirm presence of both psTolA3 and psTolB in the complex, both of which were found to be present. This confirmed that not only did psTolA and psTolB interact, but also, like in the *E.coli* proteins, it is the 3rd domain of TolA that is involved in the interaction with TolB.

Having confirmed that psTolA3 crosslinks with psTolB to form a complex, as found with *E.coli* homologues, the next step was to obtain thermodynamic data for this interaction, and thus an equilibrium binding constant. However, psTolB is an unstable protein, and this proved to be problematic. All attempts to obtain thermodynamic data for the psTolA3-psTolB interaction through ITC have met with failure. The *Pseudomonas* proteins typically become insoluble during the titration. Numerous alternative buffers, as well as a range of protein concentrations, pH's and temperatures were attempted to improve solubility and data quality, to no avail. Further alternative experiments were conducted (see section 4.2.3 below for details) in an attempt to address the issues with psTolB.

4.2.2 Interactions between TolA domain 3 and TolB are specific

In the previous chapter and section it has been shown that *E.coli* and *Pseudomonas aeruginosa* TolA domain 3 interact with TolB. To further investigate the nature of this TolA3-TolB interaction, and to determine if the interaction is specific to the cognate binding partners, or if any non-cognate complexes formed between these potential binding partners two *in vitro* experiments were performed. Firstly to capture any potential non-cognate complexes, formaldehyde crosslinking was employed, and secondly thermodynamic data were collected via ITC.

psTolA3-eTolB and eTolA3-psTolB were incubated together (pH range 6.5-8.5, along with appropriate positive and negative controls), and once any complex had been captured by formaldehyde crosslinking and excess cross-linker quenched, the samples were loaded onto 13% SDS-PAGE gels for analysis (figure 4.9).

Despite crosslinks between cognate binding partners in positive controls, no non-cognate complex was found between either TolA3/TolB (figure 4.9), suggesting that the interaction between TolA3 and TolB is species specific. Although the disordered N-termini of TolB have a general consensus sequence between the *E.coli* and *Pseudomonas aeruginosa* proteins, this consensus is not sufficient to drive an interaction with TolA3.

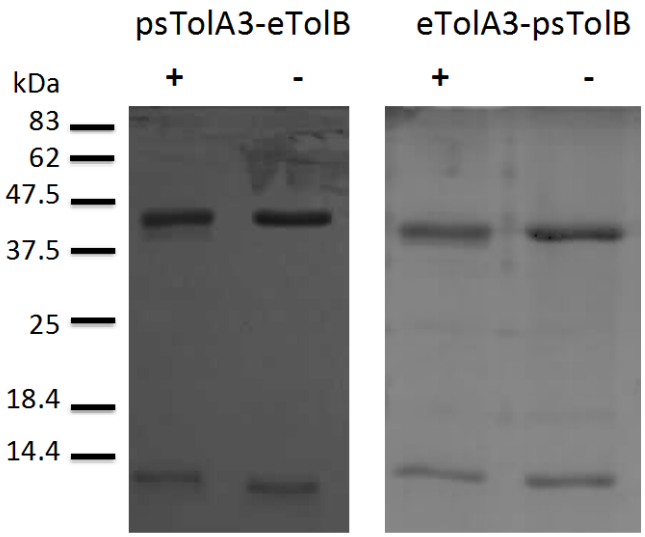


Figure 4.9 Formaldehyde crosslinking experiment attempting to capture non-cognate complex formation between *E.coli* and *Pseudomonas aeruginosa* TolA3/TolB, *in vitro*. Each protein at 10 μM (final) was incubated in presence of formaldehyde crosslinker (presence of crosslinker denoted by "+" symbol, "-" indicates control sample with no

crosslinker), as described in section 2.3.15. Following quenching of the reaction, samples were analysed on 13% SDS-PAGE gels, which were subsequently Coomassie blue stained. Crosslinking performed in the presence of appropriate positive controls (not shown).

To confirm these findings, ITC experiments were performed (at a range of pHs, from 6.5-8.5), titrating psTolA3 into eTolB. Due to the insoluble nature of psTolB, no titrations of eTolA3 into psTolB were attempted.

Again, as found with crosslinking experiments, no heats of binding were detected between psTolA3 and eTolB, suggesting that no complex is formed between these 2 proteins. This again suggests that the TolA3-TolB interaction is specific.

4.2.3. A synthetic peptide of the N-terminus of *E.coli* TolB can recapitulate the *E.coli* TolA/TolB interaction *in vitro.*

It has been previously shown that the *E.coli* TolA-TolB interaction is dependent on the N-terminus of TolB, and that this N-terminus of *E.coli* TolB must be in a disordered conformation to bind eTolA3 (Bonsor et al. 2009). Furthermore, as reported in section 3.3.2, when the sequence known to be involved in eTolA3 binding is engineered onto a disordered region on an alternate protein, that the sequence alone is sufficient to bind eTolA3, *in vitro*. To investigate if this sequence in isolation was sufficient to interact with eTolA3 a disordered peptide of the sequence, EVRIVIDSGVDS was synthesised (peptides produced by either Activeotec, Cambridge, or Pepceuticals, Nottingham). A peptide of this sequence alone was insoluble, and so the sequence was modified with the addition of 4 extra residues. The final sequence of the peptide was EVRIVIDSGVDSWKKK. The red sequence originates from the disordered N-terminus of eTolB. Black residues denote additional sequence, of which the extra tryptophan residue

was used to aid quantification of the concentration of peptide and the 3 C-terminal lysines were added are to maintain the solubility of the peptide.

To determine if *E.coli* TolA3 can interact with *E.coli* TolA binding peptide, and to determine if the binding peptide interacts with *E.coli* TolA3 in the same binding site as *E.coli* TolB, a formaldehyde crosslinking competition assay was performed. Briefly, eTolA3 was incubated with both eTolB, and also increasing concentrations of eTABp. Any complex formed between these binding partners was crosslinked with formaldehyde, and analysed on 13% SDS-PAGE gels (figure 4.10).

As shown in figure 4.10, a complex formed between eTolA3 and eTolB (as expected), but also as the concentration of peptide is increased, the amount of eTolA3-eTolB complex decreases, indicating that eTABp is competing with eTolB for binding on eTolA3. This also indicates that eTABp and eTolB have the same binding site, and that eTABp interacts with eTolA3 in a specific manner. Despite this, it is a consideration that at higher contrations of peptide, the ratio of protein to peptide is in the order of 1:40, and thus it is possible that some non-specific binding event may occur. Having confirmed that *E.coli* TolA3 interacts with the binding peptide with formaldehyde crosslinking, this interaction was quantified with isothermal titration calorimetry (figure 4.11).

As is shown in figure 4.11, there are heats indicating an interaction between eTABp and eTolA3. Following subtraction of heats of dilution, and fitting data to single site binding model (Origin 7.0, Microcal/GE Healthcare), a K_d of approximately 40 μM (± 7 μM from 5 replicates) was estimated. This figure was almost identical to that of the native eTolA3-eTolB interaction, although the thermodynamics are different. The thermodynamics of eTolB titrated with eTolA3 are +2.92 kcal mol^{-1} and +30 cal K^{-1} mol^{-1} compared with -2.71 kcal mol^{-1} and +11 cal K^{-1} mol^{-1} for eTABp titrated into eTolA3 (table 4.2). The

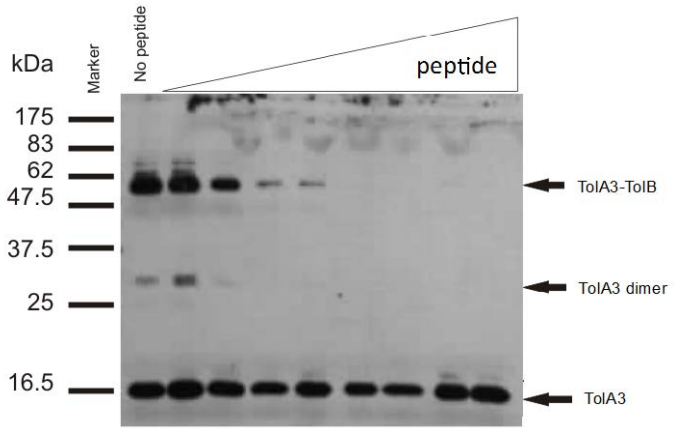
native eTolA3-eTolB interaction is endothermic and entropically driven as the N-terminus is (presumed) to be in a state of dynamic equilibrium between ordered and disordered state. Conversely, the eTolA3-eTABp interaction is exothermic and does not require energy as the peptide is already in a state of disorder and does not have to make the transition from ordered to disordered state.

Components	**ΔH (kcal mol^{-1})**	**ΔS (cal K^{-1} mol^{-1})**	**N**	**K_d (μM)**
eTolB-eTolA3	+ 2.92 (±0.02)	+ 30.0 (±0.1)	1.00 (±0.03)	43 (±2)
eTolA3-eTABp	- 2.71 (±0.04)	+ 11 (±0.1)	0.98 (±0.01)	40 (±4)

Table 4.2 Comparison of thermodynamics and affinities of eTolA3-eTolB/eTABp complex formation

Finally, the interaction between eTolA3 and eTABp was characterised in terms of temperature dependence. At 25°C K_d was determined as 70 μM (± 6 μM), and at 30 °C K_d was 113 μM (± 9 μM),. Experiments performed using ITC200 instrument (Microcal/GE Healthcare) using Origin 8.0 software and single site binding model.

A



B

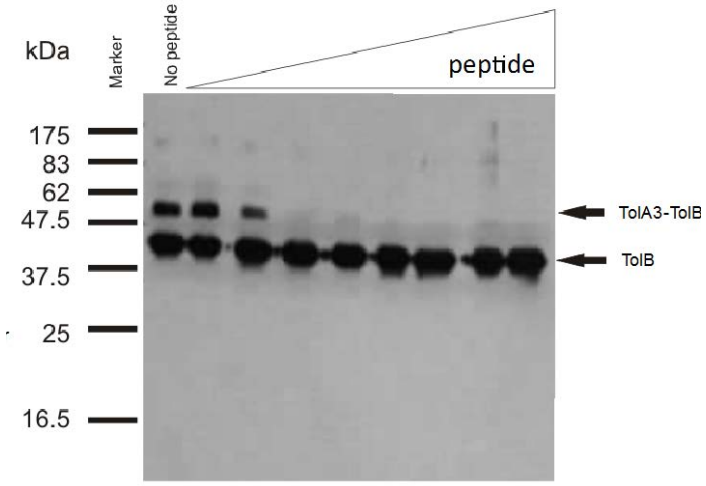


Figure 4.10 Formaldehyde crosslinking experiment showing competition of eTABp and eTolB for binding with eTolA3, *in vitro*, detected by Western blotting on 13% SDS-PAGE gels. (A) eTolA3 (10 μM) incubated with eTolB (μM), detected with anti-TolA (*E.coli*) antibody. eTolA3 and eTolB complex is incubated both in the absence (no peptide lane) and increasing concentration of eTABp (from 10 μM to 400 μM) from left to right. Note the progressive loss of TolA3-TolB complex, with complete abolition of complex at 150 μM peptide concentration. (B) As (A) except anti-TolB (*E.coli*) antibody used for detection.

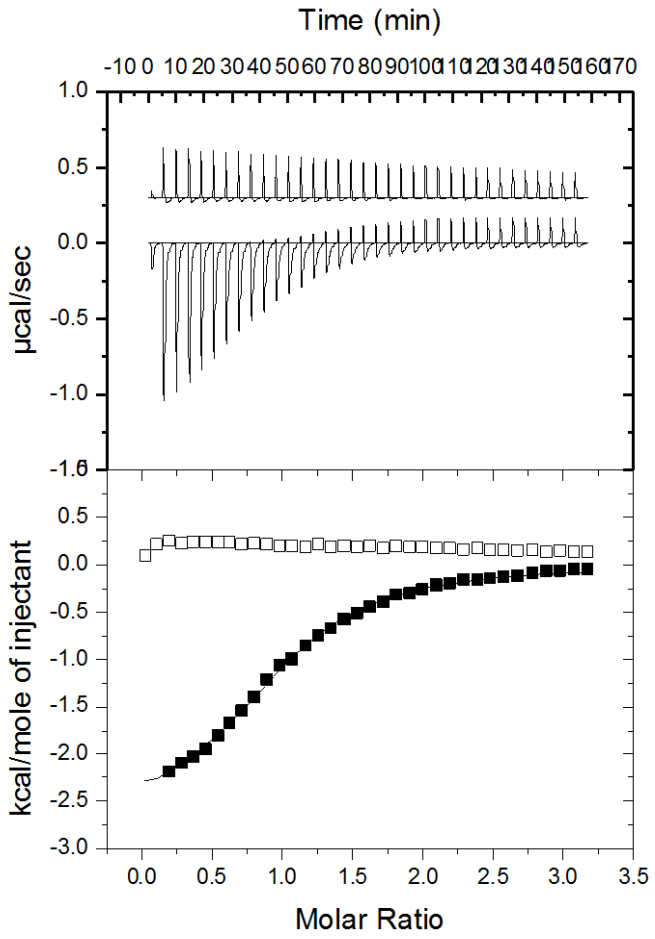


Figure 4.11 ITC titration data for *E.coli* TolA domain 3 interaction with *E.coli* TolA binding peptide. A stock of 3mM eTABp was titrated in 34 8 µl injections against 220 µM eTolA3. K_d was estimated as 40 µM, with a 1:1 stoichiometry of peptide:protein, ΔH of - 2.7 kcal mol^{-1} and ΔS 11.1 cal K^{-1} mol^{-1}. Data fitted to single site binding model using Origin 7.0 (Microcal/GE Healthcare). Titration conducted with VP ITC instrument (Microcal/GE Healthcare) at 20 °C with 270 s spacing between each injection of eTABp. Protein dialysed prior to experiment in 50 mM Hepes, 50 mM NaCl, pH 7.5. Peptide stock dissolved in same buffer.

4.2.4 Challenging the *in vivo* TolA-TolB interaction with the synthetic TolA binding peptide.

To determine if the synthetic *E.coli* TolA binding peptide was capable of binding *E.coli* TolA domain 3 *in vivo,* it was hoped that the peptide would be transported across the outer membrane (Yeaman & Yount 2003, Housden et al. 2010) and interact with *E.coli* TolA, thus abolishing the TolA-TolB interaction through competition and causing the *tol* phenotype. JM83 cells were incubated in various conditions (in the presence or absence of 2% SDS, and either grown in the presence of peptide, or supplemented with peptide prior to plating on LB-Agar) and visually assessed (figure 4.12) for presence of Tol phenotype.

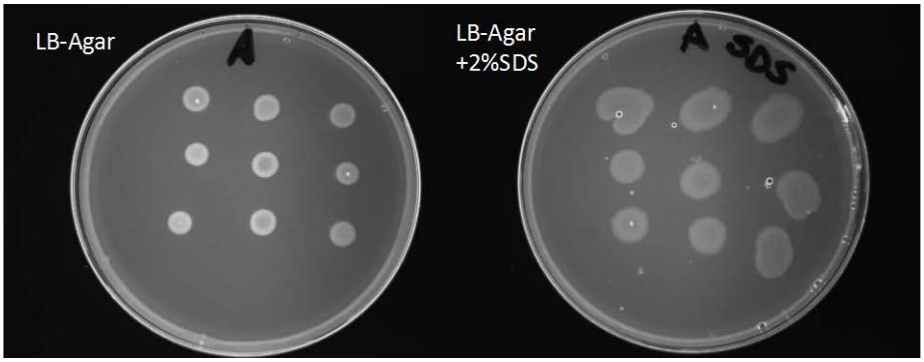


Figure 4.12 Challenging *E.coli* JM83 cells with *E.coli* TolA binding peptide. JM83 cells grown to an OD_{600} of approx. 0.6, then incubated in the following conditions; allowed to continue growing for 1 hour at 37 °C then plated in triplicate on LB-Agar and LB-Agar supplemented with 2% SDS (top row), allowed to continue growing for 1 hour at 37 °C, 500 µM eTABp was added just prior to plating in triplicate on LB-Agar and LB-Agar supplemented with 2% SDS (middle row), 500 µM eTABp was added to cells, which were grown for a further hour at 37 °C, before being plated in triplicate on LB-Agar and LB-Agar supplemented with 2% SDS.

As shown in figure 4.14, in all conditions, no Tol phenoptype was detected. Due to the hydrophobic nature of the peptide, and lack of outer membrane porin recognition sequence it is possible that the peptide is unable to cross the outer membrane of the cell and enter into the periplasm to interact with the Tol system.

4.2.5 A synthetic peptide of the disordered N-terminus of *E.coli* TolB does not interact with *E.coli* TolB lacking disordered N-terminus.

Having established that the *E.coli* TolA binding peptide binds *E.coli* TolA domain 3 and abolishes *E.coli* TolA/B interaction *in vitro*, I set out to determine if this peptide could interact with *E.coli* TolB Δ34 mutant (that lacks the N-terminal strand) *in vitro.* Wild type eTolB has a pocket in which the hydrophobic N-terminus sits when in the ordered conformation (figure 4.13), and as the Δ34 mutant lacked this disordered N-terminus, I attempted to detect any complex that formed if the synthetic peptide bound the pocket via ITC.

No heats of binding were detected for eTABp (1mM) titrated into eTolB Δ34 mutant (50 μM) at pH's 6.5 – 8.5 at a range of temperatures (15°C - 37°C) In addition, crosslinking experiments were performed under similar conditions, no complex between eTABp and eTolB Δ34 mutant was found.

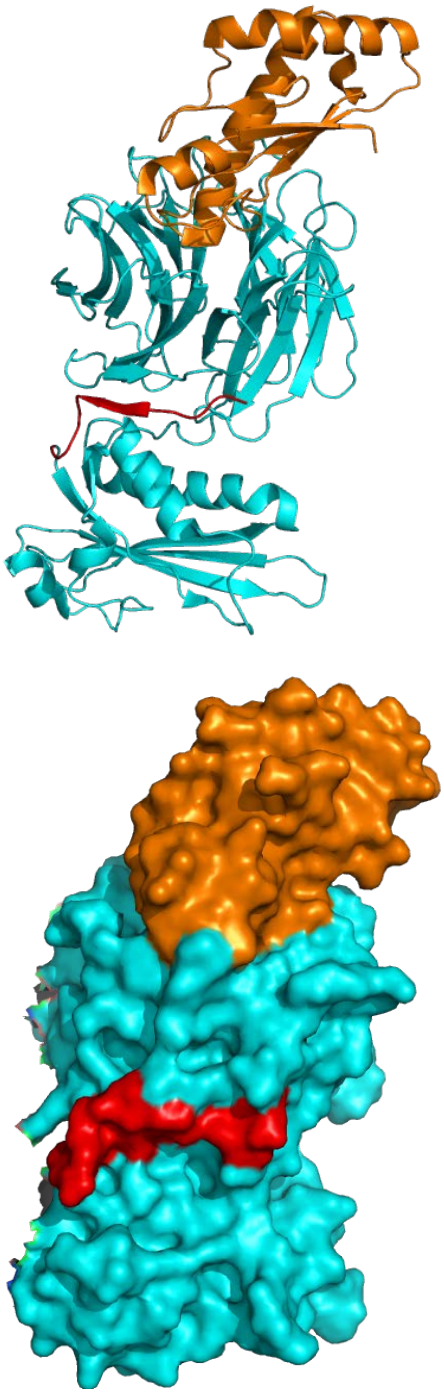


Figure 4.13 *E.coli* Pal binding *E.coli* TolB causes TolB's N-terminus to take up ordered conformation. Pal (gold), TolB (cyan), TolB's 12 N-terminal residues (red). PBD ID: 2W8B. (Bonsor et al. 2009).

4.2.6 Defining the *E.coli* TolB binding site on *E.coli* TolA domain 3 with protein crystallography

One of the primary aims of this project has been to define the TolB binding site on TolA3 and, if possible, obtain a structure for the complex of TolB-TolA3. Previous attempts (by both myself, and others in the lab) to solve the structure of the *E.coli* TolB-TolA3 complex with protein crystallography have been attempted, to no avail. In addition, further strategies have been followed, including *E.coli* TolA-TolB fusion proteins (Bonsor 2009) to obtain this complex, but have met with failure. Therefore, a new strategy to obtain structural information on the *E.coli* TolA-TolB complex was devised, using the *E.coli* TolA binding peptide as a surrogate for *E.coli* TolB in binding *E.coli* TolA3. Initial screens used eTolA3 at concentrations ranging from 10-150 mg/ml, preformed in a complex with eTABp in 5-10 fold molar excess. This complex was then screened against Peg-Ion 1&2 (YSBL), Hampton 1&2, Index (Hampton Research), Morpheus and PACT (Molecular Dimensions). Identical screens were also set up with eTolA3 alone. Results of these attempts to crystallise the complex are shown in figure 4.14.

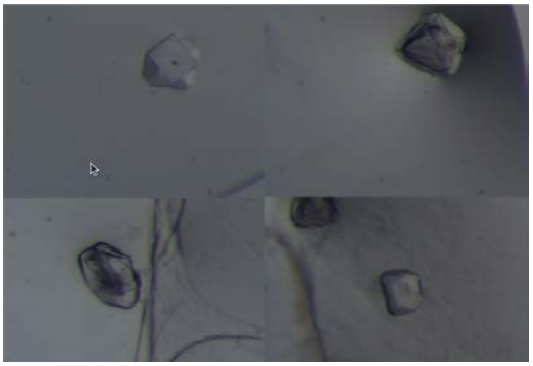


Figure 4.14 Crystallisation of eTolA3-eTABp complex eTolA3 (10, 20, 30 and 40 mg/ml, clockwise from top left panel) in complex with eTABp (5-fold molar excess) in 0.2 M NaF, 20% PEG3350 (PACT screen), after 9 months at constant temperature (22 °C).

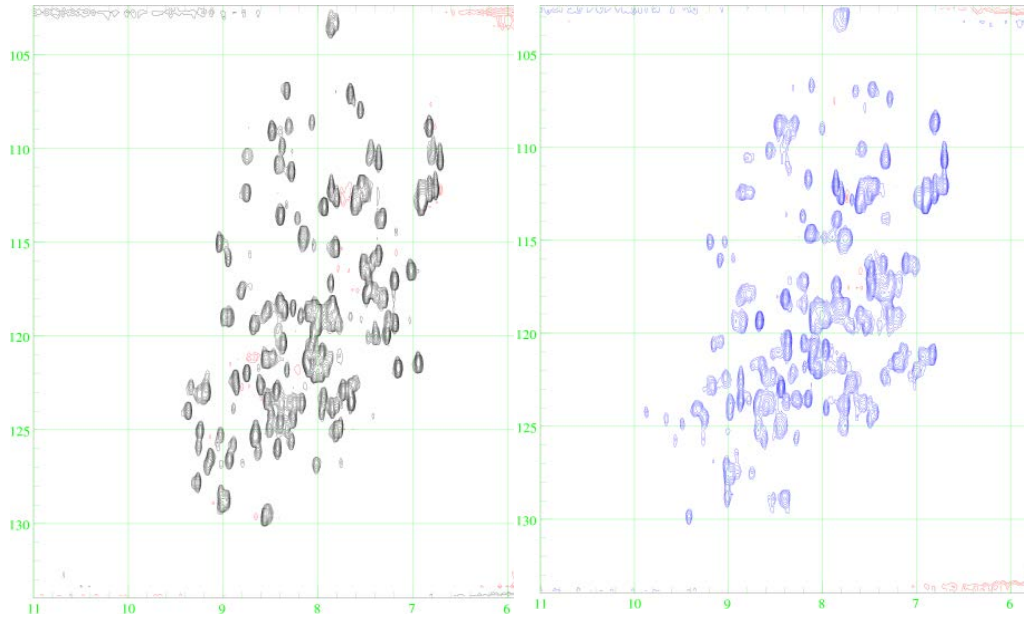
Crystals were removed from well (only 1 crystal grew in each condition) and screened by Justyna Wojdyla on Rigaku MicroMax 007HF generator with an RAXIS IV++ imaging plate detector (York Structural Biology Laboratory). Unfortunately no diffraction was detected for any of the crystals. No other conditions contained crystals, and attempts at replicating crystals failed. *E.coli* TolA domain 3 alone did not crystallise in any condition.

4.2.7 Defining the *E.coli* TolB binding site on *E.coli* TolA domain 3 with nuclear magnetic resonance spectroscopy

As crystallography had failed as a means of obtaining a structure of *E.coli* TolA3 bound to the binding peptide, nuclear magnetic resonance spectroscopy was attempted. Previously a solution NMR structure had been published *E.coli* TolA domain 3 (Deprez et al. 2005) and use of NMR to study the interactions of *E.coli* TolA and colicins has also been documented, using residues perturbations to map binding sites onto TolA (Hecht et al. 2010, 2011).

Spectra for both ^{15}N eTolA3 and ^{15}N eTolA3 – unlabeled eTABp complex were collected using a Bruker AVII 700MHz spectrometer (Dept. of Chemistry, University of York). ^{15}N eTolA3 (pAK108 construct) was purified from BL21 (DE3) cells grown in M9 minimal media supplemented with ^{15}N ammonium chloride as described in section 2.3.24. It was subsequently purified with a yield of 1.5 mg/ml. Size and fold of the protein was verified by mass spectrometry (Molecular Interactions Lab, Technology Facility, University of York) and circular dichroism (protocol described in section 3.1.8.1), respectively. Both methods indicated that protein was of correct size and identical fold to unlabeled control eTolA3. As *E.coli* TolA3 is very soluble, it was hoped that the protein would be suitable for NMR experiments.

However, 2D-HSQC (^{15}N- ^{1}H) spectra collected for eTolA3 alone showed a greater number of peaks than would be expected (117 peaks expected, 125 peaks seen, excluding sidechains) (figure 4.15 A), and in the presence of the 1.1 molar equivalent peptide this issue was compounded (in excess of 140 peaks seen, excluding sidechains) (figure 4.15 B), with not only an excessive number of peaks detected, but also that compared to the eTolA3 alone spectra, virtually every peak had shifted (figure 4.15 C). Whether or not this indicated a massive conformational change, or if the proteins peak were in slow exchange between two states was unclear. It is also possible that more than one eTABp was binding a single TolA perturbing a large number of peaks. In addition, a large number of peaks clustered around the central region (around 8ppm), which is typical of an unstructured protein (Cavanagh 2007). It was thought that this TolA construct may have been at least partially unstructured, possibly explaining the increased number of peaks. A number of different constructs of different lengths of *E.coli* TolA were also attempted, however, this issue of increased number of peaks was not reconciled (Luke Hillary, personal communication).



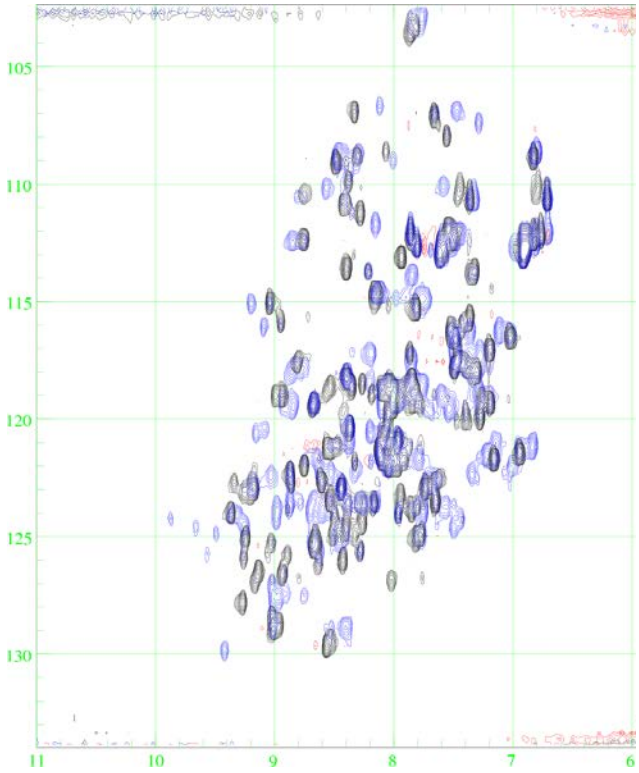


Figure 4.15 spectra of *E.coli* TolA domain 3 both alone, and in presence of *E.coli* TolA binding peptide. (A) 2D-HSQC spectrum of ^{15}N eTolA3 (pAK108 construct), (B) 2D-HSQC spectrum of ^{15}N eTolA3 in presence of 1.1x molar equivalent of eTABp, (C) Overlay of ^{15}N eTolA3 alone (black) and in presence of peptide (blue).

4.2.8 Defining the *E.coli* TolB binding site on *E.coli* TolA domain 3 with chemical crosslinking coupled with fragment mass spectrometry

Having been unable to define the TolB binding site on TolA through crystallography or NMR, a further technique was employed, inspired by work from the Rappsilber lab at the University of Edinburgh. This concerned using chemical crosslinking and fragmentation mass spectrometry to obtain information about protein complexes, and more specifically, protein-protein interfaces. Work has been published to obtain the structure and archecture of large protein complexes, the largest of which (at the time) was the 180 kDa Ndc80 complex (Maiolica et al. 2007) and more recently, the 670 kDa

RNA Polymerase II–TFIIF complex (Chen et al. 2010). Briefly, this approach employed chemically crosslinking of the target protein complex, which was subsequently digested with trypsin, and fragments from each component protein identified by mass spectrometry hopefully including additional mass from peptide fragmentation that were crosslinked together. This information could then be deconvoluted into a map of fragments from each protein that were crosslinked together, and thus it would be possible to define areas of the protein surface that were crosslinked, and in contact with one another. As the eTABp-eTolA3 protein-peptide complex could be reliably captured by chemical crosslinking, it was hoped that this approach may give information as to the protein surface that the eTABp was interacting with on *E.coli* TolA domain 3.

Thus, eTolA3 and eTABp were incubated together, and crosslinked with either formaldehyde or DSP (Dithiobis[succinimidyl propionate], also called Lomant's reagent, which is a reducible crosslinker that works in the same way as formaldehyde, through primary amines, but when reduced leaves a tell tale mass on the residues that were actually crosslinked (see appendix section 7.9 for details of crosslinking chemistry).

Crosslinked samples were separated on 13% SDS-PAGE gels (in non-reducing conditions in the case of DSP crosslinked samples) and stained with Coomassie Blue (figure 4.16). Bands corresponding to eTolA3-eTABp complex and eTolA3 alone were excised, digested in gel with trypsin and analysed with MALDI-MS (Electrospray source) (see figure 4.17 for expected trypsin digest fragments and figure 4.18 for results of actual trypsin fragments detected) by Dr Dave Ashford (University of York). Fragments from the complex that were detected were then compared to eTolA3 alone fragments (figure 4.18).

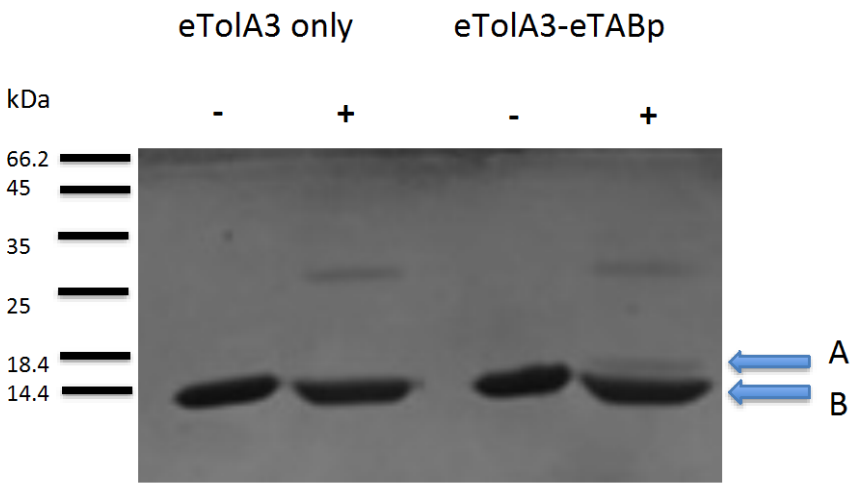


Figure 4.16 *E.coli* TolA domain 3 – *E.coli* TolA binding peptide formaldehyde cross-linking. (A) eTolA3-eTABp complex (B) eTolA3. "+" denotes presence of formaldehyde for crosslinking, "-" denotes no crosslinker.

ADDIFGELSSGKNAPKTGGGAKGNNASPA
GSGNTKNNGASGADINNYAGQIKSAIESKF
YDASSYAGKTCTLRIKLAPDGMLLDIKPEGG
DPALCQAALAAAKLAKIPKPPSQAVYEVFKN
APLDFKP

Figure 4.17 Map of theoretical trypsin cut sites for *E.coli* TolA domain 3. Red/Blue junction indicates predicted trypsin cleavage site. Predictions made using ExPASy PeptideCutter tool (Gasteiger et al. 2003).

A

ADDIFGELSSGKNAPKTGGGA**KGNNASPAGSG**
NTKNNGASGADINNYAGQIKSAIESK**FYDASSY**
AGKTCTLRIK**LAPDGMLLDIKPEGGDPALCQAA**
LAAAKLAK**IPKPPSQAVYEVFK** NAPLDFKP

B

ADDIFGELSSGKNAPKTGGGA**KGNNASPAGSG**
NTK**NNGASGADINNYAGQIK**SAIESK**FYDASSY**
AGKTCTLRIK**LAPDGMLLDIKPEGGDPALCQAA**
LAAAKLAK**IPKPPSQAVYEVFK** NAPLDFKP

Figure 4.18 Trypsin digested fragments of formaldehyde crosslinked eTolA3-eTABp complex detected by MALDI-MS. (A) eTolA3 only (residues 293 – 421), (B) eTolA3 crosslinked with eTABp. Red residues indicate those sequences detected by MALDI-MS, black represents seequences not detected by MALDI-MS. Blue represents lysine residues (K310, K316) not detected as part of tryptic peptide (although 1st lysine [K310] also not seen in eTolA3 only sample) when crosslinked eTolA3-eTABp complex is analysed with MALDI-MS.

Most of the residues detected by MALDI-MS are identical between the eTolA3 alone, and the sample crosslinked with eTABp. Although it had been hoped that a part of the eTABp would be detected by MALDI-MS, crosslinked to the eTolA3, this was not the case. It should noted however, that there was a difference in the trypsin digest pattern observed between the 2 samples, specifically regarding the second of 2 lysine residues (K316, shown in blue in figure 4.20, panel B). As formaldehyde and DSP crosslinking both crosslink via lysine residues, that this residue (K316) is lost

may indicate that due to crosslinking to this residue, trypsin is no longer able to access that specific site, and as a result of the mis-cleavage, the ionisation properties of the fragment have changed, which prevents it from being detected in MALD-MS. Although this theory is based on a negative result, it suggests that the binding peptide comes in contact with the eTolA3 at some point around K310 or K316 allowing a crosslink to form in the presence of formaldehyde or DSP. However, in the construction of the eTolA3 binding peptide, it was necessary to include an additional 3 lysine residues on the peptide's C-terminus in order to maintain the solubility of the peptide. Due to the chemistry of formaldehyde/DSP crosslinking through primary amine's (such as lysines) it is possible that non-specific crosslinks occur between the protein and peptide. It is also possible that as not all the tryptic fragments are detected in the mass spec sample for eTolA3 alone (either due them not ionizing or flying) and the same fragments are not detected for the eTolA3-eTABp complex, that the peptide may crosslink to any of these tryptic fragments and they are not detected.

4.2.9 The N-terminus of *Pseudomonas aeruginosa* TolB is involved in its interaction with *Pseudomonas aeruginosa* TolA domain 3.

Having attempted to use *E.coli* TolA-TolB to define the binding site of TolB on TolA, attention now moved to the *Pseudomonas aeruginosa* TolA-TolB interaction, with the hope that this system would be more ameniable to study.

Previously it was shown that processed psTolB interacts (and crosslinks) with psTolA3. The next step was determine, if, like the *E.coli* homologues, the N-terminus is the sole site of interaction between the 2 proteins. In addition it was to be determined if a synthetic peptide of the equivalent *Pseudomonas aeruginosa* sequence from the N-terminus of TolB can also recapitulate the interaction between psTolA3 and psTolB. A synthetic

peptide corresponding to the sequence ADPLVISSGNDRWKKK (The red sequence originates from the N-terminus of psTolB, the extra tryptophan residue is to aid quantification of the concentration of peptide, the 3 C-terminal lysines are to maintain the solubility of the peptide, and to maintain a consistent design with that of the *E.coli* TolA binding peptide. Additional peptides without these extra residues were also synthesised with both N-terminal amide and C-terminal acid or amide-capping, however, both were not soluble in any of the buffers appropriate to this work. Peptides were synthesised by Pepceuticals, Nottingham.

4.2.9.1 A synthetic peptide of the N-terminus of *Pseudomonas aeruginosa* TolB can compete with *Pseudomonas aeruginosa* TolB for binding with *Pseudomonas aeruginosa* TolA domain 3 *in vitro.*

In order to determine if the psTolA binding peptide could compete with psTolB for binding with psTolA domain 3 a simple crosslinking competition experiment was performed. psTolA3 and psTolB (10 μM final for each protein) were incubated together both in the absence, and the presence of increasing *Pseudomonas aeruginosa* TolA binding peptide (psTABp) concentrations, dissolved as a stock in 20m sodium phosphate, pH 8. Any complexes formed were captured with formaldehyde crosslinking, and analysed on 13% SDS-PAGE gels (figure 4.19).

As shown in figure 4.19, in the absence of psTABp, a complex between psTolA3 and psTolB is formed. In addition, as is found with eTolA3-eTolB-eTABp competition experiment, the psTolA3-psTolB crosslink is decreased by the presence of psTABp, and is completely abolished at higher concentrations. The concentration dependency on this abolition however is different to that of the *E.coli* proteins. It was estimated that the eTolA3-eTolB complex is abolished at 150 μM eTABp, whereas 800 μM of psTABp is estimated to completely abolish the psTolA3-psTolB complex; a 4.5 fold

increase in concentration. Thus, as the K_d for eTolA3-eTABp was estimated at 40 μM, we may therefore conclude that the K_d for psTolA-psTABp is much higher.

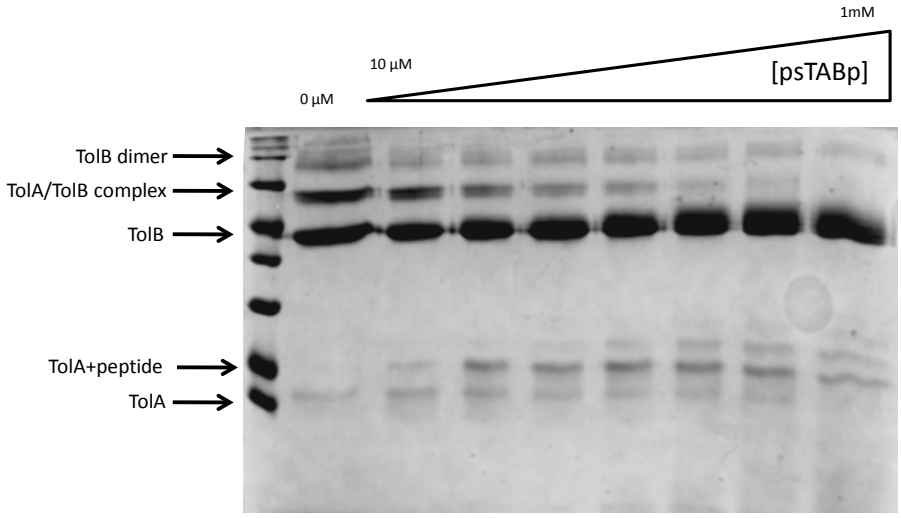


Figure 4.19 Formaldehyde crosslinking experiment showing competition of psTolB and psTABp for binding with psTolA3, *in vitro*, detected with Coomassie blue staining on 13% SDS-PAGE gel. psTolA3 and psTolB were incubated both in the absence (0 μM lane) and in increasing peptide concentration of psTABp. Note the progressive loss of the TolA/TolB complex at higher concentrations of psTABp, and introduction of TolA-peptide band.

To confirm the presence of a crosslink between the psTolA3 and psTABp, the band corresponding to this complex was excised from the gel and analysed by MALDI-MS (Adam Dowle, University of York Technology Facility), which confirmed the presence of both the psTolA3 and psTABp in complex. As ITC data for the eTolA3-eTABp complex estimated a K_d of 40 μM, an extrapolation was made estimating the K_d of psTABp-psTolA3 to several hundred micromolar, assuming a similar relationship between the TolA binding peptides and TolA3. Unfortunately, when attempting to collect

thermodynamic data for the interaction between psTABp and psTolA3 no heats of binding were found upon titration of the peptide. Large heats were detected when peptide at high concentrations was titrated into the cell, characteristic of some form of dilution induced multimeric dissociation (data not shown). These heats caused by (potentially) multimeric complexes of psTABp dissociating upon injection into the ITC cell (where peptide was at low concentration relative to the needle reservoir) would mask any heats created by the formation of a psTolA3-psTABp complex, heats that would likely be very small given the likely very weak nature of the psTolA3-psTABp interaction.

To further investigate the behaviour of the peptide, Size Exclusion Chomatography - Multi-Angle Laser Light Scattering (SEC-MALLS) using Wyatt Dawn HELEOS-II 18-angle light scattering detector and Wyatt Optilab rEX refractive index monitor linked to a Shimadzu HPLC system (University of York Technology Facility) was performed on equivalent high concentrations (as used in ITC experiments) of psTABp with Superdex peptide HR10/30 column (GE Healthcare). These data appeared to show that the psTABp was in dynamic equilibrium between monomer and dimer at concentrations above 500 μM (data not shown). Further ITC experiments were attempted titrating psTABp into psTolA3 at concentrations ranging from 250 μM to 2mM, however heats were either too weak to be detected above baseline noise (at concentrations below 500 μM, or heats caused by dimer dissociation overwhelmed any potential binding heats at concentrations above 500 μM. This prevented calculation of affinity for psTolA3-psTABp complex. Additionally, as stated in section 4.2.1.6, no thermodynamic data was obtained for the native psTolB-psTolA3 interaction.

4.2.9.2 Crystallising the *Pseudomonas aeruginosa* TolA domain 3 - *Pseudomonas aeruginosa* TolA binding peptide complex

Attempts were made at crystallising psTolA3 in the presence of psTABp to potentially define the binding site of psTABp on psTABp, using a variety of commercially available screens (Peg-Ion 1&2, Hampton 1&2, Morpheus, PACT and Index) as well as conditions reported to produce original crystal structure of free psTolA3 (5% wt/v PEG 6000, 100 mM Tris-HCl, pH 8.0, and 1 mM $ZnSO_4$) (Witty et al. 2002). Protein concentrations ranged from 1-14 mg/ml, in presence of 1-10 fold molar equivalent of peptide. No crystals grew in any condition screened, although precipitation was noted in multiple conditions. Optimisation of conditions was attempted by Justyna Wojdyla (University of York), however, to no success.

4.2.9.3 Characterising the behaviour of *Pseudomonas* TolA domain 3 in the presence and absence of *Pseudomonas aeruginosa* binding peptide with Analytical Ultra Centrifugation

In order to further characterise the behaviour of *Pseudomonas aeruginosa* domain 3, both in the presence and absence of its binding peptide, analytical ultracentrifugation experiments were performed, measuring the sedimentation velocity of the protein in the presence and absence of the peptide, particularly to determine if any aggregation was occurring. This was relevant for Nuclear Magnetic Resonance experiments on psTolA3-psTABp, discussed in the following chapter.

Thus 3 samples were prepared; psTolA3 at both low (250 µM) and high (800 µM) concentrations, as well as a sample of psTolA3 at 800 µM in the presence of 1.5 molar equivalent psTABp.

As shown in figure 4.20B, when analysed with SEDFIT (see appendix section 7.7 for details of fitting), the sedimentation co-efficient for psTolA3 at 800 μM is 1.5 S, with a single peak for it's sedimentation co-efficient, indicating that psTolA3 remains monomeric, with no suggestion of dimers or higher order complexes. When transformed with SEDFIT (figure 4.20C) to indicate molecular weight, mass is calculated as 14180 Da (± 700 Da), which when compared to the expected mass of 14495 Da is within error values. Equally, at 250 μM, similar results are found (figure 4.21), indicating protein sample is monomer with a calculated mass of 13973 Da (± 1335 Da). Finally, when studying the protein's behavior in the presence of the psTABp, it would appear as though the presence of peptide has little effect on the protein's sedimentation co-efficient. This may indicate that the psTolA3-psTABp complex is very weak and only a small population remain in complex during ultracentrifugation. This is consistent with the previous crosslinking results which suggest that the psTolA3-psTABp is a very weak interaction. Otherwise, in the presence of the peptide, psTolA3 remains monomeric, with no indication of higher order complexes forming. The peptide would appear to be aggregating, causing a smaller sedimentation co-efficient peak (as seen in figure 4.22). This is again consistent with SEC-MALLS data that suggested that the psTABp was in dynamic equilibrium between monomer and dimer. Due to the added absorbance due to the addition of psTABp in the sample, no molecular weight can be deconvoluted from sedimentation co-efficients. However, when comparing the behavior of the psTolA3 in all conditions (figure 4.23 and table 4.3), the psTolA3-psTABp sample overlays in a similar manner to psTolA3 only sample, thus it would appear to remain monomeric in the presence of the peptide.

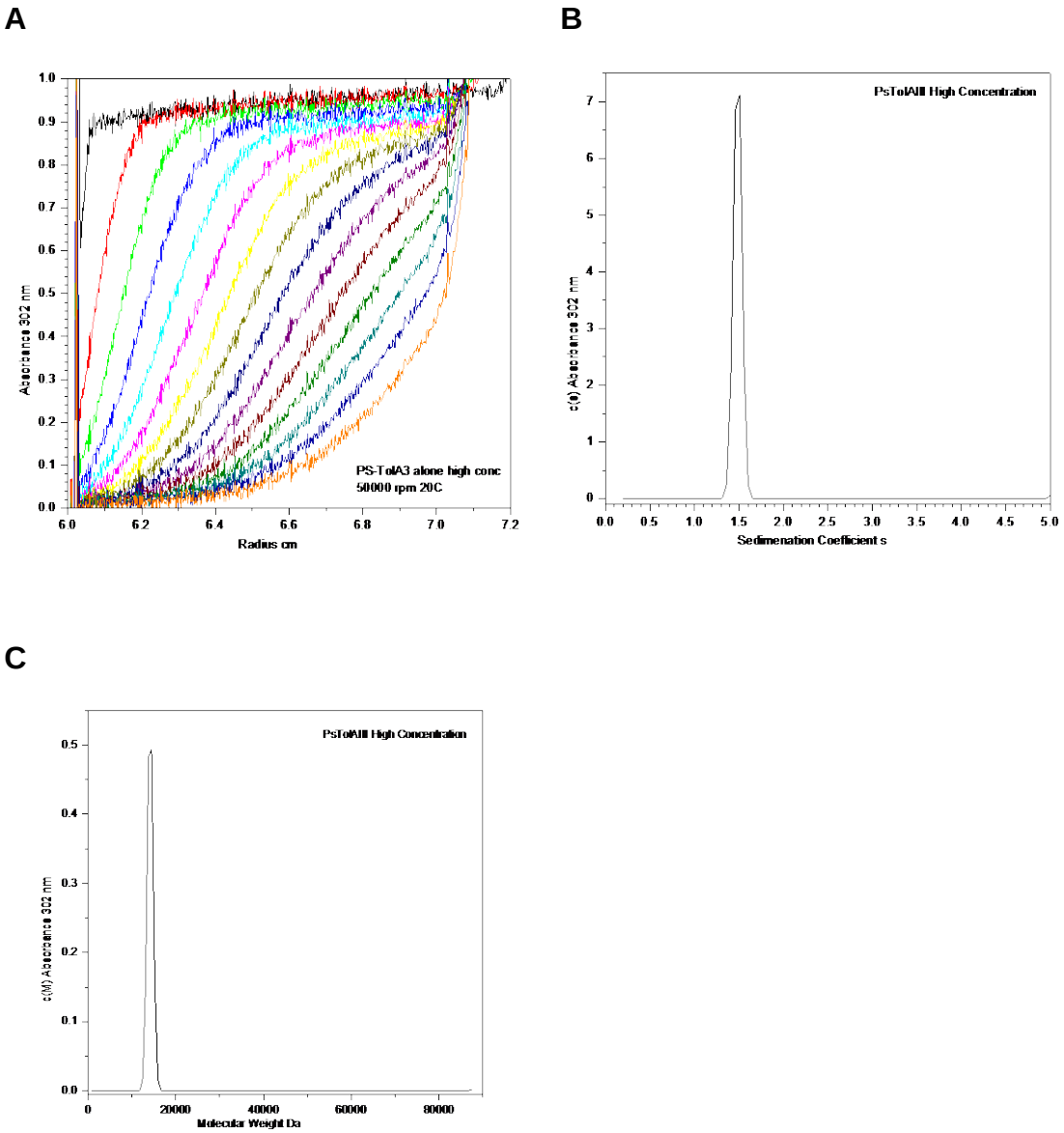


Figure 4.20 Analytical ultracentrifugation data plots for high concentration of free *Pseudomonas aeruginosa* TolA domain 3. 800 µM psTolA3 was sedimented at 50000 rpm for 10 hours at 20 °C, with absorbance measured at 302 nm, with absorbance scans collected every 180 seconds in 20 mM Sodium phosphate buffer, pH 7.5. (A) Sedimentation overlay plot with every 10th scan displayed. (B) SEDFIT c(s) analysis plot displaying sedimentation Co-efficient, S. Single narrow peak recorded, indicating a single species. Sedimentation Co-efficient recorded as 1.49 S (1.50 S when corrected for buffer density and viscosity). (C) SEDFIT s(M) molecular weight transformation plot indicating single species estimated to be 14180 Da (± 700 Da).

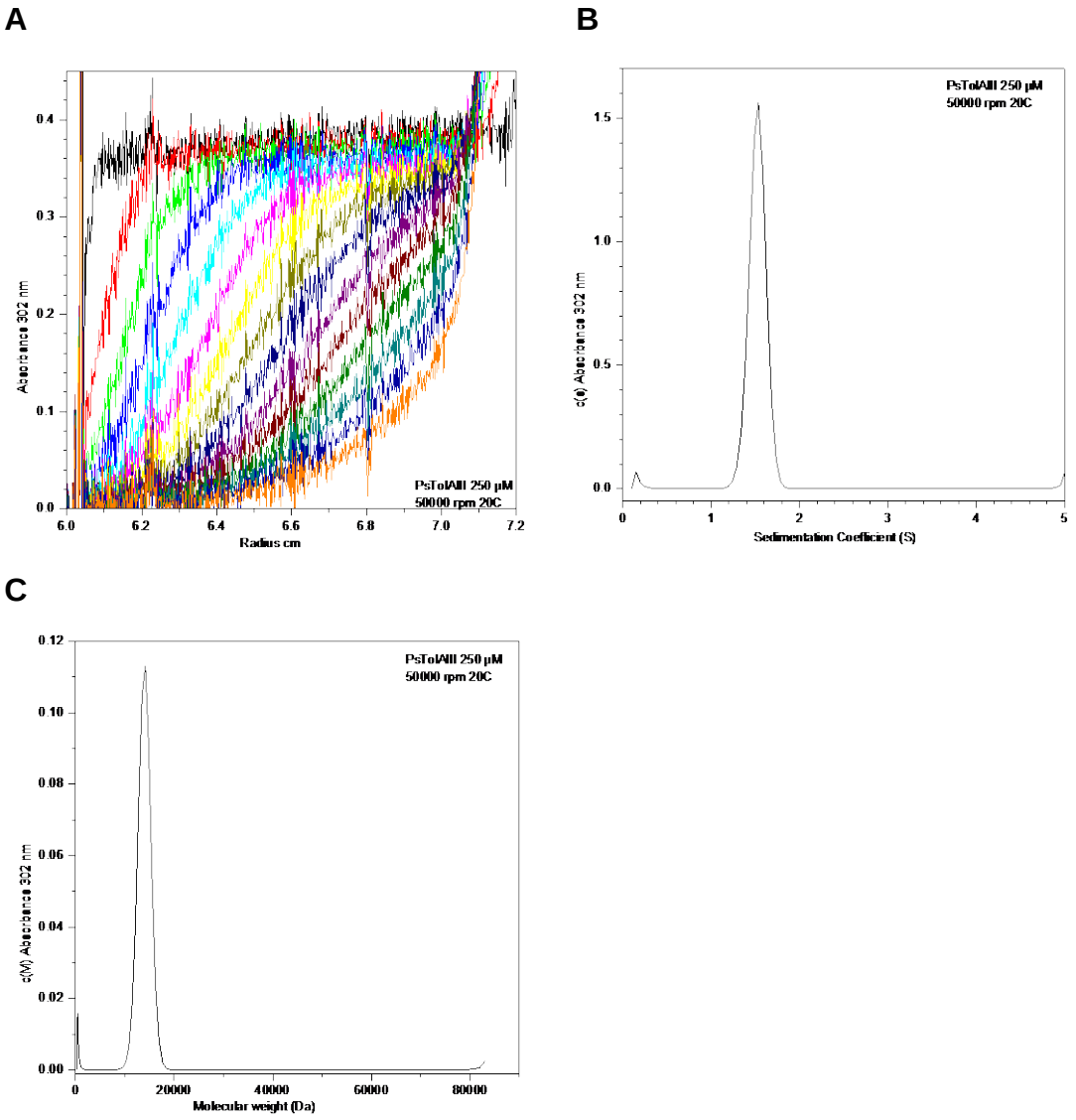


Figure 4.21 Analytical ultracentrifugation data plots for low concentration of free *Pseudomonas aeruginosa* TolA domain 3. 250 µM psTolA3 was sedimented at 50000rpm for 10 hours at 20 °C, with absorbance measured at 302 nm, with absorbance scans collected every 180 seconds in 20 mM Sodium phosphate buffer, pH 7.5. (A) Sedimentation overlay plot with every 10th scan displayed. (B) SEDFIT c(s) analysis plot displaying sedimentation Co-efficient, S. Single narrow peak recorded, indicating a single species. Sedimentation Co-efficient recorded as 1.52 S (1.53 S when corrected for buffer density and viscosity). (C) SEDFIT s(M) molecular weight transformation plot indicating single species estimated to be 13973 Da (± 1335 Da)

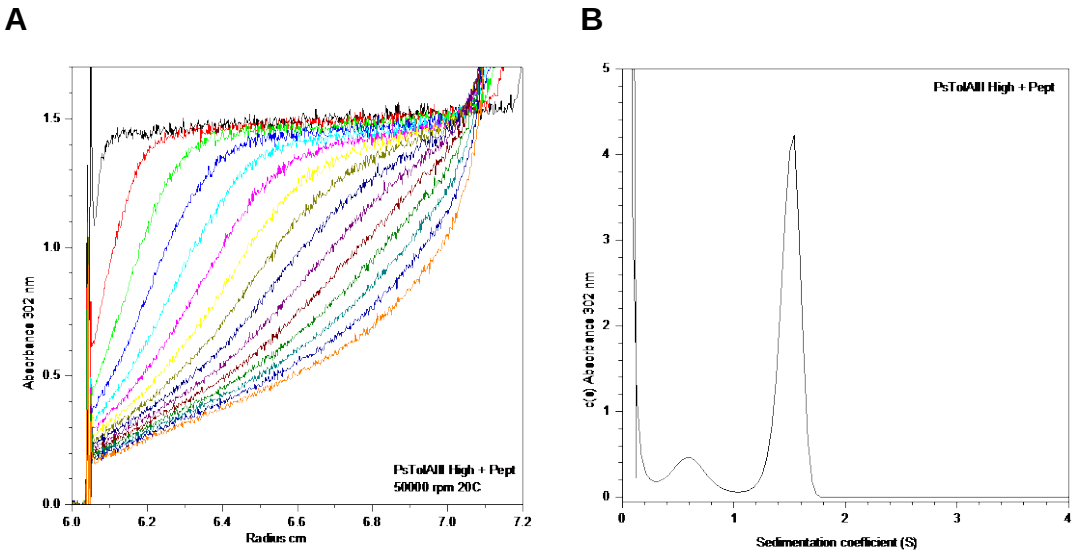


Figure 4.22 Analytical ultracentrifugation data plots for high concentration of *Pseudomonas aeruginosa* TolA domain 3 in presence of 1.5x molar equivalent *Pseudomonas aeruginosa* TolA binding peptide. 800 µM psTolA3 in the presence of 1.5 molar equivalent psTABp was sedimented at 50000rpm for 10 hours at 20 °C, with absorbance measured at 302 nm, with absorbance scans collected every 180 seconds in 20 mM Sodium phosphate buffer, pH 7.5. (A) Sedimentation overlay plot with every 10th scan displayed. (B) SEDFIT c(s) analysis plot displaying sedimentation Co-efficient, S. 2 peaks recorded, indicating a multiple species. Sedimentation Co-efficients recorded as 1.50 S (1.51 S when corrected for buffer density and viscosity) for psTolA3 and approximately 0.6 S (0.6 S when corrected for buffer density and viscosity) for psTABp, indicating that although protein is monomeric, peptide appears to be aggregating.

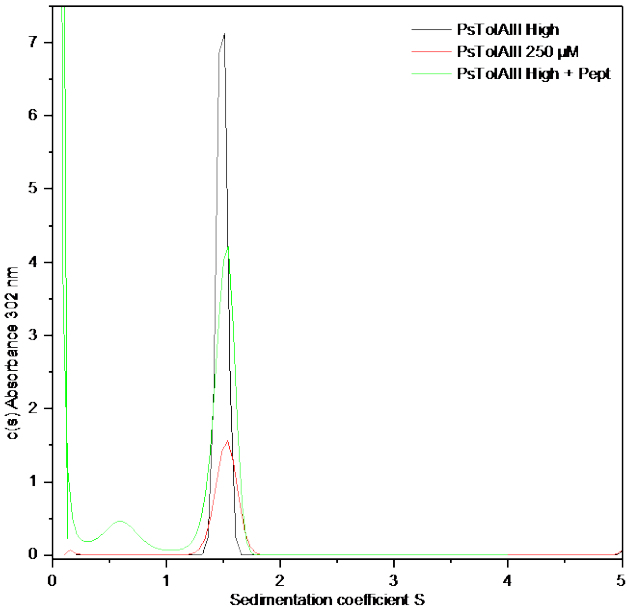


Figure 4.23 Analytic ultracentrifugation data plots for high concentration of *Pseudomonas aeruginosa* TolA domain 3 in presence of 1.5x molar equivalent *Pseudomonas aeruginosa* TolA binding peptide. Comparative plot of 800 µM psTolA3 (black trace), 250 µM psTolA3 (red trace) and 800 µM psTolA3 in the presence of 1.5 molar equivalent of psTABp (green). In all 3 conditions, psTolA3 appears to remain monomeric, and sediments with very similar Sedimentation Co-efficients, although the psTolA3-psTABp profile suggests a small shift to the right, indicating the possibility of a higher molecular weight species.

Sample	S*	$S_{20,w}$	MW (Da)
PsTolA3 (800 µM)	1.49	1.50	14180 (± 700)
PsTolA3 (250 µM)	1.52	1.53	13973 (± 1335)
PsTolA3 (800 µM) + psTABp	1.50	1.51	N/A
psTABp	0.60	0.60	N/A

Table 4.3 Comparison of Analytical Ultracentrifugation data collected for *Pseudomonas aeruginosa* TolA domain 3 and *Pseudomonas aeruginosa* TolA binding peptide. S* denotes Sedimentation Co-efficient, $S_{20,W}$ denotes Sedimentation Co-efficient corrected for buffer density and viscosity. MW denotes molecular weight prediction based on Sedimentation Co-efficients, error values in brackets.

4.2.10 The interaction of TolA binding peptides is specific to their respective TolA's.

Although it had been shown previously that the TolA binding peptides interacted with their respective TolA's, and that the TolA-TolB interaction is specific, I investigated the specificity of the TolA binding peptides for their respective TolA's. Due to the design of the peptide, 3 lysine residues were added at the C-terminus to increase solublity (peptides without any additional sequence were also synthesised, but they remained insoluble). As primary amines (such as Lysine) are active in formaldehyde crosslinking, it had been suggested that these 3 lysines were crosslinking in a non-specific manner, causing a non-native interaction with TolA3. If this were true, then all work regarding crosslinking and TolA binding peptides would be void.

To investigate this, eTolA3 was incubated with crosslinker in the presence of psTABp (and conversely, psTolA3 incubated with crosslinker in the presence of eTABp). In addition, a competition assay with eTolA3-eTolB was performed in the presence of psTABp (and visa versa). Appropriate positive controls of eTolA3-eTABp and psTolA3-psTABP were also conducted. All whole proteins were 10 µM (final), and in the presence of up to 1mM (final) peptide, and crosslinking reactions performed at pH 7.5, 8.0 and 8.5.

In formaldehyde competition assays, neither peptide could abolish the native interaction, like the class specific peptide could. In addition, no complex was found between either eTolA3 and psTABp or psTolA3 and eTABp in the presence of formaldehyde crosslinker, at high concentrations of peptide.

ITC experiments were also performed, titrating psTABp (1mM) into eTolA3 (250 µM) and eTABp (1mM) into psTolA3 (250 µM) at a variety of pH's (6.5-8.5) and temperatures (15 °C - 37 °C), again, no heats were detected indicating an interaction (data not shown).

4.3 Discussion

4.3.1 The Gram-negative TolA-TolB interaction.

The findings of this work have reported for the first time a complex between *Pseudomonas aeruginosa* TolA and TolB, and that this interaction involves the N-terminus of TolB, like that of the *E.coli* interaction. The Tol protein family is highly conserved throughout Gram-negative bacteria (Sturgis 2001), yet until the positive crosslinking data presented in this work, other than assumptions based on homology with *E.coli* proteins, it was unknown if other Gram-negative Tol proteins interacted in a similar manner to that of the *E.coli* proteins. Given that the *Pseudomonas aeruginosa* TolA domain 3 appears to interact with TolB, this suggests that this interaction may be conserved throughout Gram-negative bacteria. However, despite this positive result, there is still no thermodynamic data available for this interaction, as the *Pseudomonas aeruginosa* Tol proteins appear to be of limited solubility, particularly the *Pseudomonas aeruginosa* TolB. This limit in solubility has prevented ITC experiements from being carried out. Having suggested that the TolA-TolB interaction is conserved, it begs the question as to firstly whether or not psTolB has an intrinsically disordered N-terminus like it's *E.coli* homologue. Given the psTABp results (discussed below), this seems likely. Also, given the near identical fold of psTolA3 and eTolA3, it would suggest that psTolA3 would make a good model for study of the TolA-TolB interaction, especially as a high resolution crystal structure is available for psTolA3 (Witty et al. 2002). At the time of the inception of these experiments, no crystal structure of *E.coli* TolA3 in the absence of binding partners was available. During the preparation of this manuscript, a crystal structure of *E.coli* TolA3 both in isolation and in complex with colicin A has been published (Li et al. 2012). In their work, Li et al. also attempted to define the structure of the *E.coli* TolA3-TolB complex, through both crystallography and NMR, but were unsuccessful. This is particularly

relevant for subsequent work presented in the following chapter. Additionally, given that the *E.coli* TolA-TolB interaction is blocked by *E.coli* Pal, it raises the question as to whether or not psPal mediates the interaction of psTolA-psTolB in a similar fashion. However, unfortunately, although the psPal gene was cloned in this work it did not express, and thus protein could not be purified to test this hypothesis. Again, as psTolA3 and eTolA3 have near identical folds (Witty et al. 2002), and both psTolB and eTolB are predicted to have similar folds based on comparisons of secondary structure from circular dichroism experiments, this would seem to be likely that psPal will bind psTolB and may function as an off-switch for the psTolA-TolB interaction. In any event, proving a common interaction between both *E.coli* and *Pseudomonas aeruginosa* TolA and their respective TolB is an exciting prospect for further work in this area.

4.3.2 Recapitulating Gram-negative TolA-TolB interactions with synthetic peptides.

Having shown in chapter 3 that adding 12 extra residues of the *E.coli* TolA binding epitope of TolB onto the disordered region of colicin E9 is sufficient to generate a novel interaction between the eTolA3 and colicin E9 fusion protein, the next step was to ascertain if this sequence in isolation in the form of a synthetic peptide was sufficient to form a complex with TolA domain 3. As reported in section 4.2.3, not only is a peptide of the sequence of the *E.coli* binding epitope sufficient to create a novel interaction with eTolA3, but that it is also capable of disrupting the native eTolA3-eTolB interaction. Additionally, that when titrated against eTolA3, the thermodynamic data obtained suggested that the eTABp bound eTolA3 as tightly as the native eTolB binds eTolA3. This would suggest that although the N-terminus was known to be important in the interaction between eTolA3 and eTolB (Bonsor et al. 2009), these data show that the 12 residues on the extreme N-terminus of eTolB are the sole site of interaction with eTolA3. The

fact that a synthetic peptide representing the N-terminus of psTolB (predicted based on homology with eTolB N-terminus) also disrupted the native psTolA3-psTolB interaction further suggests that the N-terminus the sole site of interaction for the *Pseudomonas aeruginosa* proteins, and that the N-terminus of TolB may well be the sole site of interaction for any TolA3-TolB in Gram-negative bacteria. Although no thermodynamic data were obtained for the psTolA3-psTABp interaction, comparing crosslinking data with that of eTolA3-eTABp suggests that like the *E.coli* interaction, *in vitro*, psTolA3-psTABP/psTolB is a weak (high micro-molar affinity) interaction. Given the problems with attempting to capture an eTolA3-eTolB/eTABp complex by crystallography or NMR, and given that the psTolA3-psTABp interaction appears to mimic that of *E.coli*, the *Pseudomonas aeruginosa* TolA3-TABp interaction was targeted in an attempt to finally determine the TolA-TolB complex, something that has remained elusive for several decades.

4.3.3 Summary

In the work reported in this chapter I have shown that the *E.coli* TolA-TolB interaction is solely dependant on the disordered N-terminus of TolB, as evidenced by the finding that a synthetic peptide of the sequence of the N-terminus is not only capable of recapitulating the TolA-TolB interaction *in vitro*, but that the peptides interaction with TolA is of a similar affinity to that of the native TolB protein. Unfortunately, attempts to define the structure of the *E.coli* TolA-TolB complex or the site of binding on *E.coli* TolA domain 3, either through the whole proteins, or the *E.coli* TolA-TolA binding peptide complex have been unsuccessful by protein crystallography, nuclear magnetic resonance or protein-crosslinking and mass spectrometry. However, despite this set back, another, very interesting finding was reported in this chapter; for the first time (to our knowledge) we have evidence that not only do *Pseudomonas aeruginosa* TolA and TolB interact,

but that their interaction involves the N-terminus of *Pseudomonas* TolB. Given the degree of homology between *E.coli* and *Pseudomonas aeruginosa* TolA and TolB it is highly likely that the N-terminus is the sole site of interaction between *Pseudomonas aeruginosa* TolA and TolB. In addition, it has been found that the interactions between *E.coli* and *Pseudomonas aeruginosa* TolA and TolB remain specific, and no non-cognate complexes were found between them.

5. Identifying the *Pseudomonas aeruginosa* TolB binding site on *Pseudomonas aeruginosa* TolA

5.1 Introduction

As stated in chapters 3 and 4, the structure of both *E.coli* and *Pseudomonas aeruginosa* TolA3 has been reported, either by nuclear magnetic resonance spectroscopy for the *E.coli* protein (Deprez et al. 2005) or crystallography for the *Pseudomonas aeruginosa* homologue (Witty et al. 2002). Additionally, despite structures being reported for *E.coli* TolB both in isolation, and in complex with *E.coli* Pal or colicin E9 translocation domain (Bonsor et al. 2009), the structure for either *E.coli* or *Pseudomonas aeruginosa* TolA-TolB complex has yet to be reported.

Attempts at determining *E.coli* TolA-TolB complex structure through crystallography and solution NMR, as well as the *E.coli* TolA3-*E.coli* TolA binding peptide complex structure have been unsuccessful, thus attention thus turned to the *Pseudomonas aeruginosa* TolA-TolB complex. The *Pseudomonas aeruginosa* TolA and TolB have been shown interact in a similar manner to that of the *E.coli* proteins (Section 4.2.9), and given that their structures are so similar, it was hypothesised that *Pseudomonas aeruginosa* proteins could act as a model for not only the *E.coli* interaction, but given the potentially conserved nature of the TolA-B complex, all Gram-negative bacteria.

Attempts were made to crystallise the psTolA3-psTolB complex, however, due to the highly unstable nature of the psTolB, most conditions yielded nothing but heavy precipitation. Buffer screens were conducted to improve the solubility of the psTolB (as well as psTolA-psTolB complex), however no improvement was achieved. Given the unstable nature of the psTolB, it was

decided that further characterisation of the psTolA-B complex should focus on psTABp acting as surrogate for psTolB. Although no thermodynamic data had been collected for this interaction, crosslinking experiments (section 4.2.9.1) as well as preliminary NMR data on psTolA3-psTABp were convincing enough to focus on the psTolA3-psTABp complex. Attempts were made to crystallise the psTolA3-psTABp complex, however, these failed to yield crystals in any conditions (section 4.2.9.2). Having been unable to crystallise the psTolA-B complex, nuclear magnetic resonance spectroscopy was the next avenue of structural biology to be explored.

5.1.1 Pseudomonas aeruginosa TolA constructs

Having characterised the psTolA3-psTolB/psTABp interaction with chemical crosslinking (section 4.2.1.5 and 4.2.9.1), this same construct of psTolA3 (pEC4) was chosen to be isotopically labelled in this work. This construct consists of residues 226-347 of psTolA and encodes all of domain 3, as well as 40 residues of domain 2. It has a C-terminal his-tag, and has been both purified and it's fold characterised in the previous chapter (section 4.2.1.4). Apart from an additional single C-terminal residue just before the his-tag (Glu124, as a result of cloning), it is identical to the construct used in the crystallisation of psTolA3 as described by Witty et al. in 2002.

5.1.2 *Pseudomonas aeruginosa* TolA binding peptide

The psTABp is as described in section 4.2.9, and is of sequence ADPLVISSGNDRWKKK. As described in section 4.2.9.1, it is capable of forming a complex with psTolA3 (which can be captured by chemical crosslinking), and is capable of disrupting the native psTolA3-psTolB complex.

5.1.3 Nuclear Magnetic Resonance spectroscopy

Nuclear Magnetic Resonance spectroscopy works by monitoring the disruptions to the magnetic field or spin on a specific nuclei. In order for a nucleus to be probed with NMR, it must have a non-zero nuclear spin, something that is inherent to having an odd number of neutrons. Without a magnetic field applied, all individual nuclei spins are random. In an NMR magnet, when the strong magnetic field (B_0) is applied, all spins are aligned along (or against) the axis of the field and the system is said to be in equilibrium. Although all NMR active nuclei are aligned in the same axis, they still have a specific spin, termed precession (the nuclei process around the field). The frequency of this precession is dependent on the strength of the magnetic field and the gyromagnetic ratio of the nuclei (an inherent property of the nucleus) and is called the Larmor frequency. When a short power pulse of radio frequency is applied to the system, it is at the Larmor frequency, which causes an oscillation in the magnetic field or the spin of the nuclei. This is why a relatively low power radio frequency pulse can overcome the strong magnetic field and cause a change in the spin on a nucleus. This change causes an excitation in the system which decays over time, back to equilibrium. This decay signal (detected by receiver coils) is acquired (acquisition time) and is called the free induction decay (FID). In modern NMR, to visualise the spectra a Fourier transform is applied to the FID. As a single FID is so weak, it is insufficient to overcome signal to noise of the environment, and thus the multiple FIDs are collected and added together and thus their cumulative signal will be stronger. As noise is random, it will not increase in a cumulative manner as quickly as the FIDs do. The number of FIDs collected is referred to as “scans”. Following the FID, the system is allowed to return to equilibrium. The time this takes is the relaxation delay, typically several seconds for small molecules, and longer for larger ones.

The most powerful attribute of NMR is its ability to selectively probe specific nuclei by applying a specific radio frequency pulse which only effects that set of nuclei. For example, a pulse is applied to the system that only causes excitation of ^{1}H, and thus only FID signal from the ^{1}H is detected. Single nuclei excitation is the basis of a 1D experiment.

In a system where there are only two ^{1}H, if both protons are in the same chemical environment, then their behaviour will be identical. However, if the chemical environment is different in any way (such as surrounded by other nuclei which may act to shield the nucleus from the applied magnetic field), then the two nuclei will return to equilibrium in a differing way. It is this difference in the behaviour of nucleus spin returning to equilibrium which can be quantified as the chemical shift (measured in ppm). If a system has multiple ^{1}H nuclei to be probed with a radiofrequency pulse, again, if all the ^{1}H are in the same chemical environment then they will all have the same chemical shift. The chemical shift is the resonant frequency of the nucleus relative to an internal standard. However, in the case of a folded protein not only are there many hundreds of ^{1}H, they will each be in their own individual chemical environment, and thus will have different chemical shifts. However, a 10 kDa protein will have thousands of ^{1}H, and if only the ^{1}H are probed, then it is impossible to differentiate between single nuclei. An example of this is shown in figure 5.1.

Therefore it is necessary to probe different types of nuclei, so that a second or third dimension can be added to the spectra. As ^{1}H is the only highly abundant naturally occurring NMR active nucleus (over 99.9% natural abundance), it is necessary to introduce additional NMR active isotopes into the target. In the case of protein NMR, the most commonly used isotopes are ^{15}N and ^{13}C. ^{15}N and ^{13}C are only naturally abundant at 0.37% and 1.07%, respectively (Harris et al. 2001).

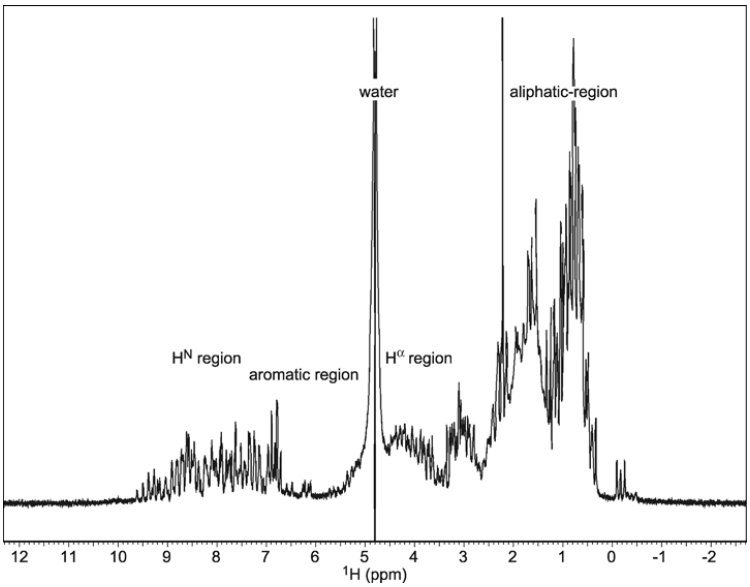


Figure 5.1 Example 1-D Proton NMR spectrum. 1D proton spectrum of ubiquitin (8.5 kDa) in 95%/5% (v/v) H_2O/D_2O solution (Breukels et al. 2011).

An isotope labelled protein sample can then probed using a specific pulse programme that only excites specific nuclei, and as a result through transfer of magnetisation it is possible to correlate nuclei spin systems that are coupled together. For example, in the case of a protein that has been labelled with ^{15}N, it is possible to correlate NH groups through transfer of magnetisation from N to H. This correlation of NH groups is the basis of ^{1}H-^{15}N Heteronuclear Single-Quantum Correlation (HSQC) experiment. This HSQC experiment has two dimensions (^{1}H, ^{15}N) and a peak is correlated to a single NH pair. As every amino acid (with the exception of Proline) has an NH group, and thus in a ^{1}H ^{15}N HSQC spectrum each peak will represent a single amino acid. In addition, $-NH_2$ groups asparagine and glutamine side chains will also be seen in the HSQC spectra as a set of double peaks, usually in the top right hand corner of the spectrum. Furthermore, $-NH_2$ side chain peaks for tryptophan are usually found in the bottom left of the spectrum. Like the ^{1}H 1D experiments, if all the NH groups are in identical chemical environment in a 2D HSQC experiment, then all the chemical shifts would be identical. However, if the chemical environments were different,

such as within a folded protein, then the chemical environment would be different, and as such, the chemical shifts would be highly dispersed. A typical ^{1}H - ^{15}N HSQC spectrum is displayed in figure 5.2.

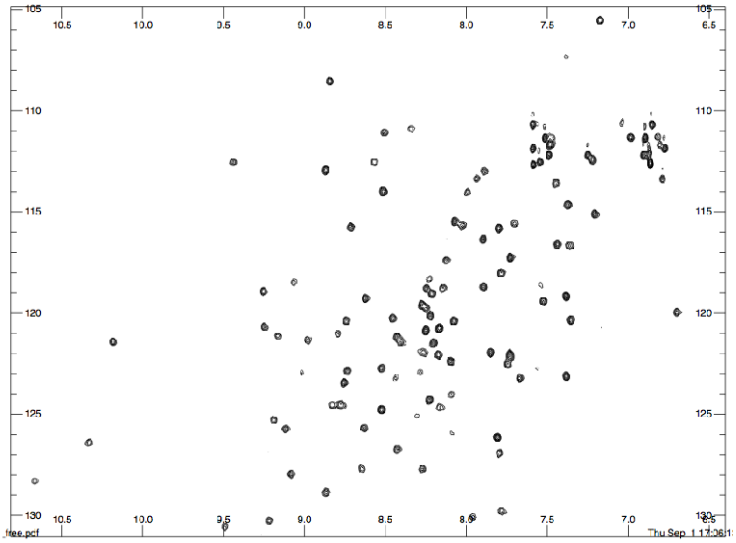


Figure 5.2 Example 2D-HSQC spectrum for a folded protein. 2D HSQC spectrum of *Pseudomonas aeruginosa* TolA domain 3 (13 kDa) in 90%/10% (v/v) H_2O/D_2O solution (this work).

In addition, if the chemical environment of an amino acid (and thus it's representative peak) changes, such as upon a binding event, this peak may enter chemical exchange. There are 3 main types of chemical exchange; fast, slow and intermediate. These regimes are dependant on the timescale of the NMR experiment, which is dependant on the strength of the magnet used. If a peak is in fast exchange, the peak will move from it's original position to a new position (figure 5.3 A). This means that the residue is in a single population, all in the new environment. Alternatively, if a residue is in 2 separate populations, chemical exchange is said to be slow (figure 5.3 B). Thus the residue will have 2 peaks, representing the 2 chemical environments. Finally, a peak may be in a hybrid of the 2 regimes, termed intermediate (figure 5.3 C). This regime is a mixture of the 2, and as a result may broaden out and be indistinguishable from the noise.

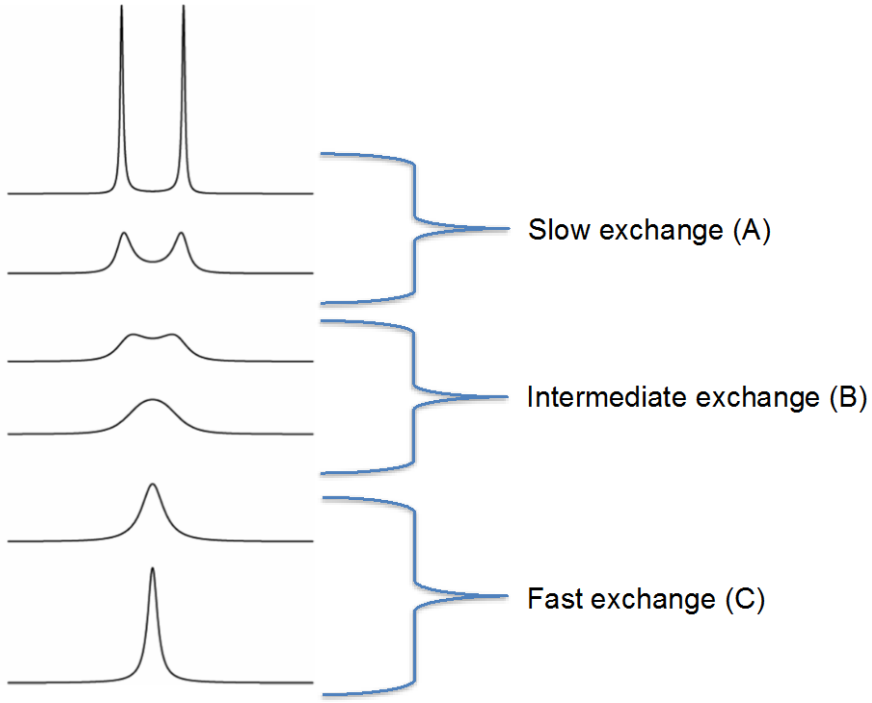


Figure 5.3 Chemical exchange regimes. In slow exchange (A), there are 2 populations and thus 2 peaks for the given residue. In intermediate exchange (B), the peak is in a hybrid situation between fast and slow exchange, and thus a weak, broad peak is populated. In fast exchange, residue is in a single population (C). (Figure adapted from Bain 2003).

Although a 2D spectra deconvolutes the spectrum to the point where it is possible to differentiate individual amino acids, for proteins greater than 30kDa this information is still insufficient to determine which peak is attributed to a specific amino acid. To do this, a 3rd dimension is added to the acquired spectra, recording the chemical shift of carbon nuclei within an amino acid. For each experiment type (HNCO, HNCACO, CBCANH, CBCACONH, see appendix section 7.7 for specific magnetisation transfer), magnetisation is transferred from one nucleus to the next. For example, in HNCACO experiments, magnetisation is transferred from the Nitrogen nucleus to it's proton, and from the C-alpha carbon nucleus of the amino

acid to the CO carbon atom (and back again), so resonances are detected in 3 dimensions; ^{1}H, ^{15}N, ^{13}C. Because the nitrogen is coupled to both its own C-alpha, and the C-alpha of the preceding residue, magnetisation transfers from the nitrogen to both C-alpha's, as well as from C-alphas to their respective COs. This means that via transfer of magnetisation, it is possible to follow the linkage of both the molecule, and it's predecessor. This type of linkage information is termed "i" for current residue, and "i-1" for preceding residue. By combining these carbon dimension experiments, it is possible to move both forward and backward from the residue of interest and identify residues that they are linked to. In addition, as each experiment type has a different scheme of magnetisation transfer, it is possible to identify individual amino acids by the transfer that occurs. For example, some amino acids have multiple carbon atoms that will lead to multiple resonances in the carbon dimension. Also, the chemical shift for Carbon nuclei varies depending on the chemical environment (residue type) that they are in, and for many residue types, these carbon chemical shifts are very distinctive, such as threonine, which has a C-alpha peak at around 60 ppm and a C-beta peak at around 70 ppm. Other distinctive carbon chemical shifts are those of alanine, which only has a single C-alpha peak at ~20 ppm and glycine with a single C-alpha at 45 ppm. These distinctive chemical shifts can act as starting points with which to walk forward or backward from, particularly if further distinctive carbon chemical shifts are found just before or afterward it (see appendix section 7.7 for full list of chemical shifts). Once peaks are assigned, this data can then be mapped back onto HSQC spectra (Cavalli et al. 2007) (Keller 2005) (Knowles et al. 1976).

5.1.4 Aims

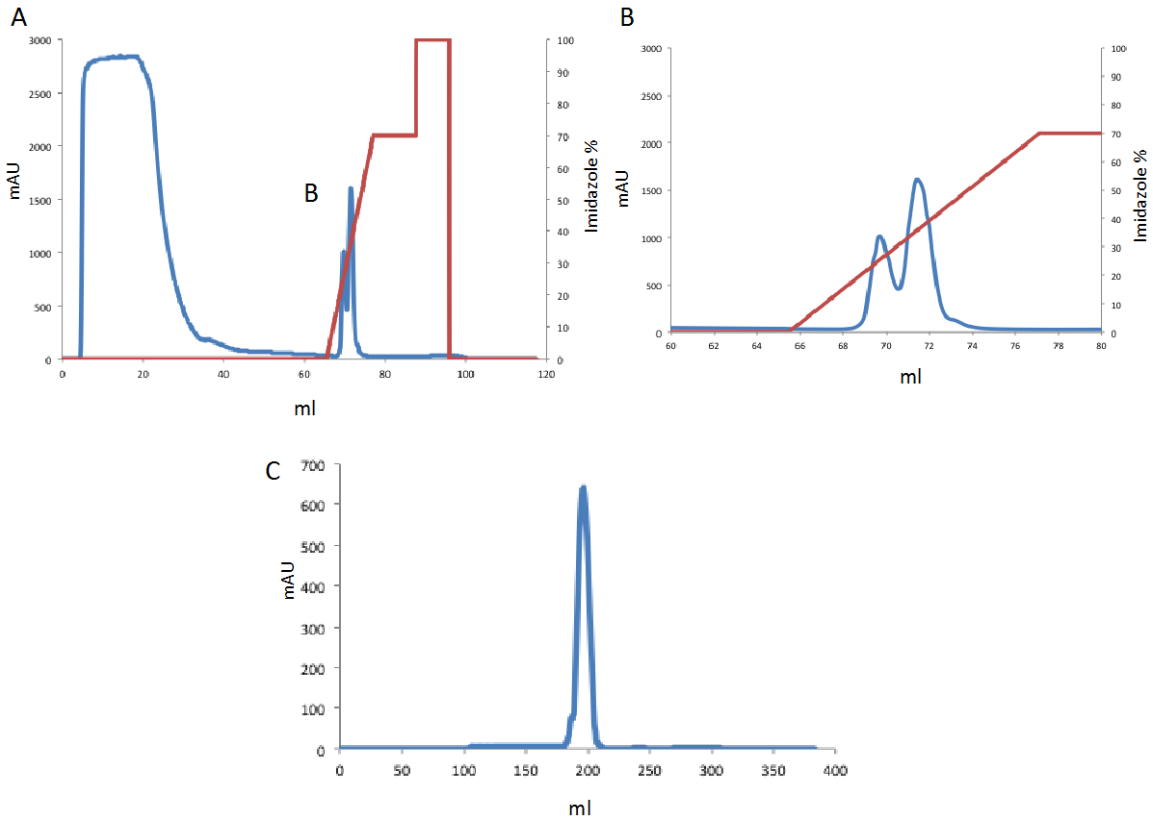
It is the aim of this chapter to investigate the binding of the *Pseudomonas aeruginosa* TolA binding peptide with *Pseudomonas aeruginosa* TolA domain 3 by NMR. psTolA3 was purified in two isotope labelled forms (^{15}N

only and $^{15}N^{13}C$). With this material, HSQC spectra were recorded, both in the absence of peptide (as a starting point of comparison, as well as to ensure protein is folded), and also in it's presence. Once it was determined that the binding of the peptide affected the pattern of peaks seen in HSQC spectra, these spectra were attempted to be fully assigned using three dimensional experiments. Once assigned, peaks that were perturbed by peptide binding were identified and mapped onto the known crystal structure (Witty et al. 2002) in an attempt to map the peptide binding site of psTolA3.

5.2 Results

5.2.1 Purification of ^{15}N labeled *Pseudomonas aeruginosa* TolA3 (pEC4 construct)

His-tagged *Pseudomonas aeruginosa* TolA domain 3 (psTolA3) was purified from BL21 (DE3) cells transformed with pEC4, grown in 1.5L of M9 minimal media supplemented with ^{15}N Ammonium Chloride (Cambridge Isotope Laboratories). Protein was purified as described in section 4.2.1.2, with the following difference; cells were grown to OD_{600} of ~0.6 (usually 6-8 hours following inoculation), induced with 1 mM IPTG, then grown for a further 16 hours. Steps of purification are shown in figure 5.4. A typical protein yield of 4 mg/L of culture was obtained. Mass spectrometry of protein indicated that protein was of the expected size, assuming isotope labelling had been successful and uniform (>95%).



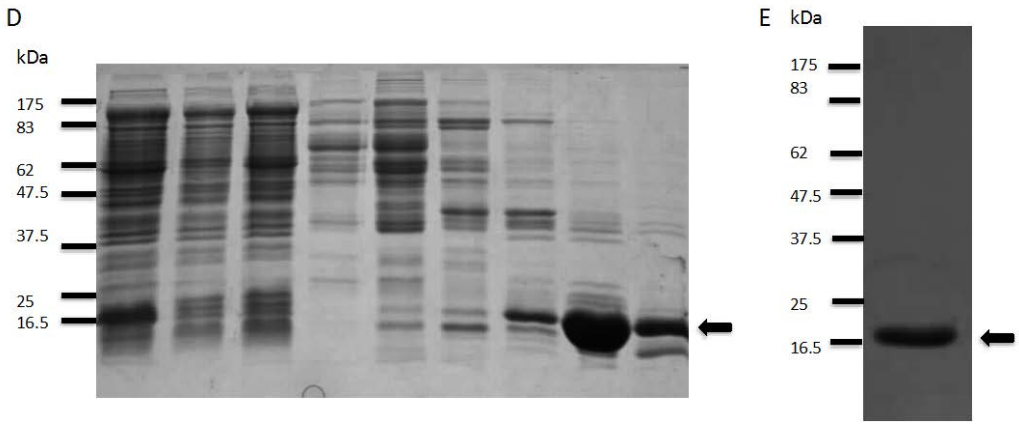


Figure 5.4 Purification of ^{15}N C-terminal His-tagged psTolA3 (pEC4 construct) (A) 280 nm Absorbance profile from Ni-NTA affinity chromatography. (B) denotes expanded view of elution peak. Protein (indicated by arrow) eluted with 0-500 mM Imidazole elution over 10 column volumes. (C) Gel-filtration histogram of eTolA3 on Superdex 75 26/60 column. psTolA eluted between 180 and 210 ml. (D) 16% SDS-PAGE gel to verify presence of psTolA3 in imidazole elution fraction (indicated by arrow). (E) 16% SDS-PAGE gel to verify presence of pure psTolA3 (indicated by arrow) after gel filtration.

5.2.2 Purification of ^{15}N ^{13}C labeled *Pseudomonas aeruginosa* TolA3 (pEC4 construct)

His-tagged *Pseudomonas aeruginosa* TolA domain 3 (psTolA3) was purified from BL21 (DE3) cells transformed with pEC4, grown in 1.5 L of M9 minimal ‘media supplemented with ^{15}N Ammonium Chloride and ^{13}C Glucose (Cambridge Isotope Laboratories). Protein was purified as described in section 4.2.1.2, and in all other respects as 5.2.1. Purified protein is shown in figure 5.5. A typical protein yield of 4 mg/L of culture was obtained. Mass spectrometry of protein indicated that protein was of the expected size, assuming successful and uniform labelling (>95%).

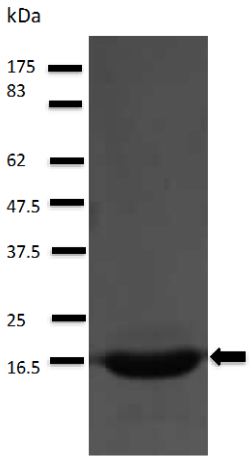


Figure 5.5 Purified ^{15}N ^{13}C C-terminal His-tagged psTolA3 (pEC4 construct) 16% SDS-PAGE gel to verify presence of pure psTolA3 (indicated by arrow) after gel filtration.

5.2.4 *Pseudomonas aeruginosa* TolA domain 3 is folded and stable, and it's stability can be monitored by Nuclear Magnetic Resonance

2D ^{1}H-^{15}N HSQC spectra were collected for singly labelled (^{15}N) psTolA3 to determine if the spectra corresponded to a folded protein, and if so, as a starting point from which to compare peptide bound spectrum. 2D HSQC showed a dispersive peak pattern consistent with that of a folded protein (figure 5.6). CD spectra of protein was also collected and compared to previously acquired CD spectra for psTolA3, and found to be almost identical (section 4.2.1.4).

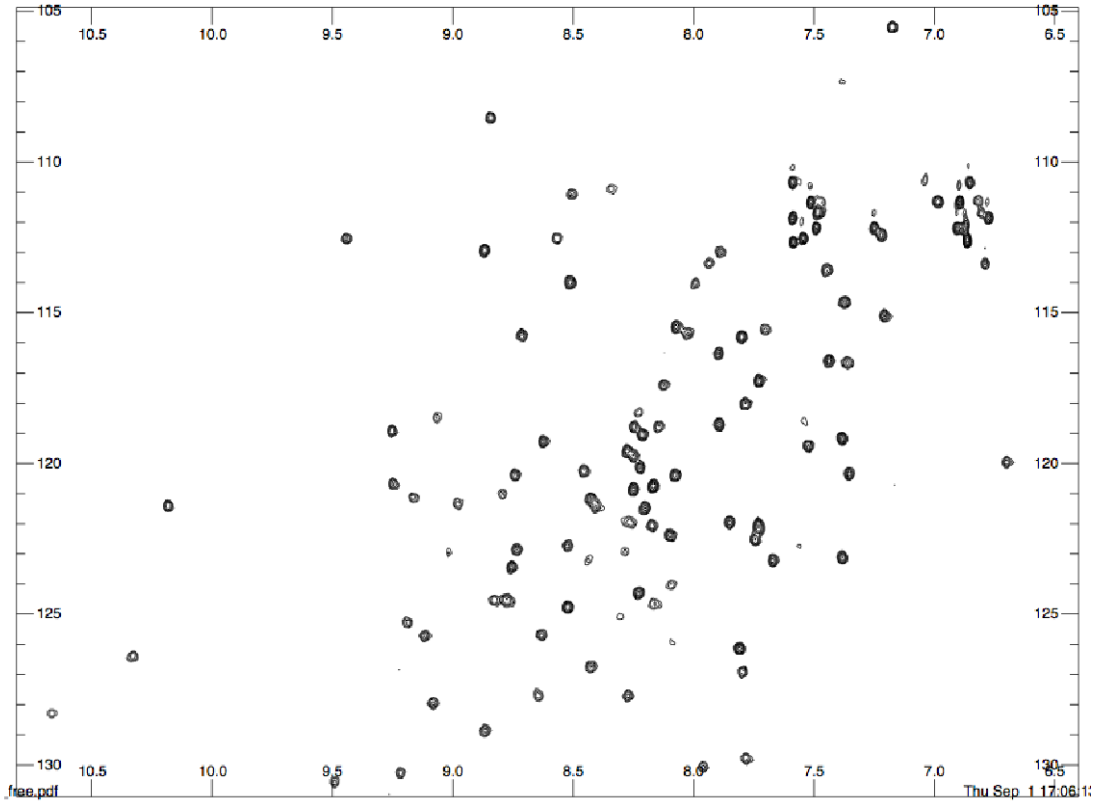


Figure 5.6 2D HSQC spectrum of ^{15}N *Pseudomonas aeruginosa* TolA domain 3. HSQC spectra collected for 220 µM psTolA3, recorded in 20 mM sodium phosphate, pH7.5, at 25 °C.

In the HSQC spectra of psTolA3 alone, 147 peaks were picked and counted. Approximately 30 were judged by their position in the spectra (as well as presence of corresponding doublet peaks) to be sidechain peaks. psTolA3 has 130 residues in total (although the protein has 7 proline residues that are not present in HSQC spectra due to lack of amide protons, as well as the N-terminal residue which will not be seen due to the protonation of the terminal amino group). This means that there should be 122 peaks corresponding to protein backbone resonances, and considering there are 147 peaks, discounting the 30 peaks which are likely to correspond to sidechains, there are approximately 117 peaks in the spectra out of the expected 122.

In addition, HSQC spectra were collected before and after each 3D experiment acquisition set, which indicated that the psTolA3 protein (at ~800 µM) was stable for over 3 weeks at 298K/25 °C. No change in position, number or intensity of peaks was detected over this period. In addition, samples of psTolA3 protein at ~800 µM were analysed on 16% SDS-PAGE gels at regular intervals following 3D spectra acquisition, the protein appeared to be intact with no hint of degradation products (figure 5.7).

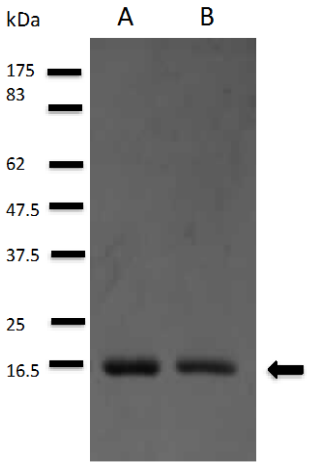


Figure 5.7 ^{15}N ^{13}C labeled *Pseudomonas aeruginosa* TolA3 did not degrade during 3D data acquisition. Double labelled psTolA3 prior to (A) and after over 3 weeks (B) at 298K/25 °C inside NMR magnet.

5.2.5 Binding of *Pseudomonas aeruginosa* TolA binding peptide perturbs a population of peaks in *Pseudomonas aeruginosa* TolA domain 3 HSQC spectrum

Having confirmed that the *Pseudomonas aeruginosa* TolA domain appeared to be folded and stable, I set out to ascertain if the binding of the *Pseudomonas aeruginosa* TolA binding peptide affected the spectra of the psTolA3 protein. The use of synthetic peptides study protein residue perturbations has been used in the past great success. Work in 1992 used 1D ^{1}H NMR to map the binding of the inhibitory region of troponin I,

represented by a synthetic peptide, onto the C-terminal domain of troponin C (Slupsky et al. 1992). More recently, in 2012, work was published on the use of short synthetic peptide dimers to mimic the Cro protein of bacteriophage λ, that not only causes similar NMR chemical shift perturbations *in vitro*, but that *in vivo* also bound to DNA operator sequences with similar selectivity to the Cro protein, potentially paving the way for novel synthetic transcription factors (Mazumder et al. 2012).

HSQC spectra were acquired for *Pseudomonas aeruginosa* TolA domain 3 at 220 μM incubated in the presence of 1.5 molar equivalent (330 μM) *Pseudomonas aeruginosa* TolA binding peptide. This spectra (figure 5.8) showed a similar number of peaks to that of the unbound spectra (144 peaks in bound spectra compared with 147 peaks in unbound), however, when the two spectra were overlaid and compared (figure 5.9), subtle differences were distinguishable in that a small population of peaks (approximately 10-20) were perturbed by the presence of the peptide. These perturbed peaks appeared to be in a variety of exchange regimes, with some peaks appearing, some disappearing, and some shifting position (see appendix section 7.7 for details on types of exchange regime). The majority of other peaks remained at constant intensity and position.

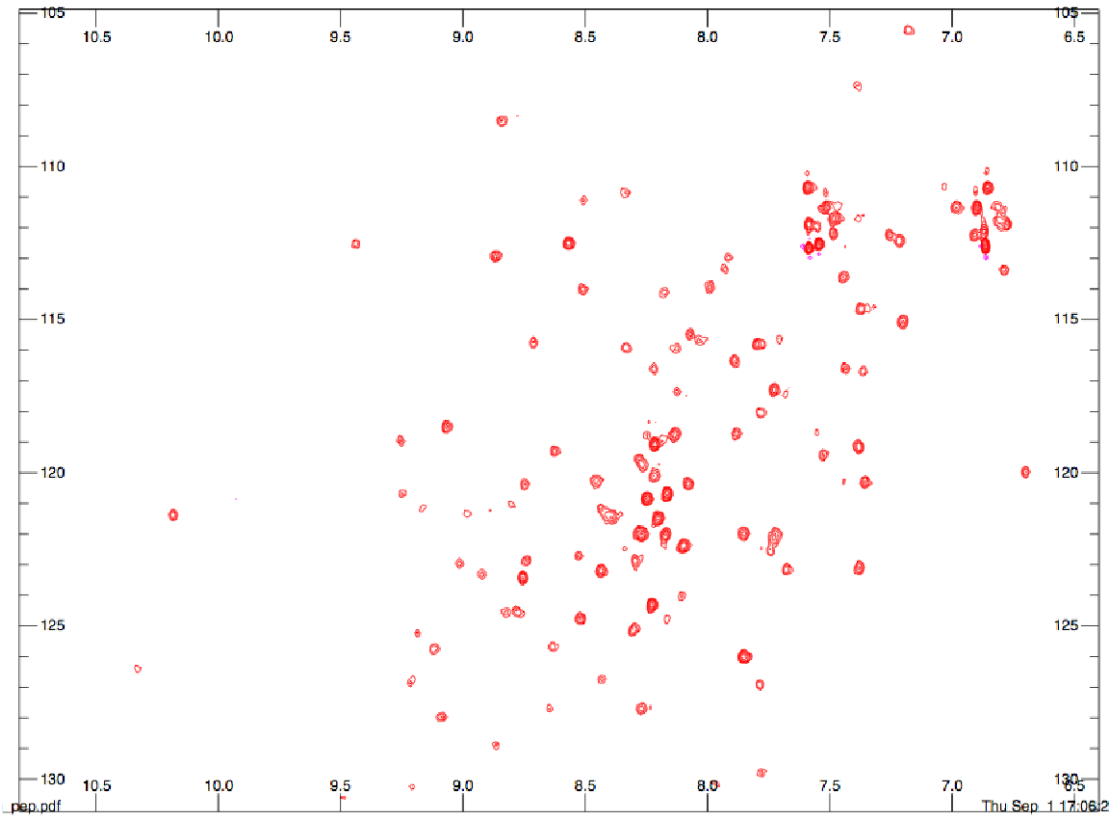


Figure 5.8 2D HSQC spectrum of ^{15}N *Pseudomonas aeruginosa* TolA domain 3 in complex with *Pseudomonas aeruginosa* TolA binding peptide. HSQC spectra collected at 700 MHz for 220 µM psTolA3 in complex with 1.5 molar equivalent psTABp, recorded in 20 mM sodium phosphate, pH 7.5, at 25 °C.

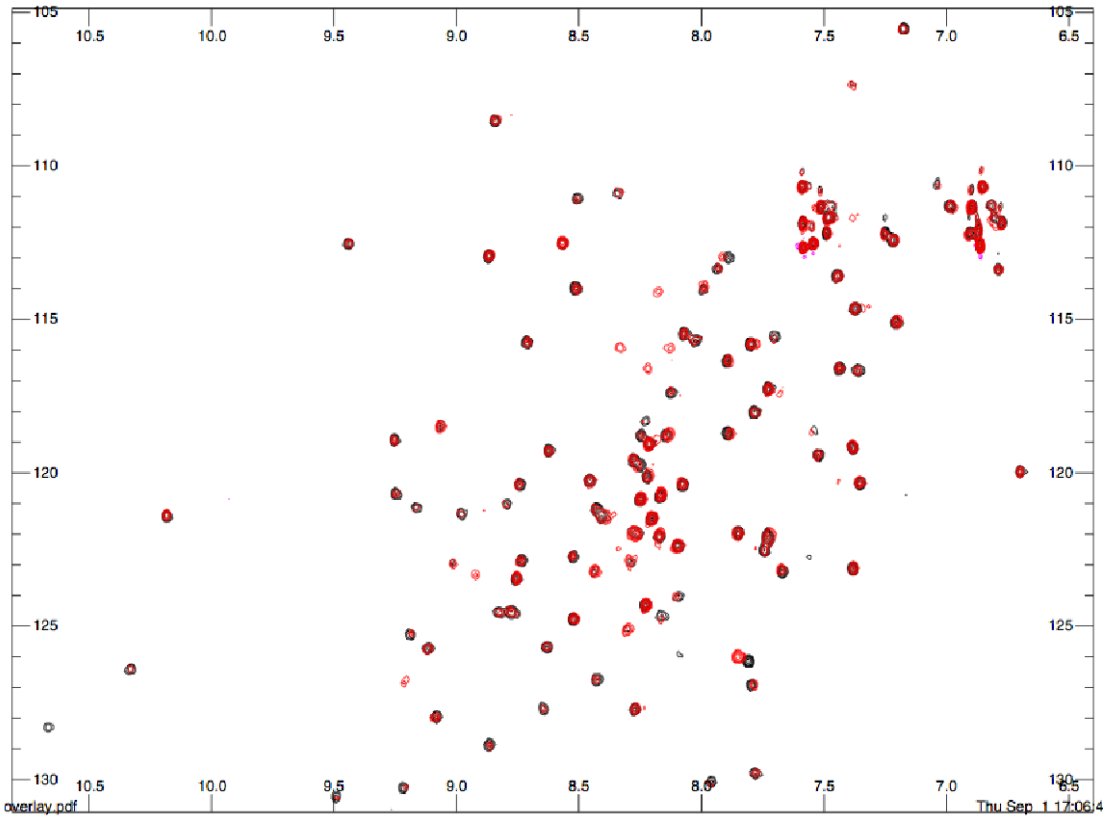


Figure 5.9 2D HSQC overlay spectra of ^{15}N *Pseudomonas aeruginosa* TolA domain 3 and ^{15}N *Pseudomonas aeruginosa* TolA domain 3 in complex with *Pseudomonas aeruginosa* TolA binding peptide. HSQC spectra collected at 700 MHz of free 220μM psTolA3 (black) and in complex with 1.5x molar equivalent (red). Spectra recorded in 20mM sodium phosphate, pH7.5, at 25 °C.

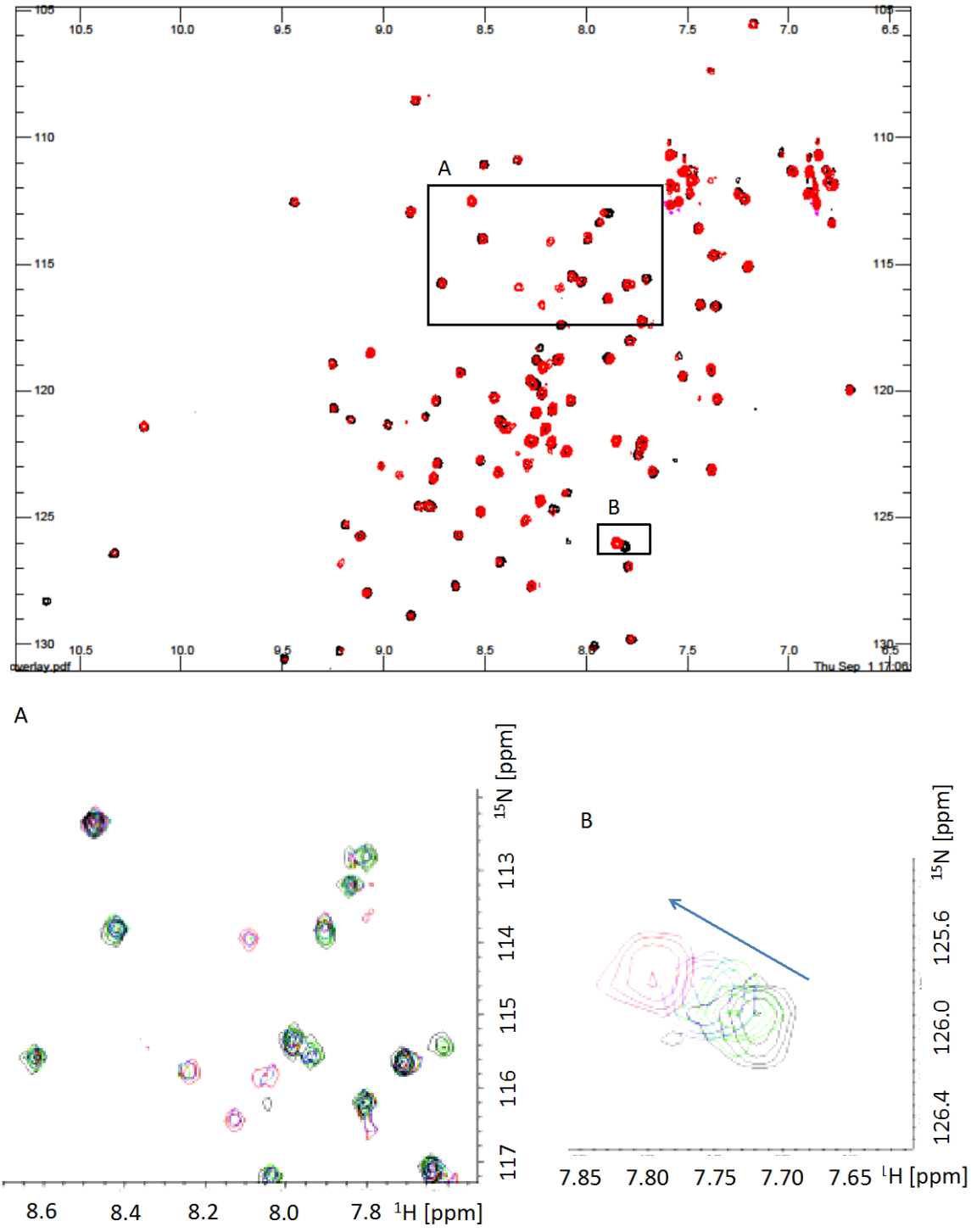


Figure 5.10 *Pseudomonas aeruginosa* TolA binding peptide perturbs *Pseudomonas aeruginosa* TolA domain 3 resonance peaks. (A) shows different populations of peaks; those that disappear in the presence of peptide (slow/intermediate exchange), those that appear in the presence of

peptide (slow/intermediate exchange), and those that remain the same. (B) Movement of single peak in fast exchange from left to right as a function of peptide titration.

In order to define the binding site of the peptide on psTolA3, a series of titrations at increasing concentration were performed. This would allow identification of the first peaks to be affected by the presence of the peptide, and thus the likely binding site of the peptide. In addition, it would also to ensure that increasing concentrations of peptide did not perturb further peaks, and thus the protein was saturated with peptide. psTolA3 was incubated for 10 mins at 25 °C with increasing concentrations of psTABp (0.2, 0.5, 0.7, 1, 1.5, 2, and 5 molar equivalent of psTolA3) and HSQC spectra acquired for each peptide concentration. As shown in figure 5.9, approximately 20 distinct peaks in total were found to be affected by peptide binding (at 1.5 molar equivalent peptide). At 2 molar equivilents of peptide, no further peaks appeared to be affected by peptide binding. At 5 mol. eqv. of peptide, quality of spectra had decreased significantly due peak broadening and was unusable.

In addition, it was noted that the very intense peak at approximately 7.7 ppm in proton dimension and 126 ppm in nitrogen dimension appeared to be in fast exchange during the peptide titration (figure 5.10B). Whereas other peaks decreased in intensity, this peak stayed at uniform intensity, and its position moved from bottom right (at position 7.7, 126 ppm) to top left (position 7.8, 125.7 ppm). However, this particular peak is located at a position that the C-terminal residue peak is usually found (Cavanagh 2007). The C-terminal peak is also usually of high intensity. In this particular construct, the 6xHis tag is located at the C-terminus, and thus it is likely that this intense peak corresponds to the extreme C-terminal histidine residue. As this peak seems to be affected by peptide binding, this is either due to a global conformational change in the protein, or that the peptide is associating

with the C-terminus of the protein. Whether or not this suggests the location of the peptide binding site is on a particular face of the protein or not is unclear, a full assignment of peaks was completed to identify those other peaks that are involved in peptide binding, to ultimately identify the peptide binding site.

5.2.6 Assigning backbone residues of unbound ^{15}N ^{13}C *Pseudomonas aeruginosa* TolA domain 3

To fully assign protein residues of HSQC, the following 3 dimensional (^{1}H, ^{15}N, ^{13}C) spectra were collected from ^{15}N ^{13}C labelled psTolA3 using the following experimental types; HNCO, HNCACO, CBCANH, and CBCACONH. Acquisition parameters for each experiment are reported in materials and methods. Once acquired, spectra were processed and phased using NMRDraw (Delaglio et al. 1995). Spectra were then converted to Azara format using NMRPipe (Delaglio et al. 1995) and CCPN Analysis (Vranken et al. 2005) was used to analyse and assign residues from carbon sequential walks (figure 5.11, see appendix section 7.7 for details of experiment types and sequential backbone assignment).

Of 130 residues in the protein, 7 prolines are present, and thus no peak is seen in HSQC. In addition, 1 peak is also lost as N-terminus, leading to 122 residues to be assigned. There were 147 peaks in total in HSQC spectra, of which 27 were confirmed as sidechains. Therefore 119 peaks were counted and picked in HSQC spectra (not counting sidechain peaks), indicating that 3 peaks were missing from HSQC spectra. Of the 119, 110 peaks were unambiguously assigned to protein residues using sequential walk assignment method using combination of CBCACONH with CBCANH and HNCO with HNCACO. As 110 resonances were assigned, statistically, the HSQC spectra is over 90% assigned. 12 resonances remain unassigned which are the following: Ala2, Ser28, Leu29, Leu36, Val37, Ser38, Asn49,

Gly50, Arg101, Lys116, Glu118, Asp119, Leu120. These resonances were unassignable due to either the quality of 3D spectra, or degenerate nature of carbon chemical shifts, leading to difficulty in differentiating between residue types. Additional 3D spectra (^{15}N NOESY and HCCCONH experiments, as well as Trosy-based pulse sequence of CBCANH) were acquired in an attempt to assign missing residues, however they failed to yield any further assignments. Assignments were mapped onto HSQC spectra, and these are shown in figure 5.12A/B. In addition, assigned residues were mapped onto psTolA3 crystal structure to show assignment coverage using PyMOL 0.99 (Delano Scientific) (figure 5.13). Assignments for residues 2 to 40 which are shown as a long alpha-helical strand in the psTolA3 crystal structure are found to be in the central (8.0 to 8.5 ppm in proton dimension) region of the HSQC spectra, a region that resonances consisting disordered/unstructured protein areas are usually clustered (Cavanagh 2007). Additionally, these peaks appear to be very intense in 3D experiments. This would suggest that rather than being alpha-helical as in the crystal structure (Witty et al. 2002), the stretch from residues 2 to 40 is unstructured. As stated previously, this same region appears to be unstructured in the solution NMR structure of *E.coli* TolA domain 3 (Deprez et al. 2005).

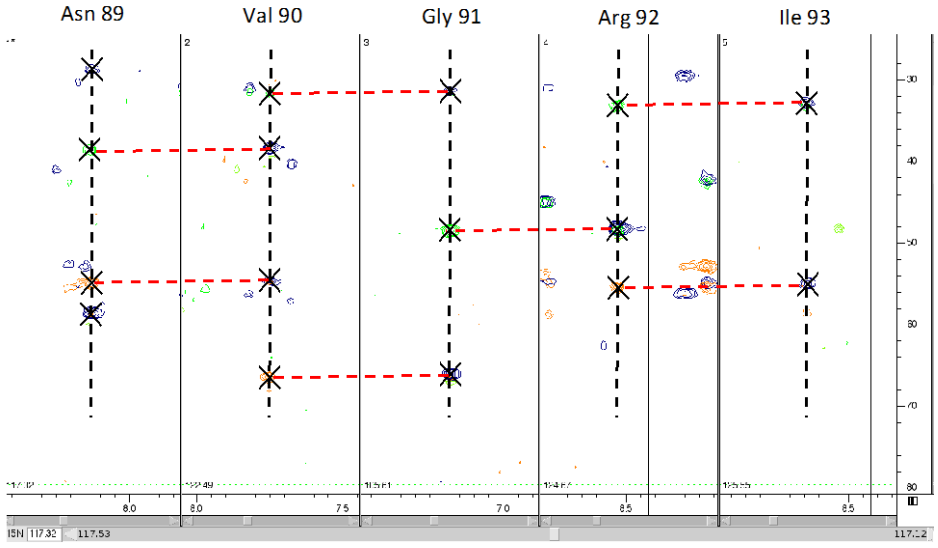


Figure 5.11 Example of sequential walk method used in assigning *Pseudomonas aeruginosa* TolA3 NMR spectrum.

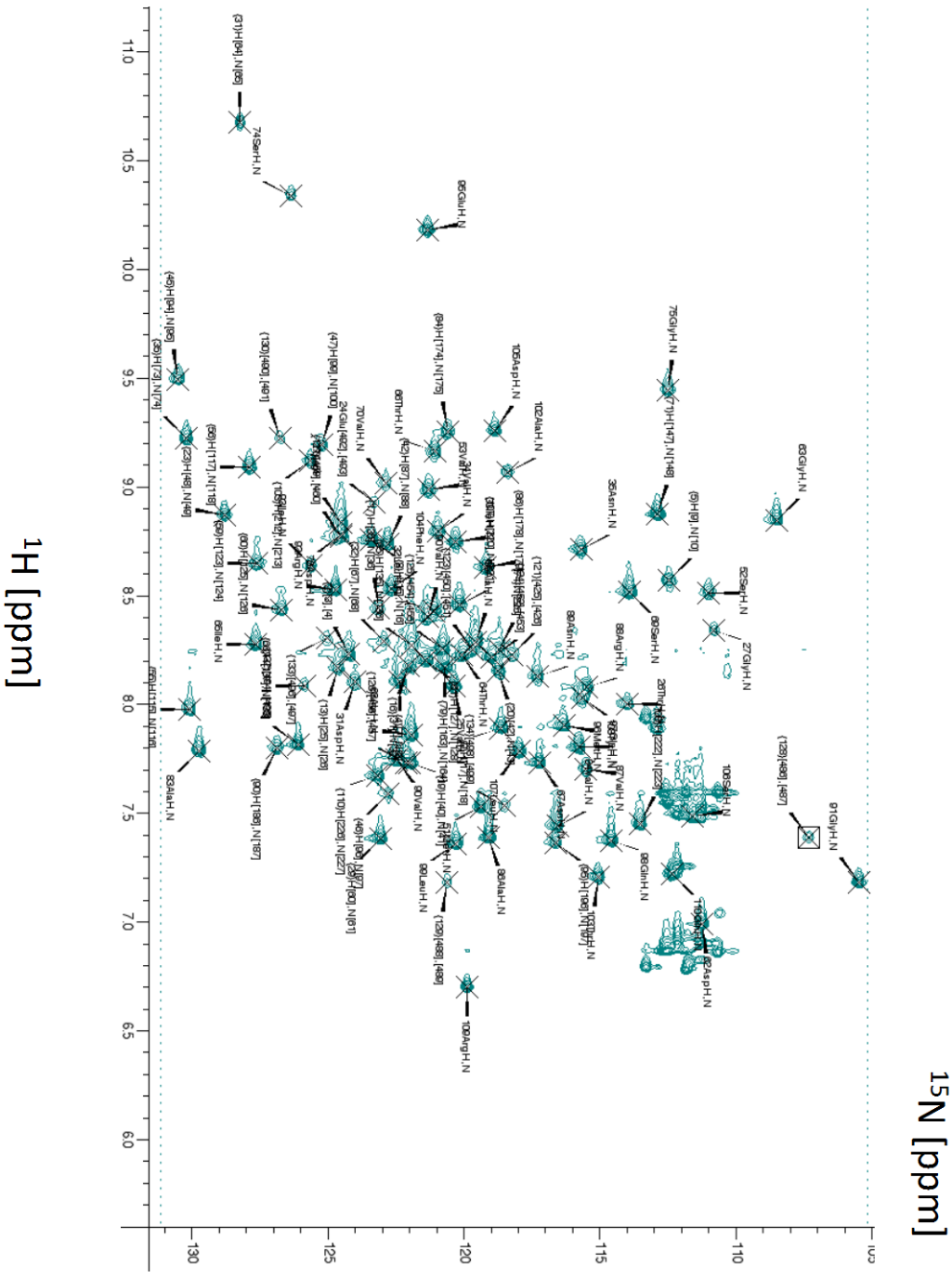


Figure 5.12A 2D HSQC spectra of ^{15}N *Pseudomonas aeruginosa* TolA domain 3 labelled with residue assignments.

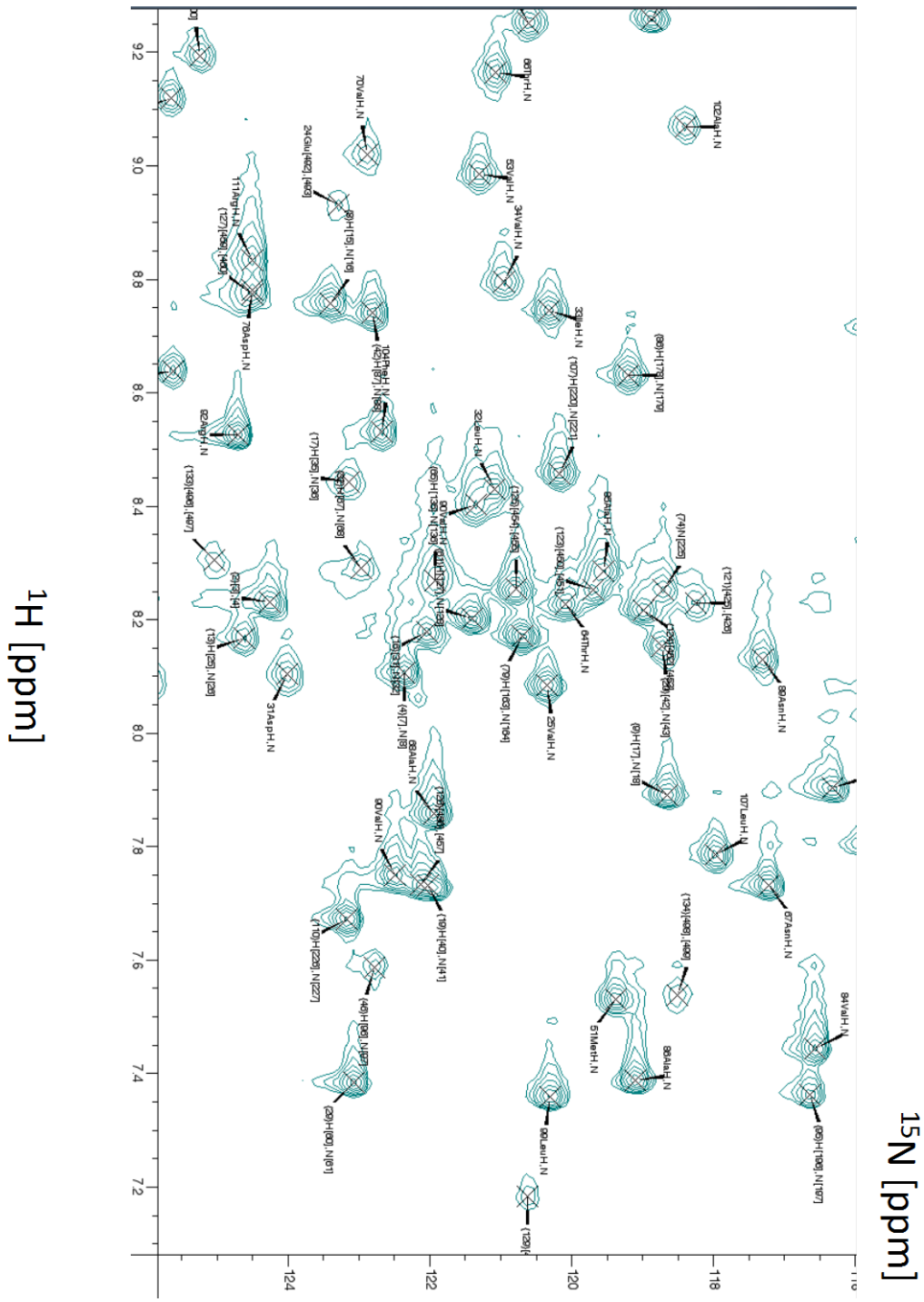


Figure 5.12B Zoomed in view of central (7.2 – 9.2 ppm) region of 2D HSQC spectra of ^{15}N *Pseudomonas aeruginosa* TolA domain 3 labelled with residue assignments.

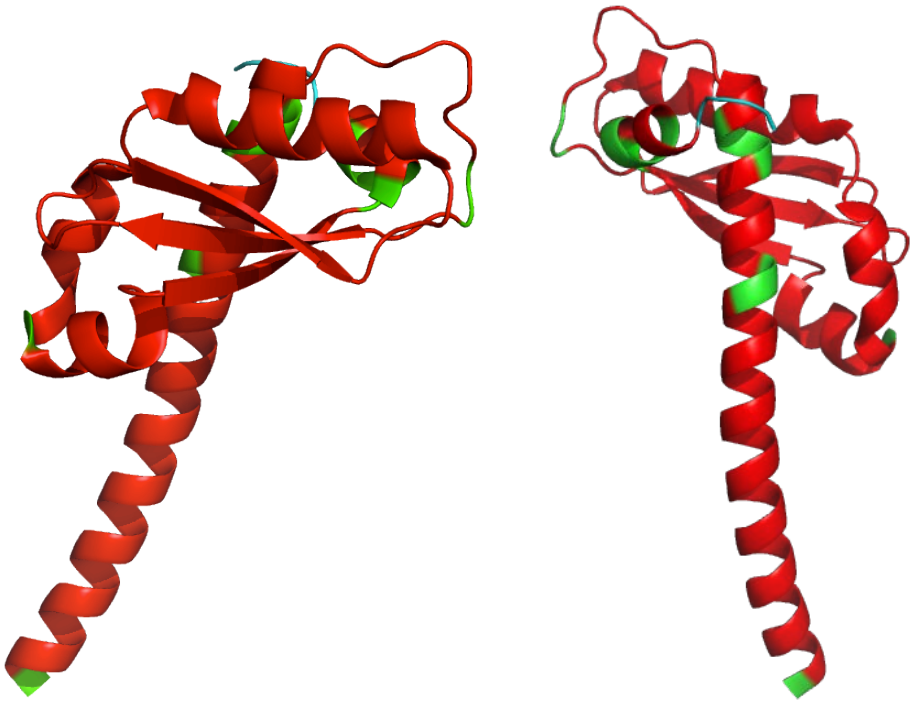


Figure 5.13 Assigned residues of unbound *Pseudomonas aerugionsa* domain 3 from NMR mapped onto crystal structure. psTolA3 (PDB ID: 1LR0) displayed in ribbon format with assigned residues (red) and unassigned residues (green) mapped onto structure to indicate assignment coverage. Cyan indicates C-terminal 6xhis tag.

5.2.7 Predicting and comparing *Pseudomonas aeruginosa* TolA domain 3 chemical shifts with Sparta+ and ShiftX2.

To verify the backbone resonance assignments for the unbound form of *Pseudomonas aeruginosa* TolA domain, as the crystal structure for this protein was already known, chemical shift predictions were made. The PDB co-ordinates (1LR0) were modified to include hydrogen atoms using MolProbity online software (Chen et al. 2010), and subsequently analysed with both Sparta+ (Shen et al. 2010) and ShiftX2 (Han et al. 2011). Sparta+ and ShiftX2 use a neural network algorithms to make chemical shift predictions based on a database chemical shifts for known protein structures (Shen et al. 2010, Han et al. 2011). When comparing predicted and

observed chemical shifts it was noted that Sparta+ yielded significantly better predicted chemical shifts (figure 5.14 for proton and figure 5.15 for nitrogen) when compared to predicted residues. Expected error values were calculated from the standard deviation of predicted errors divided by the square root of the number of shifts. All predictions in both dimensions, with the exception of Gln15 in the Nitrogen dimension were within expected error values, and thus predictions were considered to be of high confidence. For proton dimension, the average root mean square distance between predicted and observed values was 0.412 ppm and for nitrogen dimension it was 2.750 ppm. Given that most ^{1}H shifts are in the range of 7.0 – 9.5 ppm (a range of 2.5ppm), this would indicate that there is approximately 16% error in the proton dimension. In the nitrogen dimension, given that most shifts are between 110 – 130 ppm (a range of 20 ppm), this would give an error of approximately 14%. It should be noted that for residues 1-40 in the proton dimension, the agreement between observed and predicted chemical shifts is poor in comparison to the rest of the protein. While the crystal structure these residues are structured (alpha-helical), in the NMR spectra it would appear as though this stretch unstructured due to their location between 8.0 and 8.5 ppm in ^{1}H dimension. Thus, as the chemical environment is different than expected, the chemical shift prediction will be different. It is possible that these residues are structured in the crystal structure due to crystal packing.

Given that the predicted chemical shifts match well within the expected parameters of prediction with those observed experimentally, it gave good confidence that the backbone assignments were correct.

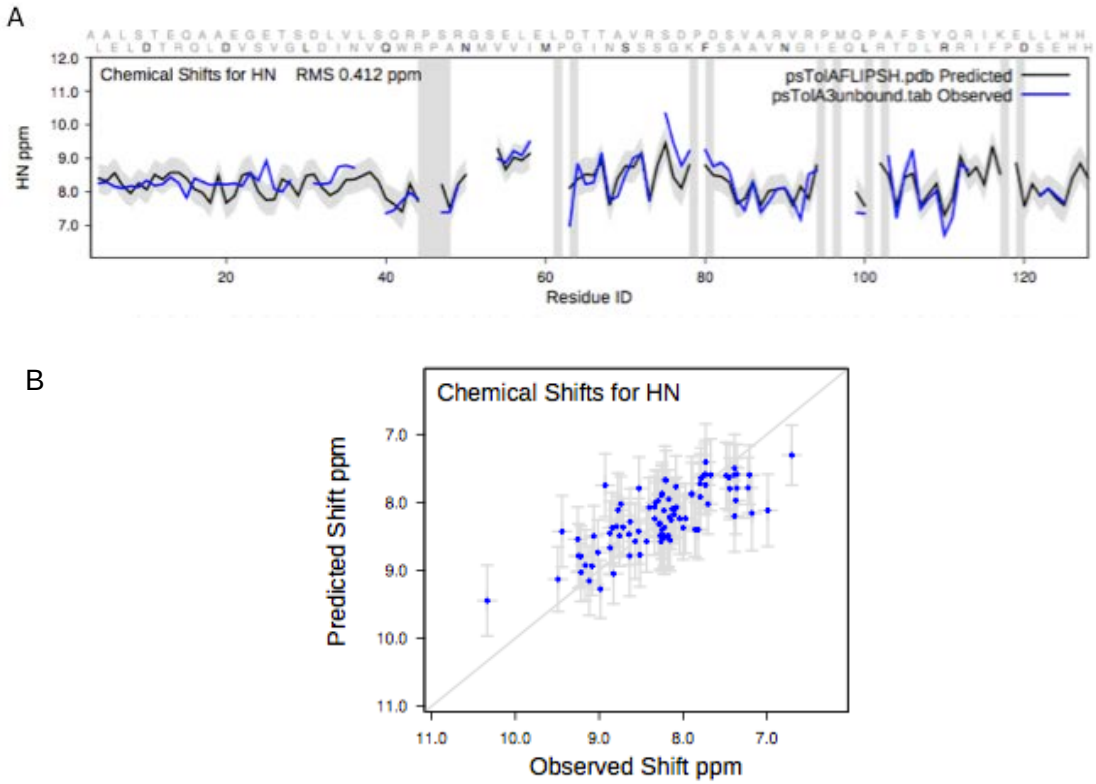


Figure 5.14 Graphical representation of Sparta+ proton chemical shift predictions compared with observed chemical shifts. (A) Predicted chemical shifts (black) against observed (blue). (B) Comparison of predicted and observed absolute chemical shifts.

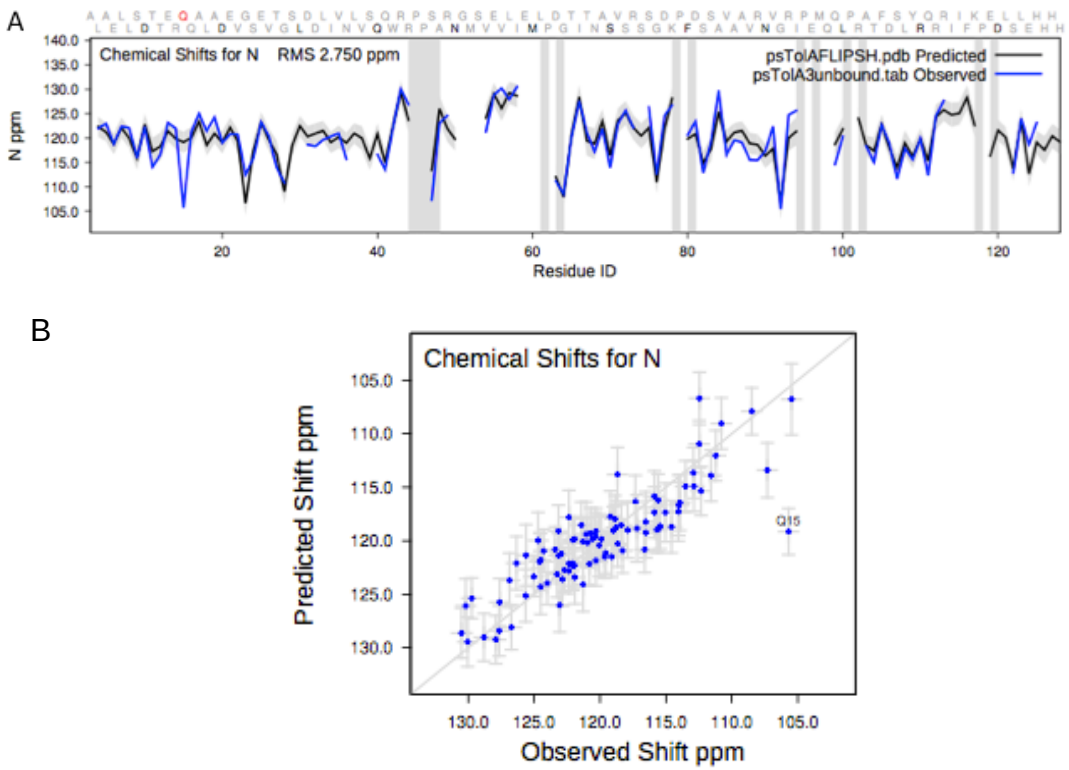


Figure 5.15 Graphical representation of Sparta+ nitrogen chemical shift predictions compared with observed chemical shifts (A) Predicted chemical shifts (black) against observed (blue). (C) Comparison of predicted and observed absolute chemical shifts.

5.2.8 Assigning bound *Pseudomonas aeruginosa* TolA domain 3 in complex with *Pseudomonas aeruginosa* TolA binding peptide

Having assigned greater than 90% of peaks found in the HSQC spectra of unbound psTolA3, a further set of 3 dimensional spectra were acquired on ^{15}N-^{13}C psTolA3, using the following experiment types; HNCO, HNCACO, CBCANH, and CBCACONH. As the majority of peaks were unaffected by binding, the unbound assignments were used as a basis for bound assignments, although all bound backbone assignments were verified through sequential walks.

144 peaks were picked in bound HSQC spectra, of which 21 were sidechains. Like the unbound spectra, there were 122 residues to be assigned, of which 103 were assigned unambiguously using unbound assignments as a basis with sequential carbon linkage walks to confirm connectivity. The 19 unassigned residues were either not present in the unbound assignment (12 residues), or their identity could not be confirmed by independent sequential walks. The bound spectra are therefore 85% assigned (figure 5.16A/B).

The quality of the 3D bound spectra was worse than that of the unbound spectra. It is possible that at higher concentrations of peptide (i.e. when the protein is saturated) then some form of dimerisation/ multimerisation/ aggregation occurs, causing the reduction in spectra quality. Although it was possible to confirm through sequential walks the assignments of the bound spectra based on the unbound, no new assignments were obtained, i.e. the 12 unassigned resonances from the unbound psTolA3 (Ala2, Ser28, Leu29, Leu36, Val37, Ser38, Asn49, Gly50, Arg101, Lys116, Glu118, Asp119, Leu120). Those peaks that were perturbed in the presence of the peptide clustered into 2 groups; those that were perturbed at lower concentrations of peptide (potentially the initial site of interaction between peptide and protein as they seemed to cluster between residues 53 and 57) and those that were perturbed at higher concentrations of peptide. These additional residues may be perturbed at higher concentrations of peptide for a number of reasons. Firstly, it is possible that as a result of peptide binding protein there is a conformational change in the psTolA3. Secondly, it is possible that the peptide is either binding to a second, lower affinity site on the protein, or it is binding to the protein in a non-specific manner. Residues affected by binding, and the peptide concentration at which they are perturbed is displayed in figure 5.17.

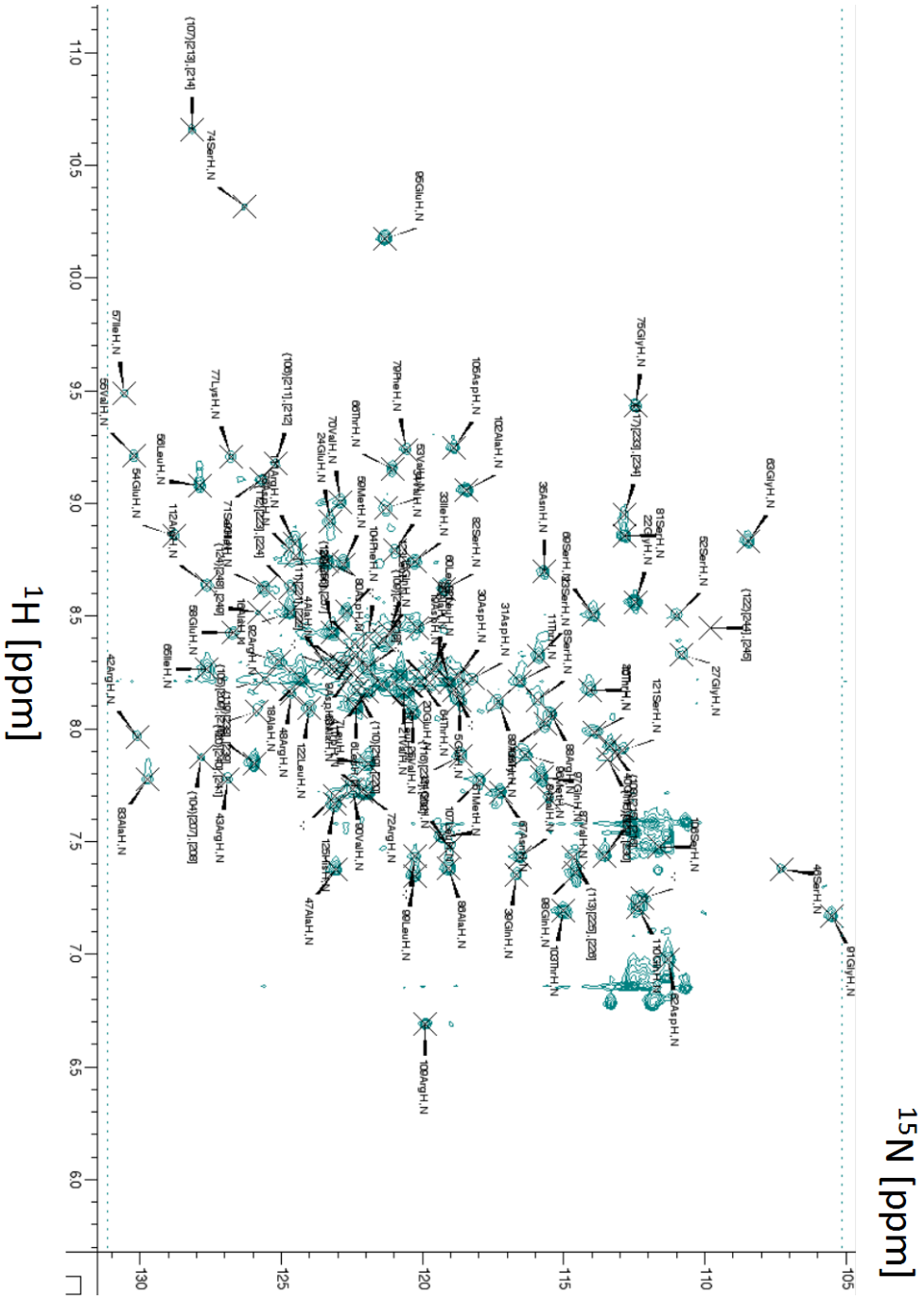


Figure 5.16A 2D HSQC spectra of ^{15}N *Pseudomonas aeruginosa* TolA domain 3 bound to *Pseudomonas aeruginosa* TolA binding peptide labelled with residue assignments. psTABp at 2 molar equivalent to psTolA3.

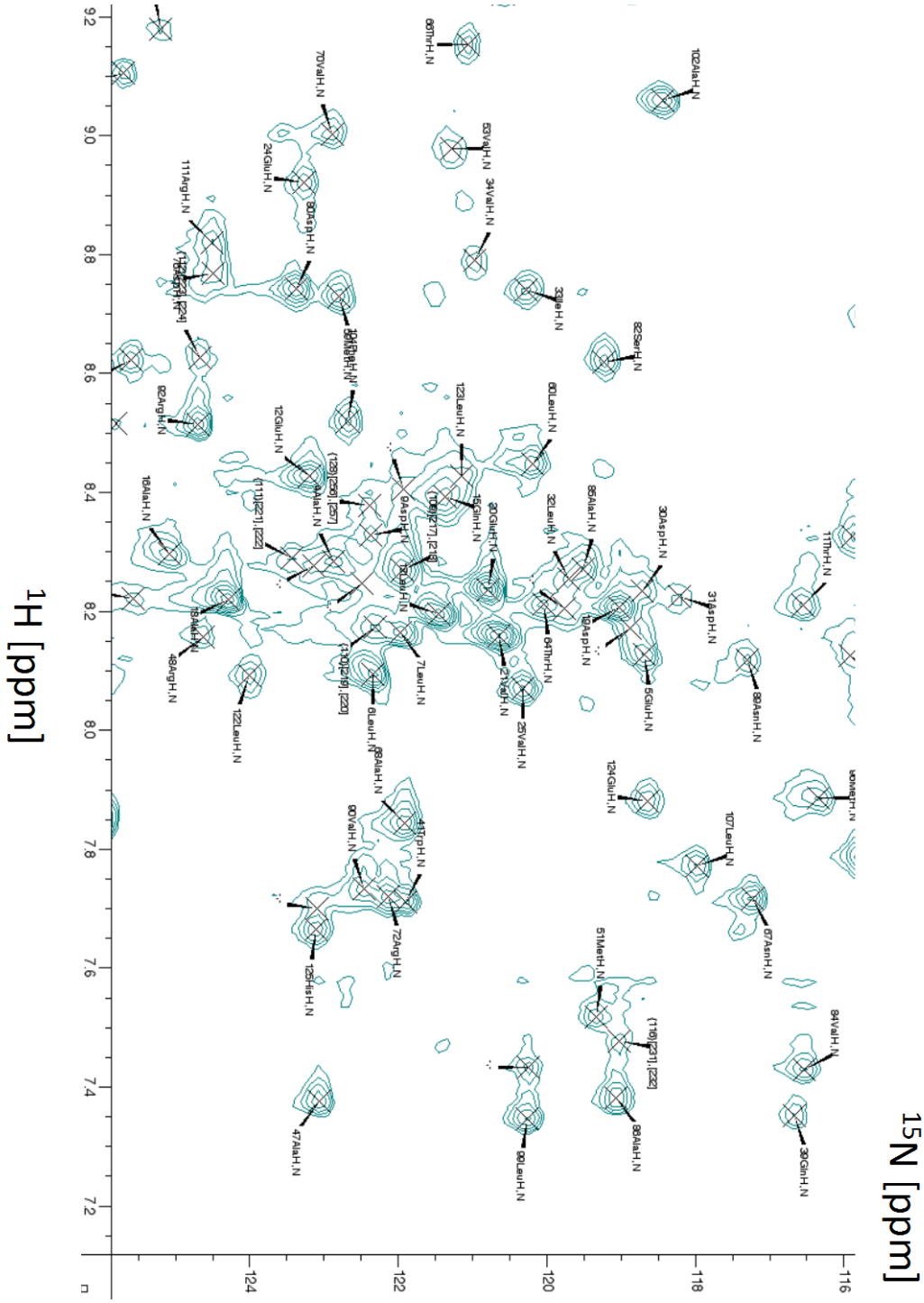


Figure 5.16B Zoomed in view of central (7.2 – 9.2 ppm) region of 2D HSQC spectra of ^{15}N *Pseudomonas aeruginosa* TolA domain 3 bound to *Pseudomonas aeruginosa* TolA binding peptide labelled with residue assignments. psTABp at 2 molar equivalent to psTolA3.

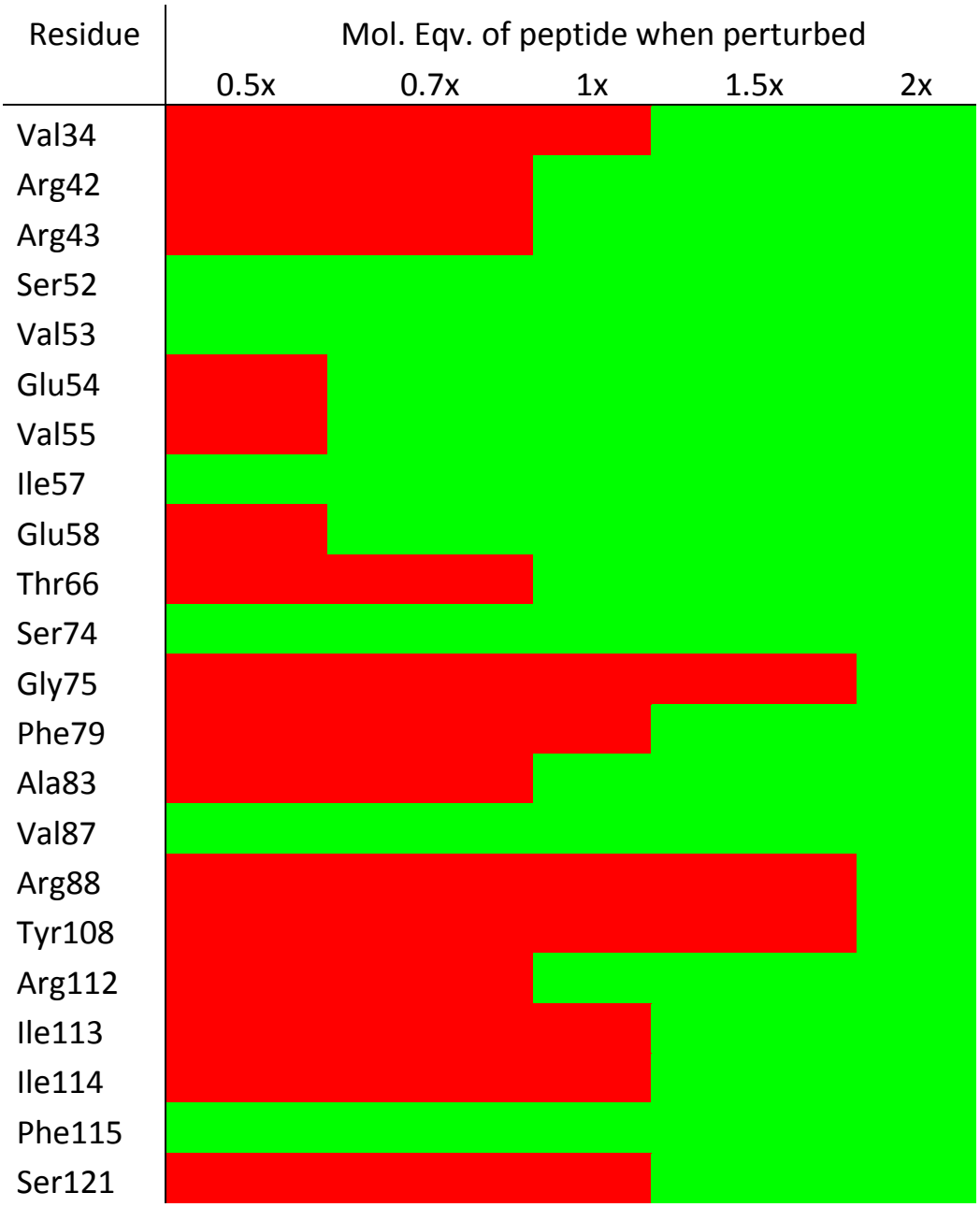


Figure 5.17 Residues perturbed by *Pseudomonas aeruginosa* TolA binding peptide. Red colour indicates residue not perturbed in presence of peptide, green indicates that it is.

5.2.9 Mapping residue assignments perturbed by *Pseudomonas aeruginosa* TolA binding peptide onto *Pseudomonas aeruginosa* TolA domain 3 crystal structure.

Having confirmed that the peptide perturbs a small population of peaks, and that those peaks seem to cluster between psTolA3 residues 53-57 (see table 5.1 for protein secondary structure), the residues affected by peptide binding were mapped onto psTolA3 crystal structure (PBD ID 1LR0, (Witty et al. 2002)).

Residue number	Assignment number	Secondary structure
228 – 265	2 – 39	1st alpha helix
278 – 284	52 – 58	1st beta strand
288 – 297	62 – 71	2nd beta strand
302 – 315	76 – 89	2nd alpha helix
319 – 323	93 – 97	3rd alpha helix
326 – 332	100 – 106	4th alpha helix
334 – 340	108 – 114	3rd beta strand
342 – 347	116 - 121	5th alpha helix

Table 5.1 *Pseudomonas aeruginosa* TolA domain 3 (residues 226 – 347) secondary structure features (Uniprot ID: P50600).

Upon the addition of 0.7 mol. eqv. psTABp, residues Val53 to Ile57 are perturbed. These residues map (figure 5.18) onto a β-sheet forming a “small pocket”. Additionally, residues in close proximity to this β-sheet are also perturbed. This would suggest that assigned residues Val53 to Ile57 are part of the binding interface between psTolA3 and psTABp. Adding up to 2 mol. eqv. psTABp, more residues were perturbed, and when mapped onto the crystal structure (figure 5.19), they are found to be not only highly dispersed

around the protein, but primarily in both in isolated positions and loop regions, suggesting that these perturbations are as a result of a conformational change in the psTolA3 upon binding of psTABp. It is also possible that these dispersed perturbations are as a result of multiple additional peptide bindings, potentially of a non-specific nature.

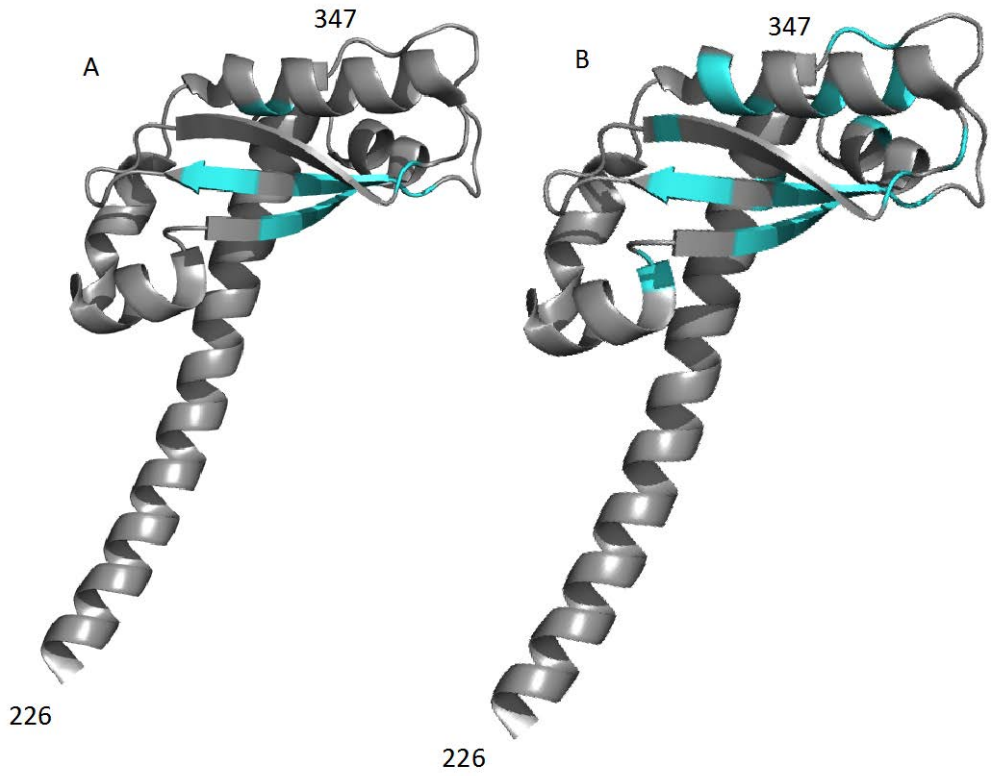


Figure 5.18 *Pseudomonas aeruginosa* TolA domain 3 residues perturbed by *Pseudomonas aeruginosa* TolA binding peptide mapped onto secondary structure of *Pseudomonas aeruginosa* TolA domain 3 crystal structure. (A) 0.7 molar equivalent (B) 2 molar equivalent. Residues perturbed by binding coloured cyan. Unperturbed residues grey. PDB ID: 1LR0.

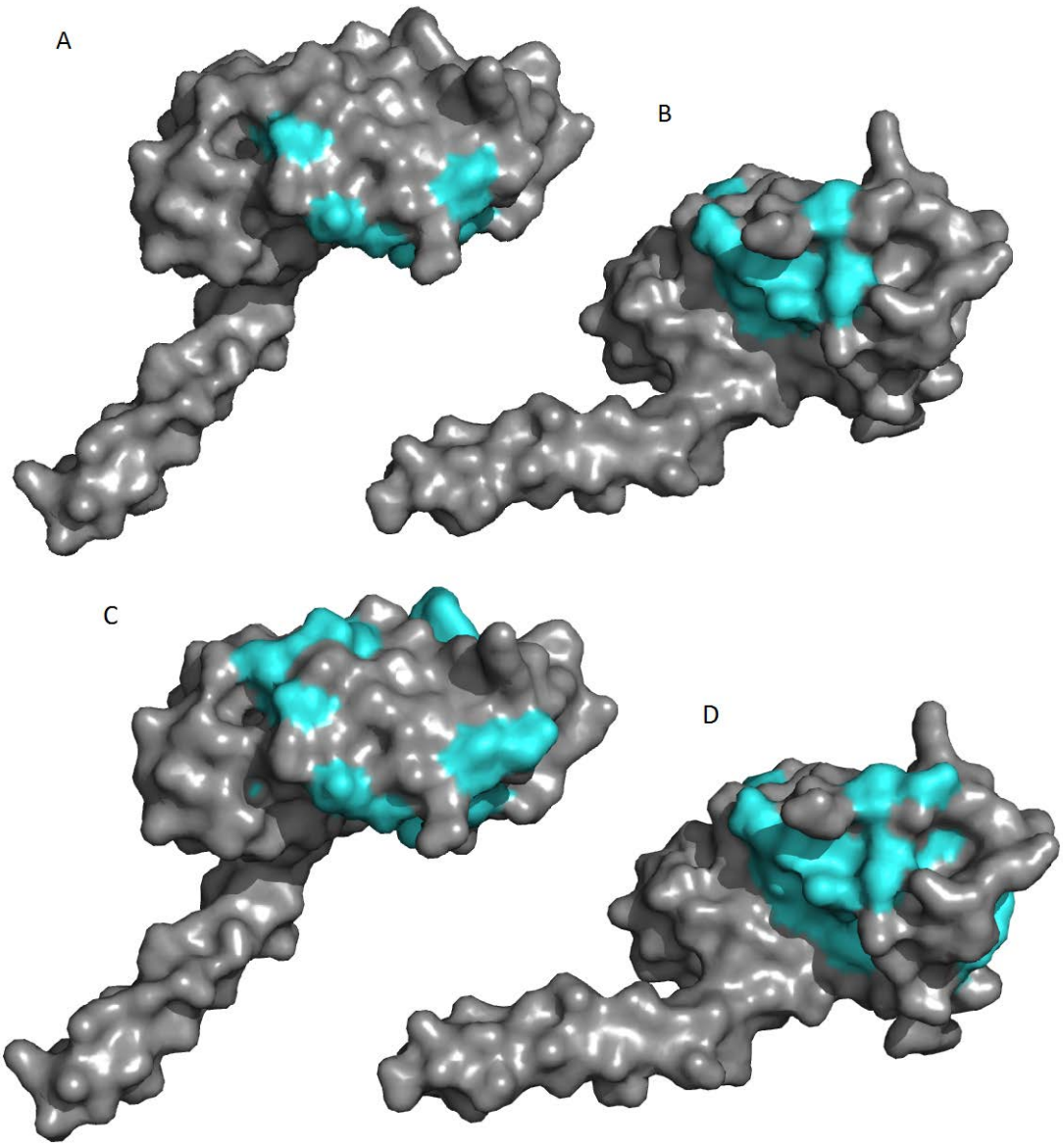


Figure 5.19 *Pseudomonas aeruginosa* TolA domain 3 residues perturbed by *Pseudomonas aeruginosa* TolA binding peptide mapped onto molecular surface of *Pseudomonas aeruginosa* TolA domain 3 crystal structure. (A) 0.7 molar equivalent (B) 2 molar equivalent. Residues perturbed by binding coloured cyan. Unperturbed residues grey. PDB ID: 1LR0.

As seen in figure 5.19A/B when mapped onto a molecular surface of crystal structure, at low concentration of peptide, a line of perturbed residues cluster

within a pocket, of 19 Å in length. Given that the peptide is 16 residues long (if in elongated linear conformation), it would suggest that the protein residues clustered around Val53 to Ile57 are the peptide binding site. As seen in figure 5.19C/D all further perturbations at higher concentrations of peptide are distributed throughout the protein, and thus are likely to occur due to a conformational change in the psTolA3 due to association of peptide.

5.3 Discussion

This NMR perturbation mapping data is the first indication of any structural information regarding the TolA-TolB interaction of Gram-negative bacteria. Although the *E.coli* TolA binding site on *E.coli* TolB had been defined as the intrinsically disordered N-terminus (Bonsor et al. 2009), the TolB binding site on TolA has not been reported, in either *E.coli, Pseudomonas aeruginosa* or any other Gram-negative bacteria. Until very recently no crystal structure had been reported for *E.coli* TolA. All structural information for the protein had been based on either an NMR solution structure of eTolA3 (Deprez et al. 2005), or the crystal structure of psTolA3 (Witty et al. 2002). The recently published *E.coli* TolA3 crystal structure will be discussed in the following chapter. Despite much research, no structural information for the TolA-TolB complex has been reported. In this work (as well as by previous members of this lab), attempts were made to crystallise *E.coli* TolA3 alone, *E.coli* TolA3-TolB complex, *E.coli* TolA3-TolB-colicin E9 T-domain complex, *E.coli* TolA3 fusion proteins that included eTolB disordered N-terminus fused via a flexible linker to TolA3), *E.coli* TolA3-TolA binding peptide complex as well as *Pseudomonas aeruginosa* TolA3-TolA binding peptide complex. All of these approaches have met with failure. This is likely due to the highly soluble nature of the *E.coli* TolA3, which even at 150mg/ml concentration could not be crystallised, coupled with the very weak nature of the TolA-TolB interaction, which is reported to be in the order of 10-40 μM for *E.coli,* and

estimated to be in the order of several hundred μM for *Pseudomonas aeruginosa* Tol proteins, *in vitro*.

This mapping appears to show that the psTABp (and thus by analogy the potentially disordered N-terminus of psTolB) interacts with the 1st beta strand of psTolA3 (and surrounding residues on the same face), as these are the residues that are first perturbed by the binding of the psTABp. Following this, at higher concentrations of peptide, further chemical shift perturbations are found in other regions of the protein, however, these are in isolated patches, suggesting that rather than caused by a binding event, these perturbations are caused by a small conformational change in the protein in response to binding.

When assigning backbone resonances to the HSQC spectra it became apparent that most of the N-terminal residues of psTolA3 were found in the middle section of the spectra (between 8.0 – 8.5 ppm), typically where peaks corresponding to disordered regions of proteins are found (Cavanagh 2007). This suggested that whereas in the crystal structure, the N-terminus of psTolA3 was of alpha helical structure, in the NMR sample, it was not. This is not without precedence, as the solution NMR structure published for *E.coli* TolA domain 3 also showed it's N-terminus to be in a disordered conformation (Deprez et al. 2005).

As the peak (which likely corresponds to the C-terminal histidine) was affected by binding, an additional psTolA3 construct was made. In this constuct the his-tag was moved to the opposite end of the protein (onto the long N-terminal strand, shown as alpha-helical in the crystal structure, but suggested to be disordered in these NMR spectra) and was also made to be thrombin cleavable. However, this construct (pEC16) was unstable, and proved to be problematic during purification. Additionally, a tagless variant of psTolA3 was made earlier in this work (pEC1), that was identical to the

construct used in this NMR assignment (pEC4), however, as shown in section 4.2.1.1, this protein was very difficult to purify and of low final yield. As a result an isotope labelled version of the protein was not attempted.

5.3.1 Summary

In the two previous chapters, attempts to define either the structure of the TolA-TolB complex from *E.coli* have not been successful. Although it was shown that *Pseudomonas aeruginosa* TolA and TolB interact with one another, I was unable to crystallise this complex, either with both whole proteins, or with the *Pseudomonas aeruginosa* TolA binding peptide as a surrogate for TolB. In chapter 5, I have used NMR to define the binding site of *Pseudomonas aeruginosa* TolB (represented by the TolA binding peptide) on TolA. An isotopically labelled *Pseudomonas aeruginosa* TolA was found to be stable in NMR experiments, and, when in the presence of the binding peptide, approximately 10-15% of it's residues are perturbed, with some residues being perturbed at lower concentrations of peptide to others. Following 3D spectra acquisition, over 90% of the unbound *Pseudomonas aeruginosa* TolA domain 3 was assigned (85% for the bound form) and once the residues that were perturbed by peptide were identified, they were mapped onto the published crystal structure of *Pseudomonas aeruginosa* TolA. This showed that the binding site on TolA is likely to be along the 1st beta strand (residues Val53 – Ile57) within a narrow pocket. This binding site is within the same location as bacteriophage g3p and colicin N proteins bind *E.coli* TolA, likely mimicking the native interaction between TolA and TolB.

6. General discussion

This work has set out to characterise and define the interaction between *E.coli* and *Pseudomonas aeruginosa* TolA and TolB. This interaction is vital to the stability of the outer membrane, and without either one of these components, the integrity of the membrane is compromised (Lloubes et al. 2001). This interaction spans the periplasm allowing communication between the inner and outer membranes. This interaction is also subverted by colicins to facilitate their entry into the bacterial cell.

Numerous functions for the Tol system have been suggested, with varying degrees of evidence. The Tol system may have a role in maintaining the structure of the bacterial cell membranes through an architectural function; forming a bridge network aiding envelope cohesion from the inner membrane, across the periplasm and onto the outer membrane (Henry et al. 2004; Cascales et al. 2007). Other potential functions have been inferred from a variety of evidence, including a role in cell division, as Tol-Pal complex has been localised to junction points between the outer and inner membrane and may act in drawing the peptidoglycan layer in contact with inner and outer membrane (Gerding et al. 2007), assembly and localisation of lipopolysaccharide (Gaspar et al. 2000) and localisation and assembly of outer membrane proteins, in particular porins (Cascales et al. 2007). It has also been reported that Tol-Pal may be involved in the transfer of structural components from the inner to the outer membrane, which again is likely to be tied to the *tol-pal* mutants lacking proper outer membrane stability (Lloubes et al. 2001; Goemaere et al. 2007). In particular, lipopolysaccharide (LPS), a core component of the outer membrane, has been reported to be at lower levels in the outer leaflet of the outer membrane in *tol-pal* mutant strains of *E.coli* (Gaspar et al. 2000; Lloubes et al. 2001; Vines et al. 2005; Gerding et al. 2007). Other work has reported that in *tolA* mutants, LPS biosynthesis is not reduced, rather a reduction was seen in the levels of LPS

in the outer leaflet of the outer membrane, likely as a result of the LPS post-synthesis processing and transportation (Gaspar et al. 2000).

In my own opinion it seems likely that the Tol proteins are involved in the catalysis/remodelling of carbohydrates or the transport of carbohydrates from the inner to outer membranes. Given that TolB is a beta-propeller protein, and that beta-propellor proteins are often enzymes or transporters it would suggest that TolB has one of these functions. In bacteria, several beta-propeller proteins, such as sialidase in *Salmonella typhimurium* and glucose dehydrogenase in *Acinetobacter calcoaceticus* are known to bind carbohydrates (Jawad & Paoli 2002). The organisation of the Tol system also lends itself to a role in transport of ligands from the inner to the outer membrane. Energised by TolQ and TolR, TolA when in complex with a ligand bound TolB (within it's beta-propeller domain) transports the TolB across the periplasm, delivering it to the outer membrane, where TolB then binds Pal, which not only causes the dissociation of TolA from TolB, but in Pal binding in the beta-propeller domain, causes dissociation of the ligand. In addition, rather than delivering a ligand to the outer membrane, as TolB associates with Pal at the peptidoglycan layer, it is possible that the transported ligand is component of peptidoglycan, such as NAG, NAM, or a branched derivate. Further evidence to support this theory comes from the similarities between the Tol and Ton systems. The TonB-ExbB-ExbD inner membrane complex is an energised transporter of Vitamin B12 and iron sideophores (Braun et al. 1993) and given that TolQRA is organised in a nearly identical manner it would seem likely that TolQRA provides energy to TolB for some form of transport to from the inner to the outer membrane. This however, as yet there is no evidence to support this theory and suggest that TolB binds any specific ligand, other than the protein interactions reported above.

It has been documented that colicin's appear to mimic native protein-protein interactions in order to subvert the system to allow uptake of the colicin into the cell (Bonsor et al. 2009). In the case of translocation of colicin E9 into the periplasm the translocation domain of the colicin threads through OmpF to contact *E.coli* TolB binds with the beta-propeller domain of eTolB. This interaction mimics the interaction of ePal with eTolB, wherein ePal also binds within the beta-propellor domain. However, rather than promote order in the N-terminus of the eTolB as ePal does, it promotes disorder, driving the interaction with eTolA (Bonsor et al. 2009). It has been commonly assumed that this is not a unique factor to colicin E9, rather all colicins may mimic a native interaction of Tol proteins. Thus, as colicin A, colicin N and Bacteriophage G3P interact with TolA independently of TolB, it has been assumed that one of these proteins may mimic the native TolA-TolB interaction. However, although colicin N and bacteriophage interact with TolA on the same face (Lubkowski et al. 1999) (Hecht et al. 2009), and in the same binding site, conversely, ColA interacts with TolA on the opposite side (Hecht et al. 2010). When comparing the perturbation mapping of the psTABp on psTolA against the crystal structures of eTolA3 interacting with either ColA or g3p, it would appear as though at high concentrations of peptide there are two clusters of chemical shift perturbations, one for each side corresponding to either ColA or g3p binding (figure 6.1, 6.2). However, when considering perturbations that occur at lower concentrations of peptide, given that they cluster along residues 52 to 58 (as well as surround residues), this would suggest that the psTABp binding is similar to that of g3p (and thus ColN) as they too interact on the same face as residues 52-58 present. This would again suggest that like colicin E9, both colicin N and g3p are mimicking a native interaction of the Tol system. It should however be noted that although ColN interacts with eTolA on the same surface as psTABp binds psTolA3, it has been reported that ColN causes a large conformational change in eTolA, potentially unfolding it (Hecht et al. 2009). Whether or not the unfolding of eTolA by ColN is unique (and a result of

ColN's translocation method) is unknown, but it seems likely as no other evidence of unfolding has been reported for eTolA when binding either eTolB or other colicins. Thus it is likely that although the native TolA-TolB interaction has been hijacked by ColN and g3p and they both bind in a similar pocket to that of TolB, their subsequent effects on TolA are different. In a recent development, a group investigating the *E.coli* TolA domain and its interactions with colicin A and TolB have confirmed these findings (Li et al. 2012). Work by Li et al has reported for the first time a crystal structure for *E.coli* TolA domain 3. This was achieved through the reductive methylation of surface lysine residues in order to make the protein more ameniable to crystallisation. In addition, Li et al. attempted to use a synthetic peptide corresponding to the disordered N-terminus of *E.coli* TolB in a similar approach to this work. Although they were unable to obtain either a crystallisable complex or an assigned NMR spectrum they have reported that the pattern of residue perturbations recorded for ^{15}N labelled eTolA3 in the presence of their eTABp was similar to the pattern of residue perturbation when the labelled eTolA3 was in the presence of colicin N. colicin N interacts with eTolA3 on the same face as bacteriophage g3p protein. Thus the work by Li et al. supports the work I have reported above (figure 6.1) in that I suggest that the binding site of TolB on TolA is identical to the binding site of colicin N and g3p (Li et al. 2012).

It should be noted that the NMR data collected for *Pseudomonas aeruginosa* TolA3-TABp interaction suggested that there may be a second site of interaction between the protein and peptide. If this is the case, then it possible that colicin A is utilising a second, lower affinity binding site, as evidenced by a small cluster of residues that were found to be perturbed at higher concentrations of peptide which sit directly underneath the colicin A binding site. If this is indeed the case, then as with colicin N and g3p potentially mimicking TolB binding in the primary higher affinity site, that colicin A may be mimicking the lower affinity site.

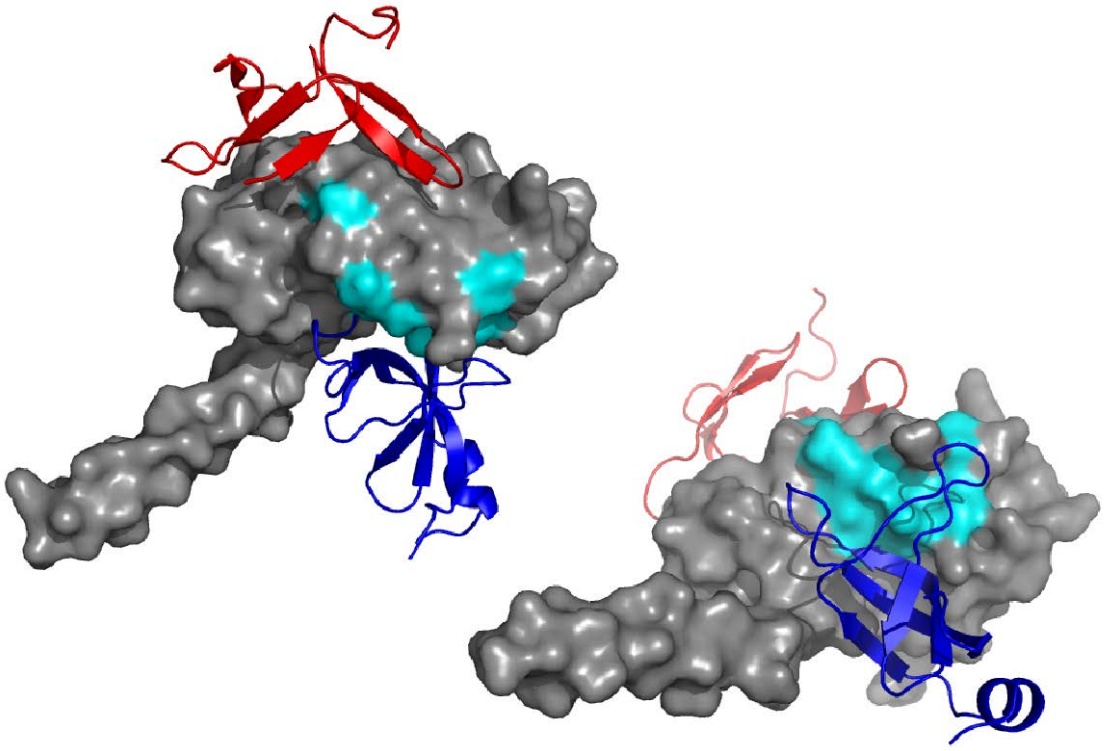


Figure 6.1 *Pseudomonas aeruginosa* TolA domain 3 residues perturbed by 0.7x molar equivalent *Pseudomonas aeruginosa* TolA binding peptide mapped onto molecular surface of *Pseudomonas aeruginosa* TolA domain 3 crystal structure docked with colicin A and Bacteriophage g3p (N1 domain) Residues perturbed by binding coloured cyan, Unperturbed residues grey, colicin A red and bacteriophage g3p (blue) PDB ID: 1LR0 (psTolA3), 1TOL (g3p), 3QDR (ColA).

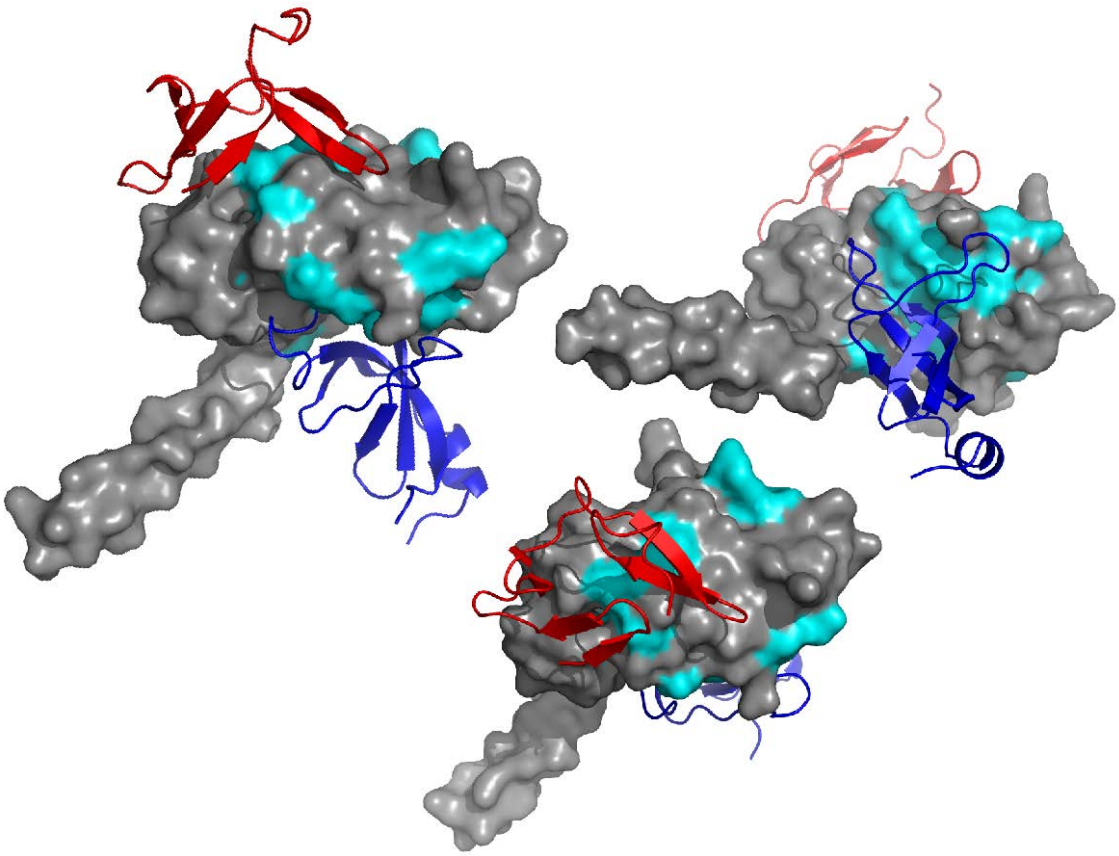


Figure 6.2 *Pseudomonas aeruginosa* TolA domain 3 residues perturbed by 2x molar equivalent *Pseudomonas aeruginosa* TolA binding peptide mapped onto molecular surface of *Pseudomonas aeruginosa* TolA domain 3 crystal structure docked with colicin A and Bacteriophage g3p (N1 domain) Residues perturbed by binding coloured cyan, Unperturbed residues grey, colicin A red and bacteriophage g3p (blue) PDB ID: 1LR0 (psTolA3), 1TOL (g3p), 3QDR (ColA).

In the preparation of this manuscript, a new structure has been reported for TolA, in *Vibrio cholera.* This TolA is a crystal structure of the complex between the CTXphi pIII N-terminal domain and the *Vibrio cholerae* TolA domain 3 (figure 6.3). When comparing this structure of the *Vibrio cholerae* TolA structure to that of *E.coli* or *Pseudomonas aeruginosa* it shares a remarkably similar fold. This lends further weight to the argument that the Tol systems, and in particular the TolA-TolB interaction is conserved throughout

Gram-negative bacteria. In addition, when comparing the CTXphi-*Vibrio cholerae* TolA structure (Ford et al. 2012) with that of the colicin A-*E.coli* TolA3 (Hecht et al. 2010), it would appear as though they share a similar site of interaction, on the opposite side to the colicin N-g3p-putative TolB site. Whether or not these two proteins are mimicking a native interaction on TolA is unknown.

Figure 6.3 Crystal structure of a complex between the CTXphi pIII N-terminal domain and the *Vibrio cholerae* TolA C-terminal domain. The filamentous bacteriophage CTX (pink, residues 3-105) interacts with TolA domain 3 from *Vibrio cholera* (green, residues 254 – 355). PDB ID: 4G7X (Ford et al. 2012).

For the future direction of work in this area, there are still a number of avenues left to explore. Althought the work of Li et al (Li et al. 2012) has attempted to identify the location of the TolB binding site on TolA through a similar synthetic peptide route as this work, they have been unable to determine the structure of this complex. As the *Pseudomonas aeruginosa* TolA was very stable for long periods of time in the NMR magnet, a future

direction could be the solvation of the psTolA3-psTABp complex through a solution NMR structure. Although this work reports that the N-terminus of *Pseudomonas aeruginosa* TolB is a site of interaction with TolA, it remains unclear as to whether or not it is in dynamic equilibrium like the *E.coli* TolB N-terminus is presumed to be in. This could be achieved through NMR and labelled *Pseudomonas aeruginosa* TolB. However, the issues that *Pseudomonas aeruginosa* TolB has in regard to solubilty and poor gene expression (despite codon optimisation for *E.coli* cells) is something that would have to addressed first. It would also be interesting to investigate if *Pseudomonas aeruginosa* Pal (also called OprL) binds TolB in a similar manner to the *E.coli* system, and potentially regulates the allosteric transition of the N-terminus from ordered to disordered state. During this work I attempted to clone and purify *Pseudomonas aeruginosa* Pal, but to without success. Like many of the *Pseudomonas aeruginosa* genes do not express in *E.coli* cells.

The ultimate unanswered question in the Tol system is to finally determine its function. Many functions have been suggested, be it a function in tethering the outer and inner membrane together, organising outer and inner membranes during cell division, or acting as a ligand transporter. Recently it was suggested that *E.coli* TolB may bind fragments of the peptidoglycan layer (Greg Papadakos, personal communication), which, if true, would suggest a role for TolB as a binding protein of peptidoglycan, assisting in either the transport or turnover. If proven, then this could yet be the breakthrough in the area of Tol biology that we have been waiting for.

7. Appendix

7.1 List of plasmids

Plasmid	Encodes	Cloning sites	Backbone
pEC1	*Pseudomonas aeruginosa* TolA (226-347)	BamHI/NdeI	pET11c
pEC2	Xanthomonas campesteris TolA (224 - 345)	BamHI/NdeI	pET11c
pEC3	Xanthomonas campesteris TolA (224 - 345)	BamHI/NdeI	pET11c
pEC4	Pseudomonas aeruginosa TolA (226-347), C-terminal 6xHis tag	BamHI/NdeI	pET21d
pEC5	Pseudomonas aeruginosa TolB (22-432), C-terminal 6xHis tag	BamHI/NdeI	pET21d
pEC6	Pseudomonas aeruginosa TolA (226-347), thrombin cleavable C-terminal 6xHis tag	BamHI/NdeI	pET21d
pEC7	Colicin E9-TolB fusion protein EC7, C-terminal 6xHis tag on Im9	KpnI/XhoI	pCS4
pEC8	Colicin E9-TolB fusion protein EC8, C-terminal 6xHis tag on Im9	KpnI/XhoI	pCS4
pEC11	Pseudomonas aeruginosa TolB (22-432), C-terminal 6xHis tag	BamHI/NdeI	pET21d
pEC12	Colicin E9-TolB fusion protein EC12, C-terminal 6xHis tag on Im9	KpnI/XhoI	pCS4
pEC13	Pseudomonas aeruginosa TolB (22-432), C-terminal 6xHis tag	BamHI/NdeI	pET21d
pEC14	Pseudomonas aeruginosa TolB (22-432), C-terminal 6xHis tag. Codon optimised for *E.coli* expression	BamHI/XhoI	pET21d

pEC15	Pseudomonas aeruginosa Pal (60-168), C-terminal 6xHis tag	NcoI/XhoI	pET21d
pEC16	Pseudomonas aeruginosa TolA (226-347), thrombin cleavable N-terminal 6xHis tag	BamHI/NdeI	pET15b
pAK108	*E.coli* TolA (293-421)	NcoI/XhoI	pET21d
pDAB18	*E.coli* TolB (22-430)	NcoI/XhoI	pET21d
pCS4	Colicin E9 (1-582) and Im9 (1-86, with C-terminal 6xHis tag)	NdeI/XhoI	pET21a

7.2 List of primers

EC12FWD: AAGTGGTGGTGCTTCTGATGGTTCAGGATGGAGTTCGGAAAATAACC

EC12REV: GGTTATTTTCCGAACTCCATCCTGAACCATCAGAAGCACCACCACTT

EC78KPNIFWD: TTCCGGTGGTGGCTCGGGTACCGGCGGTAATTTGTCAG

EC78KPNIREV: CTGACAAATTACCGCCGGTACCCGAGCCACCACCGGAA

PTOLB_EC5_TEV_FWD1: GAATTCGTTCAGGTAAGGGGACCAGGAAGG

PTOLB_EC5_TEV_REV1: GTGGTGGTGGTGGTGGTGCTCGAGTCCCTGGAAGTAGAGATTCTC

PTOLA3_EC6_FWD: GGATCCGCATTGGCCGAGTTG

PTOLA3_EC6_REV: CTCGAGTCACAGACTCAAATC

PTOLB_EC5_FWD: GGATCCATGGCCGACCCGCTGGTGA

PTOLB_EC5_REV: GAATTCGTTCAGGTAAGGGGA

PTOLB_EC5_TEV_FWD: TCCCCTTACCTGAACGAATTCGAGAATCTCTACTTCCAGGGA

PTOLB_EC5_TEV_REV: GAGAATCTCTACTTCCAGGGACTCGAGCACCACCACCACCACCAC

XANTOLA3EC3 :ACTGCGTCAGGTGGCGGCGCATATGGCGCAGCAGGCG

XANTOLA3EC4: ACACACCTCCGGATCCTTACTGATCCTGCGCGGT

XANTOLA3EC1: ACTGCGTCAGGTGGCGGCGCATATGGCGCAGCAGGCGAGCAG

XANTOLA3EC2: ACACACCTCCGGATCCCTGATCCTGCGCGGT

PSETOLA3EC1: GGCTGCCGAGGACAAGCATATGCGGGCATTGGCCGAGTTGC

PSETOLA3EC2: TTTTTTGGATCCTTACAGACTCAAATCCTCCGGTTT

7.3 List of protein sequences

pAK108 (*E.coli* TolA, residues 293-421)

ADDIFGELSSGKNAPKTGGGAKGNNASPAGSGNTKNNGASGADINNYAG
QIKSAIESKFYDASSYAGKTCTLRIKLAPDGMLLDIKPEGGDPALCQAALAA
AKLAKIPKPPSQAVYEVFKNAPLDFKP

pDAB18 (*E.coli* TolB, residues 22-430)

MEVRIVIDSGVDSGRPIGVVPFQWAGPGAAPEDIGGIVAADLRNSGKFNPL
DRARLPQQPGSAQEVQPAAWSALGIDAVVVGQVTPNPDGSYNVAYQLVD
TGGAPGTVLAQNSYKVNKQWLRYAGHTASDEVFEKLTGIKGAFRTRIAYVV
QTNGGQFPYELRVSDYDGYNQFVVHRSPQPLMSPAWSPDGSKLAYVTFE
SGRSALVIQTLANGAVRQVASFPRHNGAPAFSPDGSKLAFALSKTGSLNLY
VMDLASGQIRQVTDGRSNNTEPTWFPDSQNLAFTSDQAGRPQVYKVNIN
GGAPQRITWEGSQNQDADVSSDGKFMVMVSSNGGQQHIAKQDLATGGV
QVLSSTFLDETPSLAPNGTMVIYSSSQGMGSVLNLVSTDGRFKARLPATD
GQVKFPAWSPYL

pEC1 (*Pseudomonas aeruginosa* TolA, residues 226-347)

MALAELLSDTTERQQALADEVGSEVTGSLDDLIVNLVSQQWRRPPSARNG
MSVEVLIEMLPDGTITNASVSRSSGDKPFDSSAVAAVRNVGRIPEMQQLPR
ATFDSLYRQRRIIFKPEDLSL

pEC4 (*Pseudomonas aeruginosa* TolA, residues 226-347, C-terminal 6x His tag)

MALAELLSDTTERQQALADEVGSEVTGSLDDLIVNLVSQQWRRPPSARNG
MSVEVLIEMLPDGTITNASVSRSSGDKPFDSSAVAAVRNVGRIPEMQQLPR
ATFDSLYRQRRIIFKPEDLSLLEHHHHHH

pEC14 *(Pseudomonas aeruginosa* TolB, residues 22 – 432, C-terminal 6x His tag)

MADPLVISSGNDRAIPIAVVPFGFQGGNVLPEDMSNIIGNDLRNSGYFEPLP
RQNMISQPAQASEVIFRDWKAVGVNYVMVGNIVPAGGRLQVQYALFDV
GTEQQVLTGSVTGSTDQLRDMSHYIADQSFEKLTGIKGAFSTKMLYVTAER
FSVDNTRYTLQRSDYDGARPVTLLQSREPIVSPRFSPDGRRIAYVSFEQKR
PRIFIQYVDTGRREQITNFEGLNGAPAFSPDGNRLAFVLSRDGNPEIYVMDL
GSRALRRLTNNLAIDTEPFWGKDGSTLYFTSDRGGKPQIYKMNVNSGAVD
RVTFIGNYNANPKLSADEKTLVMVHRQQGYTNFQIAAQDLQRGNLRVLSN
TTLDDSPTVAPNGTMLIYATRQQDRGVLMLVSINGRVRIPLPTAQGDVREP
SWSPYLNLEHHHHHH

7.5 Surface Plasmon Resonance immobilisation

To couple the protein of interest to the sensor chip surface thiol coupling was used. Coupling occurs through reactive thiol-disulfide exchange with the introduced cysteine at the N-terminal end of *E.coli* TolA domain 2 (the ligand) and the prepared surface of the C1 sensor chip. A C1 chip has a simple flat carboxylated surface with no dextran matrix. This chip was chosen due to its lack of dextran coated surface, as described above. For coupling to occur, the C1 surface must have reactive disulfide groups introduced onto it and then coupling is mediated by maleimide reagents, resulting in a thioether bond between the ligand and the chip surface (figure 7.1).

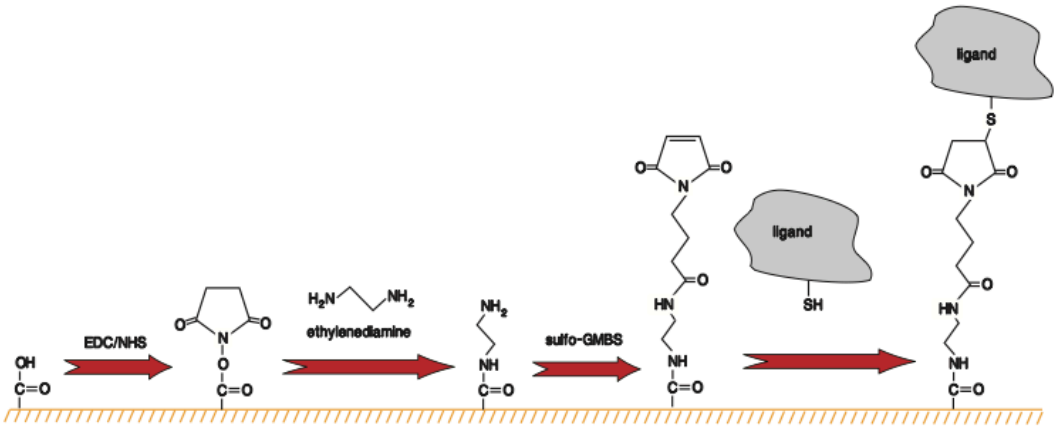


Figure 7.1 ligand immobilisation onto C1 chip. The carboxylated C1 chip surface is activated by 1-ethyl-3-[dimethylaminopropyl]carbodiimide (EDC) and modified by N-hydroxysuccinimide (NHS) to form a NHS-ester, which is further modified with ethylenediamine through amine-coupling, presenting a free amine group. This free amine reacts with N-[*Y*-maleimidobutyryloxy]-sulfosuccinimide ester (Sulfo-GMBS) which forms an irreversible thiol couple with the cysteine of the ligand. (Figure adapted from GE Healthcare Biacore Surface Sensor Handbook, 2008).

7.6 ITC single site binding model

During an ITC experiment, a protein sample (P) is placed into the ITC cell, which has a working volume (V_0). A ligand (L) is injected into it this cell, which remains at constant calorimetrically detected volume (V_0) due to the displacement of volume into the communication tube (ΔV), through which the ligand has been titrated.

As the total amount of P does not change during the experiment,

$$P_t{}^0 V_0 = P_t V_0 + \frac{(P_t{}^0 + P_t)}{2} \Delta V$$

To account for the volumetric change after each injection, Origin uses two equations; P_t and L_t:

The total concentration of P in V_0 after each injection can thus be calculated,

$$P_t = P_t^0 \left(\frac{1 - \frac{\Delta V}{2V_0}}{1 + \frac{\Delta V}{2V_0}} \right)$$

In addition, and assuming that all L stays in V_0 , the concentration of L after each infection can be calculated thus,

$$L_t{}^0 V_0 = L_t V_0 + \left(\frac{L_t}{2} \right) \Delta V \quad \text{as} \quad L_t = L_t^0 \left(1 - \frac{\Delta V}{2V_0} \right)$$

The single site binding model is expressed in the following form,

L+P = LP

With the binding constant being represented by,

$$K_A = \frac{[LP]}{[L][P]}$$

The binding constant can also be expressed in terms of Θ, the fraction of sites occupied by L,

$$K_A = \frac{\Theta}{(1 - \Theta)[L]}$$

This can be rearranged to give the concentration of free ligand, [L],

$$[L] = \frac{\Theta}{(1 - \Theta)K_A}$$

The total amount of ligand (L_t) can be calculated thus,

$$L_t = [L] + n\,\Theta\,P_t$$

Which, when combined with equation for [L] and rearranged into a quadratic, gives the following equation,

$$0 = \Theta^2 - \Theta\left(1 + \frac{L_t}{nP_t} + \frac{1}{nK_AP_t}\right) + \frac{L_t}{nP_t}$$

Q, the total heat content of the solution in the cell (V_0) when only partially saturated Θ is expressed as (where ΔH is the molar heat of binding),

$$Q = n\,\Theta\,P_t\,\Delta H\,V_0$$

If Θ is substituted with the term from the above equation (deriving by solving the above quadratic equation), then Q can be expressed as,

$$Q = \frac{n P_t \Delta H V_0}{2}\left[1 + \frac{L_t}{nP_t} + \frac{1}{nK_A P_t} - \sqrt{\left(1 + \frac{L_t}{nP_t} + \frac{1}{nK_A P_t}\right)^2 - \frac{4L_t}{nP_t}}\right]$$

The heat after *i*th injection, Q(i), can be calculated by using hypothetical values for *n*, ΔH and K_A. The heat released during the *i*th injection can be calculated as the difference between the Q(i) and the Q(i-1) (previous injection) thus,

$$\Delta Q(i) = Q(i) + \frac{\delta Vi}{V_0}\left[\frac{Q(i) + Q(i-1)}{2}\right] - Q(i-1)$$

This equation accounts for the fact that V_0 after the i-1 injection will not be the same as the V_0 following the *i*th injection. It will however, contribute approximately 50% heat as the *i*th volume (V_i), which remains V_0. Initially values for *n*, ΔH and K_A are assumed and from this the Q(i) and ΔQ(i) for each injection are calculated. The experimental data is then fitted based on these theoretical calculated heat differences. The experimental data is then fitted again using fit optimisation based on the Levenberg-Marquardt algorithm (also known as damped least squares method) until no further improvement in fit is seen (Microcal 2004).

7.7 Analytical Ultracentrifugation correction (SEDFIT)

The SEDFIT program calculates the estimated uncorrected sedimentation coefficient (S*) distribution of the sample by fitting the predicted sedimentation behaviour of trial sedimentation coefficient distributions to the

experimental data collected for the sample of interest (Schuck 2000). The uncorrected sedimentation coefficient are then corrected for buffer density and viscosity to zero protein concentration ($S^0_{20,w}$). The density and viscosity correction is expressed thus,

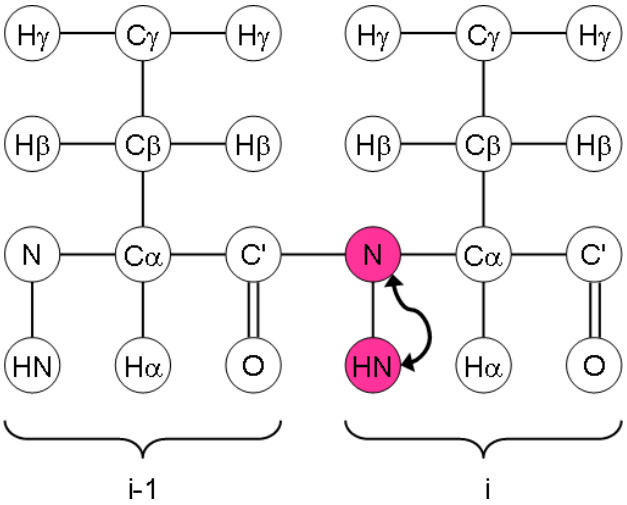
$$S_{20,w} = S^*(1 - v\,\rho_w)\,\eta_b / (1 - v\,\rho_b)\,\eta_w$$

Where S* is the uncorrected sedimentation coefficient, v is the partial specific volume of the protein, ρ is the density, η is the viscosity, and w and b refer to water and buffer respectively.

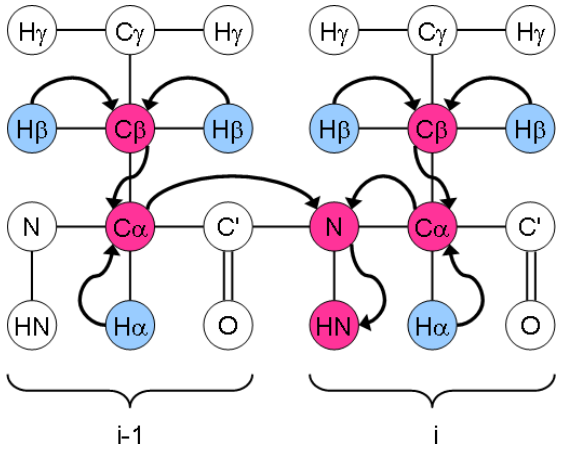
7.8 NMR experiments types

(All figures reproduced from http://www.protein-nmr.org.uk/solution-nmr/spectrum-descriptions)

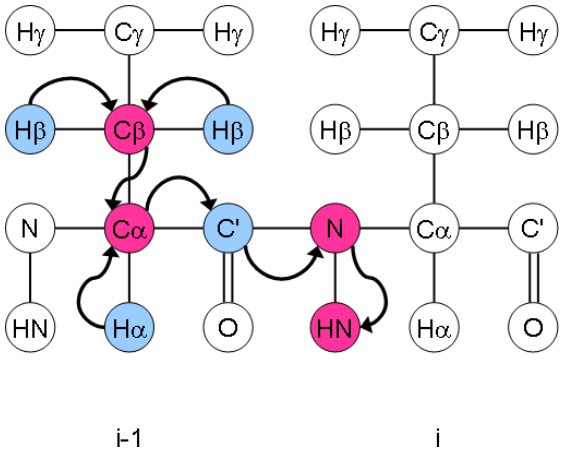
^{1}H-^{15}N HSQC



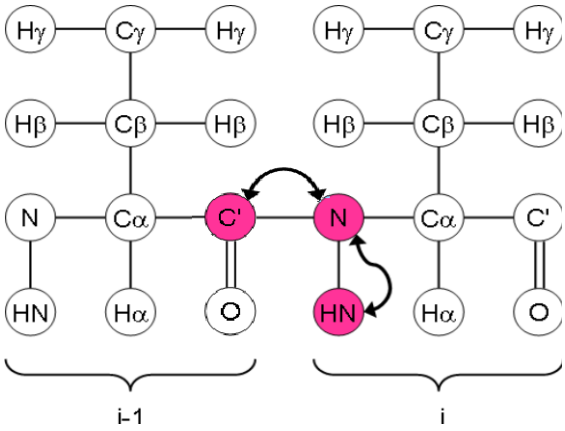
CBCANH



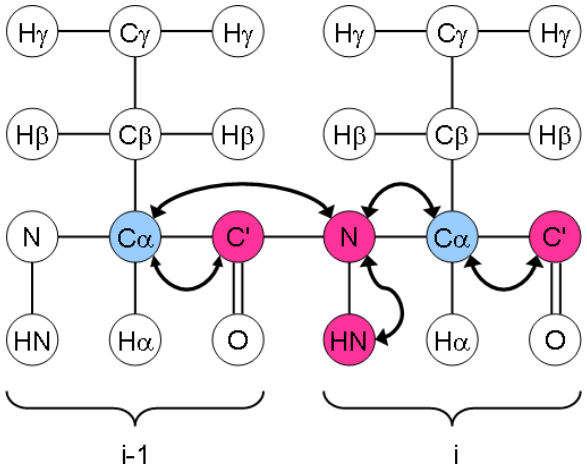
CBCACONH



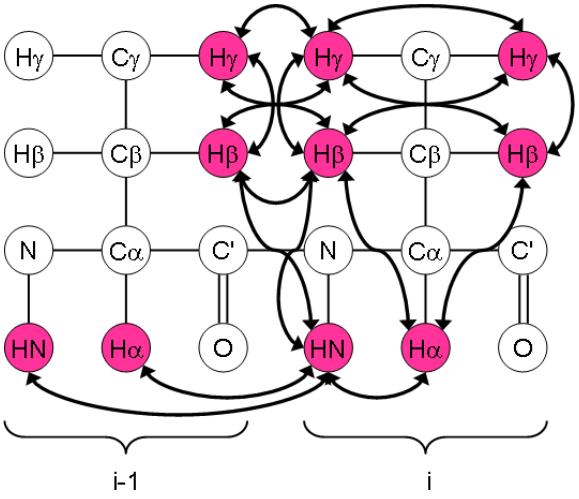
HNCO



HNCACO



H-H NOESY



7.9 Formaldehyde Crosslinking

Formaldehyde crosslinking is a method of forming a covalent (not non-thermostable) bond between polypeptides. A brief description of crosslinking chemistry is displayed below (Figure 7.2), and structure of DSP is displayed in figure 7.3.

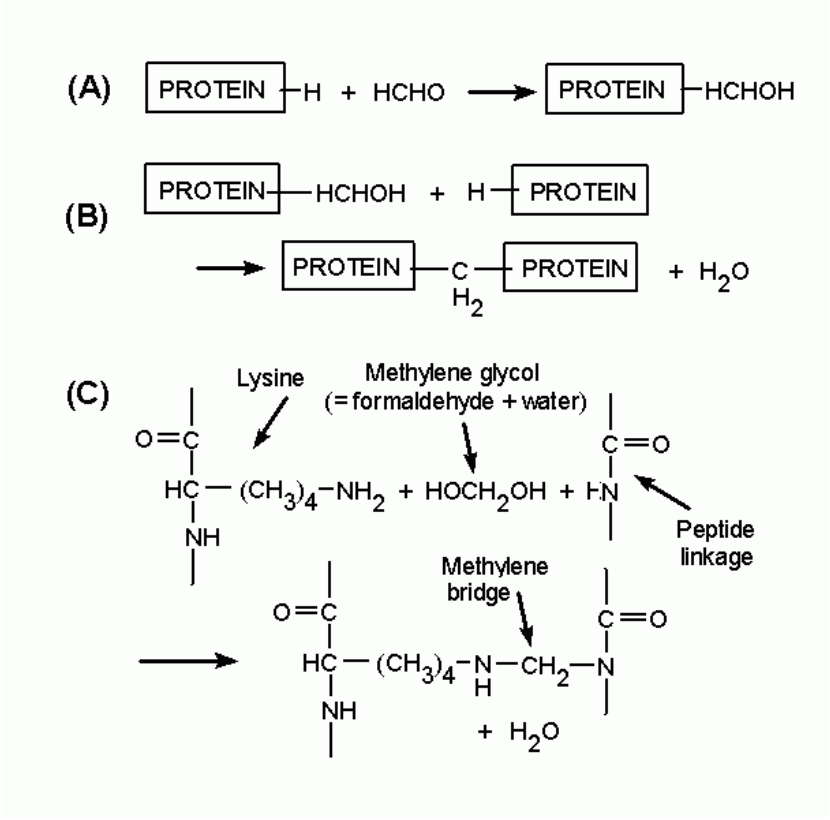


Figure 7.2 Reaction scheme of protein crosslinking with formaldehyde. Protein is modified by the addition of formaldehyde (A), which, which in the presence of another protein (B) forms a methylene bridge. Specifically, primary amines (such as Lysine) react with formaldehyde and and nearby peptide to form methylene bridge and thus form crosslink (C). Figure adapted from Kiernan 2000.

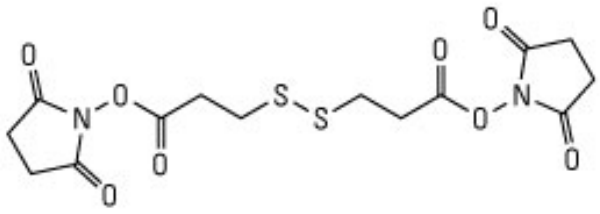


Figure 7.3 Dithiobis(succinimidylpropionate) [DSP]

List of abbreviations

ampR ampicillin resistance
ATP adenosine tri-phosphate
AUC Analytical ultracentrifugation
bp base pair
BSA bovine serum albumine
c.v. colume volume
calc calculated
CCCP carbonylcyanide m-chlorophenylhydrazone
CD circular dichroism (spectroscopy)
cps capsular polysaccharide
dNTPs deoxynucleotide triphosphates
DSP dithio-bis(succinimidyl propionate)
ePal *E.coli* Pal
eTABp *E.coli* TolA binding peptide
eTolA(3) *E.coli* TolA (domain 3)
eTolB *E.coli* TolB
ESI electrospray ionization
GlcNAc N-acetyl-β-D-glucosamine
HMM Hidden Markov Model
Hsp heat shock protein
IM inner membrane
IPTG isopropyl-β-D-thiogalactopyranoside
ITC isothermal titration calorimetry
kb kilo base (pair)
kDa kilo Dalton
KDO 3-deoxy-manno-2-octulosonic acid
LB lysogeny broth
LPS lipopolysaccharide
MALDI matrix assisted laser desorption/ionization
min minutes
MS mass spectrometry
MurNAc N-acetylmuramic acid
M_W molecular weight
NMR nuclear magnetic resonance (spectroscopy)
OD_{600} optical density at 600 nm
OM outer membrane
OMP outer membrane protein

ORF/orf open reading frame
P promoter
PAGE polyacrylamide gel electrophoresis
Pal peptidoglycan associated lipoprotein
PCR polymerase chain reaction
PEG polyethyleneglycol
PG peptidoglycan
pmf proton motive force
PMSF phenylmethanesulfonyl fluoride
psPal *Pseudomonas aeruginosa* Pal
psTABp *Pseudomonas aeruginosa* TolA binding peptide
psTolA(3) *Pseudomonas aeruginosa* TolA (domain 3)
psTolB *Pseudomonas aeruginosa* TolB
res. residue
rmsd root mean square deviation
SDS sodium dodecyl sulfate
SEC-MALLS size exclusion chromatography-multi angle laser light scattering
SPR surface plasmon resonance
SRP signal recognition particle
TBS Tris-buffered saline
TMD trans-membrane domain
TPR tetratricopeptide repeat

References

Akama, H., T. Matsuura, et al. (2004). "Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in Pseudomonas aeruginosa." *The Journal of Biological Chemistry* **279**(25): 25939-25942.

Andrea, L. D., & L. Regan (2003). "TPR proteins: the versatile helix". *Trends in Biochemical Sciences* **28**(12): 655-656

Andrews, S. C., A. K. Robinson, et al. (2003). "Bacterial iron homeostasis." *FEMS Microbiology Reviews* **27**(2-3): 215-237.

Andya, J., Y. Shen, et al. (2005). "Recombinant humanized monoclonal antibody size distribution analysis by size exclusion chromatography and sedimentation velocity analyzed by sedfit." *Abstracts of Papers of the American Chemical Society* **229**: U199-U199.

Appel, R. D., A. Bairoch, et al. (1994). "A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server." *Trends in Biochemical Sciences* **19**(6): 258-260.

Avramis, V. I., E. V. Avramis, et al. (2009). "Immunogenicity of native or pegylated E. coli and Erwinia asparaginases assessed by ELISA and surface plasmon resonance (SPR-biacore) assays of IgG antibodies (Ab) in sera from patients with acute lymphoblastic leukemia (ALL)." *Anticancer Research* **29**(1): 299-302.

Bach, D. R., D. Talmi, et al. (2011). "Automatic relevance detection in the absence of a functional amygdala." *Neuropsychologia* **49**(5): 1302-1305.

Bain, A. D. (2003). "Chemical exchange in NMR." *Progress in Nuclear Magnetic Resonance Spectroscopy* **43**(3,Äì4): 63-103.

Bell, C., & M, Lewis. (2000) “A closer view of the conformation of the Lac repressor bound to operator.” *Nature Structural Biology* **7**(3): 209-214.

Bernadac, A., M. Gavioli, et al. (1998). "Escherichia coli tol-pal mutants form outer membrane vesicles." *Journal of Bacteriology* **180**(18): 4872-4878.

Bernstein, A., B. Rolfe, et al. (1972). "Pleiotropic properties and genetic organization of the tolA,B locus of Escherichia coli K-12." *Journal of Bacteriology* **112**(1): 74-83.

Bohm, G., R. Muhr, et al. (1992). "Quantitative analysis of protein far UV circular dichroism spectra by neural networks." *Protein Engineering* **5**(3): 191-195.

Bonsor, D. A. (2009). "Colicin E9 translocation and the role of the Tol-Pal system." *PhD thesis (York)*.

Bonsor, D. A., I. Grishkovskaya, et al. (2007). "Molecular mimicry enables competitive recruitment by a natively disordered protein." *Journal of American Chemical Society* **129**(15): 4800-4807.

Bonsor, D. A., O. Hecht, et al. (2009). "Allosteric beta-propeller signalling in TolB and its manipulation by translocating colicins." *Embo Journal* **28**(18): 2846-2857.

Bonsor, D. A., N. A. Meenan, et al. (2008). "Colicins exploit native disorder to gain cell entry: a hitchhiker's guide to translocation." *Biochemical Society Transactions* **36**: 1409-1413.

Bos, M. P., V. Robert, et al. (2007). "Biogenesis of the gram-negative bacterial outer membrane." *Annual Reviews Microbiology* **61**: 191-214.

Bouveret, E., L. Journet, et al. (2002). "Analysis of the Escherichia coli Tol-Pal and TonB systems by periplasmic production of Tol, TonB, colicin, or phage capsid soluble domains." *Biochimie* **84**(5-6): 413-421.

Braun, V. and C. Herrmann (1993). "Evolutionary relationship of uptake systems for biopolymers in Escherichia coli: cross-complementation between the TonB-ExbB-ExbD and the TolA-TolQ-TolR proteins." *Molecular Microbiology* **8**(2): 261-268.

Braun, V., S. I. Patzer, et al. (2002). "Ton-dependent colicins and microcins: modular design and evolution." *Biochimie* **84**(5-6): 365-380.

Breukels, V., A. Konijnenberg, et al. (2011). "Overview on the use of NMR to examine protein structure." *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **Chapter 17**: Unit17 15.

Bui, N. K., J. Gray, et al. (2009). "The peptidoglycan sacculus of Myxococcus xanthus has unusual structural features and is degraded during glycerol-induced myxospore development." *Journal of Bacteriology* **191**(2): 494-505.

Bulieris, P. V., S. Behrens, et al. (2003). "Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide." *Journal Biological Chemistry* **278**(11): 9092-9099.

Capaldi, R. A. and R. Aggeler (2002). "Mechanism of the F(1)F(0)-type ATP synthase, a biological rotary motor." *Trends in Biochemical Sciences* **27**(3): 154-160.

Cascales, E., A. Bernadac, et al. (2002). "Pal lipoprotein of Escherichia coli plays a

major role in outer membrane integrity." *Journal of Bacteriology* **184**(3): 754-759.

Cascales, E., S. K. Buchanan, et al. (2007). "Colicin biology." *Microbiology and Molecular Biology Reviews* **71**(1): 158-229.

Cascales, E., M. Gavioli, et al. (2000). "Proton motive force drives the interaction of the inner membrane TolA and outer membrane pal proteins in Escherichia coli." *Molecular Microbiology* **38**(4): 904-915.

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

Cascales, E., R. Lloubes, et al. (2001). "The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB." *Molecular Microbiology* **42**(3): 795-807.

Cavalli, A., X. Salvatella, et al. (2007). "Protein structure determination from NMR chemical shifts." *Proceedings of the National Academy of Sciences of the United States of America* **104**(23): 9615-9620.

Cavanagh, J. (2007). Protein NMR Spectroscopy: Principles And Practice, Academic Press.

Chan, Y. C., J. L. Wu, et al. (2011). "Cloning, purification, and functional characterization of Carocin S2, a ribonuclease bacteriocin produced by Pectobacterium carotovorum." *Bmc Microbiology* **11**.

Chauleau, M., L. Mora, et al. (2011). "FtsH-dependent Processing of RNase Colicins D and E3 Means That Only the Cytotoxic Domains Are Imported into the Cytoplasm." *Journal of Biological Chemistry* **286**(33): 29397-29407.

Chen, V. B., W. B. Arendall, 3rd, et al. (2010). "MolProbity: all-atom structure validation for macromolecular crystallography." *Acta crystallographica. Section D, Biological crystallography* **66**(Pt 1): 12-21.

Chen, Y. R., T. Y. Yang, et al. (2011). "Delineation of the translocation of colicin E7 across the inner membrane of Escherichia coli." *Archives of Microbiology* **193**(6): 419-428.

Chen, Z. A., A. Jawhari, et al. (2010). "Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry." *EMBO Journal* **29**(4): 717-726.

Cho, U. S., M. W. Bader, et al. (2006). "Metal bridges between the PhoQ sensor domain and the membrane regulate transmembrane signaling." *Journal of*

Molecular Biology **356**(5): 1193-1206.

Chu, B. C., R. S. Peacock, et al. (2007). "Bioinformatic analysis of the TonB protein family." *Biometals* **20**(3-4): 467-483.

Clavel, T., J. C. Lazzaroni, et al. (1996). "Expression of the tolQRA genes of Escherichia coli K-12 is controlled by the RcsC sensor protein involved in capsule synthesis." *Molecular Microbiology* **19**(1): 19-25.

Click, E. M. and R. E. Webster (1997). "Filamentous phage infection: required interactions with the TolA protein." *Journal of Bacteriology* **179**(20): 6464-6471.

Collins, R. F., K. Beis, et al. (2007). "The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in Escherichia coli." *Proceedings of the National Academy of Sciences of the United States of America* **104**(7): 2390-2395.

Cooper, M. A. (2002). "Optical biosensors in drug discovery." *Nature Reviews. Drug Discovery* **1**(7): 515-528.

Cowan, S. W., R. M. Garavito, et al. (1995). "The structure of OmpF porin in a tetragonal crystal form." *Structure* **3**(10): 1041-1050.

Coyette, J. and A. van der Ende (2008). "Peptidoglycan: the bacterial Achilles heel." *FEMS Microbiology Reviews* **32**(2): 147-148.

Dabitz, N., N. J. Hu, et al. (2005). "Structural determinants for plant annexin-membrane interactions." *Biochemistry* **44**(49): 16292-16300.

Daley, D. O., M. Rapp, et al. (2005). "Global topology analysis of the Escherichia coli inner membrane proteome." *Science* **308**(5726): 1321-1323.

Danese, P. N. and T. J. Silhavy (1998). "Targeting and assembly of periplasmic and outer-membrane proteins in Escherichia coli." *Annual Reviews in Genetics* **32**: 59-94.

Dartigalongue, C., H. Nikaido, et al. (2000). "Protein folding in the periplasm in the absence of primary oxidant DsbA: modulation of redox potential in periplasmic space via OmpL porin." *EMBO journal* **19**(22): 5980-5988.

Deatherage, B. L., J. C. Lara, et al. (2009). "Biogenesis of bacterial membrane vesicles." *Molecular Microbiology* **72**(6): 1395-1407.

Delaglio, F., S. Grzesiek, et al. (1995). "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." *Journal of Biomolecular NMR* **6**(3): 277-293.

Dennis, J. J., E. R. Lafontaine, et al. (1996). "Identification and characterization of the tolQRA genes of Pseudomonas aeruginosa." *Journal of Bacteriology* **178**(24): 7059-7068.

Deprez, C., L. Blanchard, et al. (2002). "Macromolecular import into Escherichia coli: the TolA C-terminal domain changes conformation when interacting with the colicin A toxin." *Biochemistry* **41**(8): 2589-2598.

Deprez, C., L. Blanchard, et al. (2000). "Assignment of the 1H, 15N and 13C resonances of the C-terminal domain of the TolA protein of Escherichia coli, involved in cell envelope integrity." *Journal of Biomolecular NMR* **18**(2): 179-180.

Deprez, C., R. Lloubes, et al. (2005). "Solution structure of the *E.coli* TolA C-terminal domain reveals conformational changes upon binding to the phage g3p N-terminal domain." *Journal of Molecular Biology* **346**(4): 1047-1057.

Derouiche, R., R. Lloubes, et al. (1999). "Circular dichroism and molecular modeling of the E. coli TolA periplasmic domains." *Biospectroscopy* **5**(3): 189-198.

Drew, D., D. J. Slotboom, et al. (2005). "A scalable, GFP-based pipeline for membrane protein overexpression screening and purification." *Protein Science : A Publication of the Protein Society* **14**(8): 2011-2017.

Duan, K., E. R. Lafontaine, et al. (2000). "RegA, iron, and growth phase regulate expression of the Pseudomonas aeruginosa tol-oprL gene cluster." *Journal of Bacteriology* **182**(8): 2077-2087.

Eckert, B. and F. X. Schmid (2007). "A conformational unfolding reaction activates phage fd for the infection of Escherichia coli." *Journa of Molecular Biology* **373**(2): 452-461.

Elford, B. C., G. M. Cowan, et al. (1995). "Parasite-regulated membrane transport processes and metabolic control in malaria-infected erythrocytes." *The Biochemical Journal* **308 (Pt 2)**: 361-374.

Figueiredo, T. A., R. G. Sobral, et al. (2012). "Identification of genetic determinants and enzymes involved with the amidation of glutamic acid residues in the peptidoglycan of Staphylococcus aureus." *PLoS Pathogens* **8**(1): e1002508.

Fiuza, M., M. J. Canova, et al. (2008). "The MurC ligase essential for peptidoglycan biosynthesis is regulated by the serine/threonine protein kinase PknA in Corynebacterium glutamicum." *The Journal of Biological Chemistry* **283**(52): 36553-36563.

Fivash, M., E. M. Towler, et al. (1998). "BIAcore for macromolecular interaction."

Current Opinion in Biotechnology **9**(1): 97-101.

Ford, C. G., S. Kolappan, et al. (2012). "Crystal Structures of a CTX{varphi} pIII Domain Unbound and in Complex with a Vibrio cholerae TolA Domain Reveal Novel Interaction Interfaces." *The Journal of Biological Chemistry* **287**(43): 36258-36272.

Fratarnico, P. M., X. H. Yan, et al. (2011). "The complete DNA sequence and analysis of the virulence plasmid and of five additional plasmids carried by Shiga toxin-producing Escherichia coli O26:H11 strain H30." *International Journal of Medical Microbiology* **301**(3): 192-203.

Gabrielson, J. P., T. W. Randolph, et al. (2007). "Sedimentation velocity analytical ultracentrifugation and SEDFIT/c(s): limits of quantitation for a monoclonal antibody system." *Analytical Biochemistry* **361**(1): 24-30.

Garcia-Herrero, A., R. S. Peacock, et al. (2007). "The solution structure of the periplasmic domain of the TonB system ExbD protein reveals an unexpected structural homology with siderophore-binding proteins." *Molecular Microbiology* **66**(4): 872-889.

Garcia-Pino, A., S. Balasubramanian, et al. (2010). "Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity." *Cell* **142**(1): 101-111.

Garinot-Schneider, C., C. N. Penfold, et al. (1997). "Identification of residues in the putative TolA box which are essential for the toxicity of the endonuclease toxin colicin E9." *Microbiology* **143** (9): 2931-2938.

Gaspar, J. A., J. A. Thomas, et al. (2000). "Surface expression of O-specific lipopolysaccharide in Escherichia coli requires the function of the TolA protein." *Molecular Microbiology* **38**(2): 262-275.

Gasteiger, E., A. Gattiker, et al. (2003). "ExPASy: the proteomics server for in-depth protein knowledge and analysis." *Nucleic Acids Research* **31**(13): 3784-3788.

Ge, L., J. Yan, et al. (2012). "Three-dimensional paper-based electrochemiluminescence immunodevice for multiplexed measurement of biomarkers and point-of-care testing." *Biomaterials* **33**(4): 1024-1031.

Gerard, F., M. A. Brooks, et al. (2011). "X-Ray Structure and Site-Directed Mutagenesis Analysis of the Escherichia coli Colicin M Immunity Protein." *Journal of Bacteriology* **193**(1): 205-214.

Gerding, M. A., Y. Ogata, et al. (2007). "The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli." *Molecular Microbiology* **63**(4): 1008-1025.

Germon, P., M. C. Ray, et al. (2001). "Energy-dependent conformational change in the TolA protein of Escherichia coli involves its N-terminal domain, TolQ, and TolR." *Journal Bacteriology* **183**(14): 4110-4114.

Goemaere, E. L., E. Cascales, et al. (2007). "Mutational analyses define helix organization and key residues of a bacterial membrane energy-transducing complex." *Journal of Molecular Biology* **366**(5): 1424-1436.

Gokce, I., E. M. Raggett, et al. (2000). "The TolA-recognition site of colicin N. ITC, SPR and stopped-flow fluorescence define a crucial 27-residue segment." *J of Molecular Biology* **304**(4): 621-632.

Granseth, E., D. O. Daley, et al. (2005). "Experimentally constrained topology models for 51,208 bacterial inner membrane proteins." *Journal of Molecular Biology* **352**(3): 489-494.

Gsponer, J., H. Hopearuoho, et al. (2006). "Geometry, energetics, and dynamics of hydrogen bonds in proteins: structural information derived from NMR scalar couplings." *Journal of the American Chemical Society* **128**(47): 15127-15135.

Guihard, G., P. Boulanger, et al. (1994). "Colicin A and the Tol proteins involved in its translocation are preferentially located in the contact sites between the inner and outer membranes of Escherichia coli cells." *Journal of Biological Chemistry* **269**(8): 5874-5880.

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

Guo, H., W. Yi, et al. (2008). "Current understanding on biosynthesis of microbial polysaccharides." *Current Topics in Medicinal Chemistry* **8**(2): 141-151.

Han, B., Y. Liu, et al. (2011). "SHIFTX2: significantly improved protein chemical shift prediction." *Journal of Biomolecular NMR* **50**(1): 43-57.

Hands, S. L., L. E. Holland, et al. (2005). "Interactions of TolB with the translocation domain of colicin E9 require an extended TolB box." *Journal of Bacteriology* **187**(19): 6733-6741.

Hansen, A. L. and L. E. Kay (2011). "Quantifying millisecond time-scale exchange in proteins by CPMG relaxation dispersion NMR spectroscopy of side-chain carbonyl groups." *Journal of Biomolecular NMR* **50**(4): 347-355.

Hardy, K. G., G. G. Meynell, et al. (1973). "Two major groups of colicin factors: their evolutionary significance." *Molecular and General Genetics* **125**(3): 217-230.

Harris, R. K., E. D. Becker, et al. (2002). "NMR nomenclature: nuclear spin properties and conventions for chemical shifts. IUPAC Recommendations 2001. International Union of Pure and Applied Chemistry. Physical Chemistry Division. Commission on Molecular Structure and Spectroscopy." *Magnetic Resonance in Chemistry* **40**(7): 489-505.

Hecht, O., H. Ridley, et al. (2008). "Self-recognition by an intrinsically disordered protein." *FEBS Letters* **582**(17): 2673-2677.

Hecht, O., H. Ridley, et al. (2009). "A common interaction for the entry of colicin N and filamentous phage into Escherichia coli." *Journal of Molecular Biology* **388**(4): 880-893.

Hecht, O., Y. Zhang, et al. (2010). "Characterisation of the interaction of colicin A with its co-receptor TolA." *FEBS letters* **584**(11): 2249-2252.

Helbig, S. and V. Braun (2011). "Mapping Functional Domains of Colicin M." *Journal of Bacteriology* **193**(4): 815-821.

Helbig, S., S. I. Patzer, et al. (2011). "Activation of Colicin M by the FkpA Prolyl Cis-Trans Isomerase/Chaperone." *Journal of Biological Chemistry* **286**(8): 6280-6290.

Henry, T., S. Pommier, et al. (2004). "Improved methods for producing outer membrane vesicles in Gram-negative bacteria." *Research in Microbiology* **155**(6): 437-446.

Ho, D., M. R. Lugo, et al. (2011). "Membrane Topology of the Colicin E1 Channel Using Genetically Encoded Fluorescence." *Biochemistry* **50**(22): 4830-4842.

Holst, O. (2007). "The structures of core regions from enterobacterial lipopolysaccharides – an update." *FEMS Microbiology Letters* **271**(1): 3-11.

Holst, O., K. Bock, et al. (1995). "The structures of oligosaccharide bisphosphates isolated from the lipopolysaccharide of a recombinant Escherichia coli strain expressing the gene gseA [3-deoxy-D-manno-octulopyranosonic acid (Kdo) transferase] of Chlamydia psittaci 6BC." *European Journal of Biochemistry / FEBS* **229**(1): 194-200.

Hoogland, C., K. Mostaguir, et al. (2008). "The World-2DPAGE Constellation to promote and publish gel-based proteomics data through the ExPASy server." *Journal of Proteomics* **71**(2): 245-248.

Hoogland, C., J. C. Sanchez, et al. (1999). "Two-dimensional electrophoresis resources available from ExPASy." *Electrophoresis* **20**(18): 3568-3571.

Houben, E. N., R. Zarivach, et al. (2005). "Early encounters of a nascent membrane protein: specificity and timing of contacts inside and outside the ribosome." *The Journal of Cell Biology* **170**(1): 27-35.

Housden, N. G. and C. Kleanthous (2011). "Thermodynamic dissection of colicin interactions." *Methods in Enzymology* **488**: 123-145.

Housden, N. G., J. A. Wojdyla, et al. (2010). "Directed epitope delivery across the Escherichia coli outer membrane through the porin OmpF." *Proceedings of the National Academy of Sciences of the United States of America* **107**(50): 21412-21417.

Housden, N. G., S. R. Loftus, et al. (2005). "Cell entry mechanism of enzymatic bacterial colicins: porin recruitment and the thermodynamics of receptor binding." *Proceedings of the National Academy of Sciences of the United States of America* **102**(39): 13849-13854.

Hug, I. and M. F. Feldman (2011). "Analogies and homologies in lipopolysaccharide and glycoprotein biosynthesis in bacteria." *Glycobiology* **21**(2): 138-151.

Hug, I., B. Zheng, et al. (2011). "Exploiting bacterial glycosylation machineries for the synthesis of a Lewis antigen-containing glycoprotein." *The Journal of biological chemistry* **286**(43): 37887-37894.

Isnard, M., A. Rigal, et al. (1994). "Maturation and localization of the TolB protein required for colicin import." *Journal of Bacteriology* **176**(20): 6392-6396.

Jakes, K. S. and A. Finkelstein (2010). "The colicin Ia receptor, Cir, is also the translocator for colicin Ia." *Molecular Microbiology* **75**(3): 567-578.

Jancarik, J., R. Pufan, et al. (2004). "Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins." *Acta Crystallographica Section D* **60**(9): 1670-1673.

Jann, K., B. Jann, et al. (1965). "[Immunochemical studies of K antigens from Escherichia coli. II. K antigen from E. coli 08:K42(A):H-]." *Biochemische Zeitschrift* **342**(1): 1-22.

Jawad, Z. and M. Paoli (2002). "Novel sequences propel familiar folds." *Structure* **10**(4): 447-454.

Johne, B. (1998). "Epitope mapping by surface plasmon resonance in the BIAcore." *Molecular Biotechnology* **9**(1): 65-71.

Kampfenkel, K. and V. Braun (1993). "Membrane topologies of the TolQ and TolR proteins of Escherichia coli: inactivation of TolQ by a missense mutation in the

proposed first transmembrane segment." *Journal of Bacteriology* **175**(14): 4485-4491.

Kampfenkel, K. and V. Braun (1993). "Topology of the ExbB protein in the cytoplasmic membrane of Escherichia coli." *The Journal of Biological Chemistry* **268**(8): 6050-6057.

Karlsson, M., K. Hannavy, et al. (1993). "A sequence-specific function for the N-terminal signal-like sequence of the TonB protein." *Molecular Microbiology* **8**(2): 379-388.

Kellenberger, E. (1990). "The 'Bayer bridges' confronted with results from improved electron microscopy methods." *Molecular Microbiology* **4**(5): 697-705.

Kern, T., S. Hediger, et al. (2008). "Toward the characterization of peptidoglycan structure and protein-peptidoglycan interactions by solid-state NMR spectroscopy." *Journal of the American Chemical Society* **130**(17): 5618-5619.

Kittelberger, R. and F. Hilbink (1993). "Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels." *Journal of Biochemical and Biophysical Methods* **26**(1): 81-86.

Kiernan, J. (2000). Formaldehyde, formalin, paraformaldehyde and glutaraldehyde: What they are and what they do. *Microscopy Today* **1**: 8-12.

Kleanthous, C. and D. Walker (2001). "Immunity proteins: enzyme inhibitors that avoid the active site." *Trends in Biochemical Sciences* **26**(10): 624-631.

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

Knowling, S., A. I. Bartlett, et al. (2011). "Dissecting key residues in folding and stability of the bacterial immunity protein 7." *Protein Engineering Design and Selection* **24**(6): 517-523.

Krachler, A. M. (2009). "Functional and structural characterization of the Tol system-associated protein YbgF." *PhD thesis (York)*.

Krachler, A. M., A. Sharma, et al. (2010). "TolA modulates the oligomeric status of YbgF in the bacterial periplasm." *Journal of Molecular Biology* **403**(2): 270-285.

Krachler, A. M., A. Sharma, et al. (2010). "Self-association of TPR domains: Lessons learned from a designed, consensus-based TPR oligomer." *Proteins-Structure Function and Bioinformatics* **78**(9): 2131-2143.

Krewulak, K. D. and H. J. Vogel (2011). "TonB or not TonB: is that the question?" *Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire* **89**(2): 87-97.

Kycia, A. H., J. P. Wang, et al. (2011). "Atomic Force Microscopy Studies of a Floating-Bilayer Lipid Membrane on a Au(111) Surface Modified with a Hydrophilic Monolayer." *Langmuir* **27**(17): 10867-10877.

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

Lazzaroni, J. C., J. F. Dubuisson, et al. (2002). "The Tol proteins of Escherichia coli and their involvement in the translocation of group A colicins." *Biochimie* **84**(5-6): 391-397.

Lei, G. S., W. J. Syu, et al. (2011). "Repression of btuB gene transcription in Escherichia coli by the GadX protein." *Bmc Microbiology* **11**.

Lerner, T. R., A. L. Lovering, et al. (2012). "Specialized Peptidoglycan Hydrolases Sculpt the Intra-bacterial Niche of Predatory Bdellovibrio and Increase Population Fitness." *PLoS Pathogens* **8**(2): e1002524.

Levengood, S. K., W. F. Beyer, Jr., et al. (1991). "TolA: a membrane protein involved in colicin uptake contains an extended helical region." *Proceedings of the National Academy of Sciences of the United States of America* **88**(14): 5939-5943.

Li, C., Y. Zhang, et al. (2012). "Structural Evidence That Colicin A Protein Binds to a Novel Binding Site of TolA Protein in Escherichia coli Periplasm." *The Journal of Biological Chemistry* **287**(23): 19048-19057.

Liu, S. M., D. M. Miller, et al. (2011). "Cloning of genes encoding colicin E2 in Lactococcus lactis subspecies lactis and evaluation of the colicin-producing transformants as inhibitors of Escherichia coli O157:H7 during milk fermentation." *Journal of Dairy Science* **94**(3): 1146-1154.

Lloubes, R., E. Cascales, et al. (2001). "The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity?" *Research in Microbiology* **152**(6): 523-529.

Loftus, S. (2006). "Recruitment of TolB by the colicin E9 translocation domain." *(PhD Thesis, York)*

Lubkowski, J., F. Hennecke, et al. (1999). "Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA." *Structure* **7**(6): 711-722.

Luderitz, O., H. J. Risse, et al. (1965). "Biochemical Studies of Smooth-Rough Mutation in Salmonella Minnesota." *Journal of Bacteriology* **89**(2): 343-&.

Luderitz, O., D. A. Simmons, et al. (1965). "The immunochemistry of Salmonella chemotype VI O-antigens. The structure of oligosaccharides from Salmonella group U (o 43) lipopolysaccharides." *The Biochemical Journal* **97**(3): 820-826.

Luirink, J., G. von Heijne, et al. (2005). "Biogenesis of inner membrane proteins in Escherichia coli." *Annual Review of Microbiology* **59**: 329-355.

Magalhaes, P. O., A. M. Lopes, et al. (2007). "Methods of endotoxin removal from biological preparations: a review." *Journal of Pharmacy and Pharmaceutical Science* **10**(3): 388-404.

Majeed, H., O. Gillor, et al. (2011). "Competitive interactions in Escherichia coli populations: the role of bacteriocins." *Isme Journal* **5**(1): 71-81.

Manzo, E., A. Molinaro, et al. (2001). "A very efficient method to cleave Lipid A and saccharide components in bacterial lipopolysaccharides." *Carbohydrate Research* **333**(4): 339-342.

Maiolica, A., D. Cittaro, et al. (2007). "Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching." *Molecular and Cellular Proteomics : MCP* **6**(12): 2200-2211.

Martin, A. and F. X. Schmid (2003). "A proline switch controls folding and domain interactions in the gene-3-protein of the filamentous phage fd." *Journal of Molecular Biology* **331**(5): 1131-1140.

Masse, J. E. and R. Keller (2005). "AutoLink: automated sequential resonance assignment of biopolymers from NMR data by relative-hypothesis-prioritization-based simulated logic." *Journal of Magnetic Resonance* **174**(1): 133-151.

Mazumder, A., A. Maiti, et al. (2012). "A synthetic peptide mimic of lambda-Cro shows sequence-specific binding in vitro and in vivo." *ACS Chemical Biology* **7**(6): 1084-1094.

Medina, C., E. M. Camacho, et al. (2011). "Improved Expression Systems for Regulated Expression in Salmonella Infecting Eukaryotic Cells." *Plos One* **6**(8).

Meredith, T. C., P. Aggarwal, et al. (2006). "Redefining the requisite lipopolysaccharide structure in Escherichia coli." *ACS Chemical Biology* **1**(1): 33-42.

Meroueh, S. O., K. Z. Bencze, et al. (2006). "Three-dimensional structure of the bacterial cell wall peptidoglycan." *Proceedings of the National Academy of*

Sciences of the United States of America **103**(12): 4404-4409.

Michel-Briand, Y. and C. Baysse (2002). "The pyocins of Pseudomonas aeruginosa." *Biochimie* **84**(5-6): 499-510.

Microcal. (2004). ITC data analysis in Origin.

Mizuno, T. (1979). "A novel peptidoglycan-associated lipoprotein found in the cell envelope of Pseudomonas aeruginosa and Escherichia coli." *Journal of Biochemistry* **86**(4): 991-1000.

Mizuno, T. and M. Kageyama (1979). "Isolation and characterization of major outer membrane proteins of Pseudomonas aeruginosa strain PAO with special reference to peptidoglycan-associated protein." *Journal of Biochemistry* **86**(4): 979-989.

Morein, S., A. Andersson, et al. (1996). "Wild-type Escherichia coli cells regulate the membrane lipid composition in a "window" between gel and non-lamellar structures." *The Journal of Biological Chemistry* **271**(12): 6801-6809.

Mosbahi, K., C. Lemaitre, et al. (2002). "The cytotoxic domain of colicin E9 is a channel-forming endonuclease." *Nature Structural Biology* **9**(6): 476-484.

Muller, M. M. and R. E. Webster (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." *Journal of Bacteriology* **179**(6): 2077-2080.

Myszka, D. G., X. He, et al. (1998). "Extending the range of rate constants available from BIACORE: interpreting mass transport-influenced binding data." *Biophysical Journal* **75**(2): 583-594.

Myszka, D. G., M. D. Jonsen, et al. (1998). "Equilibrium analysis of high affinity interactions using BIACORE." *Analytical Biochemistry* **265**(2): 326-330.

Narita, S., S. Matsuyama, et al. (2004). "Lipoprotein trafficking in Escherichia coli." *Archives of Microbiology* **182**(1): 1-6.

Ng, C. L., K. Lang, et al. (2010). "Structural basis for 16S ribosomal RNA cleavage by the cytotoxic domain of colicin E3." *Nature Structural and Molecular Biology* **17**(10): 1241-1246.

Nikaido, H. (2003). "Molecular basis of bacterial outer membrane permeability revisited." *Microbiology and Molecular Biology Reviews* **67**(4): 593-656.

Pace, C. N., F. Vajdos, et al. (1995). "How to measure and predict the molar absorption coefficient of a protein." *Protein Science* **4**(11): 2411-2423.

Papadakos, G., N. G. Housden, et al. (2012). "Kinetic Basis for the Competitive Recruitment of TolB by the Intrinsically Disordered Translocation Domain of Colicin E9." *Journal of Molecular Biology*.

Parsons, L. M., F. Lin, et al. (2006). "Peptidoglycan recognition by Pal, an outer membrane lipoprotein." *Biochemistry* **45**(7): 2122-2128.

Paterson, G. K., H. Northen, et al. (2009). "Deletion of tolA in Salmonella Typhimurium generates an attenuated strain with vaccine potential." *Microbiology* **155**(Pt 1): 220-228.

Penfold, C. N., B. Healy, et al. (2004). "Flexibility in the receptor-binding domain of the enzymatic colicin E9 is required for toxicity against Escherichia coli cells." *Journal of Bacteriology* **186**(14): 4520-4527.

Pierce, M. M., C. S. Raman, et al. (1999). "Isothermal titration calorimetry of protein-protein interactions." *Methods* **19**(2): 213-221.

Postle, K. and R. A. Larsen (2007). "TonB-dependent energy transduction between outer and cytoplasmic membranes." *Biometals* **20**(3-4): 453-465.

Prieto, L. and T. Lazaridis (2011). "Computational studies of colicin insertion into membranes: The closed state." *Proteins-Structure Function and Bioinformatics* **79**(1): 126-141.

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

Pugsley, A. P. and M. Schwartz (1984). "Colicin E2 release: lysis, leakage or secretion? Possible role of a phospholipase." *EMBO journal* **3**(10): 2393-2397.

Raetz, C. R. (1978). "Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli." *Microbiological Reviews* **42**(3): 614-659.

Raetz, C. R. (1996). Bacterial lipopolysaccharide: a remarkable family of bioactive macroamphiphiles. Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology. F. C. Neidhardt. Washington, DC., ASM Press. **1:** 1035–1036.

Romer, C., S. I. Patzer, et al. (2011). "Expression, purification and crystallization of the Cmi immunity protein from Escherichia coli." *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **67**: 517-520.

Ruiz, N. D., Kahne, et al. (2006). "Advances in understanding bacterial outer-membrane biogenesis." *Nature Reviews Microbiology* **4**(1): 57-66.

Schapire, A., L, V Valpuesta, et al (2006). "TPR Proteins in Plant Hormone Signaling". *Plant Signalling and Behaviour* **1**(5): 229-30.

Schauer, K., D. A. Rodionov, et al. (2008). "New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'?" *Trends in Biochemical Sciences* **33**(7): 330-338.

Schendel, S. L., E. M. Click, et al. (1997). "The TolA protein interacts with colicin E1 differently than with other group A colicins." *Journal of Bacteriology* **179**(11): 3683-3690.

Schneider, M., M. Tognolli, et al. (2004). "The Swiss-Prot protein knowledgebase and ExPASy: providing the plant community with high quality proteomic data and tools." *Plant Physiology and Biochemistry* **42**(12): 1013-1021.

Schuck, P. (1997). "Reliable determination of binding affinity and kinetics using surface plasmon resonance biosensors." *Current Opinion in Biotechnology* **8**(4): 498-502.

Schuck, P. (1997). "Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules." *Annual Review of Biophysics and Biomolecular Structure* **26**: 541-566.

Schuck, P. (2000). "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling." *Biophysical Journal* **78**(3): 1606-1619.

Schuck, P. and P. Rossmanith (2000). "Determination of the sedimentation coefficient distribution by least-squares boundary modeling." *Biopolymers* **54**(5): 328-341.

Schwabe, B., G. Westphal, et al. (1980). "Phase-Transfer Catalyzed O-Acylation of 4-Acetyl-Phenol." *Zeitschrift Fur Chemie* **20**(5): 184-185.

Shen, Y. and A. Bax (2010). "SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network." *Journal of Biomolecular NMR* **48**(1): 13-22.

Slupsky, C. M., G. S. Shaw, et al. (1992). "A 1H NMR study of a ternary peptide complex that mimics the interaction between troponin C and troponin I." *Protein Science* **1**(12): 1595-1603.

Snyder, S., D. Kim, et al. (1999). "Lipopolysaccharide bilayer structure: effect of chemotype, core mutations, divalent cations, and temperature." *Biochemistry* **38**(33): 10758-10767.

Soelaiman, S., K. Jakes et al. (2001). “Crystal structure of colicin E3: implications for cell entry and ribosome inactivation.” *Molecular Cell* **8**(5): 1053-1062.

Spector, J., S. Zakharov, et al. (2010). "Mobility of BtuB and OmpF in the Escherichia coli outer membrane: implications for dynamic formation of a translocon complex." *Biophysical Journal* **99**(12): 3880-3886.

Sperandeo, P., R. Cescutti, et al. (2007). "Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of Escherichia coli." *Journal of Bacteriology* **189**(1): 244-253.

Sturgis, J. N. (2001). "Organisation and evolution of the tol-pal gene cluster." *Journal of Molecular Microbiology and Biotechnology* **3**(1): 113-122.

Tilby, M., I. Hindennach, et al. (1978). "Bypass of receptor-mediated resistance to colicin E3 in Escherichia coli K-12." *Journal of Bacteriology* **136**(3): 1189-1191.

Tilley, S. J. and H. R. Saibil (2006). "The mechanism of pore formation by bacterial toxins." *Current Opinion in Structural Biology* **16**(2): 230-236.

Toba, M., H. Masaki, et al. (1986). "Primary structures of the ColE2-P9 and ColE3-CA38 lysis genes." *Journal of Biochemistry* **99**(2): 591-596.

Tomich, J. M., D. Wallace, et al. (1998). "Aqueous solubilization of transmembrane peptide sequences with retention of membrane insertion and function." *Biophysical Journal* **74**(1): 256-267.

Tompa, P. (2012). "Intrinsically disordered proteins: a 10-year recap." *Trends in Biochemical Sciences* **37**(12): 509-516.

Typas, A., M. Banzhaf, et al. (2012). "From the regulation of peptidoglycan synthesis to bacterial growth and morphology." *Nature Reviews Microbiology* **10**(2): 123-136.

Vankemmelbeke, M., Y. Zhang, et al. (2009). "Energy-dependent immunity protein release during tol-dependent nuclease colicin translocation." *Journal Biological Chemistry* **284**(28): 18932-18941.

Vianney, A., M. M. Muller, et al. (1996). "Characterization of the tol-pal region of Escherichia coli K-12: translational control of tolR expression by TolQ and identification of a new open reading frame downstream of pal encoding a periplasmic protein." *Journal of Bacteriology* **178**(14): 4031-4038.

Vines, E. D., C. L. Marolda, et al. (2005). "Defective O-antigen polymerization in tolA and pal mutants of Escherichia coli in response to extracytoplasmic stress." *Journal of Bacteriology* **187**(10): 3359-3368.

Vollmer, W. (2008). "Structural variation in the glycan strands of bacterial peptidoglycan." *FEMS Microbiology Reviews* **32**(2): 287-306.

Vollmer, W. and U. Bertsche (2008a). "Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli." *Biochimica et biophysica acta* **1778**(9): 1714-1734.

Vollmer, W., D. Blanot, et al. (2008b). "Peptidoglycan structure and architecture." *FEMS Microbiology Reviews* **32**(2): 149-167.

Vollmer, W., B. Joris, et al. (2008c). "Bacterial peptidoglycan (murein) hydrolases." *FEMS Microbiology Reviews* **32**(2): 259-286.

Vranken, W. F., W. Boucher, et al. (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." *Proteins-Structure Function and Bioinformatics* **59**(4): 687-696.

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

Webster, R. E. (1991). "The tol gene products and the import of macromolecules into Escherichia coli." *Molecular Microbiology* **5**(5): 1005-1011.

Wei, Y., Z. Li, et al. (2009). "Characterization of the orf1-tolQRA operon in Pseudomonas aeruginosa." *Microbiology and Immunology* **53**(6): 309-318.

Weitzel, A. C. and R. A. Larsen (2008). "Differential complementation of TolA Escherichia coli by a Yersinia enterocolitica TolA homologue." *FEMS Microbiology Letters* **282**(1): 81-88.

Westphal, O. (1965). "Bacterial lipopolysaccharides: extraction with phenol-water and further applications of the procedure." *Methods in Carbohydrate Chemistry* **5**: 83-91.

Westphal, O., O. Luderitz, et al. (1965). "Specific polysaccharides of Enterobacteria." *Angewandte Chemie-International Edition* **4**(6): 529-&.

Whang, H. Y., O. Luderitz, et al. (1965). "Inhibition by lipoid A of formation of antibodies against common antigen of enterobacteriaceae." *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine* **120**(2): 371-374.

White, P. A., S. P. Nair, et al. (1998). "Molecular characterization of an outer membrane protein of Actinobacillus actinomycetemcomitans belonging to the

OmpA family." *Infection and immunity* **66**(1): 369-372.

Whitfield, C. (2006). "Biosynthesis and Assembly of Capsular Polysaccharides in Escherichia coli." *Annual Reviews in Biochemistry*.

Wiener, M. C. (2005). "TonB-dependent outer membrane transport: going for Baroque?" *Current Opinions in Structural Biology* **15**(4): 394-400.

Winsor, G. L., D. K. Lam, et al. (2011). “Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes” *Nucleic Acids Research*. **39**(D): 596-600

Wilken, C., K. Kitzing, et al. (2004). "Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease." *Cell* **117**(4): 483-494.

Wilkins, M. R., E. Gasteiger, et al. (1999). "Protein identification and analysis tools in the ExPASy server." *Methods in Molecular Biology* **112**: 531-552.

Witty, M., C. Sanz, et al. (2002). "Structure of the periplasmic domain of Pseudomonas aeruginosa TolA: evidence for an evolutionary relationship with the TonB transporter protein." *EMBO Journal* **21**(16): 4207-4218.

Wu, Y., S. F. Sui, et al. (1996). "Interaction of Colicin E(1) with Lipid Membranes." *Acta biochimica et biophysica Sinica* **28**(4): 346-351.

Yeaman, M. R. and N. Y. Yount (2003). "Mechanisms of antimicrobial peptide action and resistance." *Pharmacological Reviews* **55**(1): 27-55.

Zalucki, Y. M., W. M. Shafer, et al. (2011). "Directed evolution of efficient secretion in the SRP-dependent export of TolB." *Biochimica et biophysica acta* **1808**(10): 2544-2550.

Zhang, D. P., L. M. Iyer, et al. (2011). "A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems." *Nucleic Acids Research* **39**(11): 4532-4552.

Zhang, X. Y., E. L. Goemaere, et al. (2009). "Mapping the interactions between escherichia coli tol subunits: rotation of the TolR transmembrane helix." *Journal of Biological Chemistry* **284**(7): 4275-4282.

Zuber, B., M. Haenni, et al. (2006). "Granular layer in the periplasmic space of gram-positive bacteria and fine structures of Enterococcus gallinarum and Streptococcus gordonii septa revealed by cryo-electron microscopy of vitreous sections." *Journal of Bacteriology* **188**(18): 6652-6660.