THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/2049

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

The Twin Arginine Translocase Pathway: A Functional and Structural Study in *Escherichia coli* and *Synechocystis* sp. PCC6803

Gemma Zara Louise Warren

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Molecular Organisation and Assembly in Cells Doctoral Training Centre

University of Warwick

April 2009

Table of Contents

List of Figures

List of Tables

Acknowledgments

Firstly, I would like to thank Ann Dixon and Colin Robinson for their supervision and guidance throughout the course of this work. Their suggestions and direction have been invaluable in enabling me to complete this research.

Thanks go to Claire Barrett (donation of primers and advice), Andy Beevers (ATR-FTIR), Janet Crawford (peptide synthesis), Angeliki Damianoglou (LD and CD), Matt Hicks (LD, CD and donation of lipids), Cristina Matos (donation of primers and collaborative work), Joanne Oates (donation of plasmid and advice) and Ed Spence (*Synechocystis* work) for all the help and advice they have given me. I would also like to thanks Nigel Burroughs for his help at the beginning of this project. Special thanks go to Anja Nenninger for all her help with the *Synechocystis* and confocal microscopy work, and for helping to convert a mathematician into an experimental biochemist!

I would like to thank the members of the Robinson group past and present for introducing me to laboratory work and making me feel welcome. Special thanks to Nishi Vashisht for keeping us organised and making sure there were always lots of biscuits!

I would also like to thank the members of the Dixon group for all their help and many discussions, especially when I first started work in Chemistry. Thanks also go to the members of Alison Rodger's group, for letting me use their equipment and making me feel welcome on the sixth floor.

I am really grateful to everyone at MOAC for all the help and friendship over the past four and a half years, particularly the 2004 cohort. They have all helped me in so many ways, especially in providing sweet treats on Cake Day! Special thanks go to Alison Rodger, for giving me a place on the MOAC course, and always taking the time to help in any way she could. Special thanks also go to Dorothea Mangels and Monica Lucena, for keeping us all organised. I would also like to thank Adair Richards for reading my thesis and giving me many helpful suggestions.

I would like to thank my Grace Church family for all the help and prayers they have provided since moving to Birmingham. I am extremely thankful to all of my family

for all the fun they have given me throughout my life, and for always being fully behind me in whatever I have chosen to do. I am particularly thankful to my husband, Nathan, for his continued love and support through all the ups and downs that working towards a PhD brings.

Finally, but most importantly, I thank God for all the blessings He has provided for me throughout my life, and for sustaining me through my weaknesses.

"Trust in the LORD with all your heart; do not depend on your own understanding. Seek his will in all you do, and he will show you which path to take."

Holy Bible, New Living Translation, Proverbs 3:5–6

Declaration

The work presented in this thesis is original, and was conducted by the author, unless otherwise stated, under the supervision of Dr Ann Dixon and Professor Colin Robinson. It has not previously been presented for another degree.

The work was funded by the EPSRC, through the Molecular Organisation and Assembly in Cells Doctoral Training Centre.

All sources of information have been acknowledged by means of reference.

Collaborative work has been carried out contributing towards the material presented in this thesis. The extent of the collaborative work is indicated in the relevant chapters (Chapter 3 and Chapter 6). In summary, collaborative work was carried out with the following:

Anja Nenninger and Christopher Müller, Department of Biological Sciences, University of Warwick, performed work contributing to Chapter 3.

Cristina Matos, Department of Biological Sciences, University of Warwick, performed work contributing to §6.6.

Some of the work presented in this thesis has been published in the following journal:

Warren, G., Oates, J., Robinson, C. & Dixon, A. M. (2009). Contributions of the transmembrane domain and a key acidic motif to assembly and function of the TatA complex. *J Mol Biology* **388**, 122–132.

Summary

The bacterial twin-arginine translocase (Tat) pathway is able to export large folded proteins across the cytoplasmic membrane. It derives its name from the almost invariant twin-arginine motif present in the signal peptide of Tat substrates. Components of the Tat pathway have been identified in many bacteria, and in *Escherichia coli* one of these essential proteins, TatA, forms large complexes of variable size. The research presented in this thesis sought to gain insight into the Tat pathway in *E. coli* and the cyanobacterium *Synechocystis* sp. PCC6803, mainly focusing on the TatA component of the *E. coli* Tat pathway and how its various domains contribute to function and complex formation.

Green fluorescent protein tagged with a Tat signal peptide was used to compare the Tat pathway in *E. coli* and *Synechocystis*, using biochemical and bioimaging techniques. Substitution of the twin-arginine motif with lysine residues highlighted differences between the specificity of the two systems.

The transmembrane (TM) domain and amphipathic helix (APH) of *E. coli* TatA were analysed using synthetic peptides and the TOXCAT assay. Investigation of the TM domain demonstrated it is α-helical and spontaneously inserts into membranes. It interacts relatively weakly and substitution of a glutamine does not affect these interactions or function of the protein at high expression levels. In contrast, the APH only forms an α-helical secondary structure in high detergent concentration or in the presence of negatively charged lipids, and does not insert into lipid bilayers.

Analysis of the unstructured C-terminus of TatA led to the identification of an acidic motif just after the APH. Removal of this motif has a severe effect on function and complex formation of TatA. Removal of the C-terminus does not affect the proofreading abilities of TatA.

The research presented here has led to the proposal of a new model for TatA complex formation in *E. coli* – the weak interactions of the TM domains bring the monomers together, whilst interactions between the acidic motif and other regions stabilise these complexes.

Abbreviations

A	Absorbance
aa	Amino acids
ACN	Acetonitrile
Amp	Ampicillin
APH	Amphipathic
APS	Ammonium persulfate
ATP	Adenosine triphosphate
ATR-FTIR	Attenuated total reflectance Fourier transform infrared spectroscopy
AU	Absorbance units
b	Optical path length
BN PAGE	Blue-native PAGE
BS^3	Bis(sulfosuccinimidyl)suberate
BSA	Bovine serum albumin
c	Concentration
C	Cytoplasm
C-	Carboxyl-
CAM	Chloramphenicol
CAT	Chloramphenicol acetyl transferase
CD	Circular dichroism
CHI	CNS searching of helix interactions
CNS	Crystallography and NMR system
dH_2O	Distilled water
DMPC	1,2-Dimyristoyl-*sn*-Glycero-3-Phosphocholine
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DPC	Dodecylphosphocholine
DPPC	1,2-Dipalmitoyl-*sn*-Glycero-3-Phosphocholine
DPPG	1,2-Dipalmitoyl-*sn*-Glycero-3-[Phospho-*rac*-(1-glycerol)]
ECL	Enhanced chemiluminescence

EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionisation mass spectrometry
FeS	Iron-sulphur
FPLC	Fast protein liquid chromatography
GFP	Green fluorescent protein
GpA	Glycophorin A
GTP	Guanosine triphosphate
h	Hexahistidine tag
his	Hexahistidine tag
HPLC	High performance liquid chromatography
HRP	Horseradish peroxidase
IPTG	Isopropylthiogalactoside
k	Turbidity constant
LB	Luria-Bertani broth
LD	Linear dichroism
LD^T	Background turbidity dichroism
LSCM	Laser scanning confocal microscope
M	Membrane
MBP	Maltose binding protein
N-	Amino-
NMR	Nuclear magnetic resonance
OCD	Oriented circular dichroism
OD	Optical density
P	Periplasm
PAGE	Poly-acrylamide gel electrophoresis
PBS(-T)	Phosphate buffered saline (plus tween-20)
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PMF	Proton motive force
PVDF	Polyvinylidene fluoride

rpm	Revolutions per minute
RR-motif	Twin-arginine motif
s	Strep-tag II
SD	Standard deviation
SDS	Sodium dodecyl sulphate
Sec	Secretory pathway
SRP	Signal recognition particle
strep	Strep-tag II
Tat	Twin-arginine translocase
TBA	Tris-borate-EDTA
TCA	Trichloroacetic acid
TCM	Tris-calcium-magnesium
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TM	Transmembrane
TMAO	Trimethylamine-N-oxide
TorA	Trimethylamine-N-oxide reductase
UV	Ultraviolet
w/v	Weight per volume
Δ	Delta (gene knockout e.g. Δ*tatB*)
ε	Extinction coefficient
λ	Wavelength

Chapter 1. Introduction

The transport of proteins is an essential process in all living cells. A large number of proteins are not synthesised where they are required and must be directed to different cellular compartments, often having to cross one or more lipid bilayer membranes. These lipid membranes play an important role in the cell, separating subcellular compartments, maintaining ion gradients and compartment composition. However they also form a barrier which proteins need to cross in order to reach their final destination. Cells have developed complex translocation pathways in order to transport proteins across these barriers. Translocation pathways have been characterised in many different membrane systems, including those of the endoplasmic reticulum (ER), peroxisomes, chloroplasts, and the inner and outer membranes of prokaryotes (Aldridge *et al.*, 2009; Brown & Baker, 2008; Cross *et al.*, 2009; Driessen & Nouwen, 2008; Milenkovic *et al.*, 2007; Rapoport, 2007).

Despite the wide variety and complexity of transport pathways, there are general principles which are common to most systems (Schatz & Dobberstein, 1996). The first of these principles is the signal peptide, also known as the signal-, leader-, targeting-, transit- or pre-sequence. This is a sequence of amino acids, often located on the amino- (N-) terminus of the substrate, which targets the protein to the correct cellular compartment (von Heijne, 1990). Its length and composition varies between pathways, providing recognition specificity. Upon reaching its destination, the signal peptide is usually cleaved from the precursor protein (preprotein) and the mature protein is free to function. Chaperones may bind to the preprotein before translocation, keeping it in the conformation required for transport and also directing it to the membrane receptor, which recognises the signal peptide. Interactions between the membrane receptor and the signal peptide lead to association with a membrane integrated translocation pore. These transmembrane (TM) protein channels are typically hetero-oligomeric complexes and contain a hydrophilic pore through which preproteins pass. Gating of the translocation pore prevents leakage and maintains the integrity of the membrane. Translocation motors also exist to drive the transport of the preprotein through the TM protein channel. Originally it was thought that preproteins must be maintained in a largely unfolded conformation during translocation. However pathways have been discovered which are able to export folded proteins across membranes without compromising the integrity of the

membrane. One of these pathways, the twin-arginine translocase (Tat) pathway, is the subject of this study.

1.1 Protein translocation in bacteria

The structure of a Gram-negative bacterium, such as *Escherichia coli*, is relatively simple, Figure 1.1A. Bacterial cells typically have four compartments; the cytoplasm, the cytoplasmic (inner) membrane, the periplasm and the outer membrane. Proteins are synthesised in the cytoplasm, however ~20% of proteins must be targeted to one of the other three compartments or the extracellular medium via the cytoplasmic membrane (Pugsley, 1993). The majority of proteins are translocated across the cytoplasmic membrane by the general secretory (Sec) pathway. Specialised pathways to export proteins to the outer membrane also exist, reviewed in (Kostakioti *et al.*, 2005). These can either work in conjunction with cytoplasmic membrane translocation pathways such as the Sec and Tat pathways, §1.2 and §1.3, or in isolation, translocating the proteins across both the cytoplasmic and outer membrane in one step. Here, the translocation across the bacterial cytoplasmic membrane is reviewed, concentrating on the Sec and Tat pathways.

Homologues of the Sec and Tat pathways also exist in plant chloroplast thylakoid membranes. It is widely accepted that the Gram-negative, photosynthetic cyanobacteria are the ancestors of plastids in higher plants. As such they present an interesting study in protein translocation, providing a possible link between thylakoid and bacterial systems. As well as the cytoplasmic and outer membranes of a typical Gram-negative bacterium, cyanobacteria also contain an internal thylakoid membrane system, Figure 1.1B and C. Homologues of the general cytoplasmic membrane translocation pathways have been identified in cyanobacteria (Aldridge *et al.*, 2008; Fulda *et al.*, 2000; Nakai *et al.*, 1993; Nakai *et al.*, 1994; Srivastava *et al.*, 2005), however they have not been characterised like those present in *E. coli* and chloroplast thylakoid membranes.

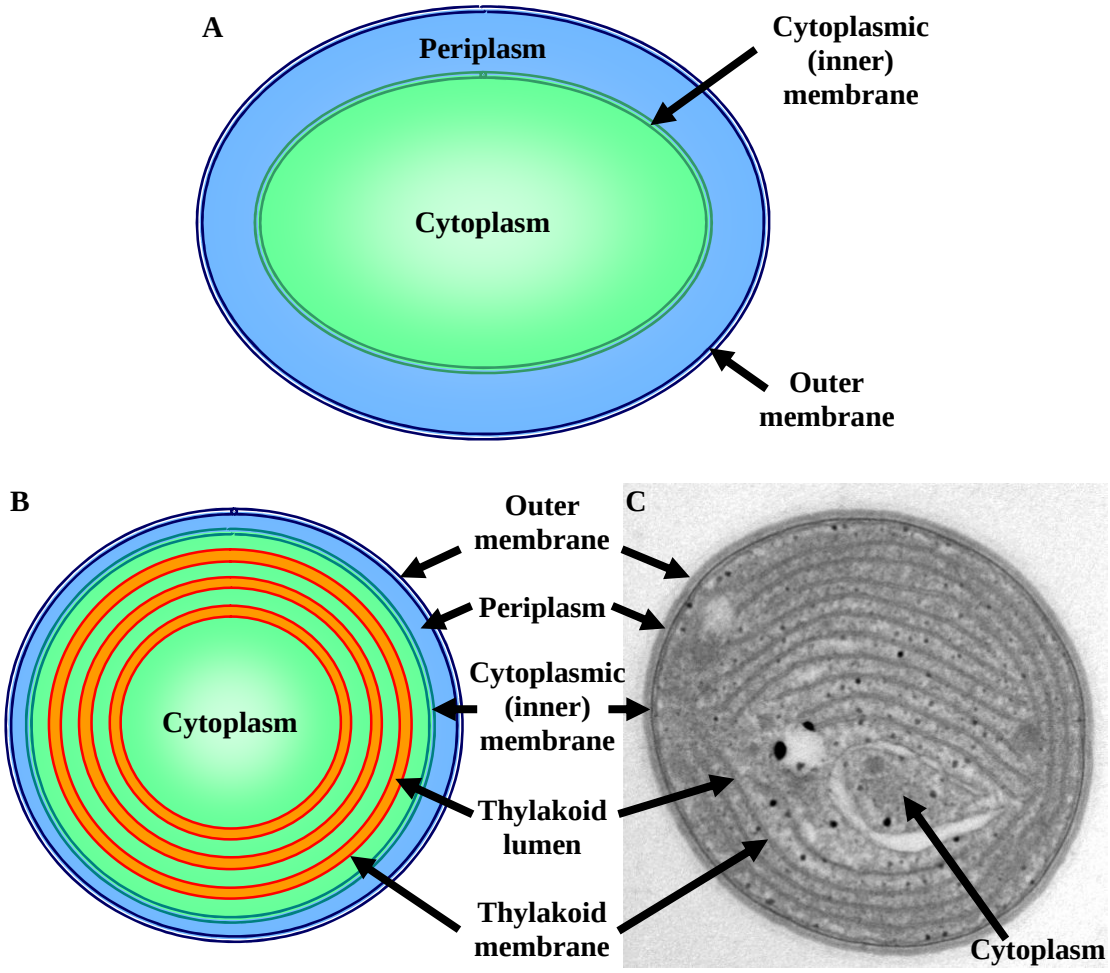


Figure 1.1 Schematics of Gram-negative and cyanobacterial cells. A. A typical Gram-negative bacterial cell consists of four compartments, indicated by arrows. **B.** As well as the four basic compartments, cyanobacterial cells also have thylakoid lumen and membranes. **C.** An electron micrograph of *Synechocystis* sp. PCC6803 shows the complexity of the cell interior.

1.2 The Sec pathway

In bacteria, the Sec pathway is responsible for the export of the majority of extracellular proteins. It is also involved in the insertion of membrane proteins into the bacterial cytoplasmic membrane, reviewed in (Xie & Dalbey, 2008). It has been identified in the cytoplasmic membrane of both Gram-negative and Gram-positive bacteria (Driessen & Nouwen, 2008), the thylakoid membrane of chloroplasts (Aldridge *et al.*, 2009), the endoplasmic reticulum (ER) (Johnson & van Waes, 1999) and archaea (Bolhuis, 2004). Components of the Sec pathway have also been identified in cyanobacteria (Nakai *et al.*, 1993; Nakai *et al.*, 1994). The pathway functions in a similar way in all systems identified so far, with proteins being exported in an unfolded conformation. In its simplest form the system consists of

protein targeting components, a motor protein and a membrane integrated translocation pore. Proteins which are destined to be exported by the Sec pathway are synthesised with a signal peptide which is recognised by protein targeting components and then directed for export. The signal peptide is on average 24 amino acids long (Cristobal *et al.*, 1999) and has a tripartite structure, composed of a positively charged N-terminal region (N-region), a hydrophobic core (H-region) and a polar C-terminal region (C-region) which normally contains a signal peptidase cleavage site. In *E. coli* the hydrophobicity of the signal peptide affects which protein targeting factors bind and hence determines if the protein is exported co- or post-translationally (de Gier *et al.*, 1998; Lee & Bernstein, 2001). Integral membrane proteins do not typically have a signal peptide and instead the hydrophobic TM domains act as signal peptides.

1.2.1 The components of the Sec pathway

In *E. coli* translocation by the Sec pathway can proceed by two different routes, depending on the interactions between the preprotein and the protein targeting components, signal recognition particle (SRP) and SecB. The bacterial SRP is much smaller than its eukaryotic homologue and consists of a 4.5 S RNA and a 48 kDa GTPase P48 or Ffh (fifty four homologue), reviewed in (Egea *et al.*, 2005). SecB is a cytosolic chaperone which, as well as being involved in the Sec pathway, has also been shown to have general chaperone activity (Ullers *et al.*, 2004; Zhou *et al.*, 2001). Depending on the hydrophobicity of the preprotein signal peptide, either SRP or SecB bind and translocation proceeds co- or post-translationally, respectively, as discussed in §1.2.2.

The motor protein, which drives translocation of the unfolded preprotein by ATP hydrolysis, is the soluble SecA protein. This is a central component of the bacterial Sec pathway and interacts with nearly all other components. Recent structural studies of SecA homologues from various bacteria have shown that it crystallises as a dimer (Papanikolau *et al.*, 2007; Sharma *et al.*, 2003; Vassylyev *et al.*, 2006; Zimmer *et al.*, 2006). It is found to localise to both the cytosol and the cytosolic membrane (Cabelli *et al.*, 1991), the latter through interactions with anionic phospholipids (Lill *et al.*, 1990) and the Sec translocation pore (Hartl *et al.*, 1990).

The Sec translocation pore in *E. coli* is composed of three integral membrane proteins, SecY, SecE and SecG, which form a stable complex denoted SecYEG (Akimaru *et al.*, 1991; Brundage *et al.*, 1990). Structural studies of SecYEG and its homologues have revealed information on the possible mechanism of translocation (Breyton *et al.*, 2002; Mitra *et al.*, 2005; van den Berg *et al.*, 2004). SecY is a 10 TM domain protein, which has its N- and C-termini located on the cytoplasmic side of the membrane. The 10 TM helices are organised into two separate domains, which are connected by a periplasmic loop to form a clamshell-like structure with a central pore. SecE, a 3 TM domain protein, holds together the two domains of SecY, whilst SecG, with a single TM domain, localises on the periphery of the complex. The structure of the SecYEG complex reveals a cavity at the periplasmic face of the membrane which is 'plugged' by a periplasmic loop of SecY. It is thought that this plug is displaced when translocation is initiated and results in an aqueous channel across the membrane (Harris & Silhavy, 1999). It is likely that the SecY clamshell-like structure also expands to allow insertion of the signal peptide and the early mature region of the preprotein in a hairpin-like structure.

1.2.2 Co- and post-translational export

There are two major routes by which proteins are directed to the Sec translocation pore in *E. coli*. In the first, which is used mainly by integral membrane proteins, SRP interacts with the signal peptide or hydrophobic TM domain of the nascent protein, directing it to be exported co-translationally, Figure 1.2A. The SRP binds to the ribosome and protein nascent chain, and then interacts with FtsY; a membrane bound protein which also interacts with SecYEG (Angelini *et al.*, 2005; Bernstein *et al.*, 1989). This interaction enables FtsY and the Ffh subunit of SRP to bind GTP. Upon GTP hydrolysis the ribosome and nascent chain complex are transferred to the SecYEG complex and the protein continues to be translated with the ribosome driving export or membrane insertion of the protein via SecYEG.

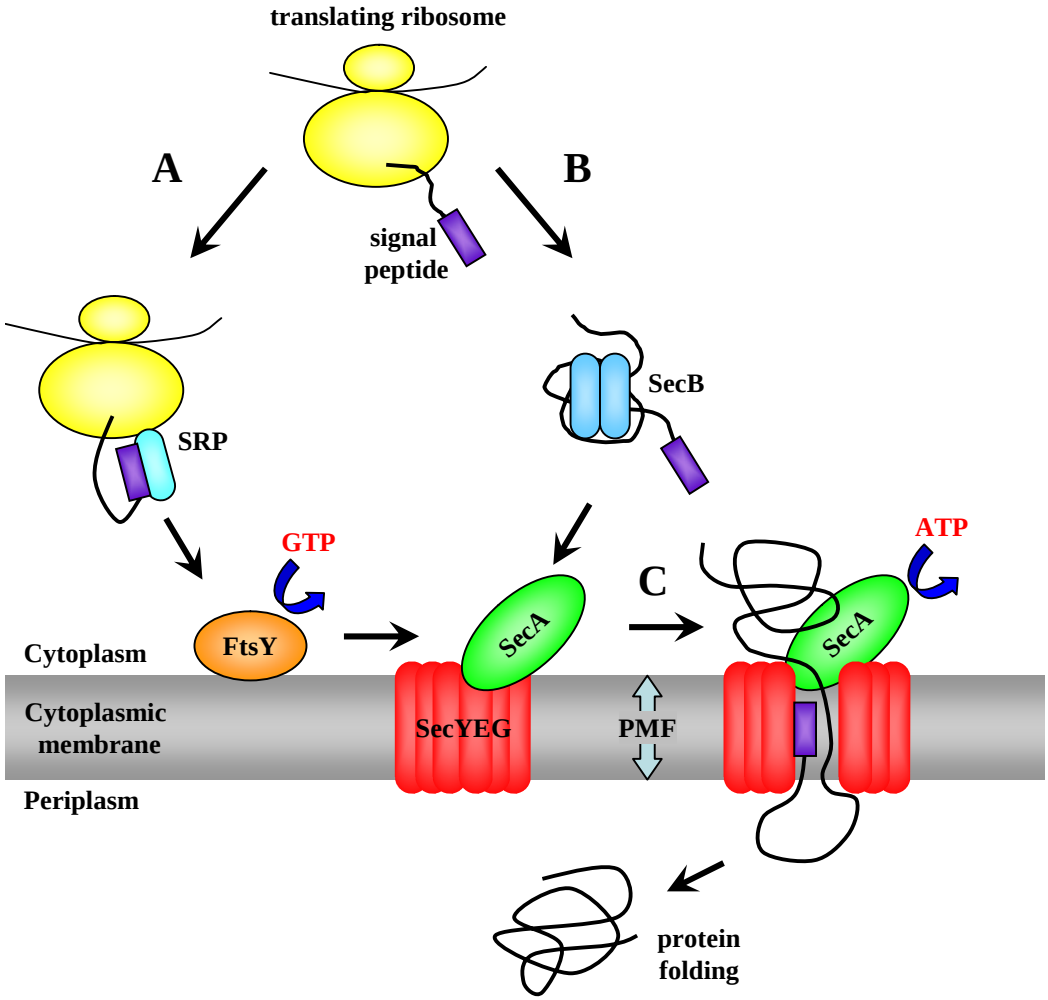


Figure 1.2 Protein translocation by the Sec pathway in bacteria. Newly synthesised proteins are directed to the Sec translocase by different routes. **A.** Co-translationally, preproteins associate with the signal recognition particle (SRP) and then associate with the Sec translocase via FtsY, which requires GTP. **B.** Post-translationally, a preprotein either associates directly with the Sec translocase or binds SecB, which keeps it in a largely unfolded state and directs it to the translocase. **C.** After association with the translocase the protein is threaded through the channel formed by the SecYEG complex, driven by ATP hydrolysis by SecA. The proton motive force (PMF) can also drive translocation. Once translocation is completed, the signal peptide is cleaved and the protein is able to fold. Figure adapted from (Driessen & Nouwen, 2008).

The second route, which is used by the majority of secretory proteins in *E. coli*, exports proteins post-translationally, Figure 1.2B. The signal peptide of the protein is recognised by SecB, which binds to the newly synthesised protein and keeps it in a largely unfolded conformation. SecB then binds to SecA, the unfolded preprotein is transferred to SecA, and SecB disassociates. The signal peptide may also bind directly to SecA, without the need for SecB. Translocation through the SecYEG pore

complex is initiated by the binding of ATP to SecA, which enables the hairpin-like structure of the signal peptide and the early mature preprotein to insert, Figure 1.2C. Upon ATP hydrolysis, the bound preprotein is released and then rebound by SecA. This rebinding results in the translocation of ~ 2–2.5 kDa of preprotein through the pore and further binding of ATP also results in translocation of a similar amount of protein. Subsequent rounds of binding and hydrolysis of SecA result in a stepwise 'pushing' of the preprotein through the SecYEG pore complex, until translocation is complete and SecA disassociates. The proton motive force (PMF) is also involved at various points of the translocation process, stimulating insertion of the preprotein (van Dalen *et al.*, 1999) and driving translocation (Driessen, 1992). Upon completion of translocation the signal peptide of the preprotein is cleaved off by a signal peptidase SPase I (Paetzel *et al.*, 2002), and the mature protein folds in the periplasm.

1.2.3 Membrane protein insertion

As well as translocating proteins across the membrane, the co-translational Sec pathway is involved in the insertion of proteins into the cytoplasmic membrane. The TM domain of a preprotein can act as a signal peptide and bind SRP, directing it to SecYEG. Just how SecYEG recognises that the protein must be inserted into the membrane is still unknown but it is believed that the hydrophobicity of the TM domain and charge distributions around its flanking region halt translocation (Duong & Wickner, 1998; von Heijne, 1986; White & von Heijne, 2008; Xie *et al.*, 2007). At this point it is likely that the SecYEG channel forms a lateral gate through which TM domains can partition into the lipid bilayer (van den Berg *et al.*, 2004). YidC, a structural homologue to Oxa1 in mitochondria and Alb3 in chloroplasts, has been shown to be involved in the insertion of membrane proteins either independently or in collaboration with the Sec pathway, reviewed in (Xie & Dalbey, 2008). The Tat pathway is also involved in the insertion of membrane proteins, though to a lesser extent, as discussed in §1.3.1.

1.3 The Tat pathway

Before the discovery of the Tat pathway, it was thought that the Sec pathway was solely responsible for the export of proteins across the bacterial cytoplasmic

membrane. It had also been shown that a Sec-type system was involved in the translocation of proteins across the thylakoid inner membrane (Laidler *et al.*, 1995; Schuenemann *et al.*, 1999; Yuan *et al.*, 1994). Export via the Sec pathway is driven by ATP and the PMF, however a subset of thylakoid lumenal proteins were identified which did not require ATP for transport but instead depended on the difference in pH between the lumen of the thylakoid and the stroma (Cline *et al.*, 1992; Klosgen *et al.*, 1992; Mould & Robinson, 1991). The identification of these proteins led to the discovery of a new Sec-independent pathway in chloroplasts termed the ΔpH pathway. Analysis of the signal sequences of proteins exported via this pathway revealed the presence and importance of a highly conserved twin-arginine (RR) motif (Chaddock *et al.*, 1995), which later led to the pathway being termed the twin-arginine translocase (Tat) pathway. Isolation of an export deficient mutant and identification of a protein involved in the ΔpH pathway, Hcf106, led to the identification of homologues in prokaryotes (Settles *et al.*, 1997). Components of the Tat pathway have since been identified in the cytoplasmic membrane of both Gram-positive and Gram-negative bacteria, archaea and cyanobacteria, with putative homologues also identified in plant mitochondria (Yen *et al.*, 2002).

Interest in the Tat pathway has grown since its discovery because of its unique features which distinguish it from other pathways. Unlike the Sec pathway, where proteins are exported in an unfolded conformation, the Tat pathway is able to export folded proteins. This has led to commercial interest with a view to secreting complex heterologous proteins which are not compatible with the Sec pathway (Brüser, 2007; Kim *et al.*, 2005). Although there are other pathways which can translocate folded proteins, namely import pathways in the peroxisome (Glover *et al.*, 1994; Walton *et al.*, 1995) and transport across the nuclear membrane in eukaryotes (Feldherr & Akin, 1990), the Tat pathway is thus far the only pathway to be identified which achieves this across a coupled membrane. Additionally, the absence of any Tat component homologues in animal and fungal cells, and the importance of Tat substrates in virulence (De Buck *et al.*, 2008) have highlighted the pathway's potential as a novel target for antimicrobial drugs.

1.3.1 The role of the Tat pathway

The Tat pathway is involved in a number of processes in the cell due to its ability to export folded proteins without compromising the integrity of the membrane. The first evidence that it is able to export folded proteins came from the identification of Tat substrates which contained cytoplasmically inserted cofactors, reviewed in (Berks *et al.*, 2003). Cofactor containing Tat substrates are involved in a variety of processes in the cell. A major group of Tat substrates are molybdenum containing enzymes, which catalyse a diverse range of redox reactions including those essential for the global carbon, nitrogen and sulphur cycles. The most common Tat substrates are iron-sulphur (FeS) proteins which are involved in generation of radicals and redox reactions. Other Tat substrates which contain cofactors include hydrogenases and copper containing proteins. To enable the cofactors to be inserted into these proteins by cytoplasmic chaperones or other factors, the protein must be substantially or completely folded. There is no evidence to suggest these cofactors can be bound in the periplasm so it is believed the proteins must be fully or largely folded when exported. In addition, prevention of the binding of cofactors has been shown to block transport of Tat substrates. When the cofactors of trimethylamine-N-oxide (TMAO) reductase (TorA) are not inserted, it is no longer exported to the periplasm (Santini *et al.*, 1998), and substitutions in the [4Fe-4S] centre of the FeS protein NrfC, which were expected to disrupt cofactor binding and folding, result in a complete block of export (Matos *et al.*, 2008). These studies also suggest some sort of quality control and proofreading, which is discussed in §1.3.2.

Further evidence for the export of folded proteins comes from the Tat dependent translocation of proteins apparently without Tat signal peptides. The *E. coli* periplasmic nickel-containing hydrogenase 2 is composed of a small subunit containing a Tat signal peptide and a large subunit which does not contain a signal peptide. The two subunits are only exported when they both bind their cofactors and form a complex, which is then targeted to the Tat pathway by the signal peptide on the small subunit in a so-called hitchhiker mechanism (Rodrigue *et al.*, 1999).

The export of heterologous proteins provides further evidence of the ability of the Tat pathway to export folded proteins. Green fluorescent protein (GFP) was demonstrated to be active in both the cytoplasm and periplasm of *E. coli* when

targeted with a Tat signal peptide (Thomas *et al.*, 2001). However, when exported via the Sec pathway, in an unfolded conformation, GFP was apparently unable to fold correctly after translocation and was not active in the periplasm (Feilmeier *et al.*, 2000). It was also demonstrated that PhoA was only transported if its native disulphide bonds, which stabilise its structure, were able to form (DeLisa *et al.*, 2003).

The ability to export folded proteins is important, and this is certainly the case for the archaeon *Halobacterium* sp. strain NCR-1, which exports the majority of its secretory proteins via the Tat pathway (Bolhuis, 2002; Rose *et al.*, 2002). The extensive use of the Tat pathway is most likely due to the need for proteins to fold in the cytoplasm and avoid aggregation caused by the virtually saturating levels of salt present in their environment.

In addition to translocation of folded proteins across membranes, the Tat pathway also demonstrates a novel mechanism for the insertion of proteins into the membrane (De Buck *et al.*, 2007; Hatzixanthis *et al.*, 2003; Summer *et al.*, 2000). In *E. coli* five predicted Tat substrates were indentified that contained a single C-terminal α-helical TM domain (Hatzixanthis *et al.*, 2003; Sargent *et al.*, 2002). Fusion of these C-terminal domains to periplasmic soluble proteins led to membrane localisation, demonstrating the Tat pathway's involvement in insertion. Bioimaging of the cyanobacterium *Synechocystis* sp. PCC6803 has also demonstrated that the predicted Tat substrate, PetC3, localises to the membrane (Aldridge *et al.*, 2008). The mechanism by which proteins are inserted is as yet unknown. It is likely that recruitment of the substrate to the membrane and translocation are two separate events since in *Legionella pneumophila* the Rieske FeS protein, PetA, was still inserted into the membrane in Tat deletion strains but in the incorrect orientation (De Buck *et al.*, 2007).

The Tat pathway has also been highlighted as playing a role in the virulence of bacteria. Homologues of components of the Tat pathway have been identified in a number of human and plant pathogens including *E. coli* O:157, *Pseudomonas aeruginosa*, *Mycobacterium tuberculosis* and *Agrobacterium tumefaciens* (Dilks *et al.*, 2003; Yen *et al.*, 2002). Tat deletion strains of these pathogens demonstrate multiple phenotypes including effects on infection, iron acquisition, biofilm

formation and response to stress conditions, reviewed in (De Buck *et al.*, 2008). In *M. tuberculosis* difficulty in obtaining *tat* mutants indicates that the Tat pathway could be essential under standard laboratory conditions (Saint-Joanis *et al.*, 2006). However, to date only one protein with a direct role in virulence has been confirmed as a Tat substrate, the phospholipase C protein (De Buck *et al.*, 2005; McDonough *et al.*, 2005; Saint-Joanis *et al.*, 2006; Voulhoux *et al.*, 2001). More research must be undertaken to understand just how the Tat pathway affects virulence and which proteins are involved. Nonetheless, studies have so far demonstrated that the Tat pathway is important for virulence and may be a good candidate for the target of novel antimicrobial drugs.

1.3.2 Proofreading and quality control

The Tat pathway is believed to be dedicated to the transport of folded proteins, often with cofactors bound, but how does the cell ensure that proteins are correctly folded and have their cofactor inserted before translocation? The transport of misassembled proteins could interfere with correctly assembled proteins, slow the translocation process down and be costly in terms of resources. The proposed solution is the existence of a proofreading or quality control mechanism, which rejects incorrectly assembled proteins, reviewed in (Sargent, 2007). The existence of such a mechanism is supported by studies demonstrating that the *E. coli* TorA was only exported upon cofactor insertion (Santini *et al.*, 1998), and PhoD, heterologously expressed with a Tat signal peptide, could only be exported when correctly folded (DeLisa *et al.*, 2003).

For a subset of cofactor binding proteins, this quality control is provided by the existence of cytosolic chaperones. This is the case for the widely studied *E. coli* TorA, whose chaperone has been identified as TorD (Ilbert *et al.*, 2003; Jack *et al.*, 2004). This chaperone binds to the signal peptide of TorA, preventing recognition of the twin-arginine motif until the cofactor is correctly inserted and TorA is ready for export. Similar chaperones have been identified for a number of cofactor binding proteins (Oresnik *et al.*, 2001; Potter & Cole, 1999) but these all appear to be substrate-specific and no binding affinity to other Tat substrates has been observed (Hatzixanthis *et al.*, 2005). The existence of a more general quality control mechanism, which recognises hydrophobic stretches, has been proposed (DeLisa *et*

al., 2003). This is supported by research which demonstrated that the export of globular proteins via the Tat pathway could be blocked by the insertion of exposed hydrophobic stretches (Richter *et al.*, 2007).

More recently, the role of the components of the Tat pathway in the rejection and subsequent degradation of unfolded proteins has been demonstrated (Matos *et al.*, 2008; Matos *et al.*, 2009). Mutated NrfC which could not correctly bind cofactors was not exported and was instead quickly degraded in wild-type cells. It was also demonstrated that the *E. coli* Tat pathway components, TatA, TatE and TatD, play a role in the rejection and degradation of the misassembled proteins, discussed in §1.3.4. This indicates that the Tat substrates interact extensively with the components of the Tat pathway before rejection and subsequent degradation.

Relatively little is still known about the exact mechanism of proofreading and quality control. Although some components involved in this process have been identified, it is still not clear how the misassembled proteins are identified. The situation is further complicated by the ability of the thylakoid Tat system to tolerate incorrectly folded protein (Hynds *et al.*, 1998), demonstrating there is still much work to be done.

1.3.3 The Tat signal peptide

Apart from a few specialised systems (e.g. type I secretion), proteins destined for export are synthesised with a signal peptide (as discussed above). Like the Sec signal peptide the Tat signal peptide has a tripartite structure, with a polar N-region, hydrophobic H-region, and polar C-region, Figure 1.3. It typically contains an Ala-Xaa-Ala (A-x-A) cleavage motif, where Xaa is any amino acid, at the end of the C-region where the signal peptide is cleaved off after translocation. Investigations into the secondary structure of the Tat signal peptides indicate that they are largely unstructured in an aqueous environment (Kipping *et al.*, 2003; San Miguel *et al.*, 2003). However, in a membrane mimicking environment the SufI signal sequence was shown to be 40% α-helical (San Miguel *et al.*, 2003). This is similar to the Sec signal peptide whose secondary structure depends on the local environment (Jones *et al.*, 1990).

Figure 1.3 The Tat signal peptide. A typical tripartite signal peptide structure applies to Tat signal peptides; a polar amino domain (N-region), hydrophobic core (H-region), and polar carboxyl domain (C-region). Tat signal peptides are distinguished by the presence of a twin-arginine containing motif at the end of the N-region. They are also typically longer and less hydrophobic than Sec signal peptides. The cleavage motif at the end of the C-region, A-x-A, where the signal peptide is cleaved off to form the mature protein is also shown.

Tat signal peptides contain distinct properties which enable cells to differentiate them from Sec signal peptides. Sequence analysis of Tat substrates revealed the existence of an almost invariant twin-arginine motif in the N-region, which gave rise to the pathway's name (Berks, 1996; Chaddock *et al.*, 1995). Further analysis of bacterial Tat signal peptides revealed a highly conserved seven amino acid motif; T/S-R-R-x-F-L-K, where x is any amino acid (Berks, 1996; Stanley *et al.*, 2000), and the other amino acids appear with a frequency of more than 50% in Gram-negative bacteria.

The RR-motif of the signal peptide is almost invariant in all Tat substrates analysed thus far. Naturally occurring Tat substrates without RR-motifs have, however, been identified – *E. coli* penicillin amidase with an RNR motif (Ignatova *et al.*, 2002), and the TtrB subunit of *Salmonella enterica* tetrathionate reductase with a KR motif (Hinsley *et al.*, 2001). The conservative substitution of a single arginine with lysine does not always block transport of substrates in *E. coli* (Buchanan *et al.*, 2001; DeLisa *et al.*, 2002; Stanley *et al.*, 2000), indicating other factors may be involved. This is supported by evidence that the residues in position -1 and +2, with respect to RR, are also important for translocation by the *E. coli* and *Bacillus subtilis* Tat systems (Mendel *et al.*, 2008). In contrast, substitution of either of the arginines with lysine blocks transport in the thylakoid Tat pathway (Chaddock *et al.*, 1995), demonstrating potential differences between the bacterial and thylakoid Tat pathways. The consensus Tat signal sequence is also different in thylakoids; R-R-x-x-Φ, where Φ is leucine, phenylalanine, valine or methionine (Brink *et al.*, 1998).

As well as the twin-arginine motif, Tat signal peptides also differ from Sec signal peptides in several other ways. Tat signal peptides are typically longer than their Sec counterparts, with an average length of 38 amino acids (Cristobal *et al.*, 1999). Sec signal peptides are also more hydrophobic, and increasing the hydrophobicity of Tat signal peptides can convert them into Sec signals (Cristobal *et al.*, 1999; Ize *et al.*, 2002). Additional basic residues in the C-region (Bogsch *et al.*, 1997; Ize *et al.*, 2002) and a proline in the H-region (Cristobal *et al.*, 1999; Ize *et al.*, 2002) have also been identified as playing a role in Tat signal peptide recognition.

Despite differences in the importance of the RR-motif in the thylakoid and bacterial Tat pathways, there is a high level of conservation of the Tat signal peptide between different species. This has been demonstrated by the ability of Tat signal peptides to direct transport when expressed heterologously in different organisms. Several studies have shown export of bacterial proteins in plants (Halbig *et al.*, 1999; Mori & Cline, 1998; Wexler *et al.*, 1998) and the signal peptide of *E. coli* TorA efficiently directs the export of GFP when expressed in the cyanobacterium *Synechocystis* (Spence *et al.*, 2003). However, not all Tat substrates are transferable as demonstrated by the lack of export of *B. subtilis* PhoD and *Zymomonas mobilis* GFOR proteins in *E. coli*.

Several bioinformatic tools have been developed to aid the identification of Tat substrates (Bendtsen *et al.*, 2005; Dilks *et al.*, 2003; Rose *et al.*, 2002; Taylor *et al.*, 2006). These programs have proved quite accurate in identifying Tat substrates, however the need to confirm predictions experimentally is also important as several proposed Tat substrates in *B. subtilis* have since been demonstrated to be Sec substrates (Jongbloed *et al.*, 2002). Nonetheless, bioinformatics can provide a useful tool in highlighting potential Tat substrates.

1.3.4 Components of the Tat pathway

The first gene to be identified and characterised in the Tat pathway was the maize *hcf106* gene (Settles *et al.*, 1997; Voelker & Barkan, 1995). Analysis of the *E. coli* genome revealed the existence of two homologous genes, *tatA* and *tatB*, and led to the identification of three membrane proteins which are necessary for the function of the Tat pathway; TatA, TatB and TatC (Bogsch *et al.*, 1998; Sargent *et al.*, 1998;

Weiner *et al.*, 1998). In *E. coli*, genes encoding these three proteins form an operon along with a fourth gene, *tatD*, Figure 1.4. Although no role in the Tat translocation process has been identified for TatD (Wexler *et al.*, 2000), recent research has highlighted its role in the quality control mechanism of the Tat pathway (Matos *et al.*, 2009). A fifth monocistronic gene, *tatE*, which is homologous to *tatA* and *tatB* has also been identified. It is believed that TatA and TatE have overlapping function or similar roles, as only deletion of both genes can stop translocation (Sargent *et al.*, 1998). TatA is expressed at much higher levels than TatE (Jack *et al.*, 2001), so it is thought that *tatE* is a cryptic gene duplication of *tatA*, and studies have concentrated on TatA. All of the *E. coli* tat genes are expressed at constitutive levels in a variety of growth conditions (Jack *et al.*, 2001), indicating that the Tat pathway is involved in a number of processes in the cell.

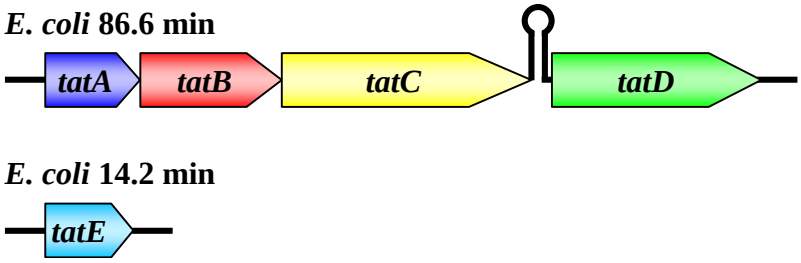


Figure 1.4 Chromosomal organisation of the genes encoding the Tat components in *E. coli*. *tatA*, *tatB* and *tatC*, which are all essential for function, are organised in an operon along with a fourth gene, *tatD*. A putative hairpin structure is seen between *tatC* and *tatD*. *tatE* is monocistronic and is most likely a cryptic gene duplication of *tatA* as the product of the genes are functionally interchangeable. Figure adapted from (Palmer & Berks, 2003).

The Tat pathway is present in many organisms, and homologues of the *E. coli* Tat components have been identified and characterised in a number of different species (Yen *et al.*, 2002). The number of components involved in the Tat pathway can vary between species. Thylakoids and Gram-negative bacteria typically have pathways consisting of two TatA/B/E homologues and one TatC homologue. The majority of Gram-positive bacteria contain pathways which only consist of one TatA/B/E homologue and one TatC homologue, with some, such as *B. subtilis*, containing

multiple Tat pathways that operate in parallel (Dilks *et al.*, 2003; Jongbloed *et al.*, 2006; Yen *et al.*, 2002).

1.3.4.1 TatA

Despite being the smallest of the three proteins involved in the *E. coli* Tat pathway, TatA is the most abundant, with ~25- and ~50-fold higher expression than TatB and TatC, respectively (Jack *et al.*, 2001). TatA has been shown to form large homo-oligomeric complexes and because of this is thought to form the translocation pore of the Tat pathway (Gohlke *et al.*, 2005; Oates *et al.*, 2005). This 89 amino acid membrane protein is predicted to have an N-terminal α-helical TM domain, followed by a short 'hinge' region, an α-helical amphipathic helix (APH) and a largely unstructured C-terminus domain, Figure 1.5 and Figure 1.6. Although predicted to have an N-out topology, the exact orientation of TatA has not been proved conclusively. Protease accessibility experiments indicate that it has an N-out topology (Porcelli *et al.*, 2002), however a recent study probing the accessibility of single cysteine mutants to thiol-reactive components concluded that it has an N-in topology (Chan *et al.*, 2007). This study also suggested a dual topology for the C-terminus of TatA, a hypothesis previously put forward in a study using TEV protease accessibility and reporter fusion constructs (Gouffi *et al.*, 2004). Both studies proposed that the APH of TatA could insert into the membrane as part of the translocation process thus giving the C-terminus a dual topology. Although it has been demonstrated that TatA expressed without its TM domain can still interact with liposomes and phospholipid monolayers (Porcelli *et al.*, 2002), no insertion into lipid bilayers has been demonstrated. Research into the *B. subtilis* TatA homologue, TatAd, demonstrated that whilst the predicted TM domain was able to insert into lipid bilayers, there was no evidence of insertion of the APH (Lange *et al.*, 2007).

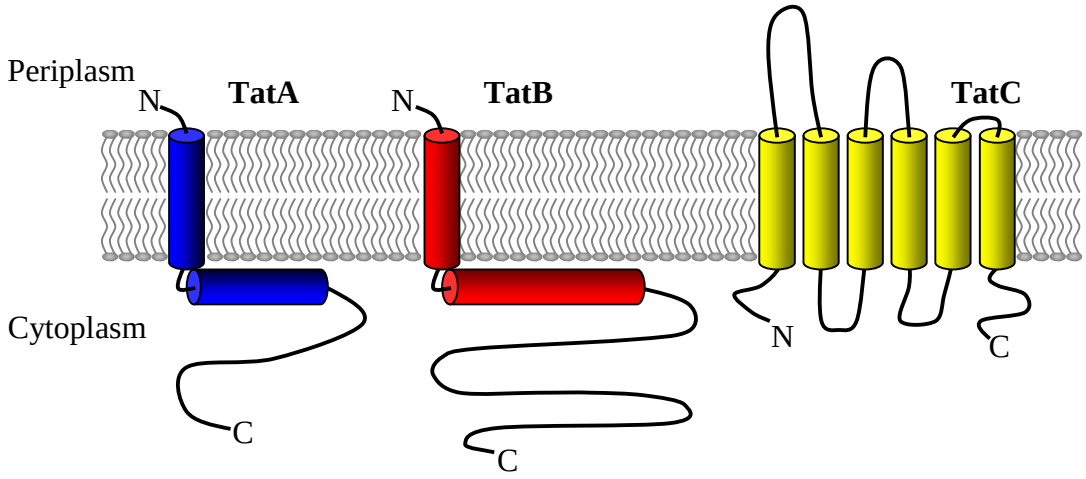


Figure 1.5 Predicted structure and topological arrangement of the *E. coli* Tat components. TatA and TatB are predicted to have a similar secondary structure; an N-terminal α-helical TM domain, followed by a short 'hinge' region, an α-helical amphipathic helix and a largely unstructured C-terminal region. They are both predicted to have an N-out topology. TatC is a six TM domain protein with both its N and C-termini located in the cytoplasm.

A number of mutagenesis studies have been performed on TatA, highlighting the importance of several residues (Barrett *et al.*, 2003a; Barrett *et al.*, 2005; Greene *et al.*, 2007; Hicks *et al.*, 2003; Hicks *et al.*, 2005). In a recent cysteine scanning study, where single amino acids in the TM domain and APH were substituted by cysteine, the majority of the substitutions in the APH lead to loss of function, indicating the sensitivity of the APH to substitution (Greene *et al.*, 2007). Several other loss-of-function substitutions also localise to the APH (Barrett *et al.*, 2005; Hicks *et al.*, 2003; Hicks *et al.*, 2005). The single substitution of a phenylalanine (Phe39), and the triple substitution of three lysines in the APH (Lys37, Lys40, Lys41) both lead to a block of export, Figure 1.6. The Phe39 substitution is dominant and blocks export even in the presence of wild-type TatA (Hicks *et al.*, 2003). The cysteine scanning study also highlighted a glutamine in the TM domain (Gln8), which when substituted with cysteine or alanine blocked function (Greene *et al.*, 2007). A glutamic acid in a similar position of the thylakoid Tha4 TM domain (Glu10) is also important for function (Dabney-Smith *et al.*, 2003).

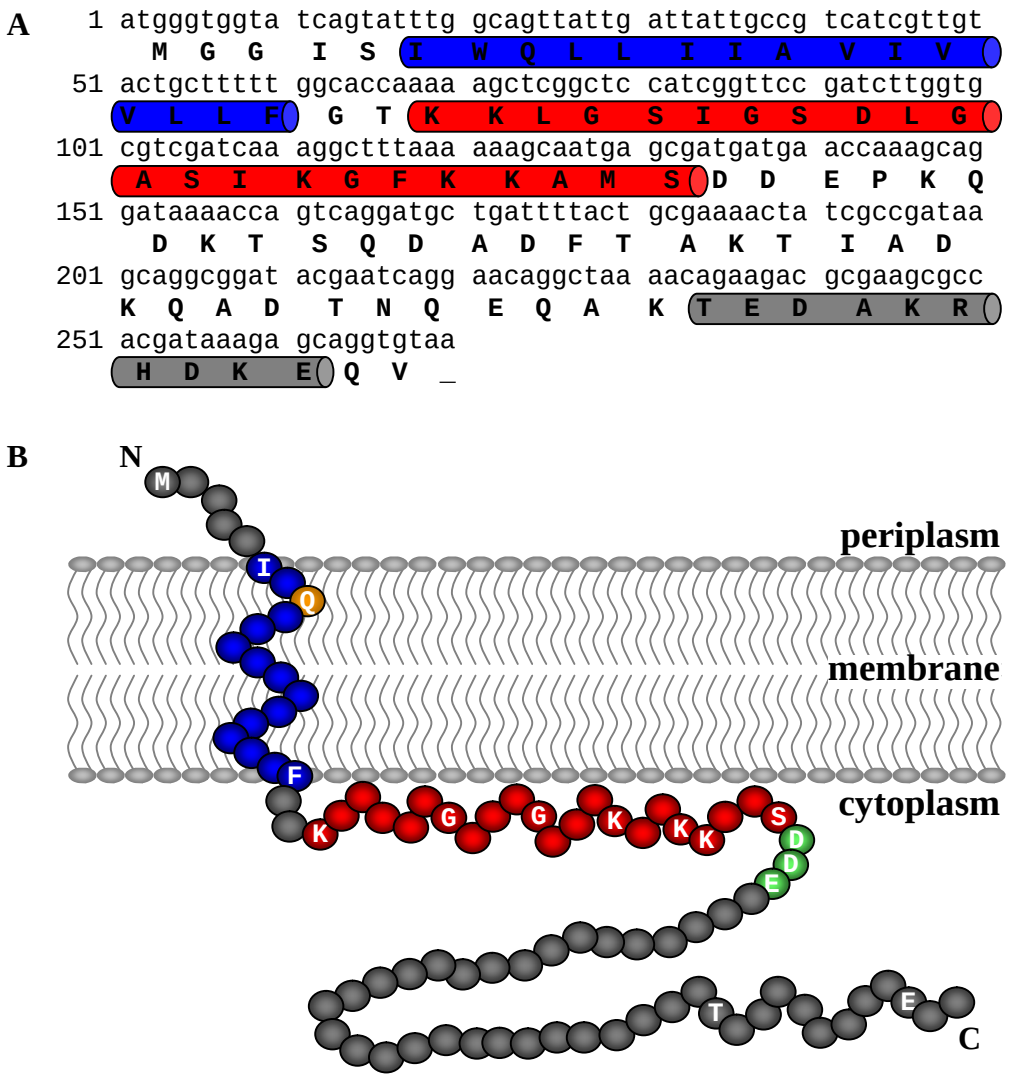


Figure 1.6 The TatA sequence and predicted secondary structure. A. Numbered sequence of *tatA* (lowercase), with the translated sequence of TatA (bold uppercase). The secondary structure predicted by PSIPRED v2.5 is shown with the α-helical domains highlighted; TM domain (blue), APH (red), C-terminal domain (grey). **B.** Schematic of TatA in the membrane with the predicted topology. The TM domain and APH are shown in blue and red, respectively. A glutamine in the TM domain which is discussed in this chapter is shown in orange, an acidic motif identified in Chapter 6 is shown in green, three lysines in the APH which are important for function, and a GxxxG motif are also labelled.

The TM domain and APH are highly conserved in TatA homologues, however the C-terminus varies widely in composition, Appendix A. The TM domain of TatA is essential for translocation and upon its deletion TatA becomes soluble (De Leeuw *et al.*, 2001; Porcelli *et al.*, 2002). Substitution of the TatA TM domain with the homologous TM domain from TatB leads to a loss of function, demonstrating that it

is not just a membrane anchor (Lee *et al.*, 2002). It is thought the TatA TM domain is involved in interactions which drive the formation of TatA complexes, §1.3.4.5 (De Leeuw *et al.*, 2001; Greene *et al.*, 2007; Porcelli *et al.*, 2002). In contrast, the C-terminus of TatA appears to be non-essential for function. Truncations by up to 40 amino acids from the C-terminus of *E. coli* TatA are still able to function (Lee *et al.*, 2002). This was also demonstrated for the thylakoid TatA homologue, Tha4, (Dabney-Smith *et al.*, 2003).

There is some evidence that as well as functioning in the membrane, TatA may be present in the cytoplasm in some bacteria. Soluble TatA homologues have been recovered from the Gram-positive bacteria *Streptomyces lividans* (De Keersmaeker *et al.*, 2005) and *B. subtilis* (Pop *et al.*, 2003). A recent study has also observed the formation of cytoplasmic 'tubes' upon the overexpression of TatA in *E. coli* (Berthelmann *et al.*, 2008). The physiological relevance of these tubes is questionable as *E. coli* TatA has previously been observed only in the membrane (Sargent *et al.*, 2001), where it is resistant to extraction by salt, alkali and urea (De Leeuw *et al.*, 2001). Another recent study demonstrated TatA tagged with yellow fluorescent protein (YFP) and expressed at native levels localises exclusively to the membrane (Leake *et al.*, 2008).

As well as being essential for function, a recent study has indicated that TatA, in conjunction with TatE, might be involved in the degradation of misfolded FeS proteins, rejected by the Tat pathway (Matos *et al.*, 2008). In wild-type cells the misassembled NrfC was quickly degraded after rejection by the Tat pathway, whereas it was stable in a *tatA/E* deletion strain.

1.3.4.2 TatB

E. coli TatB shares ~25% homology with TatA and is predicted to have the same secondary structure, Figure 1.5. At 171 amino acids and 18.4 kDa, TatB is almost double the size of TatA and has a longer APH and C-terminal domain. It is able to interact with TatC to form the translocation receptor of the Tat pathway, §1.3.4.5. Similar to TatA, the C-terminal domain of TatB is not essential for function and up to 70 amino acids can be removed from the C-terminus without affecting function (Lee *et al.*, 2002). However TatB does differ from TatA in a number of ways. Unlike

TatA, the TM domain of TatB is not believed to play a role in function and complex formation (Lee *et al.*, 2002). Although severely compromised, substitution of the TM domain of TatB with the homologous TM domain from TatA does not completely abolish function of the Tat pathway (Lee *et al.*, 2002).

Although sequence alignments reveal that several amino acids are highly conserved in TatB homologues, mutagenesis studies have so far not highlighted any residues as being essential for function (Barrett *et al.*, 2003a; Hicks *et al.*, 2003; Lee *et al.*, 2006). Only when selected substitutions were co-expressed with mutated TatA was any abolishment of function observed, indicating that TatA and TatB interact as part of the translocation process (Barrett & Robinson, 2005).

Due to the similarity of TatA and TatB it is not always possible to distinguish them in other organisms. Although in proteobacteria such as *E. coli* TatB typically has an extended APH and C-terminal domain, this is not always the case. TatA homologues are usually identified by an invariant phenylalanine and glycine in the hinge region (Phe20 and Gly21 in *E. coli* TatA), and a highly conserved phenylalanine in the APH (Phe39 in *E. coli* TatA) (Hicks *et al.*, 2003). Whereas TatB homologues normally contain an invariant glutamic acid in the TM domain (Glu8 in *E. coli* TatB) (Sargent *et al.*, 1999), and a proline after the hinge region glycine (Gly21 and Pro22 in *E. coli* TatB) (Hicks *et al.*, 2003; Weiner *et al.*, 1998). These are not defining features, as thylakoid Tha4 and Hcf106, and two identified TatA/B homologues from the cyanobacterium *Synechocystis*, Slr1046 and Ssl2823, contain a FGP sequence in the hinge region (Aldridge *et al.*, 2008).

In some bacterial systems no TatB homologue is required and the Tat pathway functions as a TatAC system (Dilks *et al.*, 2003; Yen *et al.*, 2002). This has led to the hypothesis that in these systems the TatA/B homologue may be bifunctional and takes the role of both TatA and TatB. This has been demonstrated for the TatAdCd system from *B. subtilis* (Barnett *et al.*, 2008). Expression of TatAd was able to restore function in *E. coli* strains lacking either TatA/E or TatB. For *Synechocystis*, both of the TatA/B homologues, Slr1046 and Ssl2823, could restore function in the absence of TatA/E or TatB in *E. coli* (Aldridge *et al.*, 2008), indicating the additional feature of thylakoid membranes may mean the Tat components have to fulfil multiple roles. In *E. coli* TatA mutants have been isolated which are capable of

complementing strains lacking TatB (Blaudeck *et al.*, 2005). All of the substitutions in the bifunctional TatAs localised to the extreme N-terminal region of TatA, highlighting this region as important for TatB function.

1.3.4.3 TatC

At 28.9 kDa TatC is the largest of the three Tat components in *E. coli*. Its 258 amino acids form six TM domains and insert into the membrane with both the N- and C-termini located in the cytoplasm, Figure 1.5, (Behrendt *et al.*, 2004; Jeong *et al.*, 2004). TatC interacts with TatB to form the receptor complex of the Tat pathway and is widely accepted to be the site of signal peptide recognition (Alami *et al.*, 2003; Cline & Mori, 2001). Several mutagenesis studies of TatC have been performed (Allen *et al.*, 2002; Barrett *et al.*, 2005; Buchanan *et al.*, 2002; Punginelli *et al.*, 2007), resulting in the detection of two residues important for function; a phenylalanine in the second TM domain (Phe94), and a proline in the first periplasmic loop (Pro48) which is also important for the assembly of Tat complexes.

1.3.4.4 TatD

Although *tatD* is in the same operon as *tatABC* it was originally thought that TatD was not involved in the Tat pathway. TatD deletion strains showed no effect on the translocation of Tat substrates and only DNase activity of TatD was identified (Wexler *et al.*, 2000). More recently it has been demonstrated that TatD plays a role in the degradation of misfolded Tat substrates which have been rejected for export (Matos *et al.*, 2009). Mutated FeS proteins, which could not bind their cofactors, were rapidly degraded in wild-type cells. However in *tatD* deletion strains, the mutated proteins were not degraded and remained stable for several hours. The exact role of TatD is still uncertain and it will be important to determine whether this quality control is specific to FeS proteins or part of a more general quality control mechanism.

1.3.4.5 Component organisation

In *E. coli* the three main components of the Tat pathway come together in the cytoplasmic membrane to form two distinct types of complex; a TatABC complex

and separate TatA complexes. The TatABC complex has been purified and shown to contain equal amounts of TatB and TatC, with small and variable amounts of TatA (Bolhuis *et al.*, 2001). This complex was estimated to be ~600 kDa, so clearly contains multiple copies of the Tat components. Complexes purified from a mutant where the C-terminus of TatB was fused to the N-terminus of TatC were also estimated to be ~600 kDa, indicating that TatB and TatC have a 1:1 stoichiometry (Bolhuis *et al.*, 2001). TatABC complexes from other eubacteria have also been purified and are estimated to be of a similar size (Oates *et al.*, 2003). Electron microscopy reveals they form elliptical complexes with a maximum diameter of 13 nm (Oates *et al.*, 2003). In the thylakoid Tat pathway Hcf106 and cpTatC also form a complex of ~700 kDa, however no Tha4 is found associated (Cline & Mori, 2001).

The majority of TatA is found in separate homo-oligomeric complexes that range in size from less than 100 kDa to well over 500 kDa (Gohlke *et al.*, 2005; Leake *et al.*, 2008; Oates *et al.*, 2005). Electron microscopy and random conical tilt reconstructions of purified complexes reveal they form ring-like structures with a lid structure and variable diameter (Gohlke *et al.*, 2005). A recent study using confocal microscopy to image TatA tagged with YFP at native levels observed that these complexes were mobile in the membrane (Leake *et al.*, 2008). Calculation of diffusion rates for different sized complexes and fitting to proposed models for complex conformation confirmed the complexes formed ring-like structures. This has led to the hypothesis that the TatA complexes form the translocation pore through which the Tat substrates are exported, with the heterogeneity of the complexes corresponding to the need to export different sized substrates. Indeed the calculated diameters of the TatA complexes, 30–70 Å (Gohlke *et al.*, 2005), correspond well to the size of *E. coli* Tat substrates (Berks *et al.*, 2000; Sargent *et al.*, 2002). Counter to this hypothesis, it has been shown that TatA heterogeneity is unaffected in export defective TatC mutants (Barrett *et al.*, 2005; Oates *et al.*, 2005).

The formation of such heterogeneous complexes by a single-span membrane protein is highly unusual and the mechanism of complex assembly is still unknown. The TatA complexes vary in size by multiples of 3 to 4 TatA molecules, suggesting trimeric or tetrameric subunit organisation (Leake *et al.*, 2008; Oates *et al.*, 2005). This agrees with cross-linking studies where TatA was shown to form trimers and

tetramers in *E. coli* membranes (De Leeuw *et al.*, 2001). It was recently observed that in the absence of TatB and TatC, TatA did not form large complexes, as determined by *in vivo* confocal imaging (Leake *et al.*, 2008). However, TatA did still form smaller complexes, which were estimated to contain an average of four monomers. It has been proposed that the TM domain of TatA has a key role in oligomerisation (De Leeuw *et al.*, 2001; Greene *et al.*, 2007; Porcelli *et al.*, 2002). When expressed without its TM domain, TatA purifies as a soluble monomer and loses some of its secondary structure (Porcelli *et al.*, 2002). Additionally, a glutamine in the TM domain (Gln8) has been highlighted as important for function (Greene *et al.*, 2007). Polar residues in TM domains have been shown to be involved in the oligomerisation of membrane proteins (Gratkowski *et al.*, 2001; Zhou *et al.*, 2001), as discussed in §1.4. A recent study of the thylakoid Tha4 used cross-linking to show that Tha4 oligomers associate through their TM domains in unstimulated membranes, and formed higher order oligomers when stimulated by the addition of a Tat substrate (Dabney-Smith & Cline, 2009).

1.3.5 Proposed mechanism of translocation

Despite identification of the components and complexes involved in the Tat pathway, the exact translocation mechanism is still unknown. Unlike the Sec pathway, it does not use ATP hydrolysis but instead is driven by the PMF, which is composed of a pH gradient (ΔpH) and an electric field gradient ($\Delta\psi$). After transcription, folding and insertion of any cytosolic cofactors, the signal peptide of the Tat substrate is thought to interact with the membrane (Hou *et al.*, 2006; Shanmugham *et al.*, 2006). Since the signal peptide becomes α-helical in the presence of membranes (San Miguel *et al.*, 2003), it is possible that this interaction enables the signal peptide to take on its correct secondary structure in order to interact with the Tat translocon. The first step in the translocation process is the targeting of the substrate to the Tat translocon, Figure 1.7. This is an energy-independent mechanism (Alami *et al.*, 2002; Musser & Theg, 2000b; Yahr & Wickner, 2001) and, apart from the substrate specific chaperones (§1.3.2), occurs without the assistance of additional factors (Alami *et al.*, 2002; Mould & Robinson, 1991). TatB and TatC in the TatABC complex are believed to form the substrate recognition component of the Tat translocase. The RR-motif of the substrate signal peptide initially interacts with TatC, and then TatB

interacts more extensively with the substrate (Alami *et al.*, 2003; Gerard & Cline, 2006). After recognition of the Tat substrate and binding, free TatA is recruited to the TatABC complex in a PMF-dependent manner (Alami *et al.*, 2003; Dabney-Smith *et al.*, 2006; Mori & Cline, 2002). It is widely believed that the TatA complexes form a translocation pore through which the protein is exported (Gohlke *et al.*, 2005). In this model, the C-terminus of TatA flips into the periplasm, possibly mediated by the insertion of the APH into the membrane (Gouffi *et al.*, 2004), and forms a hydrophilic channel through which the protein can pass, Figure 1.7. After translocation, TatA disassociates from TatBC and the system returns to a resting state.

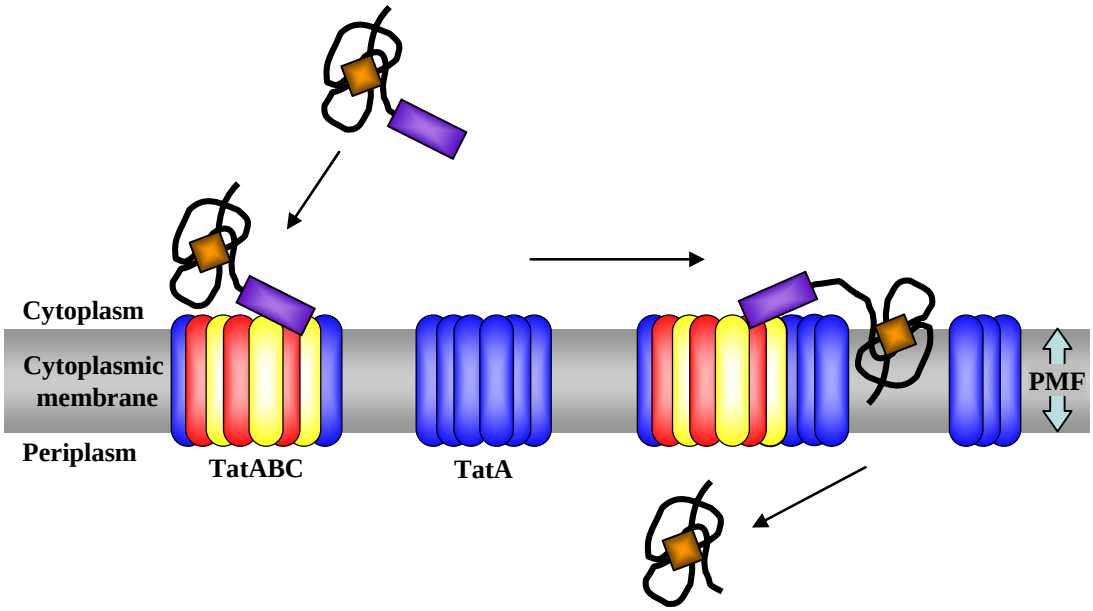


Figure 1.7 Proposed mechanism of transport via the Tat pathway. The signal peptide of the folded Tat substrate, possibly with bound substrate, binds to the TatABC complex via TatC and then TatB. TatA from separate complexes is recruited in a PMF dependent manner. The folded protein is translocated across the membrane, possibly through a pore formed by TatA. Upon translocation the signal peptide is cleaved and the mature protein is released into the periplasm.

The ability of the Tat pathway to translocate proteins using only the PMF is unusual amongst protein translocation systems. Most research has concentrated on the thylakoid Tat pathway, which was originally named the ΔpH pathway due to its reliance on the pH gradient of the PMF (Alder & Theg, 2003b; Mori & Cline, 2002;

Mould & Robinson, 1991). Since only the ΔpH is required, it has been proposed that proton transfer is coupled to protein transport and the counter movement of protons drives translocation (Alder & Theg, 2003a; Musser & Theg, 2000a). It has been calculated that on average 7.9×10^4 protons are released from the gradient per translocated substrate, which is equivalent to 10^4 ATP molecules and is ~3% of the total proton pumping capacity of chloroplasts (Alder & Theg, 2003a). Although this reliance on the ΔpH is widely accepted, one study observed that *in vivo* the thylakoid Tat pathway could be driven by the Δψ (Finazzi *et al.*, 2003). This was also observed in the bacterial Tat pathway, where it was demonstrated that two electric potentials were required at different time points in the translocation process (Bageshwar & Musser, 2007). The requirement for two Δψ suggests translocation is a two-step process. The exact driving force of the Tat pathway may also depend on the organism, as a recent study demonstrated that Tat dependent export of α-amylase in the haloarchaeon *Haloarcula hispanica* is driven by the sodium motive force (Kwan *et al.*, 2008).

1.4 Membrane protein complex formation

It has been demonstrated that two different membrane protein complexes play an important role in the Tat pathway. Understanding the formation of the TatABC and TatA complexes could reveal important structural information and provide insight into the mechanism of transport. The widely accepted model for membrane protein folding and complex formation is known as the two-stage model (Popot & Engelman, 1990; Popot & Engelman, 2000), Figure 1.8. According to this model, membrane proteins fold in two energetically distinct steps. In the first step, the TM domains of membrane proteins insert into the membrane and take on their α-helical secondary structure, which has been demonstrated by the stable insertion and secondary structure formation of TM domains when synthesised individually (Ding *et al.*, 2001; Lazarova *et al.*, 2004). In the second stage lateral helix-helix interactions drive the folding of functional, globular membrane proteins or membrane protein oligomers. This has been demonstrated by the ability of membrane proteins split into two or more fragments to assume their correct conformation (Kahn & Engelman, 1992).

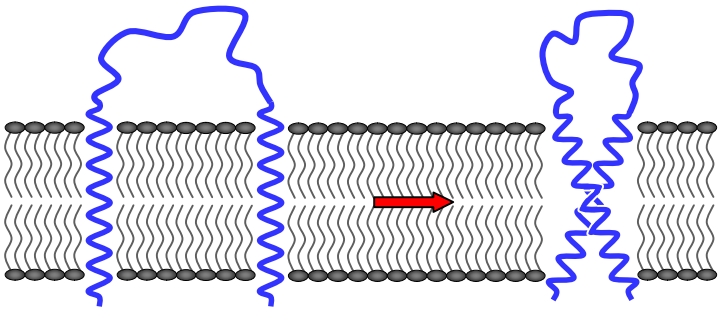


Figure 1.8 The two-stage model for membrane proteins. The TM domains of membrane proteins insert into the membrane and form a stable α-helical secondary structure. Helix-helix interactions then drive the assembly of individual proteins and protein complexes.

Several motifs have been identified that can drive the association of TM domains (Russ & Engelman, 2000; Senes *et al.*, 2000). One of the identified motifs is Gly-Xaa-Xaa-Xaa-Gly (GxxxG), two glycine residues separated by three other amino acids. The glycines in this motif appear on the same face of the helix and allow close packing of helices. The presence of hydrophilic residues in TM domains has also been shown to drive TM helix–helix association via electrostatic interactions and formation of interhelical hydrogen bonds (Dixon *et al.*, 2006; Freeman-Cook *et al.*, 2004; Zhou *et al.*, 2001). The *E. coli* TatA protein contains a glutamine in its TM domain which has been shown to be important for function, however it has not been demonstrated if it is also important for complex formation (Greene *et al.*, 2007). Additionally the APH of TatA contains a GxxxG motif, which may be important for helix packing if the APH inserts into the membrane.

Other factors have also been identified that may play a role in membrane protein complex formation. The insertion and subsequent association of TM domains may provide an environment which facilitates further folding events and stabilisation of complexes (Engelman *et al.*, 2003). The immediate environment of the TM domains is less hydrophobic than the lipid bilayer and this may assist the insertion or stabilisation of further parts of the protein.

1.5 Aims of the work presented in this thesis

The overall aims of the work presented in this thesis were to gain further insight into the specificity of the Tat pathway in *Synechocystis*, and to elucidate the importance of the different structural components of *E. coli* TatA for function and complex formation. In order to realise these aims, work was undertaken to answer the following questions:

- Are there differences in the recognition of Tat substrates by the Tat pathways in *Synechocystis* and *E. coli*?

The importance of the RR-motif in the signal peptide of Tat substrates has been analysed in both bacteria and chloroplasts. To determine if the specificity of the *Synechocystis* Tat pathway resembled that of *E. coli*, transport of a Tat substrate was analysed in both bacteria using an export assay and bioimaging.

- Does the TM domain of *E. coli* TatA play a role in function and complex formation?

A synthetic peptide of the *E. coli* TatA TM domain was analysed using a variety of biophysical techniques to determine its membrane insertion and oligomeric state. The TM domain was also analysed using the TOXCAT assay, which is used to assess interactions between TM domains (Russ & Engelman, 1999). The role of a glutamine residue in the TM domain was also analysed.

- Is the APH of *E. coli* TatA able to insert into membranes and contribute to complex formation?

The ability of the APH of *E. coli* TatA to insert into membranes was assessed using a synthetic peptide and the TOXCAT assay. The oligomeric state of the synthetic peptide was also determined.

- Do parts of the C-terminal domain of *E. coli* TatA contribute to function and complex formation?

Although not essential for function (Lee *et al.*, 2002), the C-terminus of *E. coli* TatA was analysed to identify any important motifs. Its role in the proofreading of Tat substrates was also assessed.

Chapter 2. Materials and Methods

2.1 Suppliers of reagents and chemicals

Reagents used in this study were of the highest grade available and were supplied by Fisher Scientific (UK), Sigma (UK) or VWR (UK), unless otherwise stated.

2.2 Bacterial strains

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

Strain	Genotype	Reference
DH5α	*supE44, ΔlacU169(ϕ80lacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1*	(Sambrook & Russell, 2001)
MC4100	F^- Δ*lacU*169 *araD*139 *rpsL*150 *relA*1 *ptsF rbs flbB*5301	(Casadaban & Cohen, 1979)
Δ*ABCDE*	MC4100; Δ*tatABCDE, Ara*^{r}	(Bolhuis *et al.*, 2000)
Δ*AE*	MC4100; Δ*tatAE, Ara*^{r}	(Bolhuis *et al.*, 2000)
NT326	*F−(argF-lac)U169, rpsL150, relA1, rbsR, flbB5301, ptsF25, thi-1, deoC1, ΔmalE444, recA, srlA*$^+$	(Treptow & Shuman, 1985)

The laboratory strain of *Synechocystis* sp. PCC 6803 (Rippka *et al.*, 1979), which exhibits a spreading morphology and does not utilise glucose, was used for all cyanobacteria work.

2.3 Plasmids

Table 2.2 Previously generated plasmids used in this study.

Plasmid	Details	Reference
pBAD24	Cloning Vector *araC, Amp*^{r}	(Guzman et al., 1992)
pEXT22	Cloning Vector *lacl Kan*^{r}	(Dykxhoorn *et al.*, 1996)
pBAD-ABC	pBAD24, *tatA tatB tatC* StrepII™ tag	(Bolhuis et al., 2001)
pBAD-Ah	pBAD24, *tatA* hexahistidine tag	(Bolhuis et al., 2000)
pJDT1	*araC, Amp*^{r}, TorA-GFPmut3*	(Thomas *et al.*, 2001)
3K>Q	pBAD-ABCs with Lys37, Lys40 and	(Barrett & Robinson,

	Lys41 to Gln substitutions	2005)
pCM1	pBAD24, *nrfC* hexahistidine tag	(Matos *et al.*, 2008)
pCM2	pCM1 with Cys152 and Cys155 to Ala substitutions	(Matos *et al.*, 2008)
pCM3	pCM1 with Cys168 and Cys172 to Ala substitutions	(Matos *et al.*, 2008)
pCC-KAN	New England Biolabs pMAL-c2 and –p2 vectors (pBR322 + *lacI* and Maltose binding protein)	(Russ & Engelman, 1999)
pCC-GpA	pCC-KAN with glycophorin A (GpA) transmembrane (TM) domain	(Russ & Engelman, 1999)
pCC-G83I	pCC-GpA with Gly83 to Ile substitution	(Russ & Engelman, 1999)
pCC-TatA	pCC-Kan with TatA TM from Ile4 to Lys23 inserted between the NheI and BamHI sites	J. Oates (Department of Chemistry, University of Warwick)
psbA3_L38i_TorA-GFP_omega	*Spect*^{r}*/Strep*^{r}, TorA-GFPmut3*	(Spence *et al.*, 2003)

Table 2.3 Plasmids generated in this study. The details of the primers used are given in §2.5.

Plasmid	Template plasmid	Primers used	Features
A-20h	pBAD24	tatA_NcoI_for tatA-20h_PstI_rev	TatA truncated by 20 residues from the C-terminus with a C-terminal his tag
A-40h	pBAD24	tatA_NcoI_for tatA-40h_PstI_rev	TatA truncated by 40 residues from the C-terminus with a C-terminal his tag
A-45h	pBAD24	tatA_NcoI_for tatA-45h_PstI_rev	TatA truncated by 45 residues from the C-terminus with a C-terminal his tag
A-50h	pBAD24	tatA_NcoI_for tatA-50h_PstI_rev	TatA truncated by 50 residues from the C-terminus with a C-terminal his tag
As	pBAD24	tatA_NcoI_for tatAs_PstI_rev	TatA with a C-terminal strep tag

A-40s	pBAD24	tatA_NcoI_for tatA-40s_PstI_rev	TatA truncated by 40 residues from the C-terminus with a C-terminal strep tag
A-45s	pBAD24	tatA_NcoI_for tatA-45s_PstI_rev	TatA truncated by 45 residues from the C-terminus with a C-terminal strep tag
X_A-40h	pEXT22	A_nativeRBS_for tatA-40_XbaI_rev	TatA truncated by 40 residues with a C-terminal his tag
X_A-45h	pEXT22	A_nativeRBS_for tatA-45_XbaI_rev	TatA truncated by 45 residues with a C-terminal his tag
APH-KG	pCC-Kan	tatA_K23_for tatA_G38_rev	TatA APH Lys23 to Gly38
APH-KF	pCC-Kan	tatA_K23_for tatA_F39_rev	TatA APH Lys23 to Phe39
APH-KK	pCC-Kan	tatA_K23_for tatA_K40_rev	TatA APH Lys23 to Lys40
APH-KM	pCC-Kan	tatA_K23_for tatA_M43_rev	TatA APH Lys23 to Met 43
APH-LG	pCC-Kan	tatA_L25_for tatA_G38_rev	TatA APH Leu25 to Gly38
APH-LF	pCC-Kan	tatA_L25_for tatA_F39_rev	TatA APH Leu25 to Phe39
APH-LM	pCC-Kan	tatA_L25_for tatA_K40_rev	TatA APH Leu25 to Lys40
APH-LM	pCC-Kan	tatA_L25_for tatA_M43_rev	TatA APH Leu23 to Met 43

2.4 Affinity tags

A hexahistidine (his) tag or Strep II™ tag (strep) tag (IBA, Germany) was incorporated into constructs, where indicated, to enable detection of proteins.

2.5 Primers

All primers were supplied by Invitrogen (UK). In the two tables below, which list primers used and created for this study, restriction enzyme sites are shown in **bold**, the nucleotides encoding the affinity tag are shown in *italics*, and nucleotides identical to the genomic DNA are underlined.

Table 2.4 Previously designed primers used in this study.

Primer	Sequence 5' to 3'	Source
tatA_NcoI_for	ATA**CCATGG**GTGGTATCAGTATTTG	Claire Barrett (Department of Biological Sciences, University of Warwick)
A_nativeRBS_for	GGAAACAGAATTCGAGCTC**GGTACC**CTACCACAGAGGAACATGTATGGGTGGTATCAGTATTTGGC	Cristina Matos (Department of Biological Sciences, University of Warwick)

Table 2.5 DNA primers used for the preparation of *E. coli* TatA constructs.

Primer	Sequence 5' to 3'
tatA-20h_PstI_rev	TATCCAGT**CTGCAG**CCTCCTTTA*GTGATGGTGATGGTGA TG*CGCCTGCTTATCGGCG
tatA-40h_PstI_rev	TATCCAGT**CTGCAG**CCTCCTTTA*GTGATGGTGATGGTGA TG*CTTTGGTTCATCATCGCT
tatA-45h_PstI_rev	TATCCAGT**CTGCAG**CCTCCTTTA*GTGATGGTGATGGTGA TG*GCTCATTGCTTTTTTAAAGC
tatA-50h_PstI_rev	TATCCAGT**CTGCAG**CCTCCTTTA*GTGATGGTGATGGTGA TG*AAAGCCTTTGATCGACGC
tatAs_PstI_rev	TATCCAGT**CTGCAGCCTCCT**TTA*TTTTTCAAACTGTGGG TGCGACCACGCCGA*CACCTGCTCTTTATC
tatA-40s_PstI_rev	TATCCAGT**CTGCAG**CCTCCTTTA*TTTTTCAAACTGTGGG TGCGACCACGCCGA*CTTTGGTTCATCATC
tatA-45s_PstI_rev	TATCCAGT**CTGCAG**CCTCCTTTA*TTTTTCAAACTGTGGG TGCGACCACGCCGA*GCTCATTGCTTTTTTA

tatA-40_XbaI_rev	TATCCAGT**TCTAGA**TTA*GTGATGGTGATGGTGATG*CTTTGGTTCATCATCGCT
tatA-45_XbaI_rev	TATCCAGT**TCTAGA**TTA*GTGATGGTGATGGTGATG*GCTCATTGCTTTTTTAAAGC
tatA_K23_for	TATCCAGT**GCTAGC**AAAAAGCTCGGCTC
tatA_L25_for	TATCCAGT**GCTAGC**CTCGGCTCCATCG
tatA_G38_rev	TATCCAGT**GGATCC**CGCCTTTGATCGACG
tatA_F39_rev	TATCCAGT**GGATCC**CAAAGCCTTTGATCGA
tatA_K40_rev	TATCCAGT**GGATCC**CTTTAAAGCCTTTGATC
tatA_M43_rev	TATCCAGT**GGATCC**CCATTGCTTTTTTAAAGC

2.6 DNA manipulation and cloning techniques

2.6.1 Preparation of plasmid DNA

Plasmid DNA was isolated using a QIAprep Mini-Prep kit (Qiagen, Germany). The kit was used according to the manufacturer's instructions, plasmid DNA was recovered in 50 µL dH_2O and stored at -20°C.

2.6.2 PCR

PCR was typically performed with 5 µL 10× enzyme buffer, 1 µL 10 mM dNTP mix, 1.5 µL 50 mM $MgCl_2$, 1.25 µL of each primer (from 20 µM stock), 1 µL template DNA (0.5 ng), 0.5 µL *Taq* polymerase (5 units/µL) and dH_2O to a final volume of 50 µL.

The reactions were performed in a Biometra T3 Thermocycler with the following cycling conditions:

94°C	4 min	
94°C	30 sec	35 cycles
52°C	45 sec	
72°C	2 min	
72°C	10 min	

2.6.3 Agarose gel electrophoresis

Agarose gels were prepared by dissolving 1–2% (w/v) agarose in TBE buffer (89 mM Tris; 89 mM H_3BO_3; 20 mM EDTA, pH 8.0). Ethidium bromide was added (0.5 μg/mL final concentration) to allow visualisation of DNA. DNA samples were prepared by adding 2 μL of loading dye (0.25% bromophenol blue, 40% w/v sucrose). Gels were submerged in TBE buffer and electrophoresis was performed at 100–150 V until the dye front was at least half way through the gel. DNA was visualised using a UV transilluminator and photographs of gels were taken on a gel documentation system.

2.6.4 Purification of DNA by gel extraction

DNA was excised from agarose gels and extracted using a QIAprep Gel Extraction kit (Qiagen, Germany), for large DNA fragments, or a QIAprep Minelute kit (Qiagen, Germany), for small DNA fragments. Kits were used according to the manufacturer's instructions and DNA was recovered in 30 or 10 μL dH_2O respectively.

2.6.5 Restriction endonuclease digestion of DNA

Restriction endonuclease digestion of DNA was performed according to the enzyme manufacturer's directions. Reactions were typically incubated at 37°C for 2–3 hrs, before being subjected to gel electrophoresis and purified as detailed in §2.6.4.

2.6.6 Ligation of DNA fragments

Ligation reactions were performed using T4 DNA ligase according to the manufacturer's instructions. Vector and insert DNA fragments were mixed in a 1:3 ratio and typically incubated at room temperature for 3 hours or over night at 4ºC.

2.6.7 Sequencing of plasmid DNA

Sequencing was performed by the Molecular Biology Service, University of Warwick, using Big Dye Terminator v 3.1 Chemistry and run on a 3130xl Genetic Analyser (Applied Biosystems).

2.6.8 Sequencing primers used in this study

Table 2.6 Sequencing primers used in this study.

Primer Name	Sequencing Target	Sequence (5' to 3')	Source
TatA	*tatA* in pBAD-ABCs	ACCTGACGCTTTTTATCGCAA	C. Barrett (Department of Biological Sciences, University of Warwick)
TatB	*tatB*	AAGCAGGCGGATACGAATCA	C. Barrett (Department of Biological Sciences, University of Warwick)
SEQ1	5' end of *tatC*	AACCTGCTGCGGACGCTGAA	C. Barrett (Department of Biological Sciences, University of Warwick)
SEQ5	3' end of *tatC*	ATCGGCATGGCATTCGCCTAC	C. Barrett (Department of Biological Sciences, University of Warwick)
TatA for pEXT	*tatA* in pEXT-ABC	TCGGCTCGTATAATG	C. Barrett (Department of Biological Sciences, University of Warwick)
pEXT22 SEQ_F	pEXT22	CGCTCAAGGCGCACTCCCG	C. Matos (Department of Biological Sciences, University of Warwick)
seq_pJDT1_TorA	pJDT1	GGCTAGCAGGAGGAATTCAC	A. Nenninger (Department of Biological Sciences, University of Warwick)
pCC-for	TM insert in pCC-KAN	CCTTCATCAGCCACTGTAGTGAAC	J. Oates (Department of Chemistry, University of Warwick)
pCC-rev	TM insert in pCC-KAN	CAGTTCAGCGAGACCGTTATAG	J. Oates (Department of Chemistry, University of Warwick)

2.6.9 Mutagenesis of plasmid DNA

Point mutations and deletions were introduced into plasmid DNA using *Pfu* DNA polymerase and a QuikChange Site-Directed Mutagenesis kit (Stratagene, USA), according to the manufacturer's instructions. The amplified plasmid DNA was transformed into supercompetent DH5α cells (§2.7.5)

2.6.10 Primers used for mutagenesis of plasmid DNA

Primers were designed to be used with the QuikChange Site-Directed Mutagenesis kit, according to the manufacturer's instructions. Mutations are underlined.

Table 2.7 Primers used for mutagenesis of plasmid DNA.

Construct Generated	Template	Primer	Sequence (5' to 3')	Substitution/ Deletion
RRK	pJDT1	for	CGATCTCTTTCAGGCATCACGT CGGAAGTTTCTGGCACAACTCG GCG	Arg18 to Lys
		rev	CGCCGAGTTGTGCCAGAAACTT CCGACGTGATGCCTGAAAGAGA TCG	
RKR	pJDT1	for	CGATCTCTTTCAGGCATCACGT AAGCGGCGTTTTCTGGCACAAC TCGGCG	Arg17 to Lys
		rev	CGCCGAGTTGTGCCAGAAAACG CTTACGTGATGCCTGAAAGAGA TCG	
KRR	pJDT1	for	CGATCTCTTTCAGGCATCAAAG CGGCGTTTTCTGGCACAACTCG GCG	Arg16 to Lys
		rev	CGCCGAGTTGTGCCAGAAAACG CCGCTTTGATGCCTGAAAGAGA TCG	
RKK	pJDT1	for	CGATCTCTTTCAGGCATCACGT AAGAAGTTTCTGGCACAACTCG GCG	Arg17 and Arg18 to Lys
		rev	CGCCGAGTTGTGCCAGAAACTT CTTACGTGATGCCTGAAAGAGA TCG	
KRK	pJDT1	for	CGATCTCTTTCAGGCATCAAAG CGGAAGTTTCTGGCACAACTCG GCG	Arg16 and Arg18 to Lys
		rev	CGCCGAGTTGTGCCAGAAACTT CCGCTTTGATGCCTGAAAGAGA TCG	

KKR	pJDT1	for	CGATCTCTTTCAGGCATCAAAGAAGCGTTTTCTGGCACAACTCGGCG	Arg16 and Arg17 to Lys
		rev	CGCCGAGTTGTGCCAGAAAACGCTTCTTTGATGCCTGAAAGAGATCG	
KKK	pJDT1	for	CGATCTCTTTCAGGCATCAAAGAAGAAGTTTCTGGCACAACTCGGCG	Arg16, Arg17 and Arg18 to Lys
		rev	CGCCGAGTTGTGCCAGAAACTTCTTCTTTGATGCCTGAAAGAGATCG	
D45N	pBAD-ABCs	for	GGCTTTAAAAAAGCAATGAGCAATGATGAACCAAAGCAGGATA	Asp45 to Asn
		rev	TATCCTGCTTTGGTTCATCATTGCTCATTGCTTTTTTAAAGCC	
D46N	pBAD-ABCs	for	GCTTTAAAAAAGCAATGAGCGATAATGAACCAAAGCAGGATAAAAC	Asp45 to Asn
		rev	GTTTTATCCTGCTTTGGTTCATTATCGCTCATTGCTTTTTTAAAGC	
E47Q	pBAD-ABCs	for	CTTTAAAAAAGCAATGAGCGATGATCAGCCAAAGCAGGATAAAACCAGT	Glu47 to Gln
		rev	ACTGGTTTTATCCTGCTTTGGCTGATCATCGCTCATTGCTTTTTTAAAG	
NNQ	pBAD-ABCs	for	GATCAAAGGCTTTAAAAAAGCAATGAGCAATAATCAGCCAAAGCAGGATAAAACCAGTCAGGATG	Asp45 and Asp46 to Asn, Glu47 to Gln
		rev	CATCCTGACTGGTTTTATCCTGCTTTGGCTGATTATTGCTCATTGCTTTTTTAAAGCCTTTGATC	

LLM	NNQ	for	GATCAAAGGCTTTAAAAAAGCA ATGAGCCTTCTTATGCCAAAGC AGGATAAAACCAGTCAGGATG	Asp45 and Asp46 to Leu, Glu47 to Met
		rev	CATCCTGACTGGTTTTATCCTG CTTTGGCATAAGAAGGCTCATT GCTTTTTTAAAGCCTTTGATC	
Q8A	pBAD-ABCs	for	GGGTGGTATCAGTATTTGGGCG TTATTGATTATTGCCGTC	Gln8 to Ala
		rev	GACGGCAATAATCAATAACGCC CAAATACTGATACCACCC	
Q8C	pBAD-ABCs	for	CATGGGTGGTATCAGTATTTGG TGCTTATTGATTATTGCCGTCA TCG	Gln8 to Cys
		rev	CGATGACGGCAATAATCAATAA GCACCAAATACTGATACCACCC ATG	
3K>Q + NNQ	NNQ	for	GGTGCGTCGATCCAAGGCTTTC AACAAGCAATGAG	Lys37, Lys40 and Lys41 to Gln, Asp45 and Asp46 to Asn, Glu47 to Gln
		rev	CTCATTGCTTGTTGAAAGCCTT GGATCGACGCACC	
3K>Q + LLM	LLM	for	GGTGCGTCGATCCAAGGCTTTC AACAAGCAATGAG	Lys37, Lys40 and Lys41 to Gln, Asp45 and Asp46 to Leu, Glu47 to Met
		rev	CTCATTGCTTGTTGAAAGCCTT GGATCGACGCACC	
TC-Q8A	pCC-TatA	for	GAGCTAGCATCAGTATTTGGGC GTTATTGATTATTGCCGTC	TatA Gln8 to Ala
		rev	GACGGCAATAATCAATAACGCC CAAATACTGATGCTAGCTC	
TC-Q8C	pCC-TatA	for	CGAGCTAGCATCAGTATTTGGT GCTTATTGATTATTGCCGTCAT CG	TatA Gln8 to Cys
		rev	CGATGACGGCAATAATCAATAA GCACCAAATACTGATGCTAGCT	

			CG	
TatA-19	pCC-TatA	for	CTGCTTTTTGGCACCGGGATCCTGATCAAC	Deletion of Lys23 in TatA TM of pCC-TatA
		rev	GTTGATCAGGATCCCGGTGCCAAAAAGCAG	
TatA-18	pCC-TatA	for	GTACTGCTTTTTGGCGGGATCCTGATCAAC	Deletion of Thr22 and Lys23 in TatA TM of pCC-TatA
		rev	GTTGATCAGGATCCCGCCAAAAAGCAGTAC	
TatA-17	pCC-TatA	for	CGTTGTACTGCTTTTTGGGATCCTGATCAACC	Deletion of Gly21, Thr22 and Lys23 in TatA TM of pCC-TatA
		rev	GGTTGATCAGGATCCCAAAAAGCAGTACAACG	
TatA-16	pCC-TatA	for	CATCGTTGTACTGCTTGGGATCCTGATCAAC	Deletion of Phe20, Gly21, Thr22 and Lys23 in TatA TM of pCC-TatA
		rev	GTTGATCAGGATCCCAAGCAGTACAACGATG	

2.7 Growth and maintenance of *E. coli*

2.7.1 Media

Luria-Bertami (LB) medium (10 g/L bacto-tryptone, 5 g/L yeast extract, 10 g/L sodium chloride) was used for the aerobic growth of *E. coli* in liquid media. Antibiotics were added, when necessary, to the following concentrations: Ampicillin 100 μg/mL, Kanamycin 50 μg/mL. For the growth of NT326 *E. coli* cells containing derivatives of pCC-KAN, Ampicillin was added to a final concentration of 200 μg/mL. LB medium with 16% (w/v) agar was used for the aerobic growth of *E. coli* on plates. Anaerobic media (LB supplemented with 0.5% glycerol, 0.4% Trimethylamine-N-oxide (TMAO) and 1 μM sodium molybdate) was used for the anaerobic growth of *E. coli*.

2.7.2 Maintenance

For the long term storage of *E. coli* strains and constructs generated in this study, glycerol stocks were prepared by adding 400 µL 50% glycerol to 800 µL of an overnight culture. These were frozen immediately on dry ice and stored at -80ºC.

2.7.3 Preparation of competent *E. coli* cells

A starter culture was prepared by inoculating 5 mL LB, containing antibiotics where necessary, with *E. coli* cells and incubating overnight at 37°C with shaking at 250 rpm. 10 mL LB was inoculated with the starter culture and cells were incubated at 37°C with shaking at 250 rpm until OD_{600} = 0.3–0.4. Cells were pelleted by centrifugation at 1700×g (3000 rpm, Jouan GR422) at 4°C for 10 min, resuspended in 10 mL ice cold 100 mM $MgCl_2$, and incubated on ice for 5 min. Cells were re-pelleted, resuspended in 1 mL ice-cold 100 mM $CaCl_2$ and incubated on ice for 2–24 hours before use.

2.7.4 Preparation of super competent *E. coli* cells

A starter culture was prepared by inoculating 5 mL LB with *E. coli* DH5α cells and incubating overnight at 37°C with shaking at 250 rpm. 100 mL LB was inoculated with the starter culture and cells were incubated at 25°C with shaking at 250 rpm until OD_{600} = 0.3–0.4. Cells were incubated on ice for 30 min before being pelleted by centrifugation at 1700×g, resuspended in 50 mL ice cold 50 mM $CaCl_2$ and incubated on ice for 1 hour. Cells were pelleted, resuspended in 3 mL ice cold 50 mM $CaCl_2$ plus 20% glycerol, split into 100 µL aliquots and frozen immediately on dry ice. Aliquots were stored at -80°C until needed.

2.7.5 Transformation of competent *E. coli* cells

100 µL competent cells was mixed with 1–5 ng DNA in 100 µL TCM (10 mM Tris-HCl pH 7.5, 10 mM $CaCl_2$, 10 mM $MgCl_2$) and incubated on ice for 30 min. Cells were incubated at room temperature for 10 min before adding 500 µL pre-warmed LB and incubating at 37°C with shaking at 250 rpm for 1 hour. Cells were harvested by centrifugation at 1700×g, 550 µL supernatant was removed, the pellet was

resuspended in the remaining liquid and plated on LB-agar plates containing the appropriate antibiotics. Plates were incubated overnight at 37°C.

2.8 Growth and maintenance of *Synechocystis*

2.8.1 Media

BG11 medium (Rippka *et al.*, 1979) was used for the growth of *Synechocystis* in liquid media. Spectinomycin was added to a final concentration of 50 μg/mL where needed. Cultures were grown under normal growth illumination (30 $\mu E/m^2/s$) at 30ºC with constant shaking at 150 rpm. 16% (w/v) agar was used for the aerobic growth of *Synechocystis* on plates.

2.8.2 Transformation of *Synechocystis*

The transformation protocol was based on (Porter, 1988). An exponentially growing culture was resuspended to give a final cell concentration of 10^9 cell/mL. 10 μL plasmid DNA was added to 150 μL culture and incubated under normal growth illumination at 30ºC for 1 hour. Cells were plated on BG11 plates without antibiotic and incubated under illumination for 16–18 hours before antibiotic was added under the agar to the required concentration. Transformants appeared after 7–10 days and were restreaked onto selective plates at least three times to ensure segregation.

2.9 Protein preparation

2.9.1 Plasmid induction

Typically the pBAD24 based constructs were induced with varying amounts of arabinose (0–200 μM) for 3–4 hours. For TorA-GFP export assays, *E. coli* Δ*ABCDE* contained pEXT-ABC and pJDT1. The pJDT1 plasmid was induced for 2 hours with 200 μM arabinose. Cells were then washed three times by harvesting cells and resuspending them in fresh LB medium without inducer. The pEXT-ABC plasmid was then induced for 2 hours with 1 mM IPTG.

2.9.2 Preparation of whole cell fractions

Whole cell fractions were prepared by harvesting 1 mL of a mid-exponential phase culture and resuspending in 100 µL SDS PAGE sample buffer (§2.10.1). Samples were incubated at 50ºC for 10 min before analysis by SDS PAGE (§2.10.1) or frozen on dry ice and stored at -80ºC until required.

2.9.3 Fractionation of *E. coli* cells

Cells were separated into periplasmic, cytoplasmic and membrane fractions using a procedure based on the EDTA/lysozyme/cold osmotic shock method (Randall & Hardy, 1986). Typically cells were harvested and resuspended in 1 mL chilled buffer I (100 mM tris-acetate pH 8.2; 0.5 M sucrose; 5 mM EDTA). 40 µL lysozyme (2 mg/mL) and 500 µL dH_2O was added before incubation on ice for 5 min followed by the addition of 20 µL $MgSO_4$. The spheroplasts were pelleted by centrifugation at 20,800×g (14,000 rpm Eppendorf 5417R) and the supernatant was collected as the periplasmic fraction. Spheroplasts were washed in 1 mL chilled buffer II (50 mM tris-acetate pH 8.2; 0.25 mM sucrose; 10 mM $MgSO_4$) and pelleted by centrifugation. The supernatant was discarded and the spheroplasts were resuspended in 1 mL chilled buffer III (50 mM tris-acetate pH 8.2; 2.5 mM EDTA). Spheroplasts were lysed by sonication (3× 10 sec) and membranes were separated from the cytoplasmic fraction by centrifugation at 265,000×g (70,000 rpm, Beckman TL100, TLA100.3 rotor) for 30 min at 4°C. The membranes formed a pellet and the supernatant was collected as the cytoplasmic fraction.

2.9.4 Solubilisation of *E. coli* membranes with detergent

Membranes were solubilised by resuspending in 500 µL detergent containing buffer using a syringe and needle. For blue-native PAGE (§2.10.3) membranes were solubilised in BN PAGE solubilisation buffer (50 mM bis-tris-HCl pH 7.0; 750 mM aminocaproic acid; 2% digitonin). For fast protein liquid chromatography (FPLC) membranes were solubilised in FPLC solubilisation buffer (20 mM Trizma-HCl pH 8.0, 10% glycerol, 150 mM NaCl, 2% digitonin). Membranes were incubated at 4ºC with constant rotation for 16 hours. Insoluble material was removed by centrifugation at 265,000×g (70,000 rpm) for 15 min at 4°C.

2.9.5 Trichloroacetic acid (TCA) precipitation

Protein samples were concentrated for gel electrophoresis by TCA precipitation. 1 mL TCA was added and samples were incubated on ice for 30 min before centrifugation for 15 min (12,100×g, 13,400 rpm, Eppendorf MiniSpin). The supernatant was removed and 1 mL acetone was added and incubated on ice for 5 min before centrifugation for 10 min. The supernatant was removed and samples were allowed to air dry at room temperature. Typically samples were resuspended in 80 µL SDS sample loading buffer (§2.10.1).

2.10 Protein resolution

2.10.1 SDS poly-acrylamide gel electrophoresis

SDS polyacrylamide gels were cast and run on a vertical gel electrophoresis system (CBS) according to the manufacturer's instructions, based on the method described by (Laemmli, 1970). Typically 0.75 mm gels were prepared with a separating gel (17.5% acrylogel solution (acrylogel 2.6 (40%) solution (BDH)); 375 mM Tris-HCl pH 8.8; 0.1% SDS; 0.02% APS; 0.06% TEMED) and a stacking gel (5% acrylogel solution; 125 mM Tris-HCl pH 6.8; 0.1% SDS; 0.6% APS; 0.06% TEMED).

Samples were prepared by mixing with SDS sample loading buffer (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.02% bromophenol blue, 5% β-mercaptoethanol) and boiling for 10 min. Samples to be immunoblotted with Tat antibodies were heated to 50ºC for 10 min.

A uniform running buffer (25 mM tris; 250 mM glycine; 0.1% SDS) was used for electrophoresis. Gels were run until the dye front had migrated off the gel, typically 3 hours at 35 mA or 16 hours overnight at 3 mA.

Peptides were separated on NuPAGE® 12% Bis-Tris gels using the XCell SureLock® mini-cell system and NuPAGE® MES SDS running buffer (Invitrogen), according to the manufacturer's instructions.

2.10.2 Native poly-acrylamide gel electrophoresis

Native PAGE was performed as SDS PAGE (§2.10.1), with the same equipment and solutions but without the inclusion of SDS. The gels were 10% separating gels with a 5% stacking gel.

2.10.3 Blue-native poly-acrylamide gel electrophoresis

The blue-native (BN) PAGE protocol was based on (Schägger & von Jagow, 1991) using the same equipment as SDS-PAGE. A 1.5 mm gradient gel was prepared by combining a 5% acrylamide solution (5% acrylogel; 50 mM Bis-Tris-HCl pH 7.0 (at 4ºC); 500 mM aminocaproic acid; 0.76% glycerol; 4% APS; 0.4% TEMED) with a 13% acrylamide solution (13% acrylogel; 50 mM Bis-Tris-HCl pH 7.0 (at 4ºC); 500 mM aminocaproic acid; 5% glycerol; 0.35% APS; 0.035% TEMED) in a gradient maker. A stacking gel (4% acrylogel; 50 mM Bis-Tris-HCl pH 7.0 (at 4°C), 0.5 mM aminocaproic acid; 0.7% APS; 0.07% TEMED) was added. Gradient gels were equilibrated at 4ºC before loading samples. Typically 1 µL of sample buffer (760 mM aminocaproic acid, 50 µM Bis-Tris-HCl pH 7.0, 5% Serva G Coomassie blue) was added to 10 µL of solubilised membrane.

Gels were initially run at 70 V at 4ºC with coloured cathode buffer (50 mM Tricine pH 7.0, 15 mM Bis-Tris-HCl pH 7.0 (at 4°C), 0.02% Serva G Coomassie blue) and anode buffer (50 mM Bis-Tris-HCl pH 7.0 (at 4°C)). When the gel front had migrated half way through the gel the coloured cathode buffer was replaced with clear cathode buffer (50 mM Tricine pH 7.0, 15 mM Bis-Tris-HCl pH 7.0 (at 4°C)) and the gel was run for 16 hours overnight.

After electrophoresis the gel was equilibrated for 15 min in clear cathode buffer containing 0.05% SDS before transfer to PVDF membrane.

2.10.4 Fast protein liquid chromatography (FPLC)

Typically membranes were isolated from a 50 mL *E. coli* culture and solubilised in FPLC solubilisation buffer (§2.9.4). Clarified membranes were loaded onto a Superose 6HR 10/30 gel filtration column (Amersham Pharmacia Biotech) and were eluted with FPLC wash buffer (20 mM Trizma-HCl pH 8.0; 150 mM NaCl, 10%

glycerol; 0.1% digitonin). 600 µL samples were collected and either analysed immediately or stored at -20ºC until required.

2.11 Protein detection

2.11.1 Coomassie staining

After electrophoresis, the gel was placed in fixer (50% methanol, 10% acetic acid) for 30 min, then in stain solution (10% acetic acid, 0.025% Coomassie G250) for 1 hour, followed by de-stain (20% methanol, 7% acetic acid) until the protein bands were visible. The gels were dried for long term storage.

2.11.2 Silver staining

After electrophoresis the gel was placed in fixer (50% acetone, 1.25% TCA, 0.015% formaldehyde) for 15 min, followed by washing in dH_2O (3× quick washes, 1× 5 min, 3× quick washes). The gel was then placed in 50% acetone for 5 min, followed by incubation in sodium thiosulphate solution (0.017% (w/v)) for 1 min. The gel was washed in dH_2O (3× quick washes) before being placed in the stain solution (0.26% (w/v) silver nitrate; 0.37% formaldehyde) for 8 min. The gel was washed in dH_2O (2× quick washes) and then placed in developer (0.2 mM sodium carbonate; 0.004% (w/v) sodium thiosulphate; 0.015% formaldehyde) until protein bands could be visualised, typically 20–30 sec. Stopping solution (1% acetic acid) was then added for 1–2 min. The gel was washed in dH_2O and dried for long term storage.

2.11.3 Transfer of proteins to PVDF membrane

Proteins were transferred from an acrylamide gel to PVDF membrane (Amersham Biosciences, UK) using a semi-dry Western blotting system (Sigma Aldrich, UK) with Towbin transfer buffer (Towbin *et al.*, 1979). The PVDF membrane was prepared by soaking in methanol and then in Towbin transfer buffer. Transfer was carried out by applying a constant current of 200 mA for 2 hours.

For BN PAGE, the PVDF membrane was soaked in methanol for 5 min, dH_2O for 5 min and then clear cathode buffer (§2.10.3) with 0.05% SDS for 5 min. Proteins were transferred from BN gels using the same apparatus as described above but with

clear cathode buffer with 0.05% SDS. PVDF membranes were destained after transfer by incubating in 25% methanol, 10% acetic acid with gentle agitation for 1 hour. The membranes were then washed with phosphate-buffered saline (PBS, 137 mM NaCl; 2.7 mM KCl; 10 mM Na_2HPO_4; 2 mM KH_2PO_4) containing 0.1% Tween 20 (PBS-T).

2.11.4 Detection of proteins by immunoblotting

After transfer, PVDF membranes to be immunoblotted with antibodies to the Strep II™ tag were blocked with PBS-T containing 3% BSA for at least 1 hour. The membranes were washed in PBS-T (typically 1× 15 min, 2× 5 min) before incubation in PBS-T with 6 μg/mL avidin for 10 min. The Streptactin-horseradish peroxidise (HRP) conjugate antibody (IBA) was added directly to the avidin solution according to the manufacturer's instructions and incubated, with agitation, for 2 hours.

PVDF membranes to be immunoblotted with other antibodies were blocked with PBS-T containing 5% (w/v) dried skimmed milk powder for at least 1 hour. The membranes were washed in PBS-T before incubation with PBS-T containing the desired primary antibody for 1 hour. The membranes were washed before incubation with the secondary antibody for 1 hour. The membranes were washed and immuno-reactive bands were detected using the ECL (enhanced chemiluminescence) kit (Amersham Biosciences) according to the manufacturer's instructions. X-ray films (super RX film, Fujifilm) were developed on an AGFA Curix 60 automatic developer according to the manufacturer's instructions.

Table 2.8 Antibodies used in this study.

Antibody	Concentration	Manufacturer/Source
anti-TatA	1 in 2000	Laboratory stock (Bolhuis *et al.*, 2000)
anti-TatB	1 in 1000	Laboratory stock (Bolhuis *et al.*, 2000)
anti-Maltose binding protein	1 in 2000	Sigma
Streptactin-HRP conjugate	1 in 12000	IBA

anti-hexahistidine(C-term)	1 in 5000	Invitrogen
anti-mouse IgG (H+L), HRP conjugate	1 in 5000	Promega
anti-rabbit IgG (H+L), HRP conjugate	1 in 10000	Promega

2.11.5 TMAO reductase (TorA) activity assay

The presence of TMAO reductase (TorA) was detected using a protocol adapted from (Silvestro *et al.,* 1989). Periplasmic and cytoplasmic fractions were separated by native PAGE (§2.10.2). Immediately after electrophoresis, gels were placed in 100 mM phosphate buffer pH 6.5 (saturated with nitrogen gas) with 0.25% (w/v) methyl viologen. The gel was stained blue by adding freshly prepared sodium dithonite (0.1% in 10 mM NaOH) and incubating for 15-30 min. Gels were then transferred to 100 mM phosphate buffer pH 6.5 (saturated with nitrogen gas) with 40 mM TMAO until white bands appeared. The gels were scanned as the bands developed.

2.12 TOXCAT methods

The TOXCAT assay was used to investigate the transmembrane (TM) domain and amphipathic helix (APH) of TatA in a natural membrane as described previously (Russ & Engelman, 1999). The TOXCAT assay is a system used to measure the ability of a given TM domain, in isolation from the rest of the protein, to self-associate in a natural membrane environment. A chimeric protein consisting of the α-helical TM domain of interest inserted between the N-terminal dimerisation-dependent DNA binding domain of ToxR, and the monomeric periplasmic anchor maltose binding protein (MBP) is used, Figure 2.1. The chimera is expressed in *E. coli* NT326 cells along with a chloramphenicol acetyltransferase (CAT) reporter gene under the control of the ToxR-responsive *ctx* promoter. If the TM domains interact, the ToxR domain is able to dimerise and activate the *ctx* promoter, leading to the expression of CAT. The level of CAT expression in the cells expressing the chimera is proportional to the strength of oligomerisation, and can be measured quantitatively.

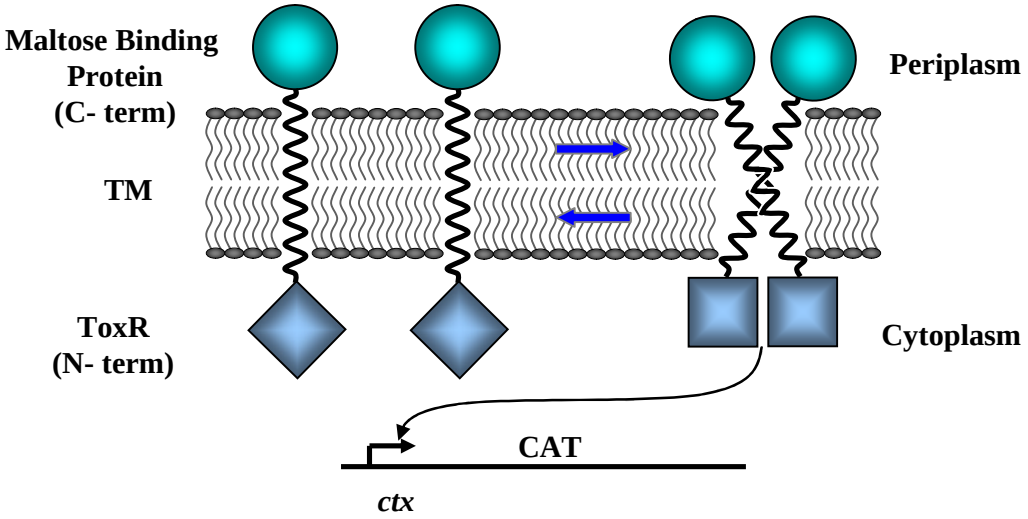


Figure 2.1 The TOXCAT assay for TM domain oligomerisation. The TOXCAT chimera, consisting of the TM domain inserted between the dimerisation-dependent DNA binding domain of ToxR (◇) and the periplasmic anchor MBP (●), is expressed in *E. coli* NT326 cells. TM domain interactions enable the ToxR domain to dimerise (▭) and activate the expression of chloramphenicol acetyltransferase (CAT) through the *ctx* promoter. Figure adapted from (Russ & Engelman, 1999).

Plasmids expressing TOXCAT chimera with different inserts were constructed, Table 2.3, and transformed into *E. coli* NT326 cells. Checks were performed to ensure the chimera was expressing and inserting into the membrane with the correct orientation. The TM domain of Glycophorin A (GpA), which is know to strongly dimerise, and its dimerisation defective mutant (G83I), were used as a positive and negative controls respectively.

2.12.1 Chimera expression, insertion and orientation checks

2.12.1.1 Expression check

Cells expressing the chimera were grown until mid-exponential stage. Cell density was normalised by taking 1 mL samples of OD_{600} = 0.1. Cells were harvested by centrifugation before being resuspended in 80 µL of SDS loading buffer and analysed by SDS-PAGE (§2.10.2)

2.12.1.2 Maltose plate assay

Overnight cultures of NT326 cells containing the plasmid encoding the TOXCAT chimera were streaked out on maltose minimal media plates (M9 salts (48 mM Na_2HPO_4; 22 mM KH_2PO_4; 8.6 mM NaCl; 18.7 mM NH_4Cl); 2 mM $MgSO_4$; 100 μM $CaCl_2$; 0.4% maltose; 15% (w/v) agar) and incubated at 37º for 2–3 days. Colonies with a correctly inserted chimera were able to grow with maltose as the sole carbon source.

2.12.1.3 Sodium hydroxide washes

Correct insertion of chimera in the membrane was confirmed using sodium hydroxide washes based on (Chen & Kendall, 1995). Cells expressing the chimera were grown until mid-exponential stage and an aliquot was harvested by centrifugation. Cells were resuspended in 100 μL lysis buffer (24 mM $MgCl_2$, 0.5 mg/mL DNase, 0.5 mg/mL lysozyme), incubated at room temperature for 1 hour, and then cooled on ice. 150 μL ice cold dH_2O was added and 125 μL was removed. 125 μL ice cold 0.1 M NaOH was added to the remaining sample, vortexed for 1 min, and centrifuged at 13,400 rpm for 15 min. The pellet contained membrane proteins and the supernatant contained soluble proteins. Soluble proteins were TCA precipitated (§2.9.5) before analysis by SDS-PAGE.

2.12.1.4 Proteolysis of spheroplasts

Spheroplasts were prepared as described (§2.9.3) and resuspended in proteolysis buffer (10 mM HEPES pH 7.6, 2 mM EDTA). The spheroplasts were then either treated with Proteinase K (to a final concentration of 0.25 mg/mL) for 30 min on ice, or broken open by 5× free-thaw cycles before treatment with Proteinase K.

2.12.2 Measurement of [CAT] activity

2.12.2.1 Disk diffusion assay

Cultures were grown to OD_{600} = 0.6 and were diluted 6× with LB. 100 μL was plated on an LB Amp agar and allowed to dry. A Whatman Grade 1 filter paper disk (diameter 30 mm) with 42 μL of 90 mg/mL chloramphenicol (CAM) in ethanol dried

onto it was placed in the centre of the plate. The plate was incubated overnight at 37°C and the clear zone of inhibition of growth around the disk was recorded.

2.12.2.2 Quantitative CAT assay

Cultures were grown to OD_{600} = 0.6, and 200 µL was removed, pelleted and resuspended in 500 µL 100 mM Tris-HCl pH 8.0. Cells were lysed by adding 20 µL lysis mix (100 mM EDTA, 100 mM DTT, 50 mM Tris-HCl pH 8.0), a drop of toluene and incubated at 30ºC for 30 min. Samples were prepared by taking 60 µL of a 10× dilution of lysed cells. The levels of CAT expression were assayed using the FAST CAT® Green (deoxy) Chloramphenicol Acetyltransferase assay kit (Invitrogen) according to the manufacturer's instructions. The fluorescence of the samples was determined using a Perkin Elmer LS50B fluorimeter. Samples were excited at 495 nm and emission at 525 nm was recorded. The fluorescence was normalised using to the positive control, GpA.

2.13 Synthetic peptide purification

Synthetic peptides were ordered from the Keck Biotechnology Resource Laboratory, Yale University, USA. The peptides were synthesised with an acetylated N-terminus and an amidated C-terminus using F-moc chemistry. Peptides were purified by reverse phase high performance liquid chromatography (HPLC) on a purpose built dual pump system (Jasco). HPLC grade solvents were used for all purification steps.

2.13.1 Purification of the TatA TM domain peptide

A peptide corresponding to the TM domain of TatA ($TatA_{TM}$) was synthesised. The sequence of the $TatA_{TM}$ peptide was GGISIWQLLIIAVIVVLLFGTKK**K**, corresponding to amino acids 2 to 24 of TatA with an additional lysine (underlined in bold) on the C-terminus. The extra lysine was added to aid solubilisation of the peptide. 5 mg of crude $TatA_{TM}$ was solubilised in 1.5 mL Trifluoroacetic acid (TFA). The solubilised peptide was loaded onto a C4 column (Jupiter C4 250 × 10 mm, 300 Å, 5 µm, Phenomenex) equilibrated in 70% buffer A (H_2O + 0.1% TFA) and 30% buffer B (Acetonitrile (ACN) + 0.1% TFA). The peptide was eluted at 2 mL/min with the gradient given in

Table 2.9, and absorbance was monitored at 280 nm. $TatA_{TM}$ typically eluted after 44 min, Figure 2.2. The presence of $TatA_{TM}$ in fractions was confirmed by mass spectrometry (§2.14.1). Fractions containing the $TatA_{TM}$ peptide were pooled and lyophilised (§2.13.3).

Table 2.9 Gradient used for the purification of the $TatA_{TM}$ peptide by HPLC.

Time (min)	% Buffer A	% Buffer B
0	70	30
15	70	30
20	50	50
50	20	80
55	0	100
65	0	100
70	70	30

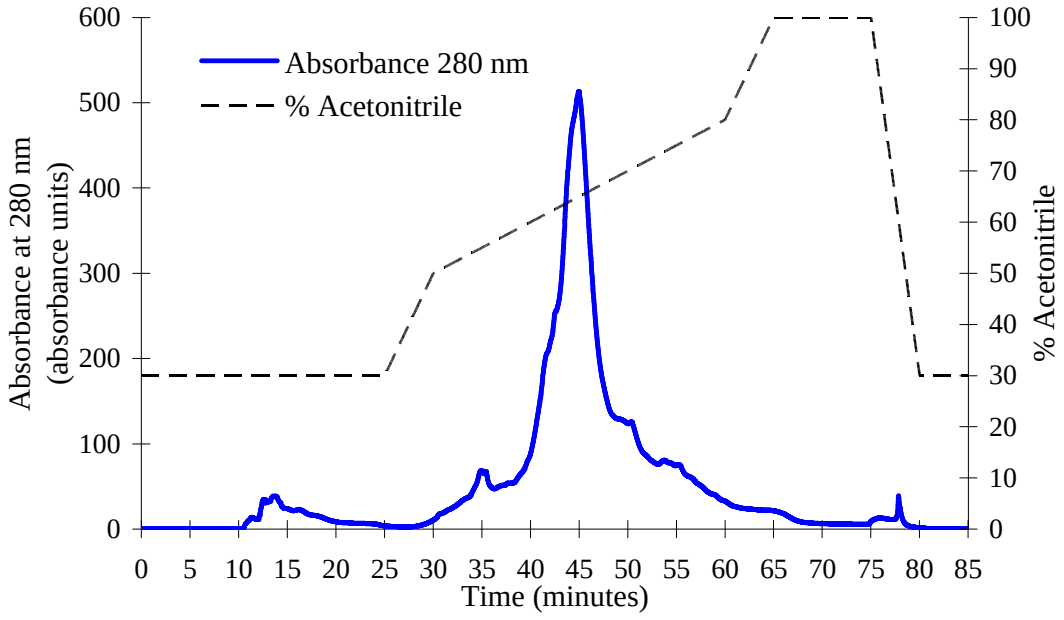


Figure 2.2 Reverse phase HPLC purification of the $TatA_{TM}$ peptide. The $TatA_{TM}$ peptide was purified by reversed-phase HPLC using a linear acetonitrile gradient with 0.1% TFA on a Phenomenex C4 column. The absorbance at 280 nm was recorded to determine when the peptide eluted.

2.13.2 Purification of the TatA APH peptide

A peptide corresponding to the TatA APH ($TatA_{APH}$), was synthesised. The sequence of the $TatA_{APH}$ peptide was: TKKLGSIGSDLGASIKGFKKAMSDDEP, corresponding to amino acids 22 to 48 of TatA. 5 mg of crude peptide was solubilised in 4 mL 10% ACN with 16 µL TFA. The solubilised peptide was loaded onto a C4 column equilibrated in 90% buffer A (H_2O + 0.1% TFA) and 10% buffer B (ACN + 0.1% TFA). The peptide was eluted at 2 mL/min with the gradient given in Table 2.10, and absorbance was monitored at 222 nm, Figure 2.3. The $TatA_{APH}$ peptide typically eluted between 35–37 min. The presence of $TatA_{APH}$ in fractions was confirmed by mass spectrometry (§2.14.1). Fractions containing the $TatA_{APH}$ peptide were pooled and lyophilised (§2.13.3).

Table 2.10 Gradient used for the purification of the $TatA_{APH}$ peptide by HPLC.

Time (min)	% Buffer A	% Buffer B
0	90	10
10	90	10
12	80	20
32	60	40
45	0	100
55	0	100
65	90	10

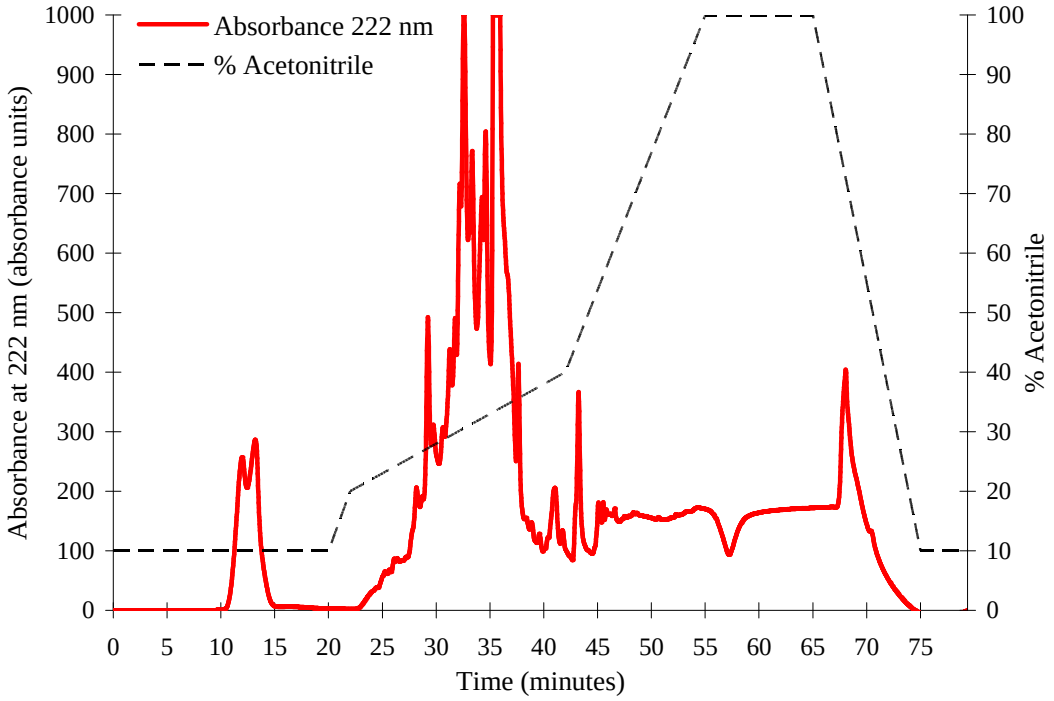


Figure 2.3 Reverse phase HPLC purification of the $TatA_{APH}$ peptide. The $TatA_{APH}$ peptide was purified by reversed-phase HPLC using a linear acetonitrile gradient with 0.1% TFA on a Phenomenex C4 column. The absorbance at 222 nm was recorded to determine when the peptide eluted.

2.13.3 Lyophilisation of purified peptides

Pooled HPLC fractions containing peptide were frozen on dry ice to form a thin layer in a round bottomed flask. The flask was briefly placed in liquid nitrogen before being attached to a vacuum. The flask was left under vacuum until all the solvent had been removed. Lyophilised peptide was dissolved in 2,2,2-trifluoroethanol (TFE), aliquoted, dried under nitrogen to remove the organic phase and then stored at -20ºC until required.

2.13.4 Peptide concentration determination

Peptide concentration was determined by measuring UV absorbance. Peptides were dissolved in organic solvent or buffer and their absorbance was measured using a spectrophotometer (Biomate, Thermo Scientific). The extinction coefficient at 280 nm of the $TatA_{TM}$ (ε_{280} = 5500 $M^{-1}cm^{-1}$) was calculated using ProtParam, and the peptide concentration was calculated using the Beer-Lambert law (1); A =

absorbance, ε = extinction coefficient ($M^{-1}cm^{-1}$), b = optical path length (cm), c = concentration (M).

$$A = \varepsilon bc \qquad (1)$$

Due to the lack of cysteine, tryptophan and tyrosine residues the $TatA_{APH}$ peptide did not absorb at 280 nm. The absorbance was measured at 205 nm and the concentration was calculated using equation (2) (Scopes, 1974).

$$c = \frac{A_{205}}{31 \times b} \qquad (2)$$

2.13.5 Detergent micelle concentration determination

The concentration of detergent micelles in a detergent solution was calculated using equation (3); m = micelle concentration (M), d = detergent concentration (M), C = critical micelle concentration (M), a = aggregation number.

$$m = \frac{d - C}{a} \qquad (3)$$

For dodecylphosphocholine (DPC) an aggregation number of 56 and a critical micelle concentration of 1 mM were used.

2.13.6 Preparation of liposomes

For 1,2-Dimyristoyl-snGlycero-3-Phosphocholine (DMPC) liposomes, the lipid and peptide were codissolved in TFE at the required ratio, typically 10:1 lipid to peptide. Samples were dried under nitrogen and then placed in a vacuum desiccator overnight to ensure complete removal of the organic phase. The lipid/peptide film was hydrated using the appropriate buffer, typically 50mM sodium phosphate buffer, pH 7.0, and subjected to five freeze-thaw cycles using an ethanol/dry ice bath and a 40ºC water bath. Finally samples were sonicated for 5 min at 40ºC.

For 1,2-Dipalmitoyl-*sn*-Glycero-3-Phosphocholine (DPPC) and 1,2-Dipalmitoyl-*sn*-Glycero-3-[Phospho-*rac*-(1-glycerol)] (DPPG) liposomes, the lipid and peptide were codissolved in chloroform and prepared in the same way as above with a 55ºC water bath.

2.14 Analysis of synthetic peptides

2.14.1 Mass spectrometry

Samples from the HPLC fractions were prepared by adding 10 µL 10% Formic acid to 90 µL of the HPLC fraction. Samples were analysed using electrospray ionisation mass spectrometry (ESI-MS) on a Bruker MicrOTOF. The positive ionisation mode was used and detection was between 50 and 3000 m/z. Spectra were typically recorded for 2 minutes, averaged and deconvoluted to determine the mass of the main species.

2.14.2 Cross-linking

Cross-linking of the peptides in different concentrations of detergent was performed with bis(sulfosuccinimidyl)suberate (BS^3) (Pierce) or glutaraldehyde (TAAB) according to the manufacturer's instructions. Typically a 40 µL reaction containing 40 µM peptide, a 50-fold excess of cross-linker and varying amounts of detergent in 20 mM sodium phosphate buffer pH 8.0, was incubated at room temperature for 30 min. Reactions were quenched with 2 µM Tris-HCl pH 7.5.

2.14.3 Circular dichroism (CD)

Peptide samples were prepared with 50 mM sodium phosphate buffer pH 7.5, 100 mM NaCl and varying amounts of DPC (Avanti Polar Lipids) to a final concentration of 80 µM. Measurements were taken in a J-815 spectropolarimeter (Jasco UK, Great Dunmow, UK) using a 1 mm path-length quartz cuvette (Starna; Optiglass Ltd., Hainhault, UK). Circular dichroism (CD) spectra were recorded between 180 and 260 nm with a data pitch of 0.2 nm, a bandwidth of 2 nm, a scanning speed of 100 nm/min, and a response time of 1 s. Four scans were averaged and the spectrum of buffer without peptide was subtracted to give the final spectra. Spectra of peptide in liposomes were also measured. The high tension of the spectra was also recorded.

2.14.4 Oriented circular dichroism (OCD)

Lipid and peptide were codissolved in TFE with final concentrations of 10 mg/mL and 1 mg/mL respectively. The lipid and peptide solution was then dried under

nitrogen on a glass slide to form an even lipid-peptide multilamellar film. To ensure complete removal of the organic phase, the glass slide was left in a desiccator overnight. To hydrate the lipid bilayer the glass slide was placed in a custom built chamber overnight with a small amount of water. The slide was placed in a custom built orientable slide holder with the multilamellar film perpendicular to the incident light. Oriented CD (OCD) measurements were taken using the same parameters as for CD (§2.14.3) with a 10 nm bandwidth. To minimise artefacts arising from linear dichroism the slide was rotated 360° in the plane perpendicular to the incident light, taking measurements at 45º increments. The measurements were averaged and the spectrum of sample without peptide was subtracted to give the final OCD spectra.

2.14.5 Linear dichroism (LD)

Liposomes with 0.5 mg/mL peptide were prepared using 50 mM sodium phosphate buffer, pH 7.0. A 5:1 lipid to peptide ratio was used for the $TatA_{TM}$ peptide analysis. Linear dichroism (LD) spectra were acquired between 180 and 360 nm at room temperature. An in-house built micro-volume Couette cell (Marrington *et al.*, 2004; Marrington *et al.*, 2005) was used; equivalent models are commercially available (Kromatek, Great Dunmow, UK). Samples were aligned by applying voltages of 3, 4 and 5 V, corresponding to 3k, 4k and 5k rpm. The acquisition parameters were the same as those for CD (§2.14.3). A spectrum of the non-aligned sample was subtracted from the aligned spectra to give the final LD signal. The spectra were corrected for light scattering due to the liposomes using the method given in (Nordh *et al.*, 1986) (equation (4) and Figure 2.4), where LD^T = background turbidity dichroism, α = constant, λ = wavelength (nm), *k* = turbidity constant.

$$LD^T(\lambda) = \alpha\lambda^{-k} \qquad (4)$$

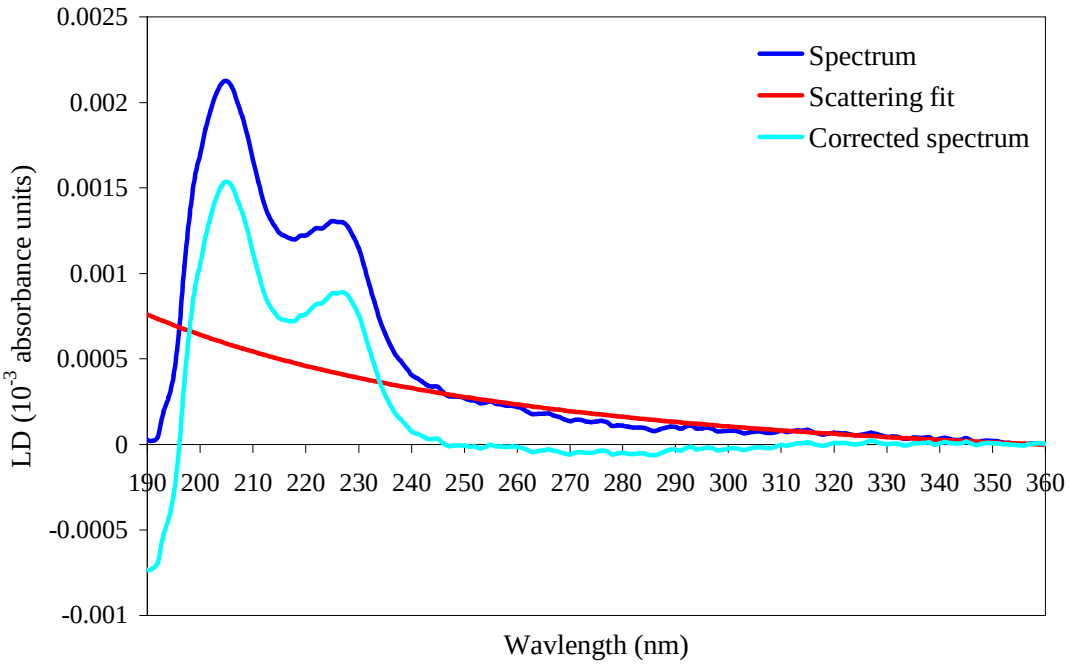


Figure 2.4 Linear dichroism light scattering correction. The method given in (Nordh *et al.*, 1986) (equation (4)) was used to calculate the light scattering due to the liposomes, which was subtracted from the spectrum.

2.14.6 Attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR)

Liposomes with a 30:1 lipid to peptide ratio were prepared with dH_2O (§2.13.6), with a final lipid concentration of 10 mg/mL. A Bruker Tensor spectrometer was used containing a high-sensitivity MCT/A detector cooled with liquid nitrogen. Water vapour and carbon dioxide were evacuated from the machine by purging using dry compressed air from an external source. The sample was deposited onto a trapezoidal germanium internal reflection element and bulk water was removed using a diffused stream of nitrogen. 1000 interferograms were recorded and their average was used for analysis. For deuterium exchange, deuterium saturated nitrogen gas was passed over the sample for 22 hours and a spectrum was recorded. To obtain information on the secondary structure of the peptide the spectra were analysed in the 1600–1700 cm^{-1} region, which corresponds to the amide I peak. The band-narrowing technique of Fourier self-deconvolution was applied as well as a baseline correction using the OMNIC spectra software (Thermo Scientific). Gaussian curves were fitted to the resulting spectra using GRAMS software (Thermo Scientific).

2.14.7 Intrinsic fluorescence

Samples with lipid were prepared as for LD (§2.14.5). Emission spectra were taken on a Perkin Elmer LS50B fluorimeter exciting at 280 nm and recording between 300 and 400 nm, with an excitation slit of 2.5, an emission slit of 2.5 and a scan speed of 50. Spectra of peptide in TFE were also recorded. The fluorescence spectra of the peptide in different environments were determined by subtracting spectra of the buffer without peptide.

2.15 Confocal microscopy

Samples were prepared for imaging by adsorbing liquid culture onto 3% agar containing growth medium. The agar was placed in an in-house built sample holder and covered with a cover slip. Imaging was performed on a laser scanning confocal microscope (LSCM) (Leica DMRE) equipped with a Leica TCS SP2 confocal unit, an acousto-optical beam splitter (AOBS®) and a 100mW argon laser. A 63× oil-immersion objective was used. For GFP, fluorescence excitation was at 488 nm and detection was between 510 and 530 nm. The *Synechocystis* chlorophyll was used as a thylakoid membrane marker. When excited at 488 nm chlorophyll fluorescence could be detected between 600 and 700 nm.

2.16 Specialist software

CHI (Crystallography and NMR system (CNS) searching of helix interactions): Model of the TatA APH (Adams *et al.*, 1995; Adams *et al.*, 1996).

Chromas: Analysis of sequencing data.

Clone manager: Design of cloning strategy and alignment of DNA sequences.

ClustalX: Alignment of protein primary sequences.

Dichroweb: Calculation of protein secondary structure from CD spectra (Whitmore & Wallace, 2008). The CDSSTR algorithm (Sreerama & Woody, 2000) was used with the SP175 reference set (Lees *et al.*, 2006).

http://dichroweb.cryst.bbk.ac.uk/html/home.shtml

GRAMS Spectroscopy: Analysis of FTIR spectra (Thermo Scientific)

ImageJ: Analysis of images (Abramoff, 2004; Rasband, 1997–2008)

OMNIC Spectra: Analysis of FTIR spectra (Thermo Scientific).

ProtParam: Analysis of protein sequences.

http://www.expasy.ch/tools/protparam.html

PSRIPRED: Protein secondary structure prediction (Jones, 1999; McGuffin *et al.*, 2000)

http://bioinf.cs.ucl.ac.uk/psipred/

TMSTAT: Analysis of TM domains to identify packing motifs (Senes *et al.*, 2000).

http://bioinfo.mbb.yale.edu/tmstat/

VMD: Visualisation of CHI results (Humphrey *et al.*, 1996).

http://www.ks.uiuc.edu/Research/vmd/

Volocity: Analysis of confocal microscopy images (Improvison, Perkin Elmer).

Chapter 3. Comparison of the Tat Pathway in *Escherichia coli* and *Synechocystis* sp. PCC6803

3.1 Introduction*

Protein export is an essential feature of living cells and several protein translocation pathways have been identified. In prokaryotes and chloroplasts the twin-arginine translocase (Tat) pathway is able to transport proteins across the cytoplasmic and thylakoid membranes, respectively. Studies suggest that the *Escherichia coli* and chloroplast Tat pathways function in a similar manner, with both having a Tat translocase consisting of three integral membrane proteins, TatA, TatB and TatC, which are also known as Tha4, Hcf106 and cpTatC in chloroplasts. Deletion of any of these components leads to a complete block of transport. The pathway was named due to the highly conserved twin-arginine (RR) motif present in the signal sequence of Tat substrates (Berks, 1996; Chaddock *et al.*, 1995). In chloroplasts substitution of either of the arginines, even with lysine, stops transport (Chaddock *et al.*, 1995; Henry *et al.*, 1997), however the *E. coli* system is less stringent and single substitutions with lysine can still support export (Buchanan *et al.*, 2001; DeLisa *et al.*, 2002; Stanley *et al.*, 2000).

Cyanobacteria pose an interesting subject when studying the Tat pathway. It is widely accepted that they are the ancestors of plastids in higher plants and as such may provide a link between thylakoid and bacterial Tat pathways. As Gram-negative bacteria they contain a plasma membrane, however, like chloroplasts, they also contain an internal thylakoid membrane, Figure 1.1B and C. Recent studies of the cyanobacterium *Synechocystis* sp. PCC6803 have suggested that the thylakoid membranes are discontinuous from the plasma membrane (Liberton *et al.*, 2006; Schneider *et al.*, 2007). Despite having two separate membrane systems the *Synechocystis* genome only contains one set of Tat apparatus, consisting of two TatA/B-like proteins and a TatC homologue (Aldridge *et al.*, 2008; Fulda *et al.*, 2000). Alignments of these proteins with the *E. coli* and chloroplast homologues reveal that they are more closely related to the plant system, however both of the

* *Work in this chapter was done in collaboration with Anja Nenninger and Christopher Müller (Department of Biological Sciences, University of Warwick). Mutagenesis and immunoblots were performed in equal collaboration and images in Figure 3.3 were taken by Anja Nenninger.*

TatA/B homologues are able to restore export in *E. coli* strains lacking either TatA or TatB (Aldridge *et al.*, 2008).

Here we have examined the need for an RR-motif in Tat substrates in both *E. coli* and *Synechocystis* with an aim to identifying any differences between the Tat pathways in the two bacteria. A chimera consisting of the signal sequence of a Tat substrate fused to green fluorescent protein (GFP) was used to compare export in *E. coli* and *Synechocystis*. Confocal microscopy has been used to follow the targeting of the GFP chimera in *Synechocystis* as this presents a powerful non-invasive tool for determining cellular location of active GFP.

3.2 Importance of the twin-arginine motif in *E. coli*

Although named the twin-arginine pathway, several previous studies have demonstrated that the RR-motif is not essential for export via the Tat pathway (Buchanan *et al.*, 2001; DeLisa *et al.*, 2002; Stanley *et al.*, 2000). Native Tat substrates have also been identified which do not contain a RR-motif (Hinsley *et al.*, 2001; Ignatova *et al.*, 2002). To enable comparison of the importance of the RR-motif in *E. coli* and *Synechocystis* it was decided to analyse substitutions in the RR-motif of the TorA-GFP construct, which consists of the signal sequence of the *E. coli* Tat substrate trimethylamine *N*-oxide (TMAO) reductase (TorA) fused to GFP (Thomas *et al.*, 2001). It has been seen previously that when expressed in *E. coli* and *Synechocystis* this construct is exported efficiently to the periplasm (Barrett *et al.*, 2003b; Spence *et al.*, 2003; Thomas *et al.*, 2001). The TorA signal peptide contains the sequence S<u>RR</u>RFL, with the two arginines of the predicted RR-motif underlined. Since there are three arginines in the signal sequence it was also decided to determine if the arginine in the +1 position, relative to the RR-motif, could compensate for the loss of the arginines in the RR-motif. Although previous studies have assessed effects of substitutions in the RR-motif on similar constructs in *E. coli* (DeLisa *et al.*, 2002), it was decided to analyse these substitutions using our system in *E. coli* as well as *Synechocystis* to allow direct comparison.

Multiple substitutions of arginine with lysine were introduced in the three arginines in the signal sequence of TorA, with all seven possible substitutions obtained (RRK, RKR, KRR, KKR, KRK, RKK, KKK). The substitutions were made by mutation of

the pJDT1 plasmid (Thomas *et al.*, 2001), which expresses TorA-GFP from the pBAD24 plasmid. DH5α *E. coli* cells expressing the constructs were fractionated into periplasmic, membrane and cytoplasmic (P, M, C) fractions, and analysed by SDS-PAGE and immunoblotting with anti-GFP antibodies, Figure 3.1. The wild-type TorA-GFP (RRR) was exported efficiently to the periplasm, as demonstrated by a band corresponding to mature-sized GFP in the periplasm. The accumulation of mature-sized GFP in the cytoplasm observed here has been seen in other studies (Barrett *et al.*, 2003b; Thomas *et al.*, 2001), and has been attributed to proteolysis of TorA-GFP when export is blocked. Substitution with lysine of one of the twin-arginines or the arginine in the +1 position (RRK, RKR, KRR, KRK, RKK) had no observable effect on the export of TorA-GFP to the periplasm, only a double substitution of the RR-motif stopped export (KKK, KKR) with no visible periplasmic band. This demonstrates that the Tat pathway in *E. coli* only needs a single arginine in the RR-motif to be present in the signal sequence for export. An arginine in the +1 position is not sufficient to compensate for the loss of the RR-motif.

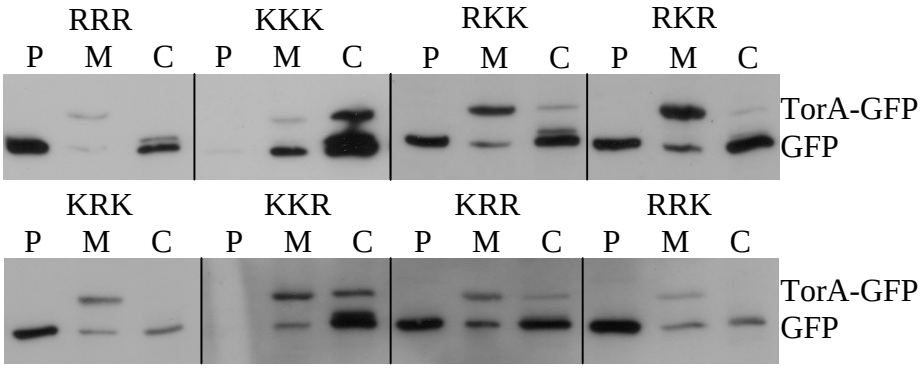


Figure 3.1 TorA-GFP export assay of the mutated TorA-GFP in *E. coli*. Immunoblots using an anti-GFP antibody of periplasmic, membrane and cytoplasmic (P, M, C) samples from DH5α cells expressing TorA-GFP from the pJDT1 plasmid (RRR) or TorA-GFP with the substitution of the RRR with lysine (KKK, RKK, RKR, KRK, KKR, KRR, RRK). Intact TorA-GFP (TorA-GFP) and clipped TorA-GFP (GFP) are indicated.

3.3 Importance of the twin-arginine motif in *Synechocystis*

In order to compare the stringency of the Tat pathway in *E. coli* and *Synechocystis*, export of TorA-GFP mutants was assessed in *Synechocystis* cells. Although TorA is not a native *Synechocystis* Tat substrate, its signal sequence directs GFP to the periplasm (Spence *et al.*, 2003). Unlike *E. coli* which can be fractionated into separate cell compartments with relative ease, fractionation of *Synechocystis* is a much more involved process (Norling *et al.*, 1998). To overcome this problem it was decided to use a bio-imaging approach as this provides spatial information on the location of the TorA-GFP *in vivo*. Since *Synechocystis* is a photosynthetic bacterium its thylakoid membranes possess an intrinsic autofluorescence. Using a confocal microscope to image an optical slice through the centre of a cell, exciting with a wavelength of 488 nm and detecting in the 600–700 nm region, results in an annular fluorescence which corresponds to the thylakoid membranes, shown in magenta in Figure 3.2A. By detecting the fluorescence due to GFP in the 510–530 nm region it is possible to determine the location of the TorA-GFP with respect to the thylakoid membranes. Due to the fluorescent nature of the thylakoid membranes, they also emit autofluorescence in the same region as GFP. It is possible to distinguish the GFP and autofluorescence by comparison with wild-type cells, and also by the localisation of the signal in the GFP region.

To obtain an accurate profile of *Synechocystis* cells expressing TorA-GFP, mixtures of wild-type and TorA-GFP expressing cells were imaged, Figure 3.2A. A signal is detected in the GFP region (shown in green) of both wild-type, indicated by an arrow, and TorA-GFP expressing cells. In wild-type cells the signal intensity is much less and cells expressing TorA-GFP can clearly be distinguished by their intense green ‘halo’. To demonstrate more clearly the location of the TorA-GFP with respect to the thylakoids the plot-profile tool in the ImageJ analysis software was used. A line can be drawn across the image and the value of each pixel along the line is plotted against position. A rectangle can also be selected; this gives the average of the line plot-profiles in the rectangle. For example the plot-profile of a rectangle height 5, width 100 will be the average of 5 line plot-profiles of length 100. By overlapping plot-profiles of the GFP and auto-fluorescence channels we can determine the relative position of the GFP. Plot-profiles of the GFP fluorescence and autofluorescence were taken of wild-type and TorA-GFP *Synechocystis*, Figure 3.2B

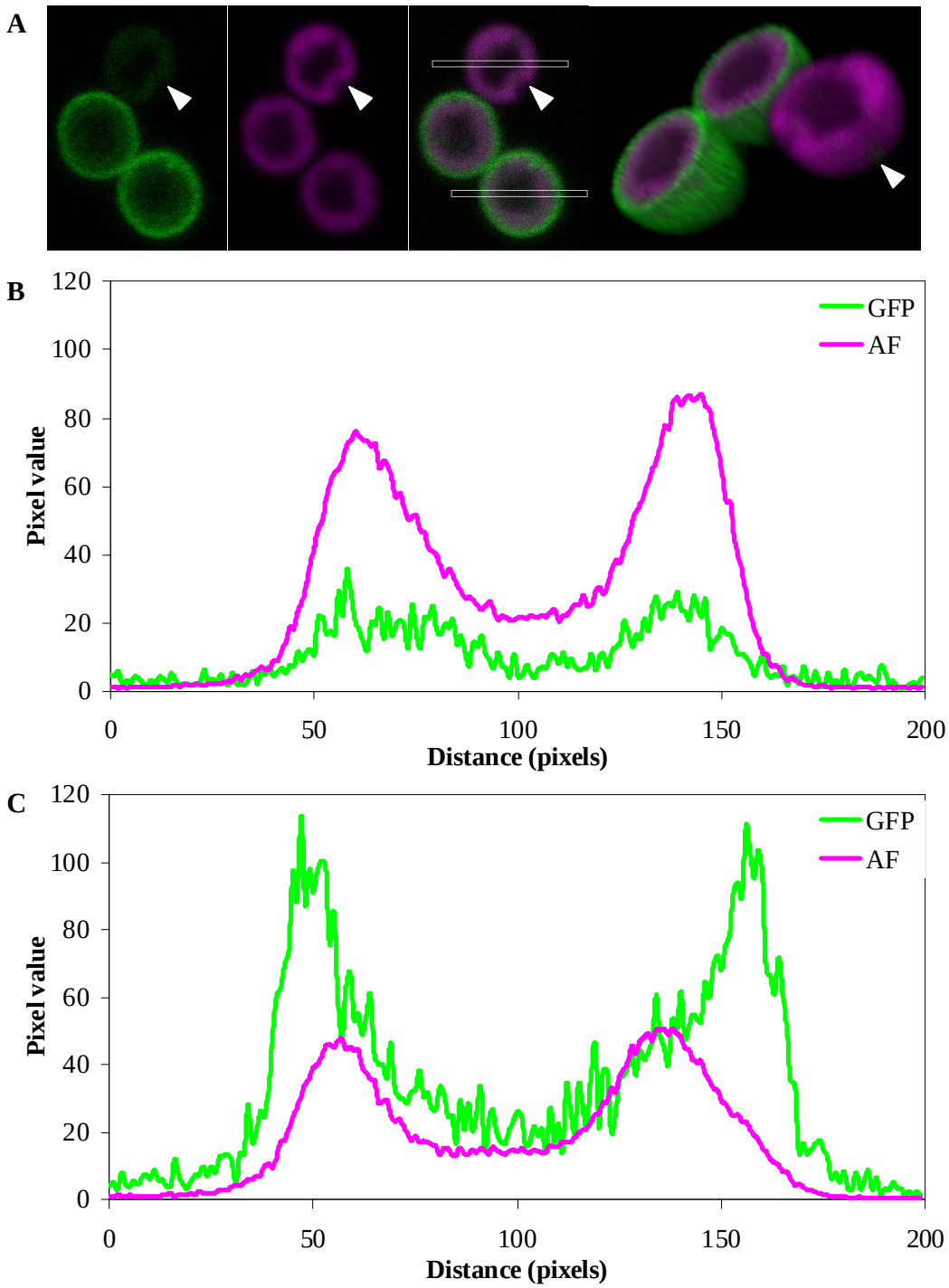


Figure 3.2 Comparison of confocal microscopy images of wild-type *Synechocystis* and cells expressing TorA-GFP. A. Confocal microscopy image of a wild-type *Synechocystis* cell (arrow) and two cells expressing TorA-GFP. Samples were excited with the 488 nm line of the argon laser, GFP emission was detected between 510 and 530 nm (shown in green), and autofluorescence was detected between 600 and 700 nm (shown in magenta). Overlays of the channels and a 3D reconstruction of a z-stack are also shown. **B.** Plot profile of a wild-type cell in the region indicated by the white box and arrow in A. The grey level in each channel was taken across the region defined by the white box and plotted against distance. **C.** Plot profile of a cell expressing TorA-GFP in the region defined by the white box in A.

and C, in the regions indicated by the white boxes in the overlay of the GFP and autofluorescence images, Figure 3.2A panel 3. The difference between the plot-profiles of the two different cells can clearly be seen by comparison of the GFP channels. In wild-type *Synechocystis* the GFP plot-profile has a very similar shape to the auto-fluorescence plot-profile, with quite a low intensity. In contrast the TorA-GFP expressing cell has a much higher GFP intensity, which does not coincide with the autofluorescence but peaks outside, indicating a periplasmic location. The three-dimensional reconstruction from a z-stack of half the cells, Figure 3.2A panel 4, clearly demonstrates the continuous periplasmic location of the TorA-GFP.

Using the same methodology as for *E. coli*, the three arginines in the signal sequence of TorA-GFP were substituted for lysine in the *Synechocystis* plasmid used previously (Spence *et al.*, 2003). Transformation of the plasmids resulted in RKR, RKK, KRR, KKR and KKK TorA-GFP mutants. Cells were imaged in the same way as above, and plot profiles of typical cells were measured, Figure 3.3. The green 'halo' typical of the TorA-GFP expressing cells was not observed for any of the mutants, Figure 3.3. Plot-profiles of the GFP channel were very similar to the autofluorescence channel, indicating that the TorA-GFP mutants were not exported, Figure 3.3. Punctate regions of GFP were observed in the thylakoid and cytoplasmic region of all the mutants. We concluded that these were aggregates of GFP, resulting from the loss of export to the periplasm.

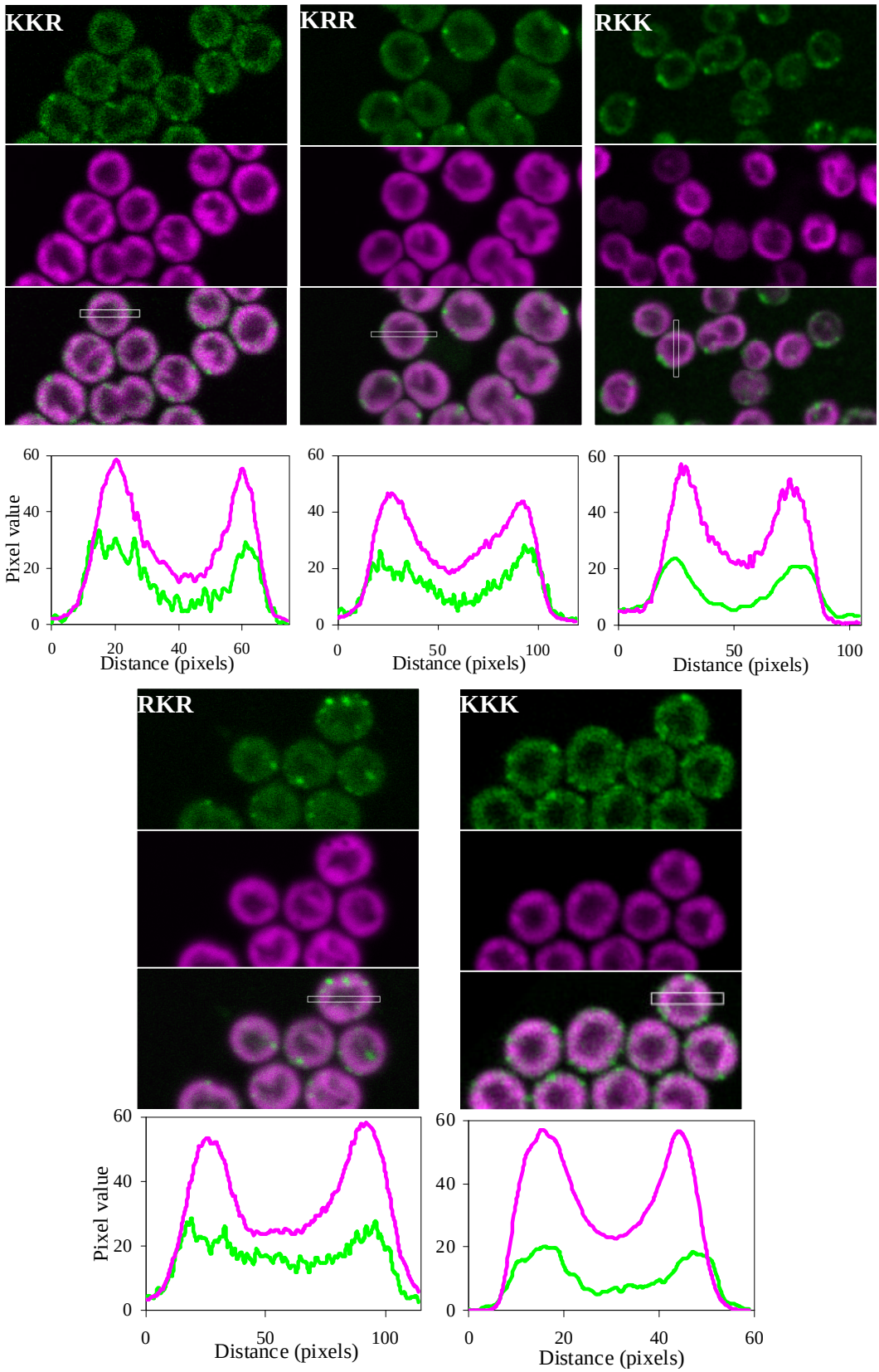


Figure 3.3 Analysis of the export of mutated TorA-GFP in *Synechocystis*. Confocal microscopy images of *Synechocystis* cells expressing TorA-GFP with the substitution of the RRR motif with lysine (KKR, KRR, RKK, RKR, KKK). The GFP signal is shown in green and the autofluorescence is shown in magenta. Overlays of the two channels are also shown. The 488 nm laser line of the argon

laser was used to excite the samples, GFP emission was detected between 510 and 530 nm, and autofluorescence was detected between 600 and 700 nm. Plot profiles of the GFP and autofluorescence channels are shown. The grey level in each channel was measured across the region defined by the white box and plotted against distance.

3.4 Discussion

Comparison of the Tat pathway from different bacteria has shown that the composition and functional specificity varies from bacterium to bacterium. Here we have sought to gain insight into the differences between the Tat pathway in *E. coli* and *Synechocystis* by analysing the export of the Tat substrate TorA-GFP.

3.4.1 *E. coli* requires a single arginine in the twin-arginine motif

To assess the importance of the RR-motif in the export of proteins via the *E. coli* Tat pathway, substitutions of RR-motif in TorA-GFP were analysed. Single substitutions in the RR-motif and the additional arginine in the +1 position (RRK, RKR, KRR, KRK, RKK) did not affect the export of TorA-GFP to the periplasm. Only complete substitution of the RR-motif (KKR, KKK) stopped export. Similar results have been seen in previous studies of TorA and other Tat substrates (Buchanan *et al.*, 2001; DeLisa *et al.*, 2002; Stanley *et al.*, 2000). It is worth noting that substitution of arginine with lysine is conservative and substitution with other, non-polar residues with a different charge have been seen to have a greater effect on translocation (Buchanan *et al.*, 2001). These results demonstrate that there is some flexibility in the recognition of the RR-motif by the *E. coli* Tat apparatus. The cause of this flexibility is as yet unidentified. One hypothesis is that other parts of the twin-arginine signal sequence compensate for the loss of an arginine. This is supported by the observation that substitution of the -1 and +2 residues, with respect to the RR-motif, can severely affect export of Tat substrates (Mendel *et al.*, 2008).

3.4.2 *Synechocystis* cannot tolerate substitutions in the twin-arginine motif

Unlike *E. coli*, the Tat pathway in *Synechocystis* could not tolerate the conservative single substitution with lysine in the RR-motif. The KRK, RKK and RKR

substitutions all stopped the appearance of the green periplasmic 'halos', indicating a loss of export. On loss of export, some GFP accumulated into small regions, as indicated by punctate spots in the green channel. There is no evidence to suggest these accumulations are physiologically relevant, and similar observations in *E. coli* have been credited to aggregates (Mullineaux *et al.*, 2006; Ray *et al.*, 2005).

The strict requirement of a RR-motif is reminiscent to that of the Tat pathway in plant thylakoid membranes. Alignments have demonstrated that the TatA/B homologues from *Synechocystis* are more closely related to the Tha4/Hcf106 proteins in plant thylakoids (Aldridge *et al.*, 2008), and phylogenetic studies cluster the TatC homologue with the plant thylakoid cpTatC (Yen *et al.*, 2002). These observations, taken together with the endosymbiotic theory for the origin of plastids in higher plants, suggest that that the Tat pathway in cyanobacteria is more closely related to plant thylakoids than *E. coli*.

Originally it was intended to use imaging techniques such as fluorescence recovery after photobleaching (FRAP) and light inducible GFP to track the movement of TorA-GFP in cells. Bleaching experiments were performed on *Synechocystis* cells expressing TorA-GFP, however due to the small size of the cells, ~3 μm, it was not possible to bleach only the TorA-GFP in the periplasm. Previous FRAP experiments on cyanobacteria demonstrated that a bleach area of 0.5–1 μm was possible (Mullineaux, 2004), meaning any bleaching would include the thylakoid membrane. Subsequent recovery in the GFP channel would therefore consist of both GFP and autofluorescence signals, and parameters would be difficult to calculate. Due to these experimental constraints it was decided to gain more information on the *E. coli* Tat pathway by studying the TatA component using biochemical and biophysical techniques, as detailed in Chapters 4 to 6.

In summary, here we have demonstrated that there are differences in the recognition of twin-arginine signal sequences between *E. coli* and *Synechocystis*. These differences suggest that the cyanobacterial Tat pathway is more closely linked to that in thylakoids than proteobacteria.

Chapter 4. Analysis of the *Escherichia coli* TatA Transmembrane Domain

4.1 Introduction

In *E. coli* three proteins have been identified as being essential for translocation via the Tat pathway; TatA, TatB and TatC (Bogsch *et al.*, 1998; Sargent *et al.*, 1998; Weiner *et al.*, 1998). The smallest of these is TatA, which is predicted to have an N-terminal α-helical transmembrane (TM) domain, followed by an amphipathic α-helix (APH) and a largely unstructured C-terminal region, Figure 1.6. Although a relatively small protein (89 amino acids, predicted molecular weight 9.6 kDa), TatA is able to form large homo-oligomeric complexes of varying size and diameter (Gohlke *et al.*, 2005; Oates *et al.*, 2005). Just how this single membrane spanning protein can form these complexes is still unknown. Previous research has shown that when expressed without its TM domain TatA purifies as a soluble monomer (De Leeuw *et al.*, 2001; Porcelli *et al.*, 2002). This, along with a study highlighting the importance of a glutamine in the TM domain (Greene *et al.*, 2007), has indirectly indicated that the TM domain is involved in the formation of homo-oligomers.

TM domains of proteins are often found to play an important role in the formation of complexes (Lemmon *et al.*, 1992; Melnyk *et al.*, 2001). Motifs have been identified which drive TM domain helix-helix interactions and facilitate complex formation (Russ & Engelman, 2000; Senes *et al.*, 2000). These motifs include the presence of polar residues, such as glutamine, in TM domains. Electrostatic interactions and interhelical hydrogen bonds between hydrophobic residues can drive helix-helix interactions and lead to the formation of stable complexes (Dixon *et al.*, 2006; Freeman-Cook *et al.*, 2004; Zhou *et al.*, 2001).

In this study the TM domain of TatA was analysed to determine if it had a role in the oligomerisation of TatA. Short TM domains can often be problematic to express and purify in sufficient quantities for subsequent analysis, so it was decided to use a synthetic peptide to analyse the TatA TM domain. Synthetic peptides of TM domains have been used in many studies and can give a good indication of the role of the TM domain in the full length protein (Dixon *et al.*, 2006; Iwamoto *et al.*, 2005; Melnyk *et al.*, 2001). Here, a variety of biophysical techniques were used to determine the secondary structure, ability to insert into lipid bilayers and propensity for self-association of the TatA TM domain peptide. Helix-helix interactions were also probed in a membrane environment using the TOXCAT assay (Russ & Engelman,

1999), which was specifically designed for the analysis of TM domain interactions. The importance of the glutamine in the TatA TM domain for function and oligomerisation was analysed using the TOXCAT assay, a functional TorA assay and blue-native electrophoresis of the TatA complex.

4.2 Analysis of a TatA TM domain synthetic peptide

To analyse the TatA TM domain *in vitro* a synthetic peptide corresponding to the TatA TM domain ($TatA_{TM}$) was synthesised as detailed in §2.13. The $TatA_{TM}$ peptide (GGISIWQLLIIAVIVVLLFGTKKK) corresponded to residues Gly_2 to Lys_{24} of TatA, with an additional C-terminal non-native lysine (underlined) to aid solubility of this extremely hydrophobic peptide. Previous studies have shown that the addition of non-native lysines do not affect the ability of strongly interacting TM domains to oligomerise (Ding *et al.*, 2001; Lazarova *et al.*, 2004; Melnyk *et al.*, 2001; Melnyk *et al.*, 2003). However there is evidence that lysines and other positively charged residues may affect the interactions between more weakly interacting helices (Iwamoto *et al.*, 2005; Prodöhl *et al.*, 2007), so their inclusion is not always appropriate. Since TatA includes two native lysines, it was decided that the addition of one further lysine should not disrupt any strong interactions.

The $TatA_{TM}$ peptide was purified using reversed-phase high performance liquid chromatography (HPLC) using a linear acetonitrile (ACN) gradient with 0.1% trifluoroacetic acid (TFA) as detailed in §2.13.1. Fractions were checked for the presence of pure peptide using electrospray ionisation mass spectrometry (ESI-MS), before being pooled and lyophilised. The major species detected by ESI-MS, Figure 4.1, corresponded to pure peptide (2651 Da), with additional peaks corresponding to sodium adducts (2673, 2695 Da). Sodium adducts have previously been observed in the purification of TM peptides (Rijkers *et al.*, 2005).

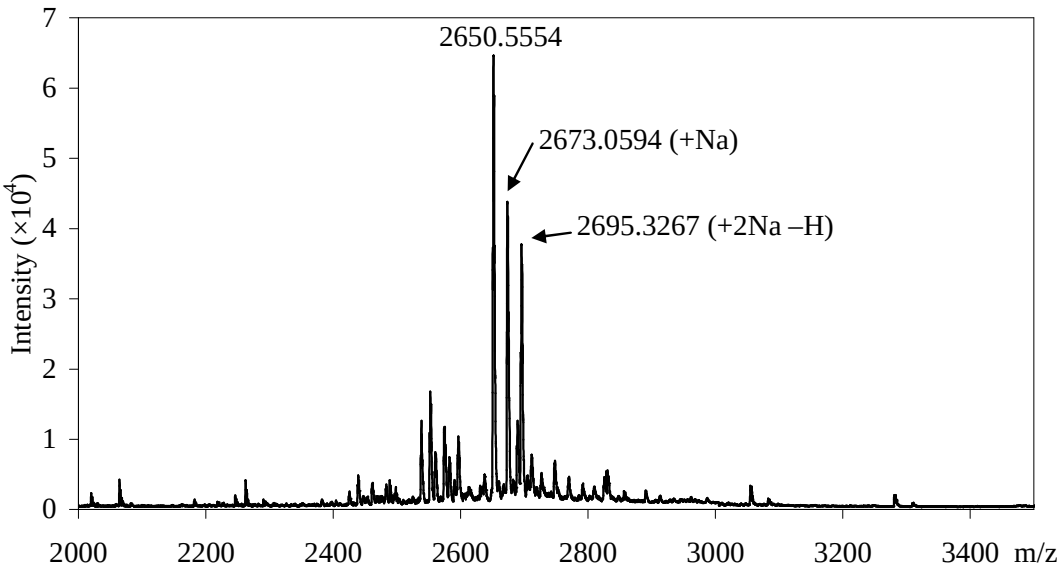


Figure 4.1 Assessment of the purified $TatA_{TM}$ peptide by electrospray ionisation mass spectrometry. Pooled fractions from reversed-phase HPLC purification of the $TatA_{TM}$ peptide were checked for purity by ESI-MS. The major peak corresponded to the $TatA_{TM}$ peptide (2651 Da), with additional peaks corresponding to sodium adducts.

4.2.1 Determination of the $TatA_{TM}$ peptide secondary structure

To determine the secondary structure of the $TatA_{TM}$ peptide in detergent micelles and lipid bilayers, circular dichroism (CD) spectra were obtained. CD spectra were recorded for peptide solubilised in buffer containing 7.5 mM dodecylphosphocholine (DPC) detergent, and peptide reconstituted into 1,2-Dimyristoyl-sn-Glycero-3-Phosphocholine (DMPC) liposomes in a 10:1 lipid to peptide ratio, Figure 4.2. Negative minima at 208 and 222 nm, characteristic of an α-helical secondary structure, can be seen in the spectra. Fitting of the data using CDSSTR (Sreerama & Woody, 2000) with the SP175 reference set (Lees *et al.*, 2006) predicts that the peptide is predominantly α-helical; ~69% α-helical in DPC and ~54% α-helical in DMPC. Although these values are lower then expected there may be some aggregation of this extremely hydrophobic peptide, which would decrease the α-helical content of the sample. Additionally, the signal is not very accurate below 195 nm due to light scattering of the sample. This may affect the results from the fitting program, which fits between 190 and 240 nm. Despite this, the data indicate that the

$TatA_{TM}$ peptide does have a largely α-helical secondary structure in both detergent micelles and liposomes, which agrees with the predicted secondary structure of TatA.

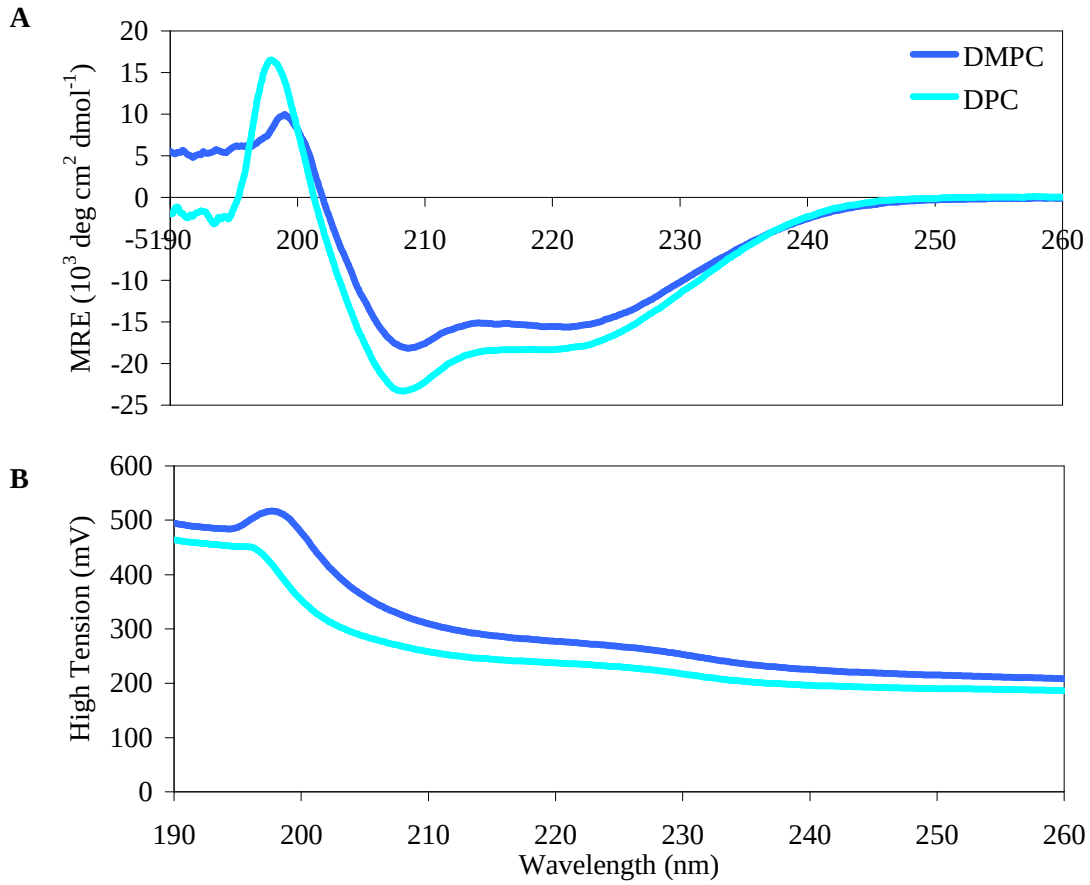


Figure 4.2 Secondary structure determination of the $TatA_{TM}$ peptide by circular dichroism. A. CD spectra of the $TatA_{TM}$ peptide were recorded for the peptide solubilised in buffer containing 7.5 mM DPC detergent (light blue), and peptide reconstituted into DMPC liposomes in a 10:1 lipid to peptide ratio (dark blue). Baselines of samples without peptide were subtracted to give the final spectra. Data are reported in mean residue ellipticity (MRE) units. **B.** The high tension signals of the CD spectra were also recorded.

4.2.2 Determination of the insertion of the $TatA_{TM}$ peptide into lipid bilayers

The exact cellular location of TatA is still a subject of some dispute. Whilst it is widely considered to be a membrane protein a recent study has claimed that it forms cytoplasmic tubules (Berthelmann *et al.*, 2008). Furthermore, the TatA homologue in

the Gram-positive *Bacillus subtilis*, TatAd, has been shown to be active in the cytoplasm and membrane of cells (Pop *et al.*, 2003). To ascertain that the $TatA_{TM}$ peptide does insert into lipid membranes, the solution-phase method of linear dichroism spectroscopy (LD), oriented circular dichroism (OCD), attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) and intrinsic tryptophan fluorescence were used.

4.2.2.1 Analysis of the $TatA_{TM}$ peptide by linear dichroism (LD)

The solution-phase method of LD (Dafforn & Rodger, 2004; Rodger *et al.*, 2002) has recently become a more viable technique for studying proteins due to advances which have enabled the use of small sample volumes (< 50 µL) (Marrington *et al.*, 2004; Marrington *et al.*, 2005). LD is a spectroscopic technique which measures the difference in absorption of light polarised parallel to an orientation direction and light polarised perpendicular to that direction (LD signal). In order to obtain a signal, samples must be able to be aligned, which means they must have an asymmetric characteristic in charge or shape. The study of membrane proteins by this method has been facilitated by the use of liposomes, which can be orientated by shear flow in a spinning Couette cell (Ardhammar *et al.*, 1998; Rodger *et al.*, 2002), Figure 4.3A. The liposomes are elongated by the flow and align parallel to the direction of flow. If the protein does not interact with the liposomes it will not align, Figure 4.3B, and there will be no difference between absorption of light polarised parallel and perpendicular, and hence no LD signal. If the protein interacts with the liposomes, an LD signal can be seen, with positive and negative maxima which result from the transition polarisations of the protein, and depend on its structure and mode of interaction. For an ideal α-helical peptide, if the peptide inserts across the lipid bilayer, perpendicular to the direction of flow, Figure 4.3C, it has been calculated that the $n \rightarrow \pi^*$ transition produces a positive maximum at ~220 nm, with the $\pi \rightarrow \pi^*$ transition producing a negative maximum at ~ 210 nm and a positive maximum at ~190 nm (Dafforn & Rodger, 2004; Rodger *et al.*, 2002). For an α-helical peptide lying on the surface of the liposome, parallel to the direction of flow, Figure 4.3D, the LD signal will be reversed, giving negative maxima at ~190 and ~220 nm, and a positive maximum at ~210 nm.

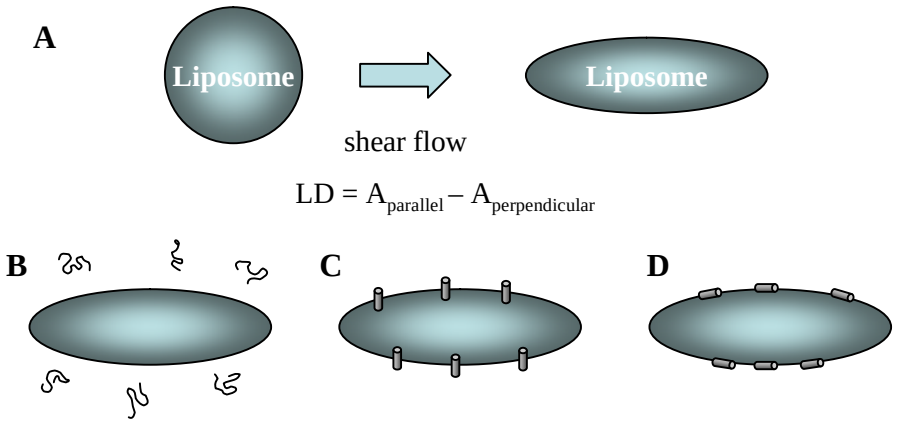


Figure 4.3 Linear dichroism of peptides in liposomes. A. In shear flow liposomes become elongated and align with the direct of flow. The LD signal is the absorbance parallel to the flow (arrow) minus the absorbance perpendicular to the flow. **B.** If the peptide does not associate with the liposomes, it does not orientate and gives no LD signal. **C.** If the peptide inserts across the liposome bilayer, it aligns perpendicular to the direction of flow. **D.** If the peptide associates with the surface of the liposome it orientates parallel to the direction of flow. Figure adapted from (Hicks *et al.*, 2008).

The interaction between the α-helical $TatA_{TM}$ peptide and DMPC liposomes, and its orientation within DMPC bilayers, was investigated by LD. DMPC liposomes were prepared containing $TatA_{TM}$ peptide and orientated in a Couette cell by spinning at three different rotation speeds; 3000, 4000 and 5000 rpm. Spectra were recorded between 190 and 360 nm and corrected for light scattering due to the liposomes, Figure 4.4. The resulting spectra all show positive maxima at ~205 and 225 nm, and a negative maximum at ~217 nm. It was assumed that the 225 nm maximum corresponded to the $n \rightarrow \pi^*$ transition and the 205 and 217 nm maxima corresponded to the $\pi \rightarrow \pi^*$ transition, thus indicating that the peptide was perpendicular to the direction of flow and inserted into the lipid bilayer, Figure 4.3C. Although the wavelengths of the maxima are shifted from those calculated for an ideal α-helix (190, 210 and 220 nm), this has been observed previously for peptides inserted into liposome bilayers. The antibiotic gramicidin inserts into membranes and gives an LD signal which has positive maxima at ~200 and 227 nm, and a negative maximum at ~213 nm (Rodger *et al.*, 2002). Shifts have also been observed for the α-helical TM domains of E5 (Oates *et al.*, 2008) and invariant chain (Ii) when they insert into lipid bilayer (A. J. Beevers, J. Oates and A. M. Dixon, unpublished work). Despite the shift in maxima, it was concluded that the $TatA_{TM}$ peptide was inserting into the

liposome membranes. The increase in intensity of the peptide LD signal as rotational speed increases is due to the increased orientation of the liposomes, showing that the LD signal also depends on alignment.

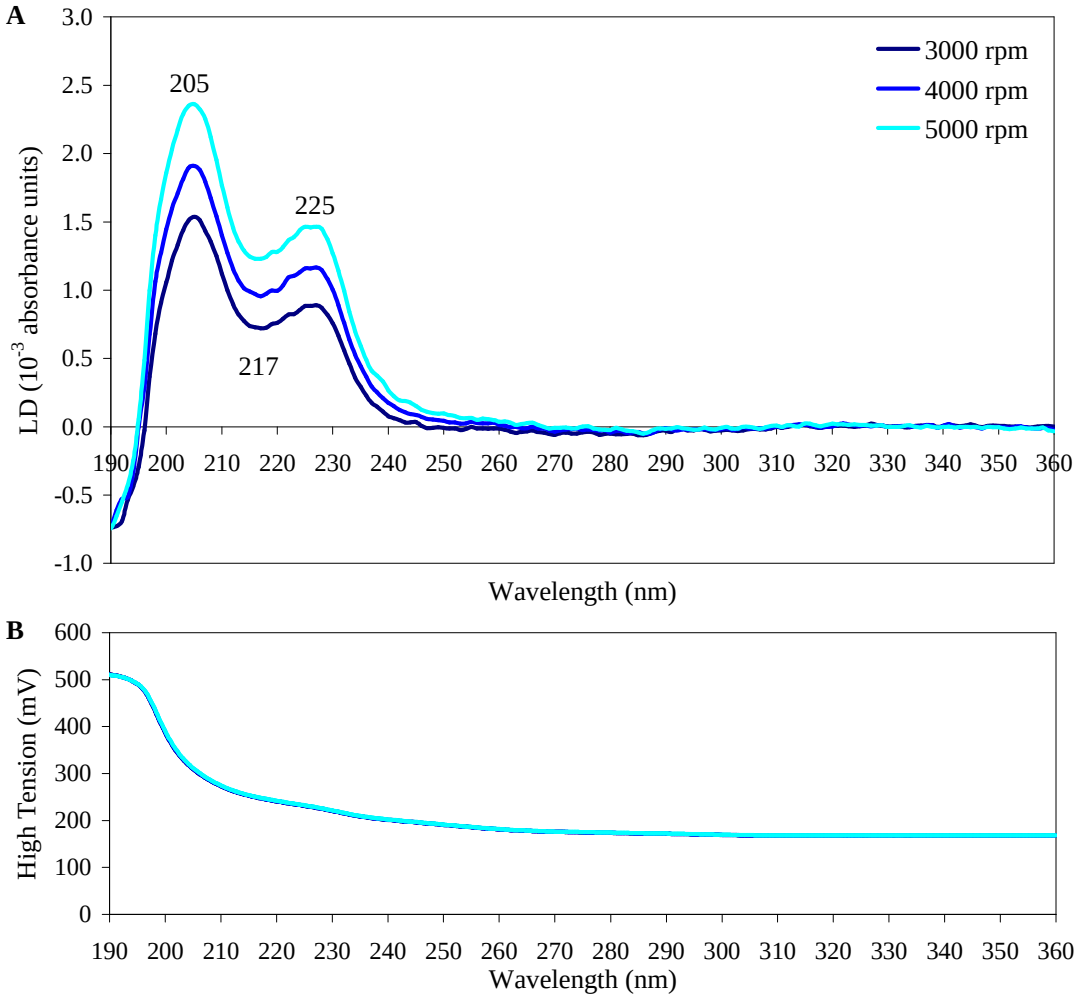


Figure 4.4 Linear dichroism spectra of the $TatA_{TM}$ peptide in DMPC liposomes to determine insertion. A. LD spectra were taken at three different rotating speeds; 3000, 4000 and 5000 rpm. A spectrum of the sample without spinning was subtracted from the spectra of the sample with spinning to give the LD signal, and then the spectra were corrected for light scattering. Numbers indicate the wavelengths of the positive and negative maxima. **B.** The high tension signals of the LD spectra were also recorded.

4.2.2.2 Analysis of the $TatA_{TM}$ peptide by oriented circular dichroism (OCD)

To confirm the insertion of the $TatA_{TM}$ peptide into DMPC lipid bilayers, the spectroscopic technique of OCD was employed. This technique has been used previously to analyse the insertion of the TM domain and APH of the *Bacillus subtilis* TatA homologue, TatAd, (Lange *et al.*, 2007). The determination of the insertion of α-helices into lipid bilayers using this technique relies on the fact that the $\pi \rightarrow \pi^*$ transition component at ~208 nm is polarised parallel to the helical axis. When an α-helical peptide inserts into a oriented lipid bilayer, light propagating along the axis of the helix (that is perpendicular to the lipid bilayer) is unable to be absorbed. Thus a reduction of the negative band at ~208 nm indicates that the peptide is inserted into the membrane.

To analyse the $TatA_{TM}$ peptide using this technique, a multilamellar film consisting of aligned lipid bilayers containing the peptide was placed on a quartz slide. The slide was positioned in the spectropolarimeter with the incident light beam perpendicular to the multilamellar film. To minimise artefacts due to sample inhomogeneity, the slide was rotated through 360°, taking measurements every 45°. The resulting spectra were averaged and a baseline of sample containing no lipid was subtracted to give the resulting OCD spectra, Figure 4.5. Comparison with the isotropic CD spectra of the peptide in liposomes, shows that the negative maximum at 208 nm is reduced, indicating that it is inserted in the lipid bilayer. This agrees with the LD analysis and the previous study of TatAd (Lange *et al.*, 2007).

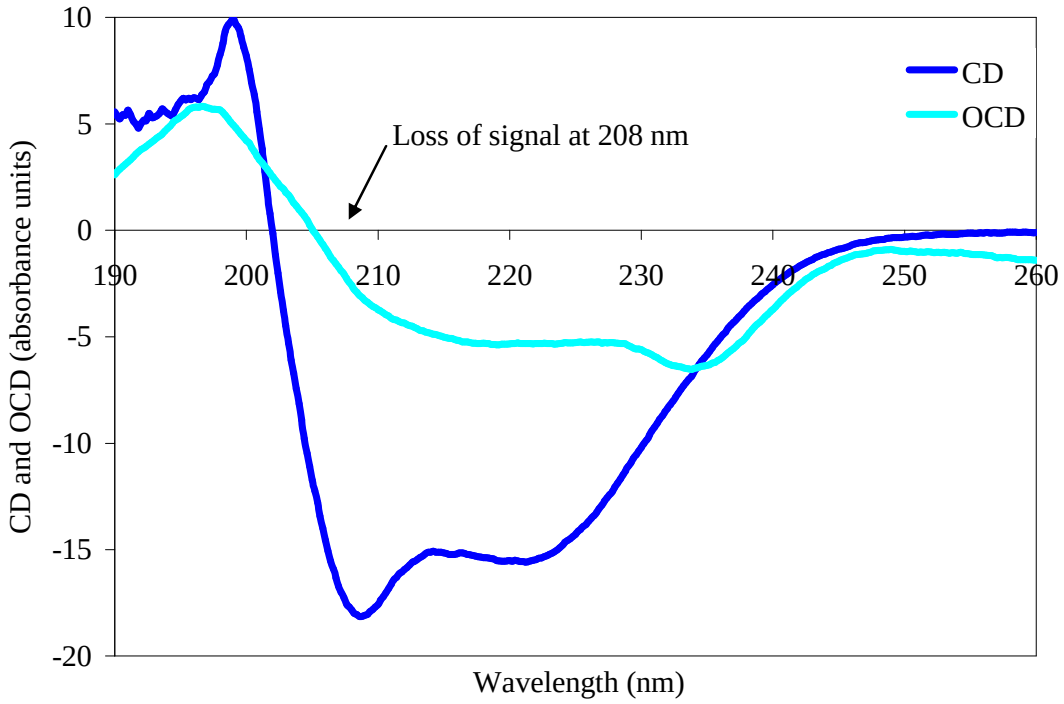


Figure 4.5 CD and OCD measurements of the $TatA_{TM}$ peptide in DMPC lipid bilayers to determine insertion. The CD spectrum (dark blue) shows a typical α-helical peptide with negative maxima at 208 and 222 nm. The loss of signal at 208 nm in the OCD spectrum (light blue) indicates insertion of the peptide into the lipid membrane.

4.2.2.3 Analysis of the $TatA_{TM}$ peptide by attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR)

The $TatA_{TM}$ peptide was analysed using ATR-FTIR, a technique which can give information on both secondary structure and insertion into lipid bilayers (Tatulian, 2003). ATR-FTIR measures the absorption of infrared light by the molecule and promotion to higher vibrational/translational energy levels. Transitions are due to stretching, bending, rocking, wagging or twisting of bonds within the molecule (Harwood & Claridge, 1997). For peptides and proteins, specific transitions of interest are the amide I and amide II vibrations of the amide bond. The amide I peak ranges from 1600 to 1700 cm^{-1} and is directly related to the backbone conformation. The amide II peak, which ranges from 1510 to 1580 cm^{-1}, results primarily from in-plane N-H bending and is often used for determination of the degree of the protein or peptide's accessibility to water since it is sensitive to deuterium exchange. For proteins in an α-helical conformation, the mean frequency of the amide I peak is 1652 cm^{-1} and for the amide II peak is 1548 cm^{-1} (Nevskaya & Chirgadze, 1976). If

the protons which contribute to the amide II peak are accessible to deuterium exchange, the amide II peak will shift due to the exchange of protons for deuterons. This gives information on the insertion of proteins into lipid membranes; protons within the lipid bilayer are protected from deuterium exchange, whilst protons exposed outside the bilayer will exchange with deuterium.

As for the OCD study, DMPC lipid bilayers containing $TatA_{TM}$ peptide were prepared, in this instance on a germanium plate. Spectra, the average of 1000 interferograms, were recorded before and after exposure to deuterium soaked nitrogen for 22 hours, Figure 4.6. Before exposure to deuterium, the FTIR spectrum of the $TatA_{TM}$ peptide revealed an amide I peak centred at 1658 cm^{-1}, which became more distinct after exposure. The peak at ~1693 cm^{-1}, although in the β-sheet region, most likely corresponds to a TFA adduct, an artefact of the peptide purification. The lack of any equivalent observable break in the 1620–1630 cm^{-1} region would indicate that at least the majority of this peak is not caused by β-sheet structure. The disappearance of the peak after 22 hours of D_2O exposure is probably due to either a removal or possible deuterium exchange of the TFA, which makes the peak shift to a lower wavenumber. The accessibility of the TFA to D_2O would suggest that it was exposed, possibly bound to the lysine residues on the C-terminus of the peptide.

To determine if the $TatA_{TM}$ peptide inserted into the DMPC lipid bilayer, the amide II peak, centred at 1544 cm^{-1}, was compared before and after exposure to deuterium. The amide II peak did not shift indicating that the protons belonging to the peptide bonds were not accessible to deuterium exchange, being shielded by the lipid membrane. Hence the peptide was inserted into the lipid bilayer. This confirms the results from the LD and OCD analyses that the $TatA_{TM}$ peptide is able to insert into the lipid bilayers.

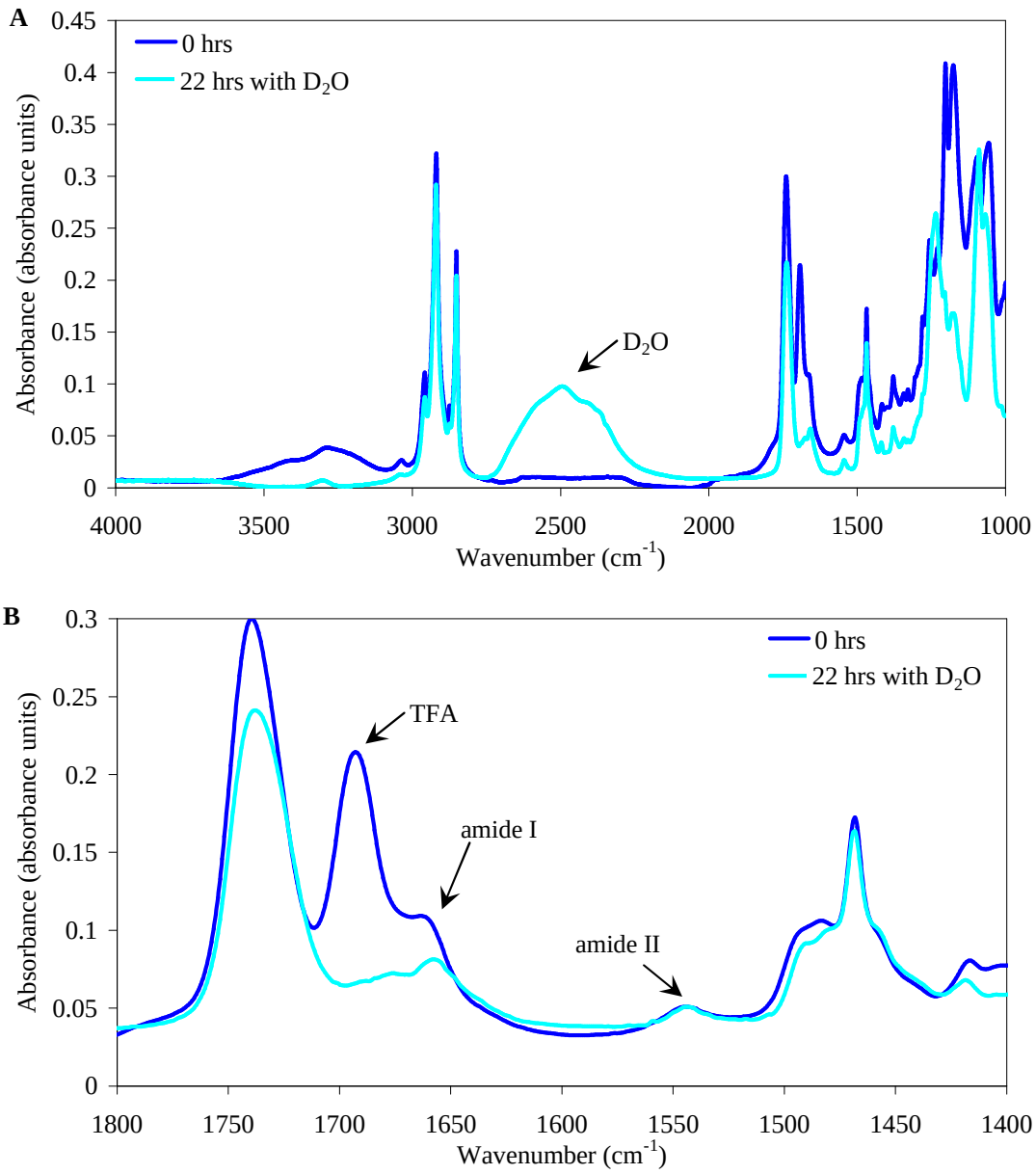


Figure 4.6 ATR-FTIR spectra of the $TatA_{TM}$ peptide in DMPC lipid bilayers. A. Spectra were taken before (dark blue) and after (light blue) exposure to deuterium soaked nitrogen for 22 hours. The peak at ~2500 cm^{-1} indicates the presence of D_2O. **B.** Overlay of the spectra focusing on the 1800–1400 cm^{-1} range shows that the amide I and II peaks do not shift after exposure to deuterium soaked nitrogen (arrows). A TFA adduct, which is an artefact from the peptide purification, is also indicated by an arrow.

To obtain an estimation of the secondary structure of the $TatA_{TM}$ peptide, the amide I peak was analysed. This is possible because the different types of secondary structures result in amide I vibrations at different frequencies, reviewed in (Tatulian, 2003). The band-narrowing tool of Fourier self-deconvolution was used on the 1700–1600 cm^{-1} region of the spectra and a baseline correction was also applied using OMNIC Spectra software. Gaussian curves were then fitted to the data using GRAMS software in order to estimate the secondary structure of the peptide, Figure 4.7. The predominant peak centred at 1657 cm^{-1} corresponded to an α-helical conformation, with additional peaks at 1691, 1676 and 1630 cm^{-1} due to β-sheet and a peak at 1648 cm^{-1} due to disordered structure. This is in agreement with CD analysis §4.2.1 where the peptide was predicted to be predominantly α-helical. The amount of β-sheet and disordered structure was larger than expected, however this may correspond to the aggregation of some of this extremely hydrophobic peptide caused by the sample preparation.

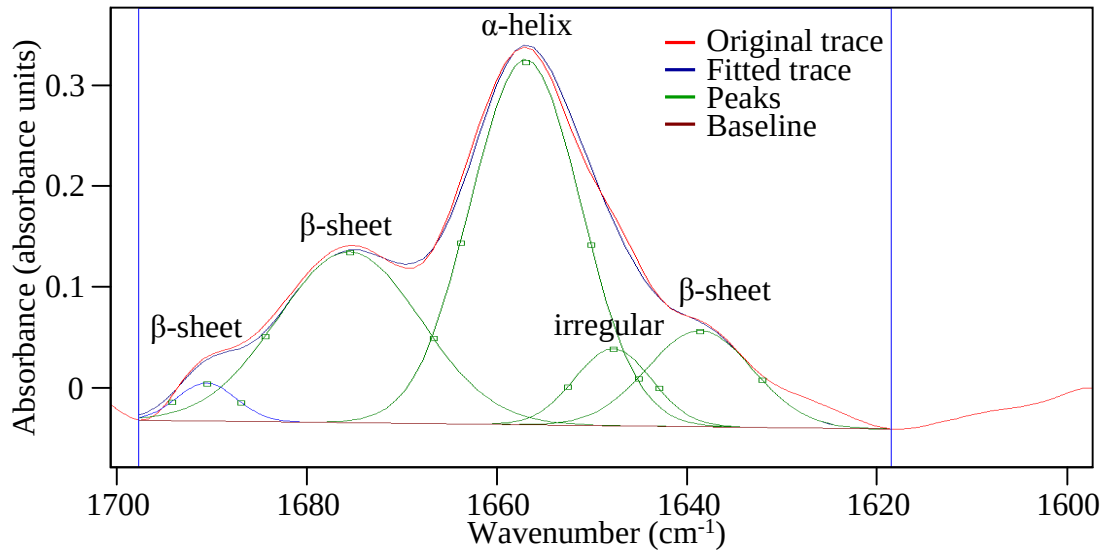


Figure 4.7 Secondary structure calculation of the $TatA_{TM}$ peptide using ATR-FTIR. The band-narrowing tool of Fourier self-deconvolution was used on the amide I peak in the 1700–1600 cm^{-1} region of the spectra after exposure to deuterium. A baseline correction was also applied. Gaussian curves were fitted to the peak and were assigned to different secondary structures according to (Tatulian, 2003).

4.2.2.4 Analysis of the $TatA_{TM}$ peptide using intrinsic fluorescence

The insertion of the $TatA_{TM}$ peptide into DMPC lipid bilayers was also confirmed by the intrinsic fluorescence properties of the native tryptophan (Trp7) in the sequence. The emission of tryptophan is strongly affected by the polarity of its local environment (Lakowicz, 2006). In an aqueous solution the fluorescence maximum of a tryptophan from an unfolded protein occurs at ~350–353 nm, shifting to ~340–342 nm for tryptophans on the surface of folded proteins and ~330–332 nm for tryptophans in a hydrophobic environment (Burstein *et al.*, 1973). Hence the insertion of tryptophan residues into a hydrophobic lipid bilayer can be detected by a blue-shift of the maximum emission peak (Christiaens *et al.*, 2002; Macek *et al.*, 1995).

Fluorescence spectra, with excitation at 280 nm, were recorded for the $TatA_{TM}$ peptide in DMPC liposomes and trifluoroethanol (TFE), and emission spectra were recorded between 300 and 400 nm, Figure 4.8. In the polar solvent TFE the emission maximum occurred at ~345 nm, indicating Trp7 was not in a completely aqueous environment. However, in the presence of liposomes the maximum shifted to ~333 nm. This blue-shift indicates that Trp7 in the $TatA_{TM}$ peptide is located in the hydrophobic lipid bilayer, confirming the results using other biophysical techniques.

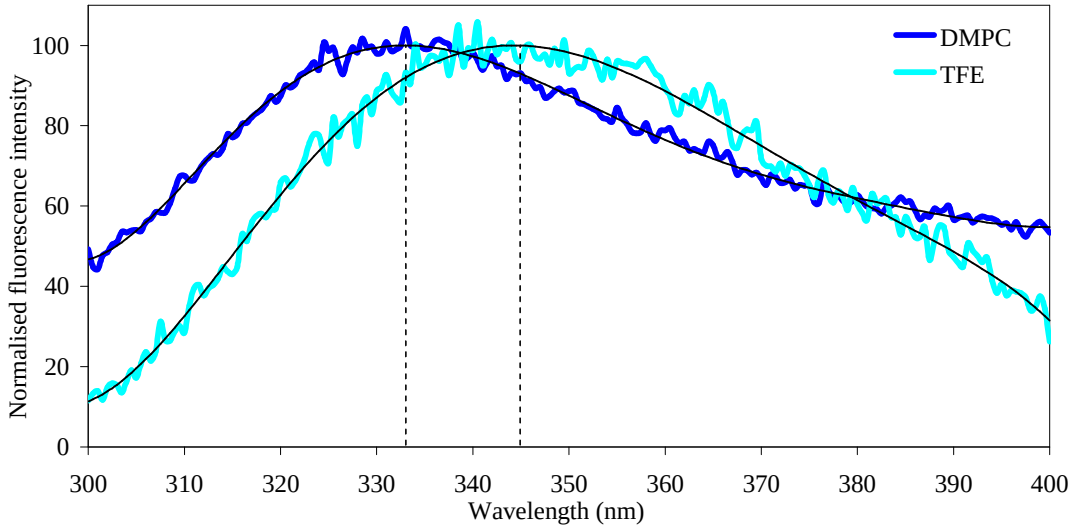


Figure 4.8 Fluorescence emission spectra of the $TatA_{TM}$ peptide in DMPC liposomes. Fluorescence spectra of the $TatA_{TM}$ peptide in DMPC liposomes (dark blue) and TFE (light blue) were recorded with excitation at 280 nm. The fluorescence of the sample without peptide was subtracted. Data were normalised and a line of best fit was calculated (black).

4.2.3 *In vitro* oligomerisation of the $TatA_{TM}$ peptide

The $TatA_{TM}$ peptide has been shown to be α-helical and is able to insert into DMPC lipid bilayers. As previously mentioned, it has been suggested that the TatA TM domain is involved in the oligomerisation of TatA (De Leeuw *et al.*, 2001; Greene *et al.*, 2007; Porcelli *et al.*, 2002). To assess the ability of the TatA TM domain to form homo-oligomers, the $TatA_{TM}$ peptide was studied using SDS-PAGE and chemical cross-linking.

4.2.3.1 Analysis of the $TatA_{TM}$ peptide oligomeric state by SDS-PAGE

In order to assess the ability of the $TatA_{TM}$ peptide to form oligomers, a range of concentrations of peptide were analysed by SDS-PAGE. This method has been widely used to assess the oligomeric state of TM domains (Jenei *et al.*, 2009; Lemmon *et al.*, 1992; Melnyk *et al.*, 2001; Oates *et al.*, 2008). When the SDS-PAGE gel was stained with Coomassie blue, a single band at ~4 kDa could be seen, Figure 4.9A. Only at the highest concentrations could a weak second band at ~6 kDa been seen. Silver stain analysis of the same gel, a technique which is 1 to 2 orders of magnitude more sensitive than Coomassie stain (Switzer *et al.*, 1979), revealed more clearly the presence of the second, higher molecular weight band, Figure 4.9B. The two bands on the $TatA_{TM}$ peptide SDS-PAGE gel, 4 and 6 kDa, do not agree exactly with the predicted monomer (***m***) and dimer (***d***) masses of 2.65 and 5.3 kDa, respectively. Similar anomalous migration has been observed for other membrane proteins (Melnyk *et al.*, 2001; Rath *et al.*, 2009; Therien *et al.*, 2001), and is thought to be caused by conformational effects and the binding of detergent. This is also seen for full length TatA, Figure 4.17B, where it runs as a band of ~15 kDa, much larger than the predicted mass of 9.6 kDa. Despite the difficulties in obtaining accurate mass approximations from SDS-PAGE, the two bands were assigned to monomer and dimer. The existence of two distinct bands implies that the exchange between the two states is slow in relation to the timescale of the experiment and is reversible (Lemmon *et al.*, 1992). This suggests that there is a monomer-dimer equilibrium for the $TatA_{TM}$ peptide.

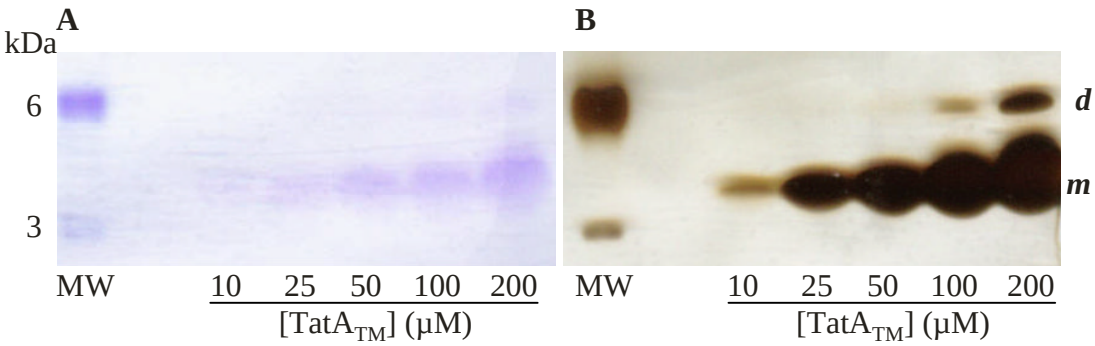


Figure 4.9 SDS-PAGE analysis of the TatA$_{TM}$ peptide to determine oligomeric state. The TatATM peptide was analysed over a range of concentrations (stated below each lane) with a molecular weight marker (M). **A.** The Coomassie stained gel revealed the peptide is mainly a monomer (*m*) (monomer MW = 2.65 kDa) with a hint of a dimer (*d*) at higher concentrations. **B.** Silver staining the same gel revealed more clearly the presence of dimers.

4.2.3.2 Analysis of the TatA$_{TM}$ peptide oligomeric state by chemical cross-linking

SDS is widely considered to denature proteins, and whilst some peptides are able to retain their secondary structure in the presence of SDS and form stable oligomers (Lange *et al.*, 2007; Melnyk *et al.*, 2001; Oates *et al.*, 2008; Therien *et al.*, 2001), TM domain interactions may be disrupted by this relatively denaturing detergent (Jenei *et al.*, 2009). In order to determine if the TatA$_{TM}$ peptide was able to form higher oligomers in less denaturing detergents such as DPC, chemical cross-linking was performed.

The peptide was solubilised in DPC detergent micelles at different peptide:detergent micelle ratios to investigate the effect of detergent concentration on oligomeric state. Cross-linking was performed with Bis(sulfosuccinimidyl)suberate (BS^3) or glutaraldehyde, which react with the NH_2 groups of lysine side chains, as well as the terminal NH_2 of the peptide, provided that they are within 11.4 Å of each other. After cross-linking, samples were analysed by SDS-PAGE and visualised by staining with silver nitrate, Figure 4.10. When no cross-linker was added, the peptide ran as a monomer and dimer (4:1 no x), as seen previously when no DPC detergent was added, Figure 4.9. When cross-linker was added, bands of higher order oligomers corresponding to up to pentamer could be seen. Due to the faintness of the bands, the

existence of higher order oligomers could not be ruled out. In all reactions the major band corresponded to monomer, with the amount of monomer increasing as the concentration of detergent increased (4:1, 1:1, 1:12.5). This agrees with previous research, which has shown that increasing detergent concentration can destabilise TM domain oligomers (Fisher *et al.*, 1999; Jenei *et al.*, 2009).

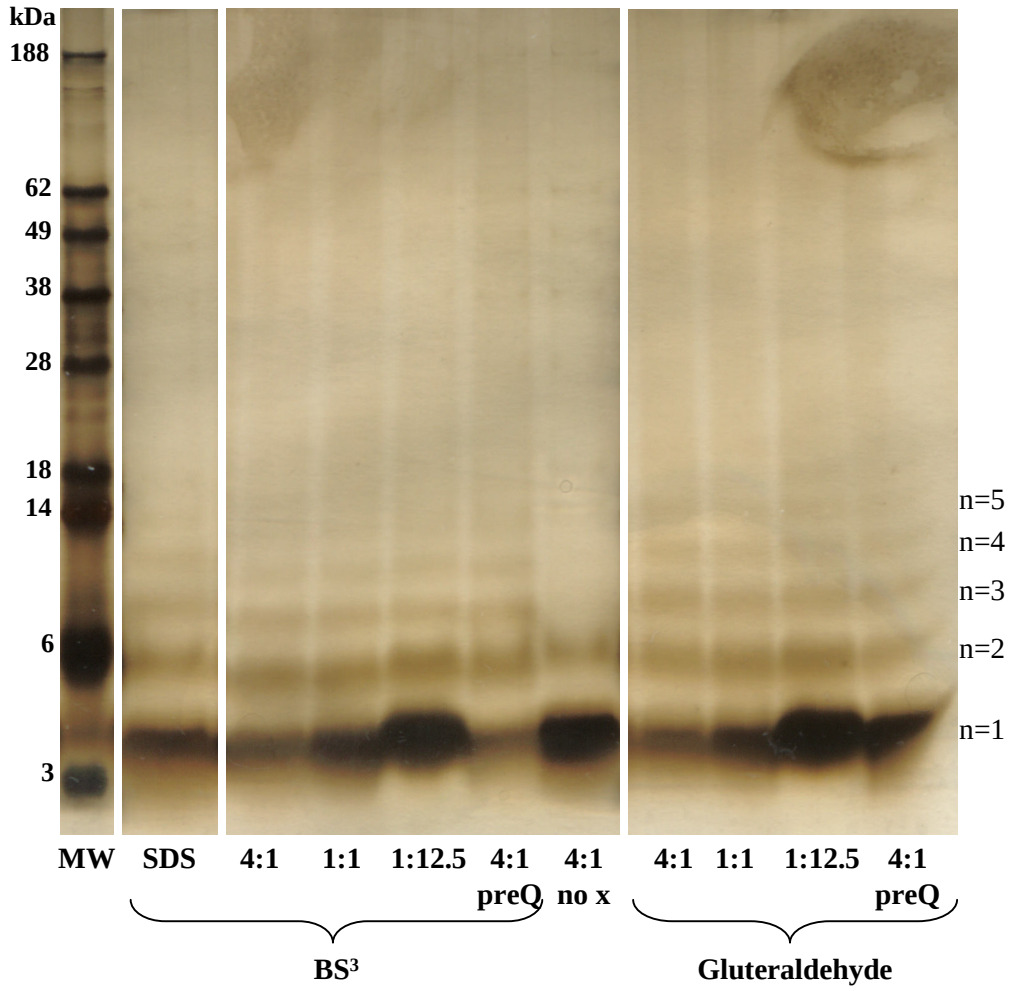


Figure 4.10 Cross-linking analysis of the $TatA_{TM}$ peptide in detergent micelles. Cross-linking was performed in DPC detergent micelles at different peptide: detergent micelle concentration ratios, indicated below each lane. Reactions were performed with BS^3 or glutaraldehyde cross-linkers, as indicated. Two reactions were quenched with 2 µM Tris-HCl pH 7.5 before addition of cross-linker (4:1 preQ), one reaction was performed without cross-linker (4:1 no x) and an additional reaction was performed in SDS. Molecular weight markers are shown in the left hand lane (MW). Proteins were visualised by staining with silver nitrate. Oligomeric states are indicated on the right (n = 1, n = 2,…).

Cross-linking was also performed in 100 mM SDS (SDS). A similar 'ladder' of higher order oligomers as seen with DPC detergent was observed. Although cross-linking in SDS can be used as a negative control (Jenei *et al.*, 2009) it was not unexpected to see these higher order oligomers since the $TatA_{TM}$ peptide is able to form SDS stable dimers. As a negative control, cross-linking was performed with samples which had been quenched before addition of cross-linker (4:1 preQ). Surprisingly higher order oligomer bands could still be visualised. This may be due to insufficient quenching. Similar 'ladders' have been seen in previous studies (Alfadhli *et al.*, 2002; Geisse *et al.*, 2002).

4.3 Analysis of the TatA TM domain using the TOXCAT assay

Analysis of the $TatA_{TM}$ peptide *in vitro* revealed the presence of dimers and higher order oligomers, but with a large population of monomers. To analyse these interactions in a natural membrane environment, as opposed to detergent micelles which can be highly denaturing (Fisher *et al.*, 2003), the TOXCAT assay was used (Russ & Engelman, 1999), described in detail in §2.12. In this system a chimeric protein consisting of the α-helical TM domain of interest inserted between the N-terminal dimerisation-dependent DNA binding domain of ToxR, and the monomeric periplasmic anchor maltose binding protein (MBP) is used. The chimera inserts into the membrane and interaction between the helical TM domain leads to the expression of chloramphenicol acetyl transferase (CAT), which can be measured. The levels of CAT expression correspond to the strength of interaction between the TM domains. The TM domain of Glycophorin A (GpA), which is known to strongly dimerise, and its dimerisation defective mutant (G83I), are typically used as a positive and negative control, respectively.

Previous studies using the TOXCAT assay (Jenei *et al.*, 2009; Langosch *et al.*, 1996; Li *et al.*, 2004) have shown that the length of the TM domain insert can affect ToxR-based transcriptional activity. For this reason, it was decided to insert different lengths of the TatA TM domain into the TOXCAT chimera, Figure 4.11. Five TOXCAT chimera with TM domain inserts of 16 to 20 residues were constructed (TM_{16}–TM_{20}) and expressed in *malE* deficient NT326 cells, which do not express any endogenous MBP. Various controls were performed to determine the expression level, correct insertion and orientation of TM_{16}–TM_{20}. The GpA and G83I TOXCAT

chimera, both of which correctly insert into the membrane, were used for comparison.

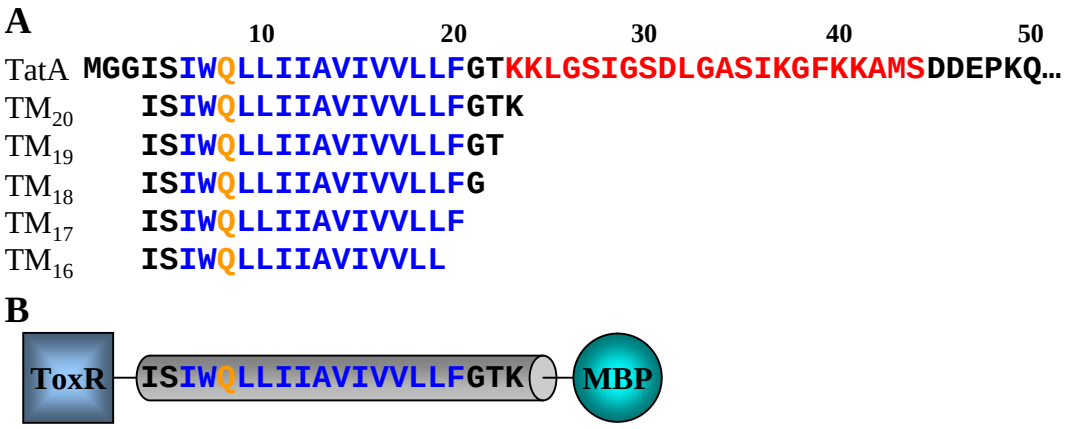


Figure 4.11 TatA TM domain TOXCAT constructs used in this study. A. The first 50 residues of the TatA primary sequence (TatA) with the predicted TM domain (blue) and APH (red); and the five different lengths of TatA TM domain inserted into the TOXCAT chimera (TM_{20}–TM_{16}). The Gln8 is also highlighted (orange). **B.** The TM_{20} TOXCAT chimera consisting of the α-helical TM domain inserted between the N-terminal dimerisation-dependent DNA binding domain of the ToxR protein (■) and the periplasmic anchor MBP (●).

To check the expression levels of TM_{16}–TM_{20}, and the positive and negative controls, Western blots of whole cells expressing the chimera were performed, Figure 4.12A. Expression levels were comparable to GpA and G83I. The correct insertion of the TOXCAT chimera in the membrane was confirmed by sodium hydroxide washes (Chen & Kendall, 1995), Figure 4.12B. Cells expressing the chimera were treated with ice-cold sodium hydroxide and then centrifuged. Cytoplasmic and periplasmic proteins as well as proteins loosely associated with the membrane were found in the supernatant (S), whilst proteins stably associated with the membrane remained in the pellet (M). Western blots of the supernatant and pellet, as well as untreated whole cells (W) reveal that GpA and TM_{20} are located exclusively in the pellet. This indicates that the chimera is inserted across the membrane. The correct orientation of the chimera was confirmed by protease sensitivity in a spheroplast assay and *malE* complementation on agar plates (Russ & Engelman, 1999). For the spheroplast assay spheroplasts (Sp) were prepared from cells expressing the TOXCAT chimera. These

were then either treated with Proteinase K (Sp') or broken open by freeze-thaw before protease treatment (B'). If the chimera is correctly inserted, MBP should be located on the outside of the spheroplast and accessible to Proteinase K. When treated with protease, the MBP is cleaved and will run as a lower molecular weight band than the intact chimera. Analysis of cells expressing TM_{20} show that the MBP is cleaved by treatment with protease, Figure 4.12C, as demonstrated by the existence of the lower molecular weight band in lanes Sp' and B'. This indicates that the TM_{20} is inserted in the correct orientation. This was also confirmed by *malE* complementation on agar plates. NT326 cells, which lack MBP, cannot transport maltose into the cytoplasm for metabolism, hence cannot grow on plates where the only carbon source is maltose. If the TOXCAT chimera is correctly inserted, the MBP is located in the periplasm and cells are able to grow on maltose. NT326 cells expressing TM_{16}–TM_{20} and GpA were able to grow on M9 agar plates containing 0.4% maltose, whilst NT326 cells not expressing the chimera were unable to grow, Figure 4.12D. These checks confirm that the TM_{16}–TM_{20} chimera are expressed at comparable levels to GpA and G83I, and insert across the cell membrane with the correct topology.

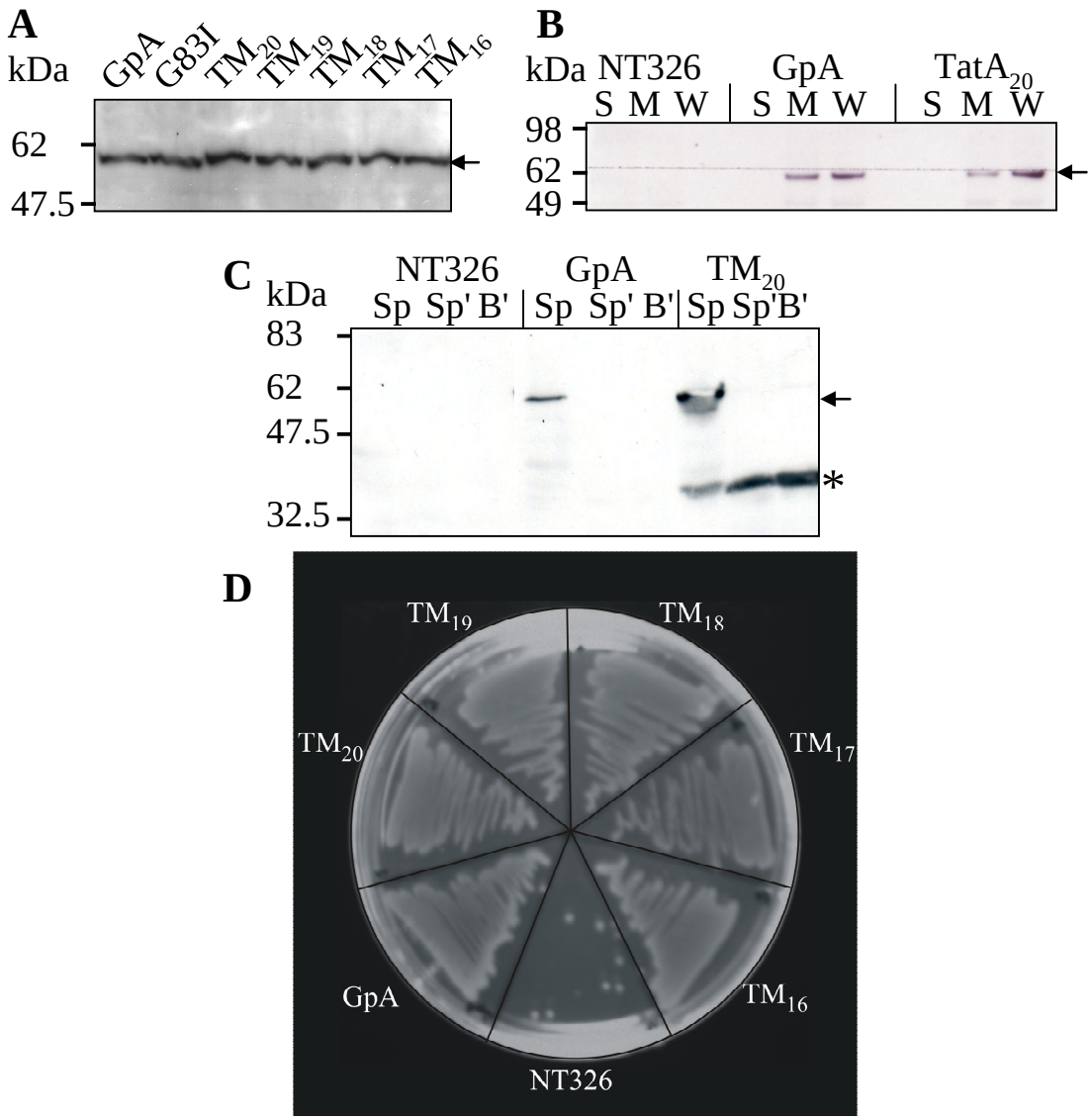
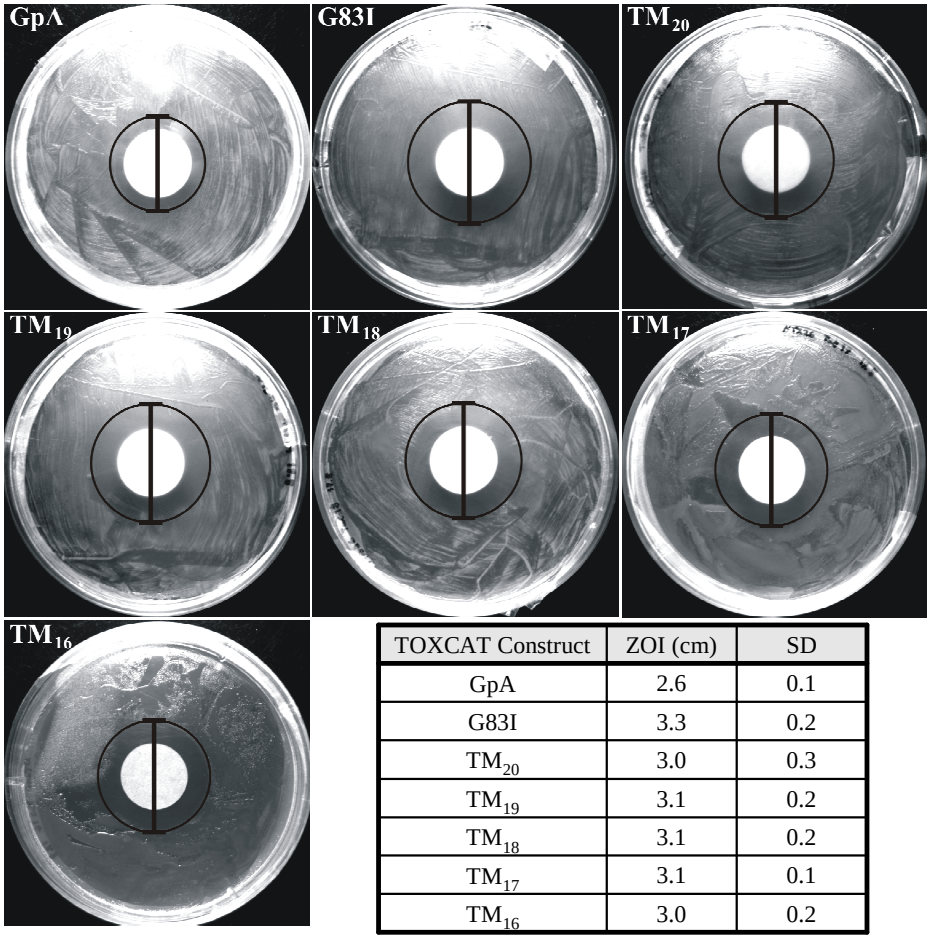


Figure 4.12 Expression, insertion and orientation checks for the TatA TM domain TOXCAT chimera. A. Immunoblot using anti-MBP antibodies of whole NT326 cells expressing the TatA TM TOXCAT constructs (TM_{16}–TM_{20}). The chimera expression levels were compared with the chimera containing the TM domain of GpA, which is known to strongly dimerise, or its dimerisation defective mutant (G83I). Arrow indicates the TOXCAT chimera. **B.** Immunoblot using anti-MBP antibodies of NaOH washes. Cells expressing the TOXCAT chimera were treated with ice-cold NaOH, then centrifuged to separate the supernatant (S), which contained cytoplasmic, periplasmic and loosely-membrane associated proteins, from the pellet (M), which contained proteins stably associated with the membrane. Untreated whole cells (W) were also analysed. Cells not expressing the chimera (NT326) were used as a negative control. Arrow indicates the TOXCAT chimera **C.** Immunoblot using anti-MBP antibodies of protease sensitivity of spheroplasts assay. Spheroplasts (Sp) of cells were treated with Proteinase K (Sp') or broken open by freeze-thaw before treatment with Proteinase K. Arrow indicates TOXCAT chimera and asterisk indicates MBP cleaved by proteolysis. **D.** *malE* complementation assay. *malE* deficient NT326 cells containing the TatA TM and GpA TOXCAT chimera were grown on M9 agar plates containing 0.4% maltose.

The levels of CAT expression in NT326 cells expressing TM_{16}–TM_{20} were initially estimated by a disk diffusion assay. Expression of CAT gives cells resistance to the antibiotic chloramphenicol (CAM). By placing a disk soaked in CAM onto the centre of an agar plate of cells, an antibiotic gradient is formed with the highest concentration closest to the disk. Only cells expressing high levels of CAT are able to grow close to the disk, hence the zone of inhibition of growth (ZOI) can give an initial indication of the CAT expression levels. The ZOI was recorded for the five TatA TM constructs and compared to the positive and negative controls of GpA and G83I, respectively, Figure 4.13. TM_{16}–TM_{20} were able to grow closer to the disk than G83I, however not as close as GpA, indicating an intermediate level of TM domain interaction.



TOXCAT Construct	ZOI (cm)	SD
GpA	2.6	0.1
G83I	3.3	0.2
TM_{20}	3.0	0.3
TM_{19}	3.1	0.2
TM_{18}	3.1	0.2
TM_{17}	3.1	0.1
TM_{16}	3.0	0.2

Figure 4.13 Disk diffusion assay for the TatA TM domain. NT326 cells expressing the TatA TM domain TOXCAT constructs (TM_{16}–TM_{20}) were grown on agar plates with a chloramphenicol soaked disk in the centre, along with the positive and negative controls of GpA and G83I respectively. The zone of inhibition of growth (ZOI) was recorded (black circles and lines) and is given in the table along with the standard deviation (SD) of different measurements.

CAT expression levels for the constructs were measured more accurately using the quantitative CAT assay as detailed in §2.12.2.2. Since TM_{17}–TM_{19} gave similar disk diffusion results, Figure 4.13, only TM_{16}, TM_{18} and TM_{20} were analysed using this method, alongside GpA and G83I as controls. The resulting [CAT] activity was normalised to GpA, Figure 4.14. All three lengths of TatA TM domain (TM_{16}, TM_{18}, TM_{20}) yielded relatively low levels of CAT in the assay ($[CAT]_{TM16}$ = 33%, $[CAT]_{TM18}$ = 27% and $[CAT]_{TM20}$ = 38% of GpA) suggesting that there is some level of interaction between TatA TM domains, but that the strength of this interaction is relatively weak when compared to the positive control, GpA. This is in keeping with the fact that GpA forms SDS stable dimers (Lemmon *et al.*, 1992), whereas the TatA TM domain is mainly monomeric in SDS.

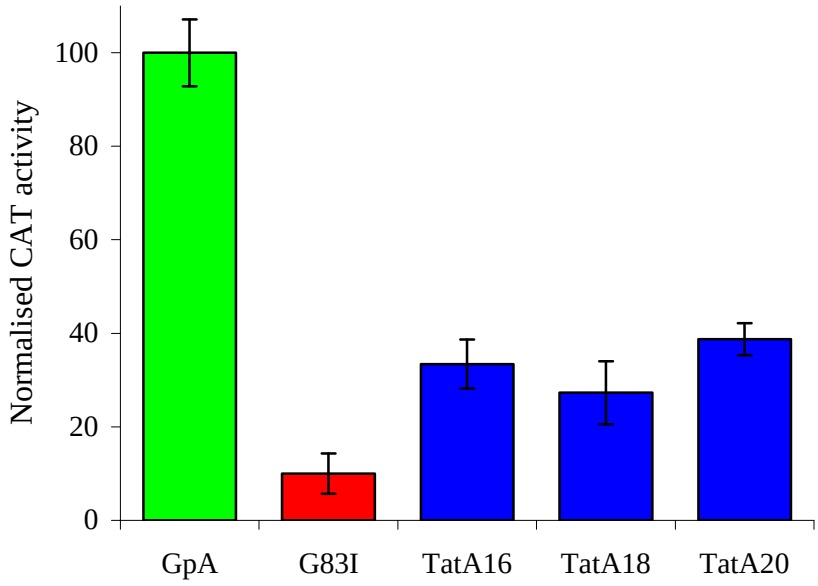


Figure 4.14 Quantitative CAT assay of TatA TM domain interactions. Levels of chloramphenicol acetyltransferase (CAT) activity in NT326 cells expressing the TOXCAT chimera with the different length TatA TM domain insert (TM_{16}, TM_{18}, TM_{20}) were assayed (as detailed in §2.12.2.2) along with GpA and its dimerisation defective mutant (G83I) as positive and negative controls, respectively. CAT activity was normalized to GpA. Error bars show the standard deviation of three separate measurements.

4.4 The role of the glutamine (Gln8) in the TatA TM domain

A recent cysteine scanning study of the TM domain and APH of TatA highlighted a glutamine residue, Gln8, in the TM domain which was important for function, Figure

4.11A (Greene *et al.*, 2007). Substitution of this residue with cysteine or alanine severely affected the export of the Tat substrate TorA to the periplasm. The presence of hydrophilic residues in TM domains, and specifically glutamine residues, has been shown to drive TM helix-helix association via electrostatic interactions and formation of interhelical hydrogen bonds (Dixon *et al.*, 2006; Freeman-Cook *et al.*, 2004). We therefore set out to test whether this residue plays a role in helix-helix interactions within the TatA complex.

4.4.1 Analysis of the TatA Gln8 using the TOXCAT assay

The role of Gln8 in TatA TM domain oligomerisation was studied by mutagenesis of this residue and subsequent remeasurement using the TOXCAT assay. The Gln8 in TM_{20}, which gave the highest [CAT] activity ($[CAT]_{TM20}$ = 38% of GpA), Figure 4.14, was substituted by cysteine (Q8C) and alanine (Q8A). Control experiments were carried out to ensure the chimera were expressed at similar levels to GpA and G83I and inserted with the correct topology as detailed in §2.12.1. The Gln8 mutants were expressed at similar levels to the TM_{20} TOXCAT chimera, Figure 4.15A, and inserted into the membrane with the correct orientation as assessed by *malE* complementation, Figure 4.15B.

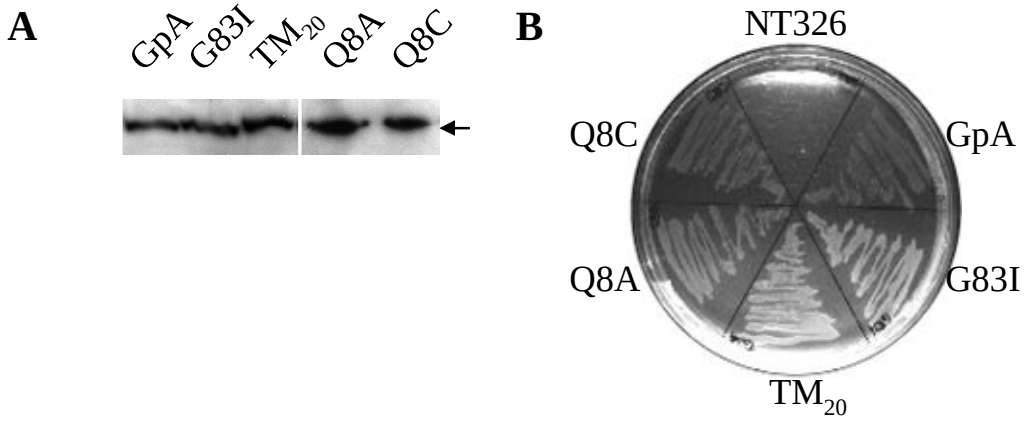


Figure 4.15 Expression, insertion and orientation checks for the Gln8 mutant TOXCAT chimera. A. Immunoblot using anti-MBP antibodies of whole NT326 cells expressing the TatA TM TOXCAT construct (TM_{20}) and the Gln8 mutants (Q8A, Q8C). The expression levels were compared with the chimera containing the TM domain of GpA, which is known to strongly dimerise, and its dimerisation defective mutant (G83I). Arrow indicates the TOXCAT chimera. **B.** *malE* complementation assay. *malE* deficient NT326 cells containing the TatA TM and GpA TOXCAT chimera were grown on M9 agar plates containing 0.4% maltose.

Levels of CAT expression were measured as detailed in §2.12.2.2. The resulting signals were normalised to GpA and compared with wild-type TM_{20}, Figure 4.16. Although these substitutions lead to a slight decrease in [CAT] activity compared to TM_{20} levels ($[CAT]_{Q8A}$ = 24% and $[CAT]_{Q8C}$ = 28%), the result is only significant within the error for Q8A.

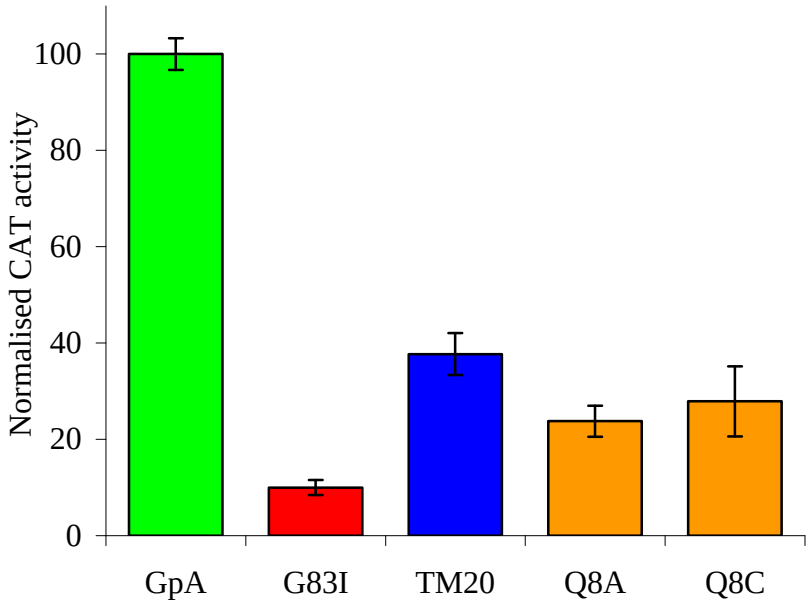


Figure 4.16 Quantitative CAT assay of Gln8 mutants. Levels of CAT activity in NT326 cells expressing the TatA TOXCAT chimera (TM_{20}) and TM_{20} with the Gln8 substituted with alanine or cysteine (Q8A, Q8C) were assayed (as detailed in §2.12.2.2) along with GpA and its dimerisation defective mutant (G83I) as positive and negative controls, respectively. CAT activity was normalized to GpA. Error bars show the standard deviation of three separate measurements.

4.4.2 Analysis of the functional role of the Gln8

Since the Gln8 substitutions did not dramatically reduce the level of TatA TM domain oligomerisation as measured by the TOXCAT assay, these substitutions were analysed in full length TatA using a different expression system from the previous study mentioned above (Greene *et al.*, 2007). The Gln8 of TatA was substituted with alanine or cysteine in the arabinose inducible plasmid pBAD-ABC, and the mutants were analysed for effects on the function of the Tat pathway.

The function of the mutated TatA was analysed by a native gel based trimethylamine *N*-oxide (TMAO) reductase (TorA) activity assay. TorA is a 90 kDa *E. coli* molybdoenzyme which is translocated across the plasma membrane by the Tat pathway. It catalyses the reduction of TMAO to trimethylamine and has been shown to be functional in the cytoplasm and periplasm, thus lending itself to analysis of the functionality of the Tat pathway (Santini *et al.*, 1998). In this study cells were separated into periplasmic and cytoplasmic fractions, and then subjected to native PAGE. The presence of TorA in the gel was confirmed by an enzyme linked redox reaction (Silvestro *et al.*, 1989). The gel was stained blue by reduced methyl viologen then incubated with TMAO. In the presence of TorA the TMAO was reduced and consequently the methyl viologen was oxidised, revealing a 'white' band.

The *E. coli tat* deletion strain (Δ*ABCDE*) cells containing pBAD-ABC or mutated pBAD-ABC were grown in the presence or absence of arabinose before fractionation. Cells were grown with or without arabinose in order to analyse any concentration-dependent effects of the substitutions. A previous study (Bolhuis *et al.*, 2001) has shown that induction with low levels of arabinose (5 µM) provides near wild-type levels of TatABC, and Tat component levels are slightly lower than wild-type levels if arabinose is omitted. An approximately 25-fold increase over wild-type levels is observed when the pBAD-ABC plasmid is induced with 50 µM concentrations, which is similar to the expression levels reported in other studies of plasmid-borne Tat proteins (Buchanan *et al.*, 2002). In this study, when grown in the absence of arabinose, the plasmid borne Tat components were expressed at very low levels, when compared to the wild-type MC4100 cells, with moderate overexpression shown in the presence of arabinose, Figure 4.17B.

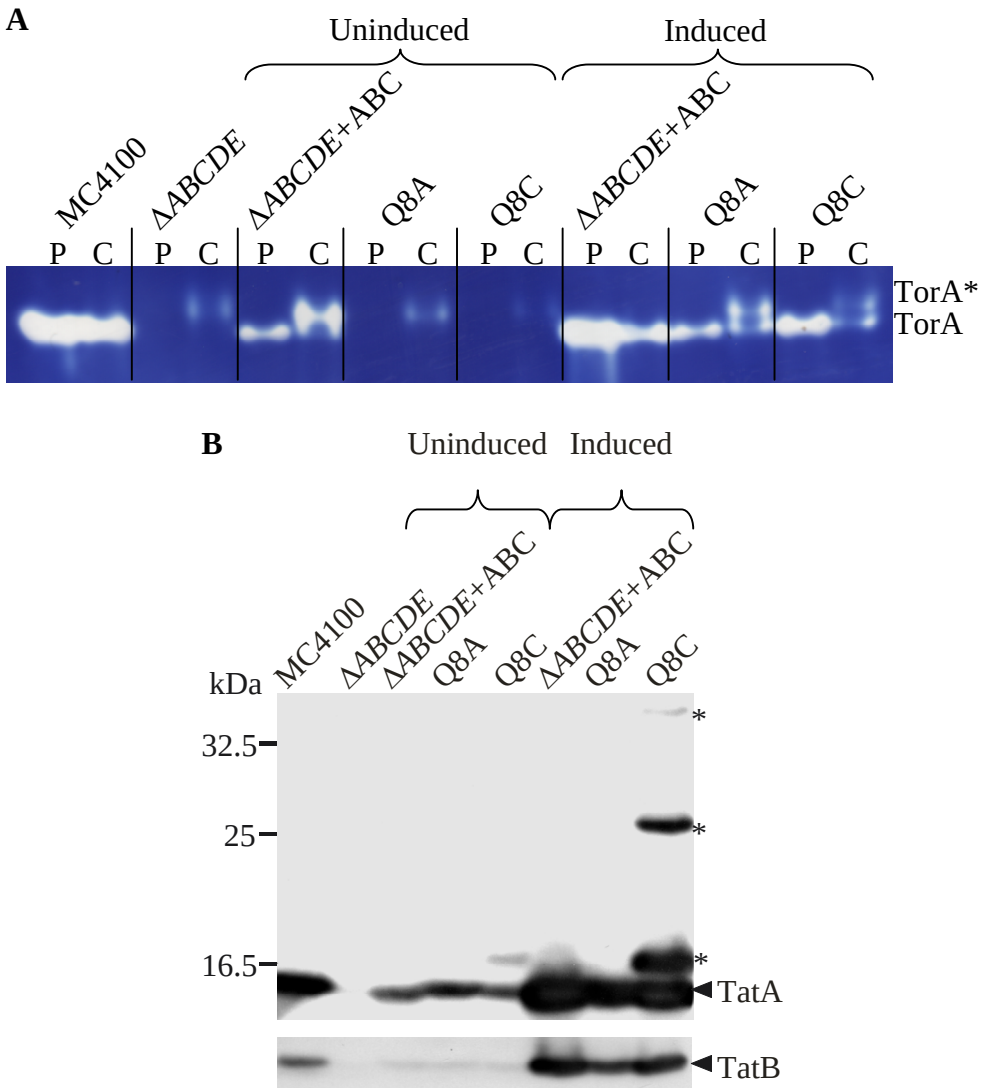


Figure 4.17 TorA export assay of the TatA Gln8 mutants at different expression levels. Periplasmic and cytoplasmic samples (P,C) were prepared from wild-type *E. coli* cells (MC4100), from Δ*ABCDE* cells, and from Δ*ABCDE* cells expressing wild-type TatABC from plasmid pBAD-ABC (Δ*ABCDE*+ABC), or the same vector with TatA Gln8 substituted with alanine or cysteine (Q8A, Q8C), grown in the absence or presence of arabinose (uninduced, induced). **A.** Native polyacrylamide gels of periplasmic and cytoplasmic samples were stained for TorA activity. White bands indicate the presence of TorA, TorA* indicates a slower migrating cytoplasmic form of TorA. **B.** Immunoblots of membrane samples using TatA and TatB antibodies. Arrows indicate TatA and TatB. Asterisks denote unidentified TatA adducts caused by the Gln8 to Cys substitution.

In the TorA assay of the control samples, Figure 4.17A, TorA is clearly apparent in the periplasm (P) of MC4100 cells and no export is evident in Δ*ABCDE* cells, where a slowly migrating form of TorA (TorA*) is observed in the cytoplasm (C) (Barrett *et al.*, 2003a; Barrett & Robinson, 2005). When arabinose was omitted from the

growth medium, resulting in low expression levels, wild-type TatA (Δ*ABCDE* +ABC) was able to support export of TorA to the periplasm, but no export was detectable with the mutated forms (Q8A, Q8C). This result agrees with the previous work, using a different expression system, which highlighted the Gln8 as being important for function (Greene *et al.*, 2007). However, when arabinose was used to induce the expression of the Tat components, the TatA Q8A and Q8C mutants were both able to support export of TorA to the periplasm, as indicated by white bands in the periplasmic lanes of Figure 4.17A. Although this is a quantitative assay, the export levels of the mutants, even when induced, seem to be reduced, indicating that the substitution is still having an effect at higher expression levels.

This result suggests that the Gln8 is important but not essential for functional activity. When the mutated TatA is expressed at higher levels, the increased concentration (~25-fold increase over wild-type) is able to compensate for the effect caused by the lack of the Gln8. This is not totally unexpected as the Gln8 mutants still oligomerise in the TOXCAT assay. A previous cysteine-scanning study (Greene *et al.*, 2007) showed that the same substitutions blocked export completely, but it is possible that the mutants were expressed at very low concentrations in that study. Furthermore, Greene *et al.* used a different assay system to that used in this study. Previous research has shown differences in the conclusions drawn from the results of these two different assays (Barrett *et al.*, 2003a).

4.4.3 Analysis of the Gln8 in the TatA TM domain by blue-native PAGE

Blue-native (BN) PAGE has been used in several studies to analyse the size of wild-type and mutant Tat complexes (Barrett *et al.*, 2005; Behrendt *et al.*, 2007; Oates *et al.*, 2005; Orriss *et al.*, 2007). By moderately overexpressing the Tat components, it is possible to gain information on the oligomeric state of TatA mutants. High protein concentrations are needed for determination of protein oligomeric state, so due to low expression levels, there is currently no technique sensitive enough to visualise TatA oligomers at wild-type levels. Moderate overexpression of the Tat components has been used in several studies of TatA (Berthelmann *et al.*, 2008; Gohlke *et al.*, 2005; Oates *et al.*, 2005).

Membranes isolated from Δ*ABCDE* cells expressing the wild-type and mutated pBAD-ABC plasmid were analysed by BN-PAGE. Immunoblots were performed with anti-TatA antibodies to detect the TatA complexes, Figure 4.18. Expression of the Tat components was at a similar level to the functional, induced samples in the TorA assay, Figure 4.17B. The 'ladder' of bands corresponding to different sized wild-type TatA complexes can clearly be seen (Δ*ABCDE*+ABC). The TatA mutants were both still able to form oligomers identical to the wild-type TatA (Q8A, Q8C), indicating that at these expression levels substitution of the Gln8 does not have an effect on complex formation.

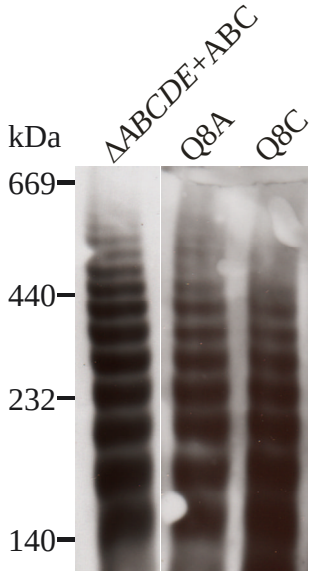


Figure 4.18 Blue-native PAGE analysis of the Gln8 mutants. Membranes solubilised in digitonin were subjected to blue-native gel analysis (as described in §2.10.3) and immunoblotted using anti-TatA antibodies. The mobility of molecular mass markers in kDa are indicated on the left.

4.5 Discussion

The work presented in this chapter sought to analyse the role of the TatA TM domain and a glutamine residue in the TM domain (Gln8) in the formation of TatA complexes. The importance of the TM domain in oligomerisation and function of TatA has been demonstrated indirectly in previous studies (De Leeuw *et al.*, 2001; Greene *et al.*, 2007; Porcelli *et al.*, 2002). However, at the start of this study there

was no direct evidence to support the hypothesis that the TatA TM domain, and more specifically the Gln8, were involved in oligomerisation.

4.5.1 The $TatA_{TM}$ peptide is α-helical and inserts into lipid bilayers

A number of approaches were used to investigate the role of the TatA TM domain. A synthetic peptide corresponding to the TatA TM domain was synthesised, purified and analysed *in vitro* using a variety of techniques. CD showed that the $TatA_{TM}$ peptide has a predominantly α-helical secondary structure in DPC detergent micelles and DMPC liposomes. The secondary structure of the $TatA_{TM}$ peptide in DMPC lipid bilayer films was confirmed by ATR-FTIR, with the appearance of amide I and amide II peaks characteristic of an α-helical structure. This agrees with the predicted secondary structure and with a study of the predicted TM domain of the TatA homologue from the Gram-positive *B. subtilis*, TatAd (Lange *et al.*, 2007).

Due to the lack of consensus on the exact cellular location of TatA from *E. coli* (Berthelmann *et al.*, 2008; De Leeuw *et al.*, 2001; Leake *et al.*, 2008; Sargent *et al.*, 2001), it was important to confirm that the $TatA_{TM}$ peptide inserted into lipid bilayers. Several approaches were used to analyse the insertion of the $TatA_{TM}$ peptide into DMPC lipid bilayers. LD and fluorescence spectroscopy demonstrated for the first time that the $TatA_{TM}$ peptide inserts into DMPC liposomes, and OCD and deuterium exchange ATR-FTIR showed that this insertion is replicated in lipid bilayer films. Although these analyses do not demonstrate that TatA is a membrane protein, it does show that *in vitro* the predicted TM domain does preferentially and spontaneously insert into membranes.

4.5.2 TatA TM domains interacts relatively weakly with one another

SDS-PAGE analyses revealed that the $TatA_{TM}$ peptide exists in a monomer-dimer equilibrium. Chemical cross-linking with both BS^3 and glutaraldehyde in less denaturing detergents stabilised the formation of these dimers and revealed the existence of higher order oligomers. Increasing the amount of detergent micelles present did not affect the formation of the higher order oligomers, however the

relative amount of monomer did increase. Detergents have been shown to have a destabilising effect on oligomers (Fisher *et al.*, 1999; Fisher *et al.*, 2003; Jenei *et al.*, 2009), with the dissociation constant increasing as the detergent concentration increases, resulting in a preference for monomeric peptide. Surprisingly, the higher order oligomers were still detectable in reactions which had been quenched with an excess of Tris-HCl buffer before addition of cross-linker. This could have been due to insufficient quenching or cross-linking occurring before the reaction had been properly quenched. Although this does bring in to question the validity of the perceived higher oligomers, their existence for both of the cross-linkers does suggest they are not artefacts, and the problem lies with the quenching of the reaction. Similar 'ladders' of oligomers have been observed using these cross-linkers (Alfadhli *et al.*, 2002; Geisse *et al.*, 2002) where the equilibrium of the banding was related to the cross-linker concentration.

Helix-helix interactions of the TM domain were also analysed in a native membrane environment using the TOXCAT assay, which has been highly effective in the study and quantification of TM interactions (Dixon *et al.*, 2006; Freeman-Cook *et al.*, 2004; Russ & Engelman, 1999; Zhou *et al.*, 2001). Varying the length of the TM domain inserted into the TOXCAT chimera had little effect for the TatA TM domain, with all chimera tested yielding low but significant levels of interaction when compared to the positive and negative controls of GpA and G83I, respectively.

4.5.3 The TatA Gln8 is not essential for function or complex formation

The effects of the glutamine in the TatA TM domain on function and complex were analysed by substitution with alanine and cysteine in the TOXCAT assay and full length TatA expressed from the pBAD24 plasmid. Although substitution of the Gln8 with alanine or cysteine in the TOXCAT assay did result in a small reduction in [CAT], it was only significant within the error for the Q8A substitution. This small drop in strength of association was not as large as seen previously for glutamine substitutions (Dixon *et al.*, 2006; Freeman-Cook *et al.*, 2004), and indicates that substitution of the Gln8 does not severely disrupt the ability of the TM domain to form homo-oligomers.

Analysis of the Gln8 mutants (Q8A, Q8C) in full length TatA revealed concentration dependent effects. At low expression levels, the TatA mutants were not able to export the Tat substrate TorA (in agreement with previous reports). However, at higher expression levels, increased concentration of TatA is able to compensate for the effects of the mutants and reinstate functional activity. These results are not in full agreement with a previous cysteine scanning study (Greene *et al.*, 2007), where substitution of the Gln8 with alanine or cysteine abolished export of TorA. These differences could be explained by the difference in expression systems and assays used. Furthermore, in the previous study it is possible that the expression levels were quite low and in this case would agree with the results presented here. Different results have also been reported for the different assay systems (Barrett *et al.*, 2003a; Weiner *et al.*, 1998) and may be due to the sensitivity of the assays. The Gln8 substitutions did not have any visible effect on the formation of oligomers, as assessed by BN-PAGE. However it is worth noting that the mutated TatA complexes analysed were functional. Due to the sensitivity of BN-PAGE, it was not possible to analyse functionally deficient mutated TatA complexes.

The results presented in this chapter suggest that although the TatA TM domain does have a propensity to interact, it is only partly involved in the formation of the wide range of TatA complexes. These results agree with a recent study of the TatA plant homologue, Tha4, where it has been shown that both the TM and C-terminal domains are involved in oligomerisation in thylakoid membranes (Dabney-Smith & Cline, 2009). Other regions of the TatA protein may be involved cooperatively with the TM domain in homo-interactions and formation of complexes. Analysis of the $TatA_{TM}$ peptide has shown that the wide variety of techniques give comparable results, thus demonstrating that these techniques are suitable for the analysis of short, very hydrophobic peptides. Combination of the data clearly show, for the first time, that the TatA TM domain preferentially inserts into lipid bilayers, and goes some way to proving that TatA is a membrane protein.

Chapter 5. Analysis of the *Escherichia coli* TatA Amphipathic Helix

5.1 Introduction

As mentioned previously, the secondary structure of *E. coli* TatA is predicted to have an N-terminal TM domain, followed by a second α-helical domain and a largely unstructured C-terminal domain, Figure 1.6. Plotting the sequence of the second helical domain on a helical wheel, Figure 5.1, reveals the amphipathic nature of the helix, with the bulk of the nonpolar residues all located on the same helical face. The role of the APH is poorly understood, and it has not been widely studied in isolation from the rest of the protein.

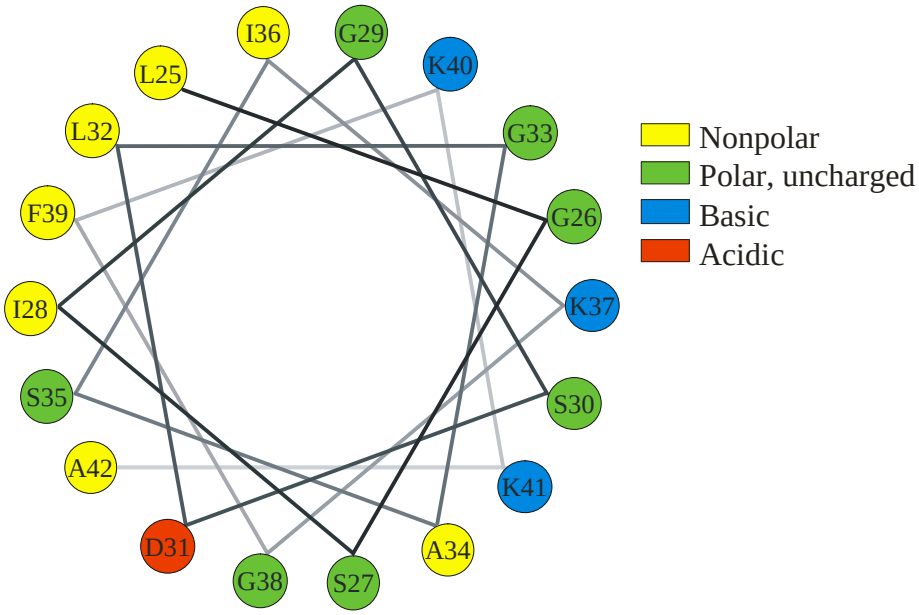


Figure 5.1 Helical wheel of the TatA APH. Residues Lys25 to Ala42 of TatA, corresponding to the predicted APH, mapped onto a helical wheel. Nonpolar residues are shown in yellow, polar uncharged residues in green, basic residues in blue and an acidic residue in red. The hydrophobic nonpolar residues are mainly located together on one face of the helix.

A recent cysteine scanning analysis has shown that the APH is very sensitive to single substitution of amino acids (Greene *et al.*, 2007), and a triple mutation of three of the lysines in the sequence also blocks export (Barrett & Robinson, 2005). Deletion of the C-terminus of TatA does not affect function, however deletion of only an additional ten residues of the APH leads to a loss of export (Lee *et al.*, 2002).

Research suggesting that TatA may have a dual topology has led to the hypothesis that the APH is able to insert into the membrane as part of the translocation process (Chan *et al.*, 2007; Gouffi *et al.*, 2004), possibly with the hydrophilic face of the helix forming a channel through which the Tat substrate can be exported. It has also been shown that TatA expressed without its TM domain is still able to interact with lipid monolayers (Porcelli *et al.*, 2002), although it was not determined if the APH was inserting into the monolayer or simply lying on the surface. The secondary structure of the APH from the *B. subtilis* TatA homologue, TatAd, has also been shown to depend on detergent and the composition of lipids present (Lange *et al.*, 2007).

Here, the *E. coli* TatA APH was characterised using *in vitro* and *in vivo* techniques. Since synthetic peptides corresponding to amphipathic helices have been used previously in order to analyse their role in isolation (Brass *et al.*, 2002; Cornut *et al.*, 1996; Srinivas *et al.*, 1992), it was decided to use a synthetic peptide of the TatA APH. Several biophysical techniques were used to analyse the behaviour of the synthetic peptide corresponding to the TatA APH in a variety of aqueous, detergent and lipid environments. The secondary structure and possible insertion of the APH into lipid bilayers were investigated. The insertion into native *E. coli* membranes was also analysed using the TOXCAT assay (Russ & Engelman, 1999).

5.2 Analysis of a TatA APH synthetic peptide

A synthetic peptide of the TatA APH ($TatA_{APH}$) corresponding to residues Thr_{22} to Pro_{48} of TatA (TKKLGSIGSDLGASIKGFKKAMSDDEP) was synthesised as detailed in §2.13. The $TatA_{APH}$ peptide was purified by reverse-phase HPLC using a linear ACN gradient with 0.1% TFA. The purity of the peptide was confirmed by ESI-MS, Figure 5.2. The major peak corresponded to pure peptide (2822 Da), with additional peaks corresponding to sodium adducts (2845, 2866 Da). These adducts were also seen in the purification of the $TatA_{TM}$ peptide §4.2.

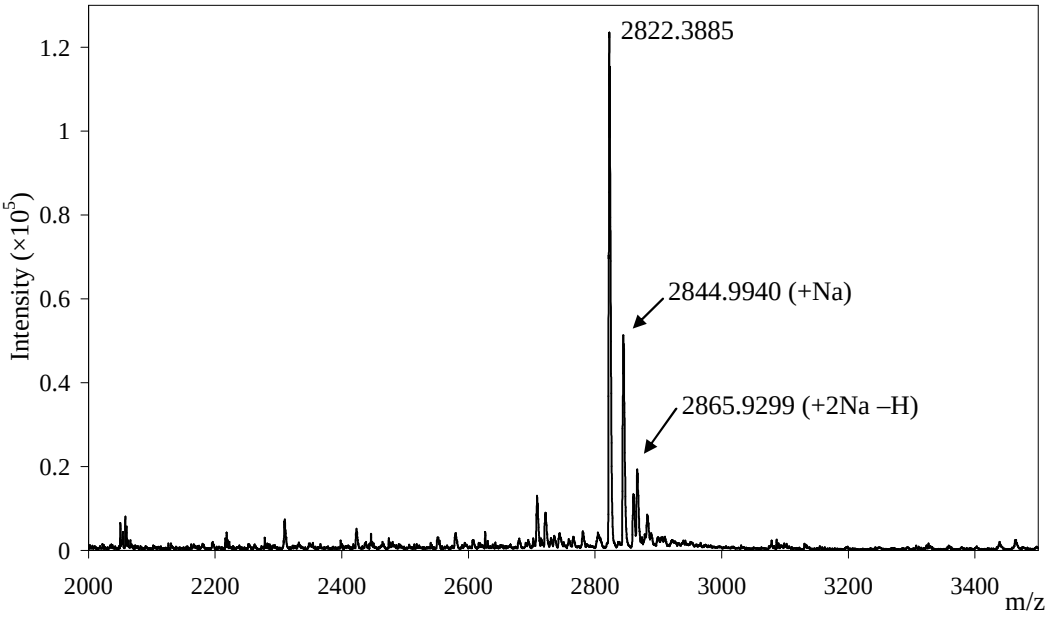


Figure 5.2 Assessment of the purified $TatA_{APH}$ peptide by electrospray ionisation mass spectrometry. Pooled fractions from reversed-phase HPLC purification of the $TatA_{APH}$ peptide were checked for purity by ESI-MS. The major peak corresponded to the $TatA_{APH}$ peptide (2822 Da), with additional peaks corresponding to sodium adducts.

5.2.1 Secondary structure of the $TatA_{APH}$ peptide

A previous study of the *B. subtilis* TatA homologue, TatAd, has shown that the secondary structure of TatAd expressed without its TM domain depends on the buffer, detergent and lipid composition (Lange *et al.*, 2007). In this study, CD was used to assess the secondary structure of our $TatA_{APH}$ peptide in a variety of solution conditions, Figure 5.3 and Figure 5.4. The $TatA_{APH}$ peptide was soluble in aqueous buffer solution, demonstrating that no detergents or lipids are needed for solubilisation of this peptide. A CD spectrum was initially acquired for the $TatA_{APH}$ peptide dissolved in sodium phosphate buffer. As shown in Figure 5.3, the 208 and 222 nm negative maxima, corresponding to an α-helical secondary structure are not present. Instead the major negative maximum is at 203 nm, indicating a predominantly random coil secondary structure in aqueous solution. Next, increasing concentrations of DPC detergent micelles were added to the peptide, Figure 5.3. The concentrations were calculated to give peptide:detergent micelle ratios as detailed in §2.13.5. For samples containing a small concentration of detergent micelles (e.g. 10:1 peptide:micelle), the secondary structure remained largely random coil.

However, as the concentration of micelles was increased (peptide:micelle = 1:1, 1:6.25), the CD spectra revealed an increase in the α-helical content of the peptide (as evidenced by the negative maxima at 208 and 222 nm). Fitting of the data using CDSSTR (Sreerama & Woody, 2000) with reference set SP175 (Lees *et al.*, 2006) reveals that the sample containing a 1:6.25 peptide:micelle ratio has an α-helical content of ~68%.

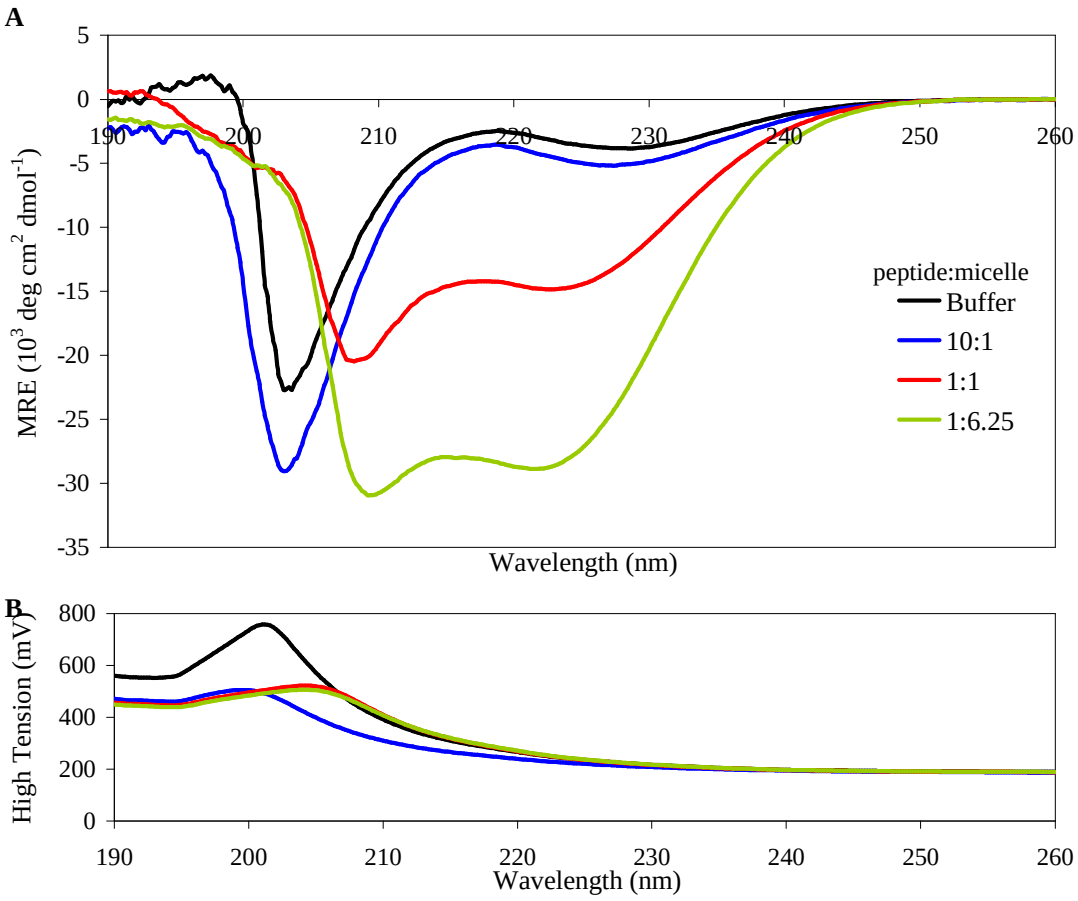


Figure 5.3 Analysis of the secondary structure of the $TatA_{APH}$ peptide in different concentrations of detergent. A. CD spectra of the $TatA_{APH}$ peptide in buffer or buffer containing different concentrations of DPC detergent calculated to give the peptide:detergent micelle ratio indicated (10:1, 1:1, 1:6.25). **B.** The high tension signals of the CD spectra were also recorded.

The *E. coli* membrane is made up of different lipids, of which the main components are the zwitterionic phosphatidylethanolamine (PE) 69%, and the negatively charged phosphatidylglycerol (PG) 19% (Ames, 1968). In the study of TatAd mentioned above (Lange *et al.*, 2007), phosphatidylcholine (PC) and PG lipids were used in the assessment of the secondary structure of a TM deletion. PC, like PE, is zwitterionic and has been widely used in the analysis of peptides (Adair & Engelman, 1994; Oates *et al.*, 2008; Sharpe *et al.*, 2002). We investigated the effect of various lipids on secondary structure by acquiring CD spectra of the $TatA_{APH}$ peptide in different mixtures of PC and PG lipids, Figure 5.4A. In the presence of DMPC, the peptide adopts a predominantly random coil conformation, as determined by the fitting program CDSSTR using the SP175 reference set, Figure 5.4C. When different ratios of 1,2-Dipalmitoyl-sn-Glycero-3-Phosphocholine (DPPC) and 1,2-Dipalmitoyl-sn-Glycero-3-[Phospho-rac-(1-glycerol)] (DPPG) were added, the peptide took on a primarily α-helical structure, Figure 5.4C. This indicates that the $TatA_{APH}$ peptide requires the presence of anionic phospholipids to fold into a helix.

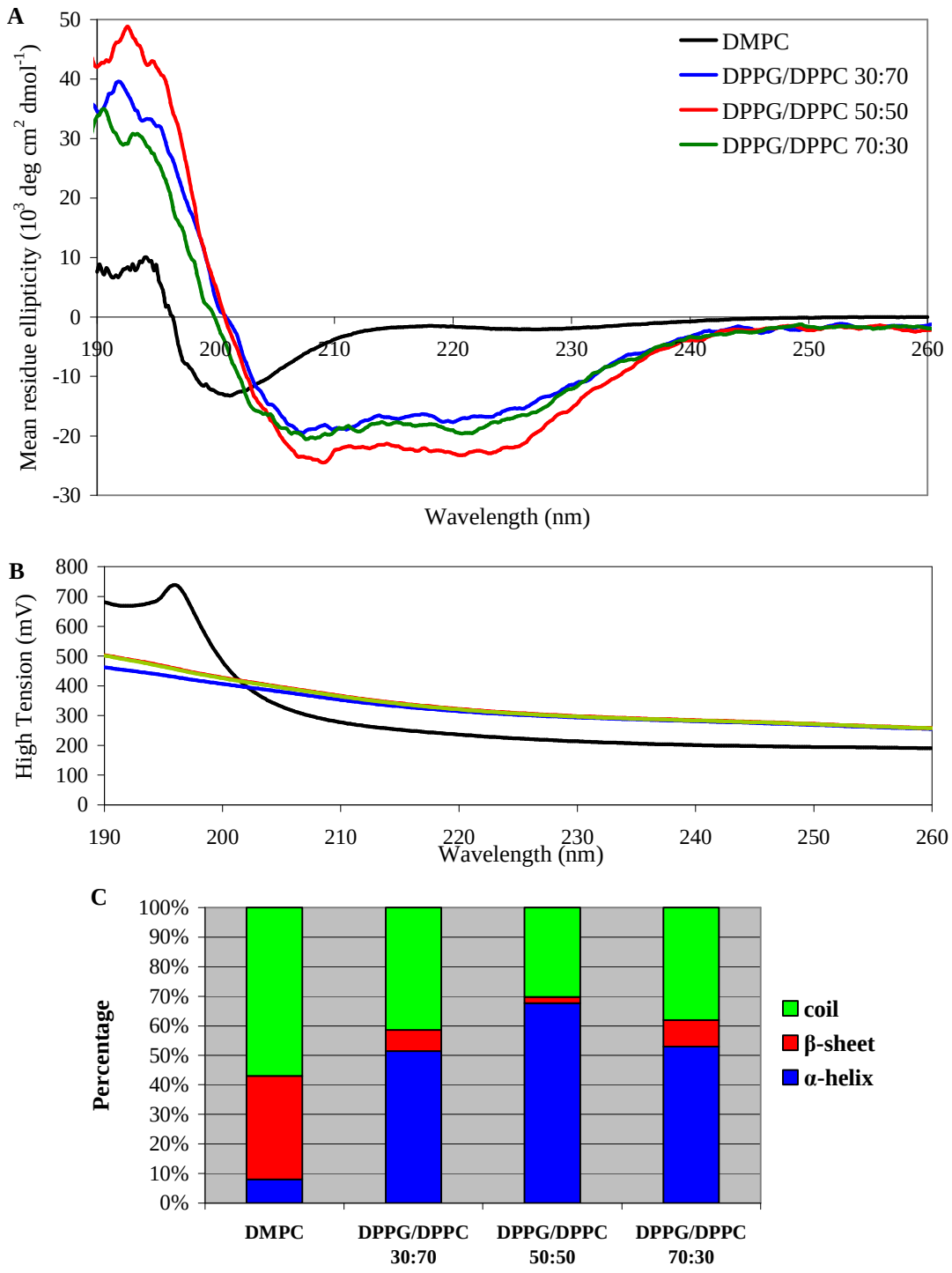


Figure 5.4 Analysis of the secondary structure of the $TatA_{APH}$ peptide in different compositions of lipid. A. CD spectra of the $TatA_{APH}$ peptide reconstituted into liposomes with different compositions of lipid; DMPC with a positively charged headgroup, and different ratios of DPPC and DPPG with positively and negatively charged headgroups respectively. **B.** The high tension signals of the CD spectra were also recorded **C.** Graph showing the percentage of α-helix, β-sheet and random coil in the secondary structure of the $TatA_{APH}$ peptide as calculated using CDSTTR using reference set SP175.

5.2.2 Analysis of membrane association of the $TatA_{APH}$ peptide

Although it has been shown that TatA expressed without its TM domain is still able to interact with lipid monolayers (Porcelli *et al.*, 2002), the mode of interaction has not been determined. Previously used methods were not able to differentiate between insertion into the membrane and association with the surface of a monolayer. Here we again used LD and ATR-FTIR to determine the nature of the interaction between the TatA APH and lipid membranes (i.e. orientation with respect to the bilayer).

5.2.2.1 LD analysis of the $TatA_{APH}$ peptide

As discussed in §4.2.2.1, LD can be used to determine if an α-helical peptide is inserted or lying on the surface of liposomes via the sign of the LD signal. Since the $TatA_{APH}$ peptide is only α-helical when there are negatively charged lipids present, Figure 5.4, liposomes were prepared with a DPPG:DPPC ratio of 70:30. This lipid composition was chosen to allow comparison with the membrane association study of the TatAd APH (Lange *et al.*, 2007). DMPC liposomes were also prepared to determine if the peptide would associate with the membrane when in a random coil conformation. Samples were aligned by spinning at two different speeds, 3000 and 5000 rpm, a spectrum of the sample without spinning was subtracted to give the LD signal and the spectra was corrected for light scattering due to the liposomes, Figure 5.5. In the presence of DMPC, no LD signal for the peptide was observed, indicating that the peptide does not associate with DMPC liposomes. Conversely, in peptide samples containing DPPG/DPPC liposomes, an LD signal was observed which increased with the speed of rotation. The increase in signal results from an increase in alignment (or orientation) and hence indicates that the observed signal results from liposome-associated peptide. Since there were no clear maxima in the LD spectra it was not possible to determine if the $TatA_{APH}$ peptide was inserted into DPPG/DPPC liposomes. However, since a signal was observed, the $TatA_{APH}$ peptide did show association with the liposomes.

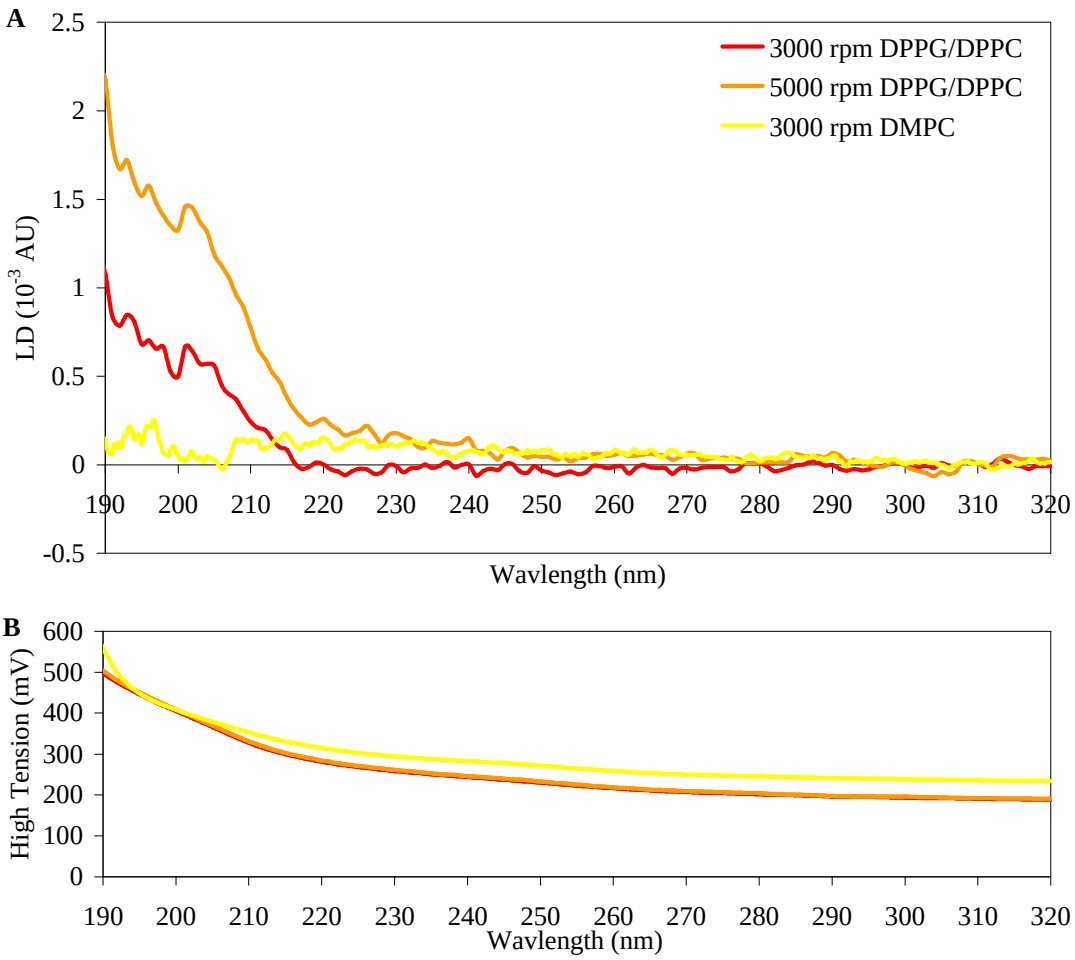


Figure 5.5 Linear dichroism spectra of the $TatA_{APH}$ peptide in liposomes. A. LD spectra were taken at two different rotating speeds; 3000 and 5000 rpm, in DMPC liposomes, or liposomes with a DPPG:DPPC ratio of 70:30. A spectrum of the sample without spinning was subtracted from the spectra of the sample with spinning then corrected for light scattering to give the LD signal. **B.** The high tension signals of the LD spectra were also recorded.

5.2.2.2 ATR-FTIR analysis of the $TatA_{APH}$ peptide

As mentioned in §4.2.2.3 ATR-FTIR can be used to determine both the secondary structure of peptides and their insertion into lipid bilayers. Since the CD analyses showed that the $TatA_{APH}$ peptide was only α-helical when both zwitterionic and negatively charged lipids were present, lipid bilayers with a DPPG:DPPC ratio of 70:30 were again prepared with peptide. Spectra were recorded before and after exposure to deuterium soaked nitrogen, Figure 5.6. In the spectra of the sample before exposure to deuterium, amide I and amide II peaks can be seen at 1667 and 1523 cm^{-1} respectively, Figure 5.6B.

To determine if the peptide inserted into the lipid bilayer the amide II peak was compared in the spectra before and after exposure to deuterium. After exposure, there is a loss of the amide II peak at 1523 cm^{-1} due to it being shifted to a lower wavenumber, where it was masked by peaks corresponding to bonds in the lipid, Figure 5.6B. The shift indicates that the protons were exchanged for deuterons and hence the stretching vibrations were shifted to a lower wavenumber due to the increased mass. The amide I peak maximum also shifted slightly from 1667 to 1649 cm^{-1}, which also indicates deuterium exchange. These shifts indicate that the peptide was lying on the surface of the lipid bilayer and was accessible to deuterium exchange.

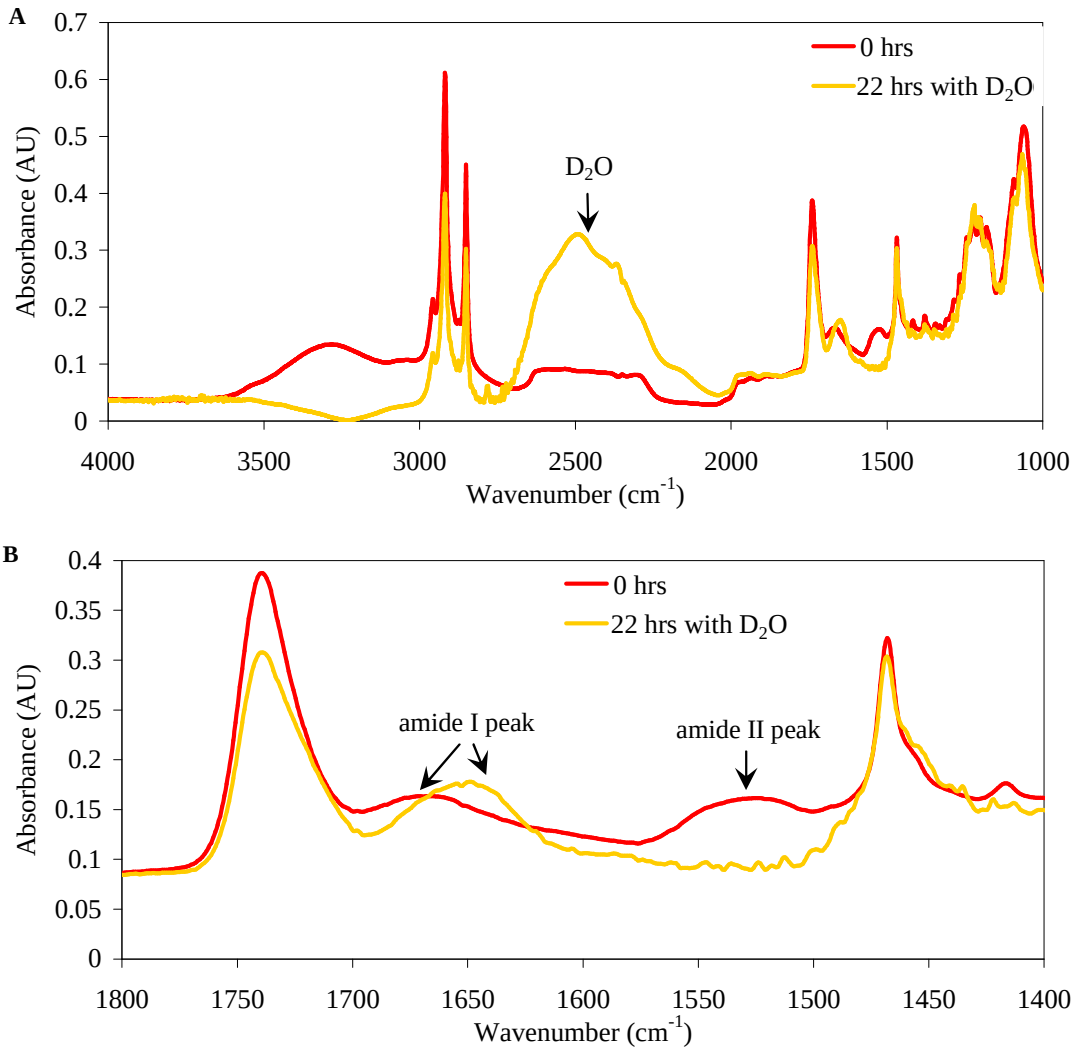


Figure 5.6 ATR-FTIR spectra of the $TatA_{APH}$ peptide in DPPG/DPPC lipid bilayers. Liposomes were prepared with a 30:1 lipid to peptide ratio, with a DPPG:DPPC lipid ratio of 70:30. **A.** Spectra were taken before (red) and after (orange) exposure to deuterium soaked nitrogen for 22 hours. The peak at ~2500 cm^{-1} indicates the presence of D_2O. **B.** Overlay of the spectra focusing on the 1800–1400 cm^{-1} region shows that the amide I and II peaks shift after exposure to deuterium soaked nitrogen (arrows), with the amide II shifting out of the 1800–1400 cm^{-1} range.

The secondary structure of the peptide in the sample was determined by peak fitting the amide I peak. In the spectrum acquired before exposure to deuterium soaked nitrogen, the amide I peak can be seen at 1667 cm^{-1}. The band-narrowing technique of Fourier self-deconvolution and a baseline correction were applied to the 1700–1600 cm^{-1} region corresponding to the amide I peak from the spectrum before exposure to deuterium using OMNIC Spectra software. Gaussian curves were fitted to the resulting peak using GRAMS software, Figure 5.7. The major peak

corresponded to an α-helical secondary structure, however there was significant amount of β-sheet present as well. Although CD estimated that the peptide was predominantly α-helical, the sample preparation for ATR-FTIR may affect the secondary structure of the peptide. Since it was not able to insert a proportion of it may have formed aggregates, which result in a β-sheet signal.

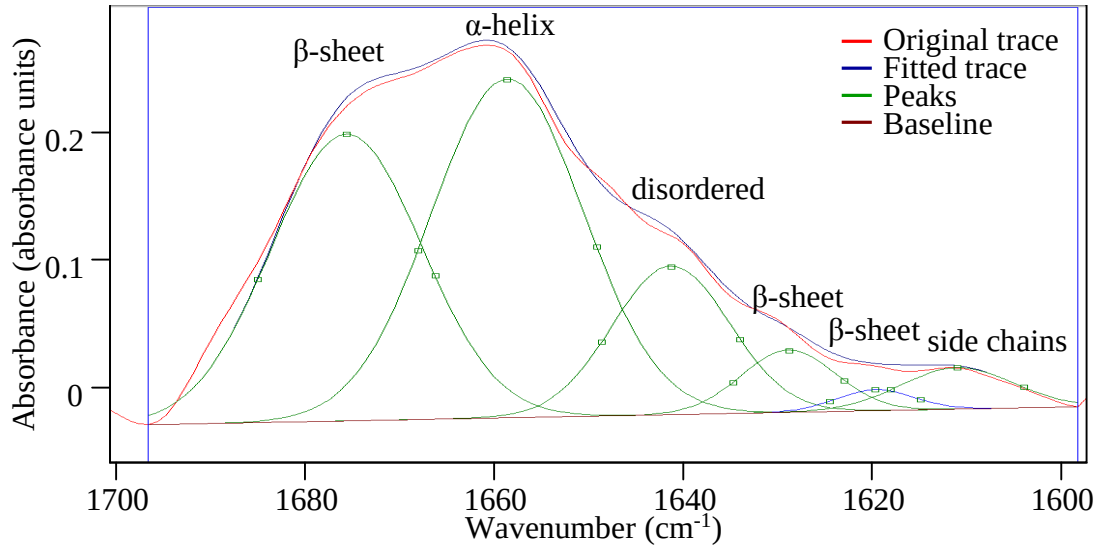


Figure 5.7 Secondary structure determination of the $TatA_{APH}$ peptide using ATR-FTIR. The band-narrowing tool of Fourier self-deconvolution was used on the amide I peak in the 1700–1600 cm^{-1} region of the spectra before exposure to deuterium. A baseline correction was also applied. Gaussian curves were fitted to the peak and were assigned to different secondary structures according to (Tatulian, 2003).

5.2.3 Oligomeric state of the $TatA_{APH}$ peptide

Several motifs have been identified in TM domains which promote helix-helix interactions by allowing close helix packing (Russ & Engelman, 2000; Senes *et al.*, 2000). One of the identified motifs is the Gly-Xaa-Xaa-Xaa-Gly (GxxxG) motif, where two glycine residues are separated by three other amino acids. Analysis of the $TatA_{APH}$ peptide using TMSTAT (Senes *et al.*, 2000) revealed the presence of a GxxxG motif, Figure 1.6, indicating that the APH may be involved in helix-helix interactions. This motif is highly conserved in TatA homologues, Appendix A (Hicks *et al.*, 2003). Substitution of either of these glycine residues with cysteine, knocks out function of TatA (Greene *et al.*, 2007), and substitution of the second glycine

with serine or aspartic acid disrupts TatA oligomerisation, as assessed by chemical cross-linking (Hicks *et al.*, 2005). To assess the ability of the $TatA_{APH}$ peptide to form oligomers, SDS-PAGE and chemical cross-linking were used.

5.2.3.1 Analysis of the $TatA_{APH}$ peptide oligomeric state by SDS-PAGE

In order to assess the ability of the $TatA_{APH}$ peptide to form homo-oligomers, a range of concentrations of peptide were analysed by SDS-PAGE. Staining of the gel with Coomassie blue revealed the presence of a single band, running at ~4 kDa, Figure 5.8A, which was assigned to monomer (*m*). When the same gel was stained using the more sensitive silver stain technique, no additional bands of higher molecular weight were seen, Figure 5.8B.

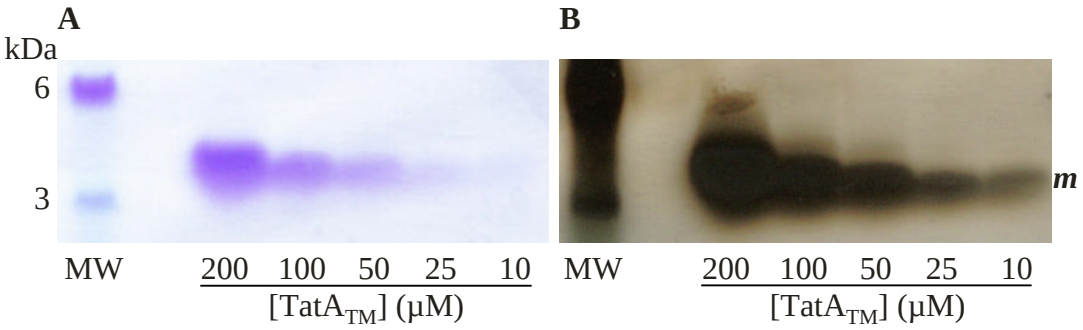


Figure 5.8 SDS-PAGE analysis of the $TatA_{APH}$ peptide. The $TatA_{APH}$ peptide was analysed over a range of concentrations (stated below each lane) with a molecular weight marker (MW). **A.** The Coomassie stained gel revealed a single band corresponding to monomer (***m***) (monomer MW = 2.65 kDa) **B.** Silver staining the same gel confirmed the presence of a single band.

5.2.3.2 Analysis of the $TatA_{APH}$ peptide oligomeric state by chemical cross-linking

As mentioned previously, §4.2.3, SDS can be a highly denaturing detergent and could disrupt any weak interactions. To detect any helix-helix interactions of the $TatA_{APH}$ peptide in less denaturing detergents (in this case DPC), chemical cross-linking with the soluble BS^3 was performed before analysis by SDS-PAGE, Figure 5.9. Since we had observed that detergent concentration affects the secondary structure of the $TatA_{APH}$ peptide, §5.2.1, cross-linking was performed in different

peptide:detergent micelle ratios with DPC detergent (10:1, 1:1, 1:6.25). A reaction was set up with no cross-linker added (1:6.25 no x) and an additional reaction was performed in 100 mM SDS (SDS), as negative controls. At low detergent concentrations, when the peptide is predominantly random coil (10:1), bands corresponding to higher order oligomers of up to four peptides could be seen (n = 1, n = 2,…). As the detergent concentration was increased, these higher order bands disappeared and the peptide existed predominantly as monomers and dimers (1:1, 1:6.25). When no cross-linker was added or the reaction was performed in SDS, only monomers were observed. Since the higher order oligomers are only seen when $TatA_{APH}$ peptide is mainly random coil, it is likely that they result from non-specific aggregation of the peptide. BS^3 reacts with the NH_2 groups of lysines, as well as the terminal NH_2 of the peptide. Since the $TatA_{APH}$ peptide contains five lysines, this increased the likelihood of non-specific cross-linking.

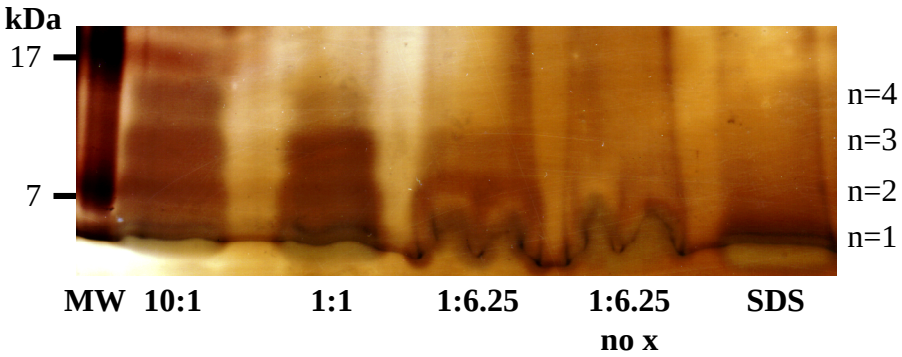


Figure 5.9 Cross-linking analysis of the $TatA_{APH}$ peptide in detergent micelles. Cross-linking was performed in DPC detergent micelles at different peptide:detergent micelle concentration ratios, indicated below each lane. Reactions were performed with BS^3 with one reaction performed without cross-linker (6.25:1 no x). An additional reaction was performed in 100 mM SDS (SDS). Molecular weight markers are shown in the left hand lane (MW). Proteins were visualised by staining with silver nitrate. Oligomeric states are indicated on the right (n = 1, n = 2,…)

5.3 TOXCAT analysis of the TatA APH

The TOXCAT system was developed to analyse the strength of helix-helix interactions in TM domains (Russ & Engelman, 1999). Although insertion of the $TatA_{APH}$ peptide into synthetic lipid bilayers had not been observed using ATR-FTIR or LD, §5.2.2, it was decided to analyse the TatA APH in a natural membrane using the TOXCAT assay. The TOXCAT chimera is optimised to assist the insertion of protein fragments into the inner membrane of *E. coli*, via the periplasmic MBP

domain and a cytoplasmic ToxR DNA binding domain. Inserting the APH sequence into this chimera may result in a 'forced' membrane insertion, and would allow us to determine if the APH has any ability to insert into membranes as previously hypothesized (Chan *et al.*, 2007; Gouffi *et al.*, 2004).

To test these hypotheses, six TOXCAT chimera containing different lengths of the APH were constructed, Figure 5.10. The GxxxG motif identified by TMSTAT, §5.2.3, is also highlighted. Different insert lengths were analysed as previous studies have shown that the length of insert can affect the strength of the TOXCAT signal (Jenei *et al.*, 2009; Langosch *et al.*, 1996; Li *et al.*, 2004). Since the APH is not predicted to form a TM domain, the insert sequence was also varied to include different proportions of hydrophilic and hydrophobic amino acids, which may affect any possible insertion.

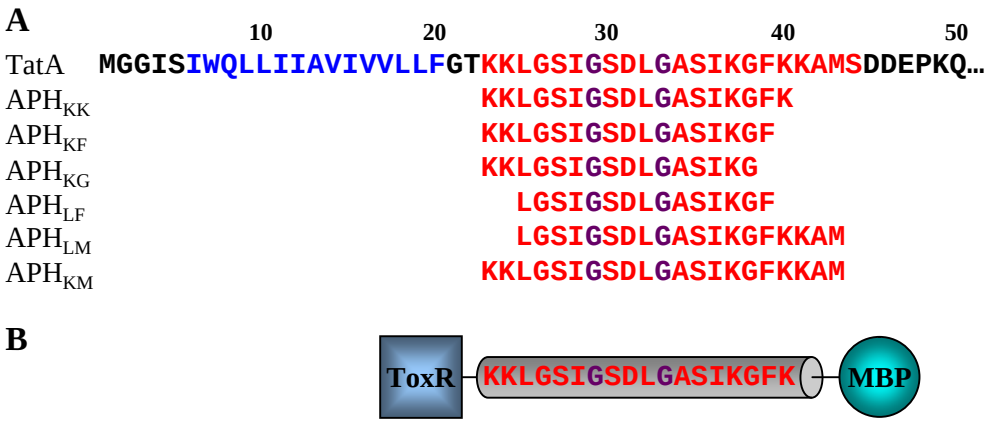


Figure 5.10 TatA APH TOXCAT constructs. A. The first 50 residues of the TatA primary sequence (TatA) with the predicted TM domain (blue) and APH (red); and the six different TatA APH sections inserted into the TOXCAT chimera (APH_{KK}, APH_{KF}, APH_{KG}, APH_{LF}, APH_{LM}, APH_{KM},). The GxxxG motif is highlighted (purple). **B.** The APH_{KK} TOXCAT chimera consisting of the α-helical APH inserted between the N-terminal dimerisation-dependent DNA binding domain of the ToxR protein (■) and the periplasmic anchor MBP (●).

Checks were performed to determine the expression levels, insertion and orientation of the APH TOXCAT chimera as detailed in §2.12.1.1, Figure 5.11. To check the expression levels, whole NT326 cells expressing the chimera were compared with the positive and negative controls of GpA and G83I, respectively, Figure 5.11A. Surprisingly the expression levels of the APH chimera (APH_{KK}, APH_{KF}, APH_{KG},

APH_{LF}, APH_{LM}, APH_{KM}) were much higher than the controls. The expression levels were also much higher than the TatA TM domain TOXCAT chimera analysed in §4.3.

Sodium hydroxide washes were performed in order to determine the localisation of the APH TOXCAT chimera, Figure 5.11B. If the chimera is inserted correctly across the membrane it should only be present in the whole cell (W) and membrane (M) fractions, not the soluble (S) fraction. Immunoblots using anti-MBP antibodies revealed that unlike the G83I construct, which is correctly inserted, the APH constructs were present in both the membrane and soluble fractions. In the whole cell and membrane fractions a second, faster migrating, unidentified band was also detected for the APH constructs. Sodium hydroxide washes of membranes remove any weakly associated proteins, with only strongly associated and integral membrane proteins remaining in the membrane fraction. This suggests that the majority of the APH chimera is strongly associated with the membrane, with a smaller but significant percentage only associating weakly with the membrane.

To establish the location of the MBP domain of the APH chimera, spheroplasts (Sp) from cells expressing the chimera were treated with Proteinase K (Sp*), or lysed by freeze-thaw before treatment with Proteinase K (B*), Figure 5.11C. If the MBP domain of the chimera is located in the periplasm, then it is accessible to Proteinase K. Upon treatment, it is cleaved from the chimera and runs at a lower molecular weight than the full length chimera. All of the APH constructs gave the same results, with the disappearance of the intact chimera band in the treated spheroplast fractions, indicating that the MBP domain of the chimera was located in the periplasm.

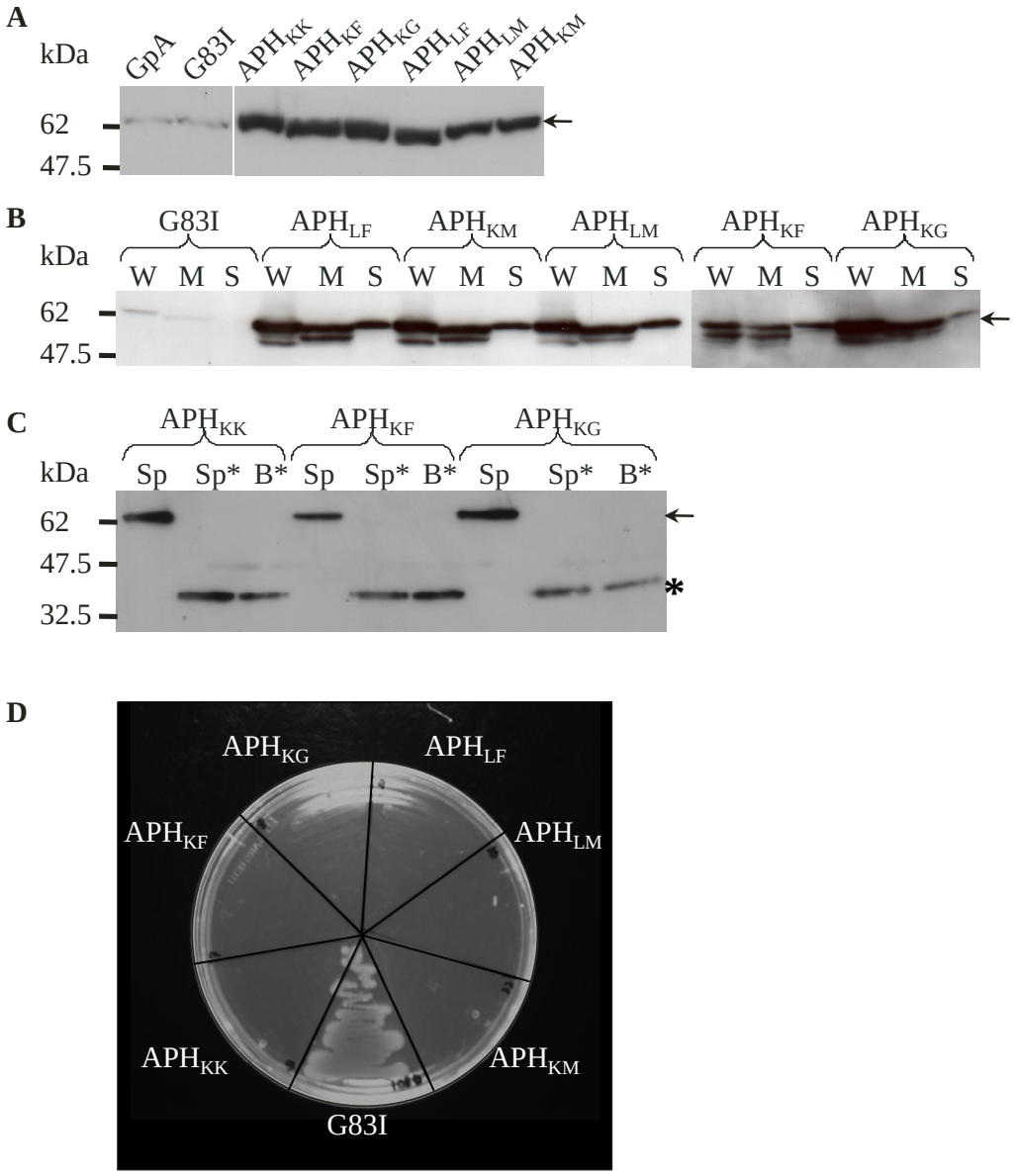


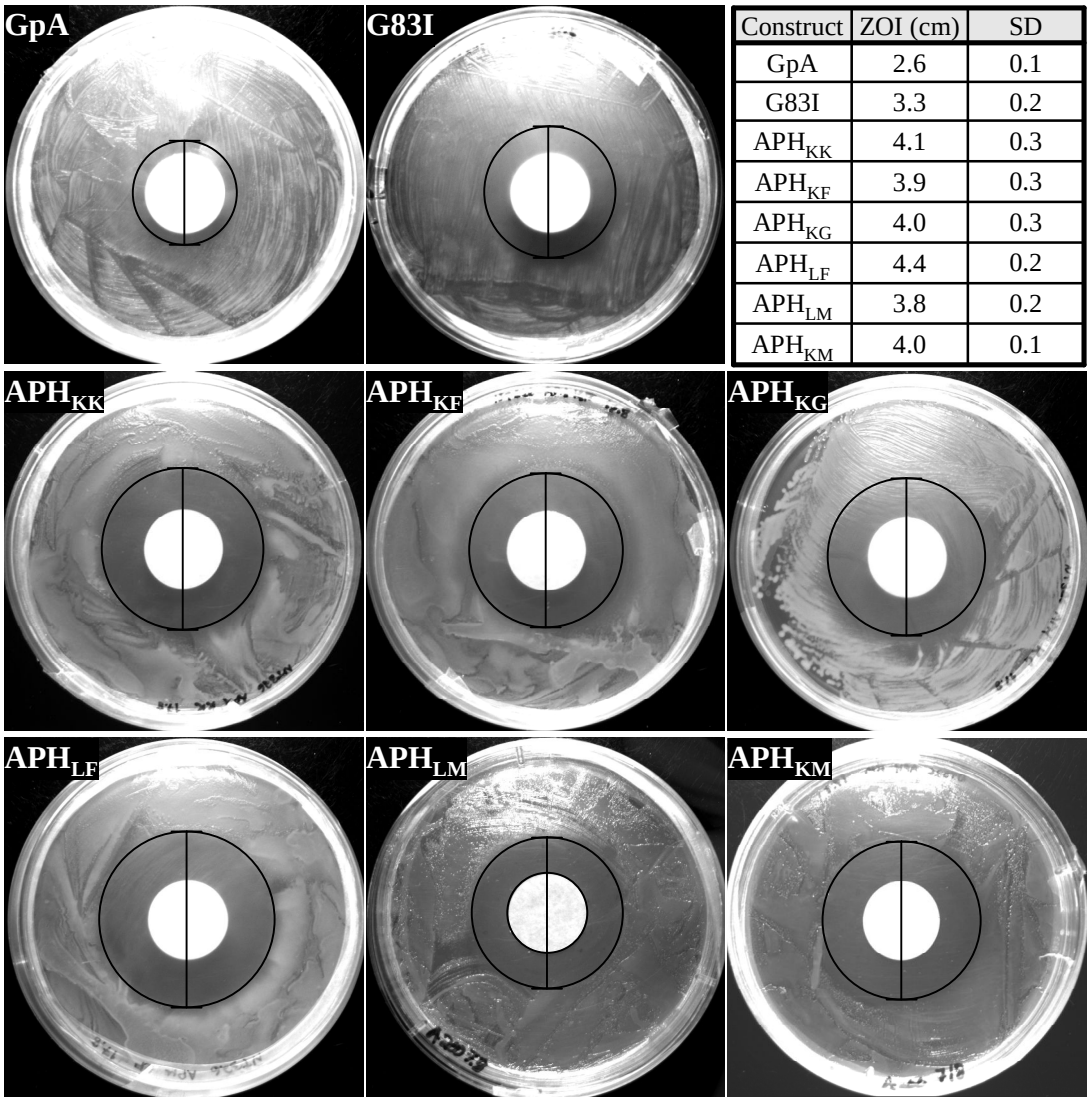
Figure 5.11 TOXCAT insertion checks for the TatA APH constructs. A. Immunoblot using anti-MBP antibodies of whole NT326 cells expressing the TOXCAT chimera with the different APH inserts (APH_{KK}, APH_{KF}, APH_{KG}, APH_{LF}, APH_{LM}, APH_{KM}), with GpA and its dimerisation defective mutant (G83I) as controls. Arrow indicates the TOXCAT chimera. **B**. Immunoblot using anti-MBP antibodies of sodium hydroxide washes. Cells expressing the TOXCAT chimera were treated with ice-cold NaOH, then centrifuged to separate the supernatant (S), which contained cytoplasmic, periplasmic and loosely-membrane associated proteins, from the pellet (M), which contained proteins stably associated with the membrane. Untreated whole cells (W) were also analysed. Cells not expressing the chimera (NT326) were used as a negative control. Arrow indicates the TOXCAT chimera. **C.** Immunoblot using anti-MBP antibodies of protease sensitivity of spheroplasts assay. Spheroplasts (Sp) of cells were treated with Proteinase K (Sp') or broken open by freeze-thaw before treatment with Proteinase K. Arrow indicates TOXCAT chimera and asterisk indicates MBP cleaved by proteolysis. **D.** *malE* complementation assay. *malE* deficient NT326 cells containing the TatA APH and G83I TOXCAT chimera were grown on M9 agar plates containing 0.4% maltose.

The cellular location of the MBP domain was also analysed by *malE* complementation on M9 minimal media agar plates containing 0.4% maltose. If the MBP deficient NT326 cells expressing the chimera are able to grow, the chimera is correctly inserted across the membrane with the MBP in the periplasm. Cells expressing the G83I chimera grew, however none of the cells expressing the APH constructs showed any growth after several days, Figure 5.11D. This indicates that the MBP domain of the APH constructs was not able to function in the periplasm.

These results suggest that the APH constructs were not able to insert correctly into the membrane.

Although the APH constructs appeared not to insert correctly across the membrane bilayer by monitoring the MBP, a disk diffusion assay was performed to compare any possible expression of CAT with the positive and negative controls, GpA and G83I respectively, Figure 5.12, thus providing information on the location of the ToxR domain. The zone of inhibition of growth (ZOI) was recorded for the controls and the APH constructs, Figure 5.12. As expected, the APH constructs did not show any CAT expression, instead displaying large ZOIs characteristic of cells not expressing the TOXCAT chimera (Russ & Engelman, 1999).

The analysis of the APH using the TOXCAT assay demonstrated, in agreement with the results from LD and ATR-FTIR, that the APH has a high propensity to associate with lipid membranes but has little or no propensity to insert across membranes. These data suggest that the previously reported hypothesis that the APH has a dual topology (Gouffi *et al.*, 2004) is not very likely.



Construct	ZOI (cm)	SD
GpA	2.6	0.1
G83I	3.3	0.2
APH_{KK}	4.1	0.3
APH_{KF}	3.9	0.3
APH_{KG}	4.0	0.3
APH_{LF}	4.4	0.2
APH_{LM}	3.8	0.2
APH_{KM}	4.0	0.1

Figure 5.12 Disk diffusion assay for the TatA APH. NT326 cells expressing the TatA APH domain TOXCAT constructs (APH_{KK}, APH_{KF}, APH_{KG}, APH_{LF}, APH_{LM}, APH_{KM}) were grown on agar plates with a chloramphenicol soaked disk in the centre. The zone of inhibition (ZOI) of growth was recorded (black circles and lines) and compared to the positive and negative controls of GpA and G83I. The standard deviation (SD) of the measurements is also shown.

5.4 Discussion

Although the TatA APH has been shown to be important for function (Greene *et al.*, 2007; Lee *et al.*, 2002), its exact role in the Tat pathway is still unknown. The aim of this study was to analyse the TatA APH in isolation using *in vitro* and *in vivo* techniques, to gain a better understanding of its role in the Tat pathway.

5.4.1 The secondary structure of the $TatA_{APH}$ peptide depends on detergent concentration and lipid composition

The first approach used to study the APH was the *in vitro* analysis of a synthetic peptide corresponding to the APH of TatA. Unlike the peptide corresponding to the TatA TM domain, studied in §4.2, the $TatA_{APH}$ peptide was soluble in aqueous buffer. However CD analysis of the $TatA_{APH}$ peptide revealed that it did not have the predicted α-helical secondary structure in aqueous buffer but instead took on a random coil conformation. The formation of α-helix could be induced by addition of increasing amounts of detergent. This behaviour could be explained by the amphiphilic nature of the peptide. Since it is soluble in the aqueous buffer the secondary structure might be relaxed, however these are not the correct conditions to induce secondary structure formation. Upon the addition of detergent, the peptide appears to fold into an α-helix on the surface of the detergent micelles. Hence only when there is an excess of detergent micelles can the majority of the peptide form the α-helical secondary structure.

Similarly, the secondary structure of the $TatA_{APH}$ peptide in the presence of liposomes varied depending on the compositions of lipids. The peptide only adopted α-helical secondary structure in the presence of anionic phospholipids. Again it was concluded that this was due to the amphiphilic nature of the peptide. When there was a charge on the liposomes, as is present in natural *E. coli* membranes (Ames, 1968), the peptide formed an α-helical secondary structure and LD confirmed that it was associating with the membrane. The requirement of anionic phospholipids indicates that there an electrostatic interaction between the peptide and the lipid bilayer, potentially between positively charged lysine residues of the APH and the negatively charged lipid headgroups. These results suggest that the peptide associates with the lipid bilayer in a two stage process: firstly the positively charged Lys residues drive association of the APH with the negatively charged bilayer; the peptide then folds into an α-helix to maximize these contacts (in an α -helix, all three lysine residues lie on the same helical face, Figure 5.1 and Figure 6.7).

Interestingly, it has previously been shown that anionic phospholipids are required in *E. coli* for the Tat pathway to function efficiently (Mikhaleva *et al.*, 1999). When the percentage of anionic phospholipids was reduced to low levels, the export of the Tat

substrate TorA was greatly reduced. This may be linked to incorrect folding of the APH of TatA, and possibly TatB, which has a similar predicted secondary structure.

5.4.2 The $TatA_{APH}$ peptide does not form SDS stable oligomers

The ability of the $TatA_{APH}$ peptide to form oligomers *in vitro* was analysed by SDS-PAGE and chemical cross-linking. Unlike the $TatA_{TM}$ peptide which forms stable dimers in SDS, no higher order oligomers were observed for the $TatA_{APH}$ peptide at any of the concentrations analysed. Higher order oligomers were observed when chemical cross-linking was performed, however the formation of these oligomers was disrupted by increasing the detergent concentration. This increase in detergent concentration corresponds to an increase in α-helical content, so the peptide did not form oligomers when in an α-helical conformation. Therefore it was concluded that these oligomers were due to non-specific interactions caused by the high number of lysine residues in the $TatA_{APH}$ peptide.

5.4.3 The TatA APH is not able to insert into synthetic lipid membranes or the inner membrane of *E. coli*

Although the $TatA_{APH}$ peptide was seen to associate with the lipid bilayers using LD, the nature of the association could not be determined due to the lack of a clear signal. However, ATR-FTIR clearly showed that the peptide was unable to insert into DPPG/DPPC lipid bilayers under the conditions used. The complete shift of the amide II peak indicated that the peptide was accessible to deuterium exchange and was therefore lying on the surface of the lipid bilayer. This agrees with an OCD study of the APH from TatAd (Lange *et al.*, 2007).

The ability of the APH to insert into *E. coli* membranes *in vivo* was assessed using the TOXCAT assay (Russ & Engelman, 1999). Although not designed for amphipathic helices, TOXCAT was used in this study to establish whether the TatA APH could be 'forced' to insert into a natural *E. coli* membrane. When the expression, insertion and orientation checks were performed for the six different TatA APH TOXCAT chimera, they suggested that while the bulk of the chimeric protein was strongly associated with the membrane, some of the chimera were found

in the soluble fraction as well. The MBP domain was correctly located in the periplasm, however it was not able to function and facilitate the take up of maltose.

One hypothesis is that the entire ToxR'-APH-MBP chimera was exported to the periplasm and remained strongly associated with the membrane. Due to the high expression levels the cell was saturated with the chimera, so some was present in the soluble fraction. The MBP may have not been able to function properly due to its uncharacteristic association with the membrane, and hence could not complement the *malE* deficient NT326 cells. In the absence of any anti-ToxR antibodies to determine the cellular location of ToxR, it was not possible to test this hypothesis. Since the ToxR'-APH-MBP chimera did not insert correctly into the membrane it was not possible to test the significance of the GxxxG motif in APH helix-helix interactions.

Although insertion of the APH was not observed in this study there may be other conditions under which it is able to insert. It has been seen that an electrical potential is required for the export of SufI by the Tat pathway (Bageshwar & Musser, 2007). This change in environment or interactions with other proteins involved in the Tat pathway may induce the APH to insert into the membrane as part of the translocation process.

The work presented in this chapter has demonstrated that the structure of the $TatA_{APH}$ peptide depends on its environment, only taking on an α-helical structure in excess detergent or in the presence of negatively charged lipids. Unlike the $TatA_{TM}$ peptide, no insertion of the $TatA_{APH}$ peptide was observed with any of the techniques used. The APH could not be 'forced' into the inner membrane of *E. coli* by inserting it into the TOXCAT chimera either. This indicates that although the APH associates with membranes it does not spontaneously insert into them. If the APH does insert into the membrane as part of the translocation process it is likely that other factors are involved which facilitate insertion.

Chapter 6. Analysis of the *Escherichia coli* TatA C-terminus and identification of an important acidic motif

6.1 Introduction

Unlike the TatA TM domain and APH, which are both highly conserved in bacterial homologues (Yen *et al.*, 2002), the C-terminus varies considerably in both composition and length, Appendix A. Secondary structure predictions of *E. coli* TatA assign the C-terminal region, starting at Ser44, as largely unstructured, Figure 1.6. The role of this region in the Tat pathway is still unknown, although it has been shown to be non-essential for function (Lee *et al.*, 2002). Truncation of *E. coli* TatA from the C-terminus has shown that up to 40 amino acids can be deleted without affecting the translocation of the Tat substrate TorA. If the role of TatA is purely in translocation of Tat substrates across the membrane, it is uncertain as to why the C-terminal region remains in all homologues when it is not needed for function. Since some Gram-positive bacteria only contain a TatA-like protein and no TatB, it is possible that the C-terminus of TatA is an evolutionary remnant, and may simply enable TatA to perform the role of TatB in a TatA/TatC system. However, contrary to this hypothesis is the observation that the C-terminus of *E. coli* TatB is also non-essential for function of the Tat pathway (Lee *et al.*, 2002). Additionally, substitutions identified in *E. coli* TatA which enable it to take on the role of TatB have all been located in the extreme N-terminus of TatA.

Recent research has suggested that TatA may have a dual role in the Tat pathway (Matos *et al.*, 2008; Matos *et al.*, 2009). It was observed that the Tat apparatus has a role in proofreading of two iron-sulphur (FeS) proteins, NrfC and NapG. TatA and TatE were highlighted as playing a part in the rejection for translocation and subsequent degradation of these two mutated FeS proteins. This apparent dual role of TatA may involve different domains of the protein.

Here, a number of different approaches were used to analyse the role of the C-terminal domain and identify important residues. Several new TatA constructs were prepared for these analyses, as shown in Figure 6.1. Analysis of a novel C-terminal truncation led to the identification of the minimal functional subunit of TatA. Examination of this truncation mutant highlighted an acidic motif, which was mutated to determine its role in function and complex formation. Possible interactions between this negatively charged region and three positively charged lysines in the TatA APH were also investigated through further mutational studies.

The TatA truncations were used to investigate the role of the C-terminus in proofreading of the FeS protein NrfC.

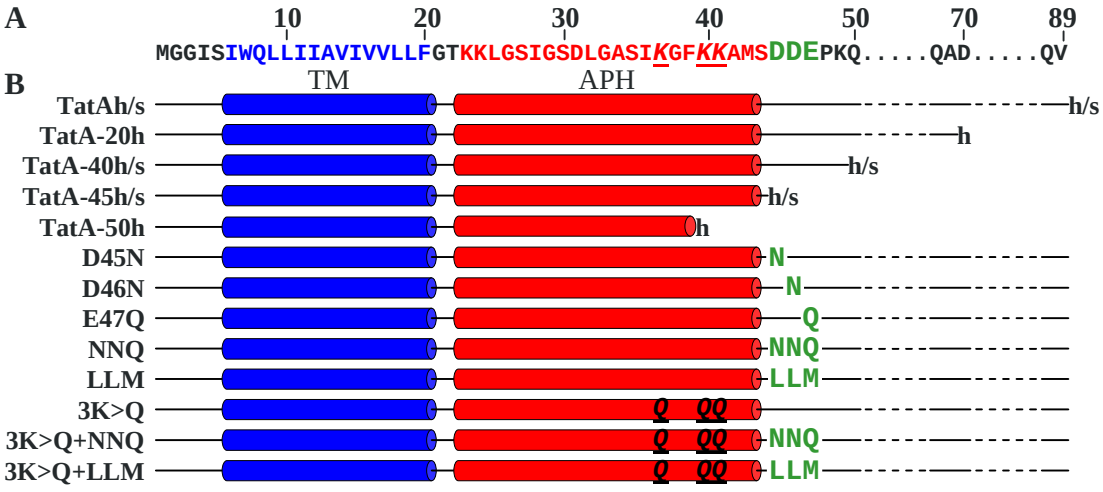


Figure 6.1 TatA sequence and constructs used in this study. A. The primary sequence of *E. coli* TatA. An acidic motif identified in this study is highlighted (green), and three lysines in the APH which have been shown to be important for function and complex formation are underlined (Barrett *et al.*, 2005). **B.** TatA constructs used in this study, with the TM domain and the APH highlighted (blue and red, respectively). Full length TatA and C-terminal truncations by 20, 40, 45 and 50 amino acids were tagged with His-6 (h) or Strep-II™ (s) tags (TatAh/s, TatA-20h, TatA-40h/s, TatA-45h/s, TatA-50h). Substitutions were introduced in the acidic motif after the APH (D45N, D46N, E47Q, NNQ, LLM), the three lysines in the APH (3K>Q) and double substitutions in both these regions (3K>Q+NNQ, 3K>Q+LLM).

6.2 Truncation analysis to identify the region of TatA essential for function

As previously mentioned, the role of the C-terminus of TatA has been analysed by sequentially truncating the protein from the C-terminus in steps of 10 amino acids (Lee *et al.*, 2002). By assaying for periplasmic activity of the Tat substrate TorA, investigators were able to determine that, although still functional when truncated by 40 amino acids, TatA lost its functional activity when truncated by 50 amino acids, suggesting the stretch of amino acids from residue 39 to residue 49 is crucial for function. To try and narrow down the essential residues in this region, it was decided to construct a novel truncation of 45 amino acids from the C-terminus, containing

either a His-6 or Strep II™ C-terminal tag, Figure 6.1B, (A-45h/s). The inclusion of a tag was necessary due to the location of the anti-TatA antibody binding epitope lying between residues 49 and 69 of TatA, Figure 6.2A. Previous research has shown that a C-terminal His-6 tag on TatA does not affect its function (Bolhuis *et al.*, 2000), however it was decided to make an additional truncation with a StrepII™ tag to test whether tags would affect function. To allow comparison of our system with the previous truncation study using a different expression system (Lee *et al.*, 2002), truncations of TatA by 20, 40 and 50 residues with a C-terminal His-6 or StrepII™ tag were also prepared, Figure 6.1B, (A-20h/s, A-40h/s, A-50h).

The expression levels of the truncations expressed from the pBAD24 plasmid in cells lacking TatA and TatE (Δ*AE*) were checked by performing immunoblots with the relevant antibodies, Figure 6.2A–C. Proteins truncated by 45 or 50 amino acids were expressed at lower levels than the other truncations, although they were still expressed at levels comparable to TatA in wild-type cells.

In order to investigate the truncations' effect on function, the TorA activity assay was used, Figure 6.2D. In the control samples, TorA activity is clearly apparent in the periplasm of wild-type MC4100 cells, as indicated by a white band in the periplasmic (P) lane, and no export is evident in Δ*AE* cells. Analysis of the tagged truncations shows that the A-20 and A-40 mutants support export of TorA, whereas the A-50 truncation mutant does not. These results agree with previous research for plasmid-borne, overexpressed TatA (Lee et al., 2002), confirming that our tagged proteins were behaving in the same manner as previously studied untagged versions. These results also demonstrate that the choice of tag does not affect the functional activity of the truncations, with both His-6 and StrepII™ truncations producing the same results.

The additional TatA mutants prepared in this study, A-45h and A-45s, in which 45 amino acids were deleted from the C-terminus, also abolished export via the Tat pathway. This absence of function identified a stretch of five residues (Asp45-Lys49) important for function.

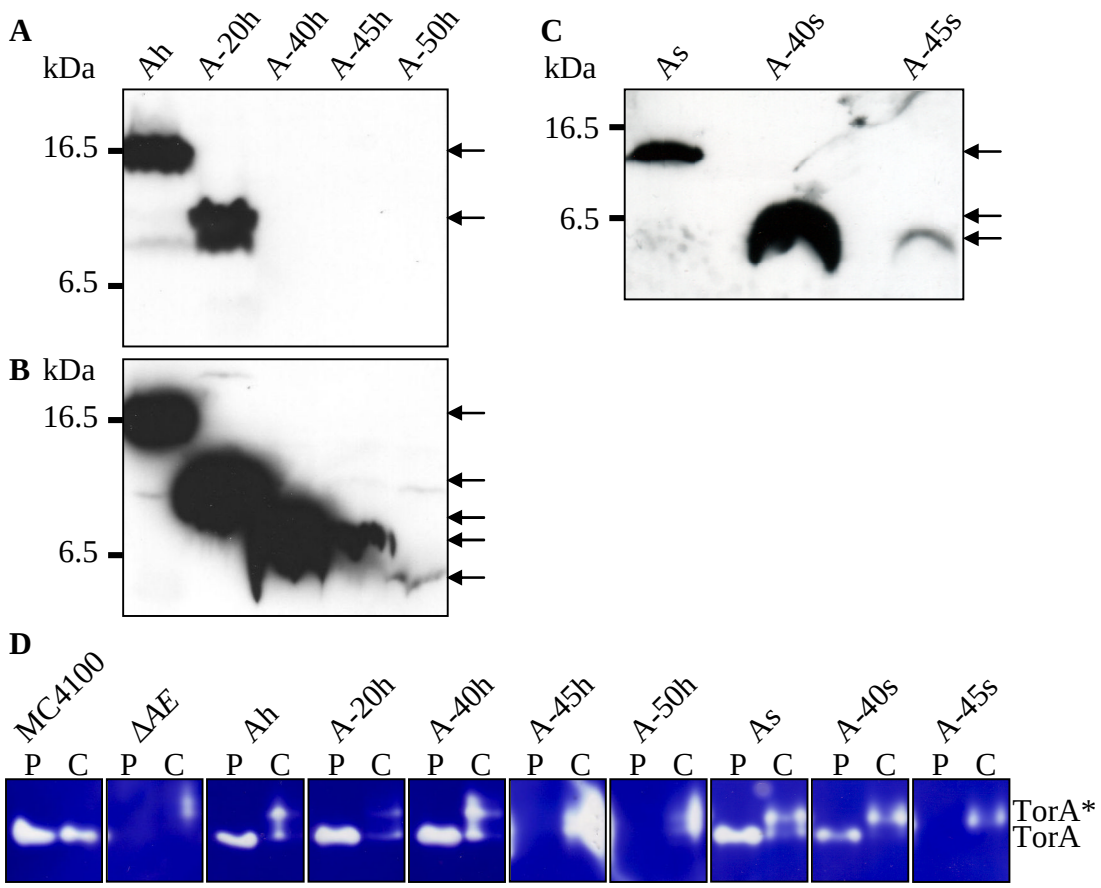


Figure 6.2 Analysis of the function of TatA truncations from the C-terminus. Immunoblots of whole Δ*AE* cells expressing full length TatA, TatA truncated by 20, 40, 45 or 50 amino acids from the C-terminus with a C-terminal His-6 tag (Ah, A-20h, A-40h, A-45h, A-50h) using (**A**) anti-TatA antibodies and (**B**) anti-His-6 antibodies. No bands are seen for A-40h, A-45h and A-50h in the anti-TatA immunoblot due to the binding epitope being removed by truncation. **C.** Immunoblot of whole Δ*AE* cells expressing full length TatA, TatA truncated by 40 or 45 amino acids from the C-terminus with a Strep II™ tag (As, A-40s, A-45s) using anti-Strep II™ antibodies. **D.** Periplasmic and cytoplasmic samples (P,C) were prepared from wild-type *E. coli* cells (MC4100), Δ*AE* cells (Δ*AE*), and Δ*AE* cells expressing the full length and tagged truncations of TatA. A native polyacrylamide gel was stained for TorA activity; white bands indicate the presence of TorA, TorA* indicates a slower migrating cytoplasmic form of TorA.

6.3 Determination of the oligomeric state of the TatA truncations

In order to assess the effect of the C-terminal truncations on the ability of TatA to form complexes of variable size, fast protein liquid chromatography (FPLC) on a Superose-6HR pre-packed gel filtration column was used. This technique uses the interactions between proteins and the gel matrix to separate proteins according to

molecular weight. Large proteins cannot enter the gel, and hence elute first, whereas smaller proteins enter the gel and elute in later fractions. This method has been used previously to assess the size of Tat complexes from *E. coli* and *B. subtilis* (Barnett et al., 2008; Bolhuis et al., 2001).

Membranes from Δ*AE* cells expressing full length TatA or TatA truncated by 40 or 45 amino acids with a His-6 C-terminal tag were isolated by fractionation and solubilised in digitonin detergent, before injection onto the column. Elution fractions were collected and immunoblots of the peak elution fractions were performed using anti-His-6 antibodies, Figure 6.3A. The relative intensity of each band was quantified using ImageJ analysis software, normalised to the maximum value and plotted against the fraction number, Figure 6.3B. The full length TatA and TatA truncated by 40 amino acids were found across a wide range of fractions, as demonstrated by bands in multiple lanes. This indicates that they are forming oligomeric complexes of variable size, which is in agreement with a previous study where untagged *E. coli* TatA was analysed (Barnett *et al.*, 2008). In contrast, the TatA truncated by 45 amino acids eluted in a much sharper peak at a later time point, suggesting it is forming a smaller, discrete complex. It is worth noting that the expression levels of the three constructs were variable, with A-45h expressed at much lower levels. Although attempts were made to normalise the amount of protein loaded onto the column, direct comparison between the elution profiles may not be accurate.

Despite the problems with expression levels, gel filtration analysis of the truncations points towards the importance of the stretch of five residues (Asp45–Lys49) for complex formation. In the absence of these residues, TatA no longer elutes across all fractions but mainly elutes as a single, lower molecular weight complex.

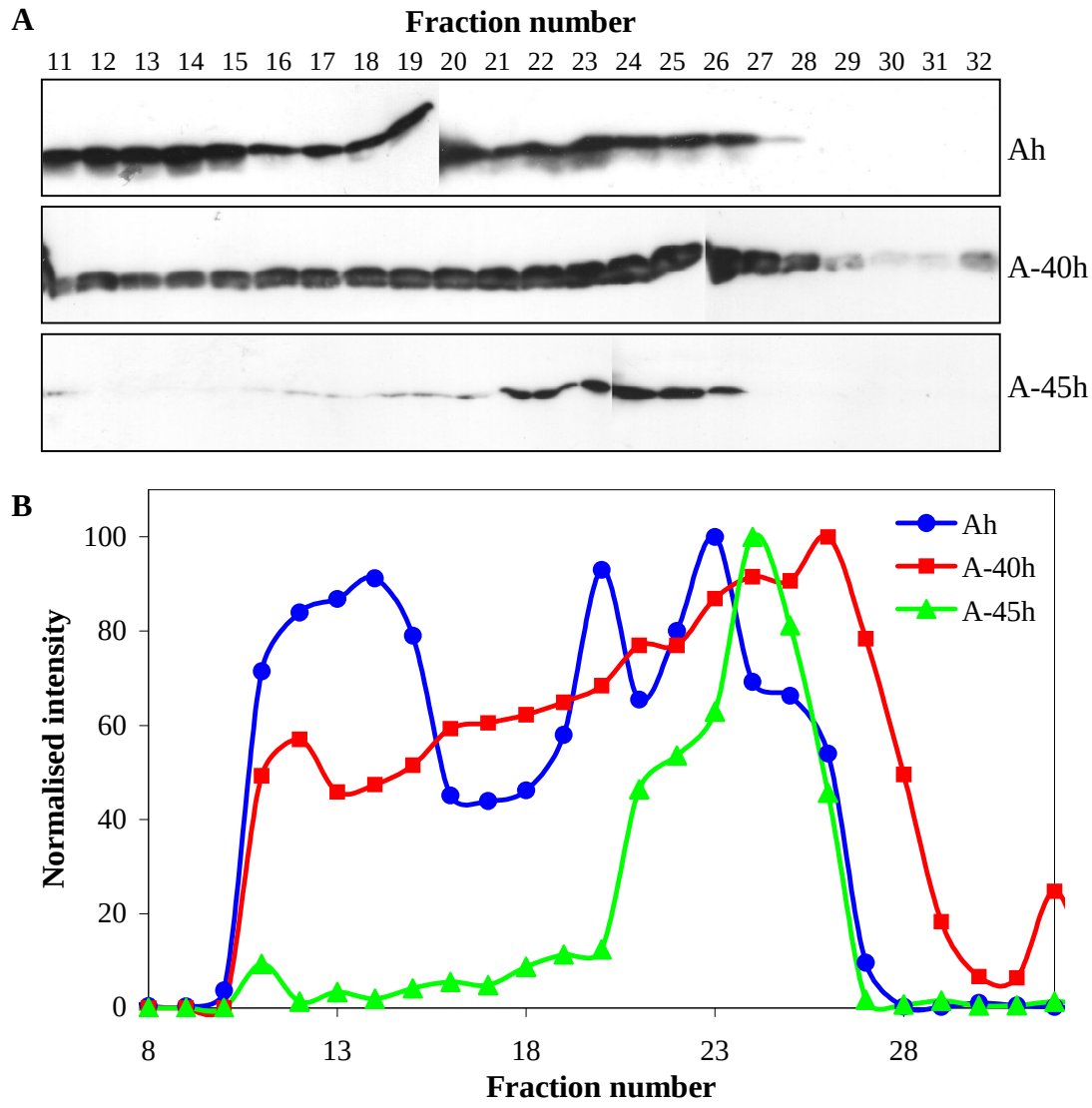


Figure 6.3 Gel filtration analysis of the complex size of TatA, TatA-40 and TatA-45 with a His-6-tag. Digitonin solubilised membranes from Δ*AE* cells expressing full length TatA or TatA truncated by 40 or 45 amino acids from the C-terminus with a C-terminal His-6 tag (Ah, A-40h, A-45h) were applied to a Superose-6HR gel filtration column. **A.** Immunoblots using anti-His-6 antibodies of the peak elution fractions for the three different TatA constructs. Ah, A-40h and A-45h are labelled. **B.** The relative intensities of the bands in the immunoblots were measured using ImageJ software, normalised to the maximum value and plotted against the fraction number, Ah (blue ●), A-45h (red ■) and A-45h (green ▲).

6.4 Identification of an acidic motif important for function and complex formation

Analysis of the five residues highlighted by the TatA truncation analyses revealed three consecutive acidic residues after the predicted APH (DDE sequence highlighted in Figure 6.1A). Alignment of TatA homologues from other Gram-negative bacteria showed this region of acidic residues straight after the APH represents a highly conserved acidic 'motif', Figure 6.4 and Appendix A. The term acidic motif is used here, not to refer to absolute conservation of the sequence DDE, but to describe a sequence of two to four acidic amino acids after the predicted APH. This acidic motif varies in composition and length but it is almost always present in the sequences analysed.

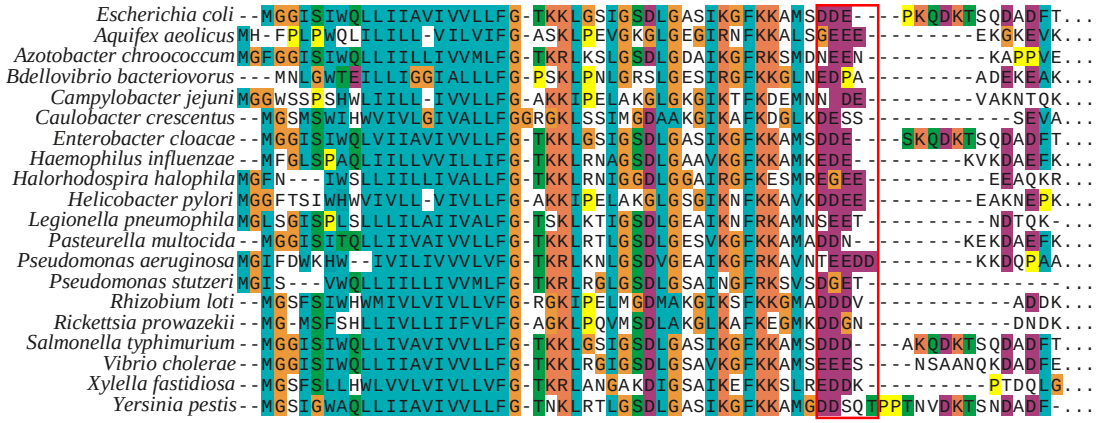


Figure 6.4 Sequence alignment of TatA homologues from Gram-negative bacteria. Sequences of TatA homologues from Gram-negative bacteria were identified and aligned using ClustalX. The alignment shows the TM domain, the APH and the region immediately following the APH. An acidic region after the predicted APH is boxed in red. The full alignment is shown in Appendix A.

The importance of the acidic motif to function and structure was analysed by mutagenesis of TatA. Each of the three residues was substituted individually with a structurally similar but uncharged residue (D45N, D46N, and E47Q), and a triple substitution was also prepared (NNQ), Figure 6.1B. Similar to aspartic acid and glutamic acid, asparagine and glutamine are able to hydrogen bond and form strong non-covalent interactions. For this reason, a second triple substitution which would

be unable to hydrogen bond to other residues was prepared (LLM), Figure 6.1B. The mutants were prepared by mutating the *tatA* gene in the pBAD-ABC plasmid.

Expression levels of the mutants expressed from the pBAD-ABC plasmid in cells lacking the Tat components (Δ*ABCDE*) were checked by performing immunoblots using anti-TatA antibodies, Figure 6.5A. The expression levels of the mutants were comparable to that of wild-type TatA expressed from the pBAD-ABC plasmid. The functional effect of the substitutions were analysed using the TorA assay, Figure 6.5B. The single substitutions do not appear to affect the function of the Tat pathway, with the majority of the TorA activity found in the periplasmic as in the wild-type positive control (MC4100). The NNQ triple mutant does seem to affect function, with the majority of the TorA activity found in the cytoplasm, and only a faint band apparent in the periplasm. The LLM mutant has an even more severe effect on function, with no TorA band appearing in the periplasm even after extended periods of time.

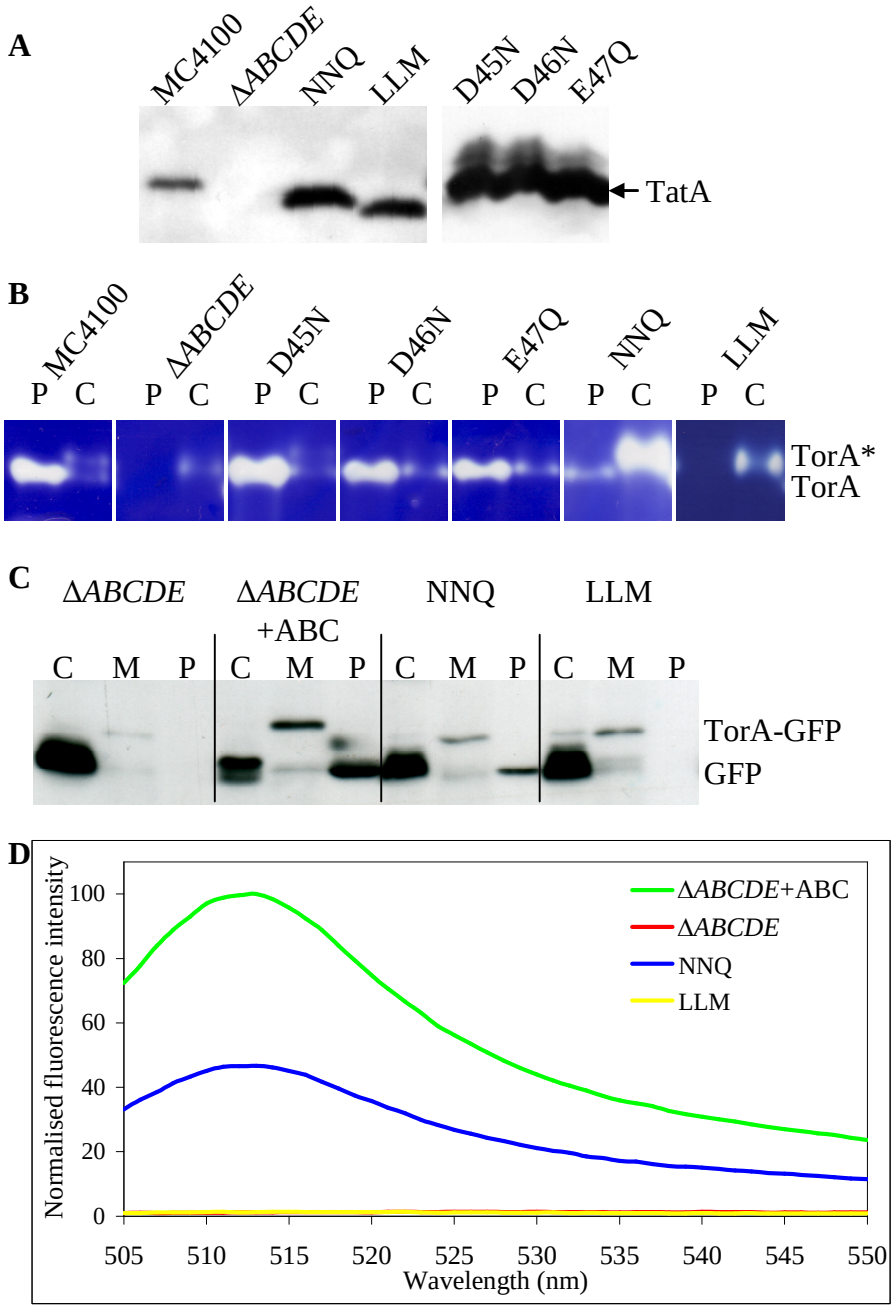


Figure 6.5 Analysis of the role of the acidic motif in TatA function. The affects of five different substitutions of the acidic motif of TatA (Asp45→Asn, D45N; Asp46→Asn, D46N; Glu47→Gln, E47Q; Asp46 and Asp46→Asn and Glu47→Gln, NNQ; Asp45 and Asp46→Leu and Glu47→Met, LLM) were analyzed by a TorA activity assay and a TorA-GFP export assay. **A.** Immunoblots of whole wild-type cells (MC4100), Δ*ABCDE* cells (Δ*ABCDE*) and Δ*ABCDE* cells expressing the mutated TatA (D45N, D46N, E47Q, NNQ, LLM) using anti-TatA antibodies to check expression levels. **B.** Cytoplasmic, membrane and periplasmic samples (C, M, P) were prepared and run on a native polyacrylamide gel, which was stained for TorA activity. White bands indicate the presence of TorA, TorA* indicates a slower migrating cytoplasmic form of TorA. **C.** Immunoblots using an anti-GFP antibody of C, M, P samples from Δ*ABCDE* cells expressing TorA-GFP from the pJDT1 plasmid (Δ*ABCDE*), Δ*ABCDE* cells expressing TorA-GFP and wild-type TatABC from the pEXT-ABC

plasmid (Δ*ABCDE*+ABC), and *ΔABCDE* cells expressing TorA-GFP and mutated TatABC from the pEXT-ABC plasmid (NNQ, LLM). Intact TorA-GFP and clipped TorA-GFP (GFP) are indicated. **D.** The fluorescence of periplasmic samples of Δ*ABCDE* cells expressing TorA-GFP from the pJDT1 plasmid and the TatA mutants was recorded from 505 to 550 nm, exciting at 488 nm. The resulting fluorescence was normalised to Δ*ABCDE*+ABC.

Since the white bands in the TorA assay develop over time, the assay can only be used to assess export qualitatively. In order to analyse more accurately the export efficiency of the triple mutants, the more quantitative TorA-GFP export assay was also used, Figure 6.5C and D. In this assay the TorA-GFP chimera, consisting of the TorA signal sequence fused to GFP, is expressed from the pBAD24 plasmid and exported via the Tat pathway. The efficiency of export can be assessed by immunoblotting using anti-GFP antibodies (Barrett & Robinson, 2005) and measuring fluorescence of the GFP using a fluorimeter. Δ*ABCDE* cells expressing the TorA-GFP chimera and wild-type or mutated TatA from the pEXT-ABC plasmid were fractionated into cytoplasmic, membrane and periplasmic fractions (C, M, P). Immunoblots revealed that for cells expressing the wild-type TatA (Δ*ABCDE*+ABC) and the TatA carrying the NNQ substitution, the Tat pathway was still able to export the TorA-GFP, as demonstrated by a mature-sized GFP band in the periplasm, Figure 6.5C. The intensity of the band for the NNQ mutant was clearly lower than that observed for wild-type TatA, indicating less efficient export. This difference was quantified by measuring the fluorescence of the periplasmic sample due to the presence of GFP. Samples were excited at 488 nm, the optimal excitation wavelength for GFP, fluorescence emission was recorded between 505 and 550 nm, and the resulting fluorescence was normalised to the positive control, Δ*ABCDE*+ABC, Figure 6.5D. The GFP fluorescence maximum at 513 nm of the NNQ mutant was 47% of the wild-type TatA, showing that the NNQ mutant is only able to export at less than 50% of wild-type levels. Immunoblots of the LLM mutant show that it completely prevents export via the Tat pathway, with no visible TorA-GFP in the periplasm. The accumulation of mature-sized GFP in the cytoplasm is due to proteolysis of TorA-GFP when export is blocked (Barrett *et al.*, 2003; Thomas *et al.*, 2001). No fluorescence was detected by the fluorimeter in the periplasm of the LLM mutant samples. These data reveal the importance of the acidic motif in the function

of TatA and demonstrate that function can be restored in at least part by the substitution of polar residues such as glutamine and asparagine.

In order to investigate the possible link between TatA function and complex formation, blue-native gel analyses were performed on the TatA acidic motif mutants, Figure 6.6. For the single mutants, D45N, D46N and E47Q, which were all functional in the TorA assay, a banding pattern identical to the wild-type TatA (Δ*ABCDE*+ABC) was observed, indicating that they do not affect the ability of TatA to form variable sized complexes. Conversely, the triple mutants, NNQ and LLM, appear to severely destabilise the formation of TatA complexes as judged by this approach. Although there are still multiple bands present, indicating oligomer formation, these are fewer, much less well defined and appear at different molecular weights to the wild-type bands. This effect might be caused by the increase in predicted pI from 5.73 for the wild-type TatA, to 8.96 for the triple mutants. For BN-PAGE the pI should ideally be no greater than 8.6 (Schägger *et al.*, 1994). Proteins with pIs of 9.3 or above result in broad banding patterns, and their slower migration results in inaccurate molecular mass assignment. However, the pI of the triple mutants is only just outside the ideal range and, although there is a broadening of bands, there are visibly fewer bands present. These observations, taken together with the absence of this effect for the single mutants, imply that the change in banding pattern for the triple mutants is not an artefact, and results from complex disruption *in vivo*. The acidic motif appears to be not only important for function but also for the correct formation of TatA complexes.

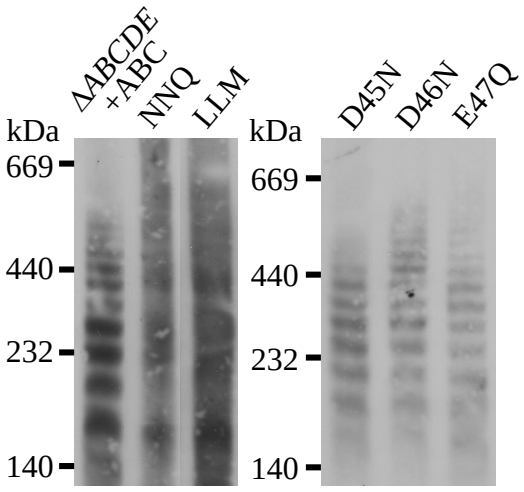


Figure 6.6 Assessment of the role of the acidic motif in TatA complex formation using blue-native PAGE. Membranes from Δ*ABCDE* cells expressing wild-type (Δ*ABCDE*+ABC) or mutated TatA (NNQ, LLM, D45N, D46N, E47Q) from the pBAD-ABC plasmid solubilised in digitonin were subjected to blue-native gel analyses and immunoblotted using anti-TatA antibodies. The mobility of the molecular weight markers is indicated.

6.5 Probing possible interactions between the acidic motif and lysines in the APH

Charged and polar regions on the surface of proteins can play an important role in stabilising protein-protein interactions, reviewed in (Sheinerman *et al.*, 2000). Since substitution of the acidic motif destabilised the formation of TatA complexes, it was proposed that this negatively charged region may interact with a positively charged region in the protein. Previous studies have highlighted the importance of three positively charged residues in the APH (Lys37, Lys40 and Lys41) (Barrett *et al.*, 2005; Barrett & Robinson, 2005). Substitution of these residues with polar, uncharged glutamine, disrupted TatA function and complex formation. Plotting the APH on a helical wheel reveals that these three residues are located on the same face of the helix, Figure 6.7A. Their proximity to one another in the predicted three-dimensional structure can be seen more clearly when the APH is modelled as an α-helical domain, Figure 6.7B. The existence of a positively charged region in the APH and a negatively charged region nearby led to the hypothesis that the there was a charge interaction between these two regions which stabilised TatA complex formation.

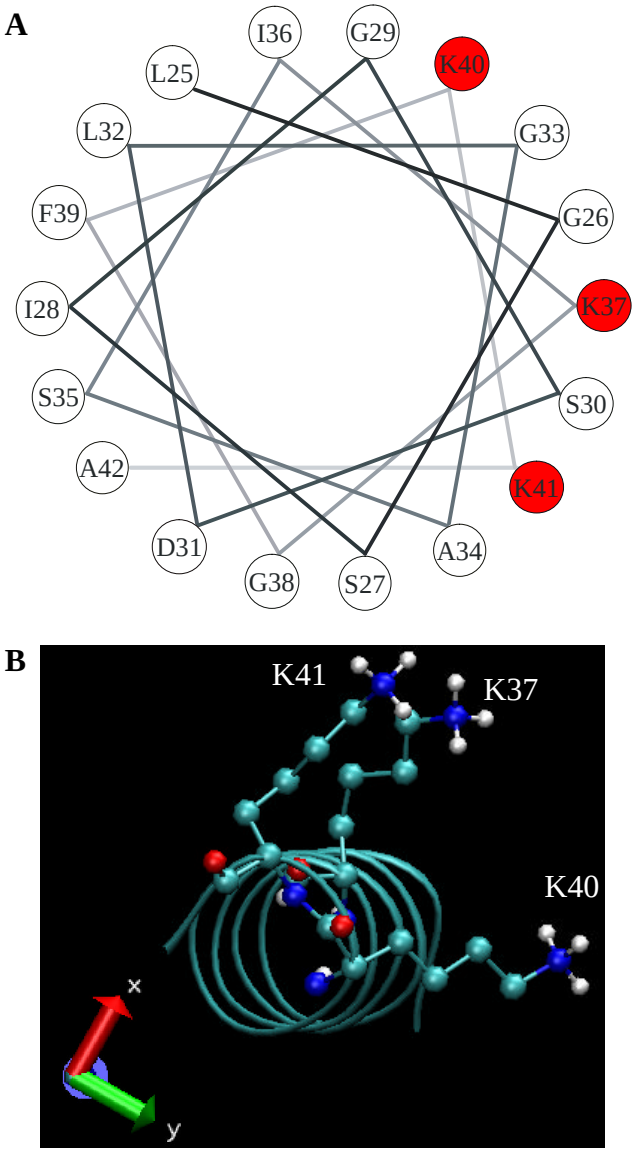


Figure 6.7 Spatial analysis of the three lysines in the TatA APH. A. Helical wheel plot of the TatA APH, the three lysines which have been shown to be important for function (Barrett *et al.*, 2005), are highlighted in red (K37, K40, K41). **B.** Three dimensional model of the TatA APH prepared using the programme CHI (CNS searching of helix interactions (Adams *et al.*, 1995; Adams *et al.*, 1996)), showing the α-helical secondary structure and the three lysines. Arrows indicate the axes.

The possible interaction between these two oppositely charged regions of TatA was probed by mutagenesis studies. As well as the triple substitution in the TatA acidic motif (NNQ or LLM), the lysines forming the positively charged region were also mutated to glutamine (Lys37, Lys40 and Lys41→Gln; 3K>Q), resulting in the mutants NNQ+3K>Q and LLM+3K>Q, Figure 6.1B. A TorA activity assay was used to assess the functional activity of these mutants, Figure 6.8A. As seen

previously (Barrett *et al.*, 2005; Barrett & Robinson, 2005) the 3K>Q substitution is not able to export TorA to the periplasm. Similarly there is no white band in the periplasm for the NNQ+3K>Q and LLM+3K>Q mutants, indicating they were also not functional. The more quantitative TorA-GFP export assay was also used to assess the functional activity of the NNQ+3K>Q and LLM+3K>Q mutants, Figure 6.8B. No mature sized GFP band was observed for either of these mutants, leading to the conclusion that they were unable to function. These results demonstrate that substitution of the positively charged region is not able to compensate for the deleterious effect of the substitution of the acidic motif, and indeed completely blocks the low levels of export seen for the NNQ mutant.

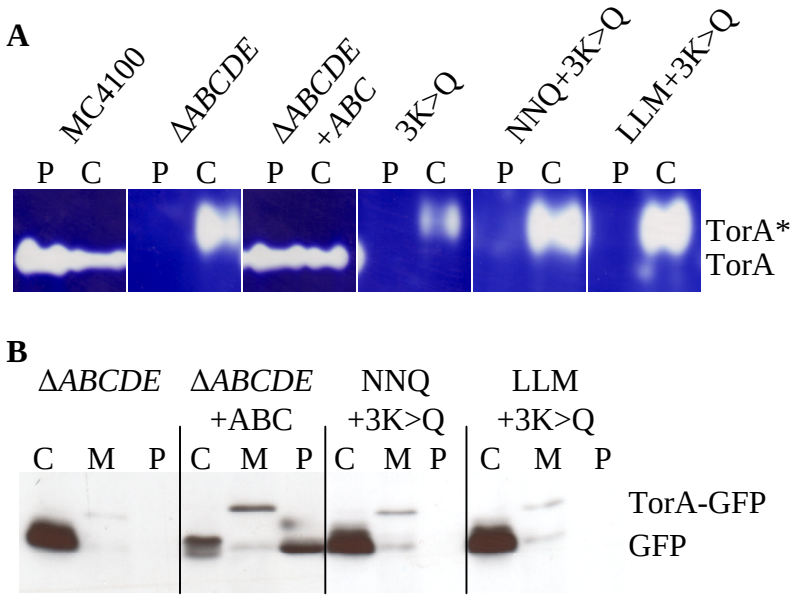


Figure 6.8 Mutation of the acidic motif and three lysines cannot restore TatA function. The effects of substitution of the acidic motif and the three essential lysines were analyzed by a TorA activity assay and a TorA-GFP export assay. **A.** Cytoplasmic, membrane and periplasmic samples (C, M, P) were prepared from wild-type cells (MC4100), Δ*ABCDE* cells (Δ*ABCDE*), and Δ*ABCDE* cells expressing wild-type TatA (Δ*ABCDE*+ABC) or mutated TatA (Lys37, Lys40, Lys41→Gln, *3K>Q*; Lys37, Lys40, Lys41→Gln, Asp45, Asp46→Asn, Glu47→Gln, *NNQ+3K>Q*; Lys37, Lys40, Lys41→Gln, Asp45, Asp46→Leu and Glu47→Met, *LLM+3K>Q*) from the pBAD-ABC plasmid and run on a native polyacrylamide gel, which was stained for TorA activity. White bands indicate the presence of TorA, TorA* indicates a slower migrating cytoplasmic form of TorA. **B.** Immunoblots using an anti-GFP antibody of C, M, P samples from Δ*ABCDE* cells expressing TorA-GFP from the pJDT1 plasmid (Δ*ABCDE*), Δ*ABCDE* cells expressing TorA-GFP and wild-type TatABC from the pEXT-ABC plasmid (Δ*ABCDE*+ABC), and Δ*ABCDE* cells expressing TorA-GFP and mutated TatABC from the pEXT-ABC plasmid (NNQ+3K>Q, LLM+3K>Q). Intact TorA-GFP and clipped TorA-GFP (GFP) are indicated.

6.6 The role of the TatA C-terminus in the proofreading of NrfC[†]

Although truncation analysis has led to the identification of an acidic motif in the TatA C-terminus, the role of the final 40 residues of TatA is still unclear. Whilst the main role of TatA appears to be in the translocation process, recent research has highlighted the importance of the Tat apparatus in proofreading and degradation of FeS proteins (Matos *et al.*, 2008; Matos *et al.*, 2009). An unidentified proofreading mechanism in the Tat pathway prevented mutations of the FeS Tat substrates, NrfC and NapG, from being exported. TatA, TatD and TatE were shown to be involved in the degradation of the substrates rejected by the Tat pathway.

Since the TatA C-terminus does not contribute to export it is possible the proofreading and degradation abilities of TatA arise from unidentified properties of the C-terminus. Consequently, removing the C-terminus of TatA might affect the proofreading abilities of the Tat pathway and stop rejection of misfolded and misassembled proteins. To test this hypothesis, the TatA truncation mutants were assessed for their ability to export wild-type and mutated NrfC, using the same methods as in previous studies (Matos *et al.*, 2008; Matos *et al.*, 2009).

NrfC forms part of the formate-dependent nitrite reductase complex and is a member of the CooF family of FeS proteins (Cole, 1996). Based on strong homology to formate dehydrogenase-N subunit β, which has had its three-dimensional structure solved (Jormakka *et al.*, 2002), the cysteine residues involved in the four predicted [4Fe-4S] centres were identified, Figure 6.9. Matos *et al.* had shown that the preprotein, pre-NrfC, with a C-terminal His-6 tag expressed from the arabinose inducible vector pBAD24 is exported via the Tat pathway. Upon export to the periplasm the signal sequence is cleaved off and it is processed to the mature size, which runs as a lower molecular weight band. The proofreading ability of the system was demonstrated by the use of two NrfC mutants, M1 and M2. In M1, two of the cysteine residues involved in FeS centre 2 were substituted with alanine, as indicated

[†] *Work in §6.6 was done in collaboration with Cristina F. R. O Matos (Department of Biological Sciences, University of Warwick). The NrfC constructs were provided by Cristina F. R. O. Matos and the immunoblots were performed in equal collaboration.*

by asterisks in Figure 6.9. In M2, one cysteine residue from each of FeS centres 1 and 2 was substituted with alanine, as indicated by arrows in Figure 6.9. Expression of these mutants in wild-type cells led to their rejection by the Tat pathway and resulted in degradation.

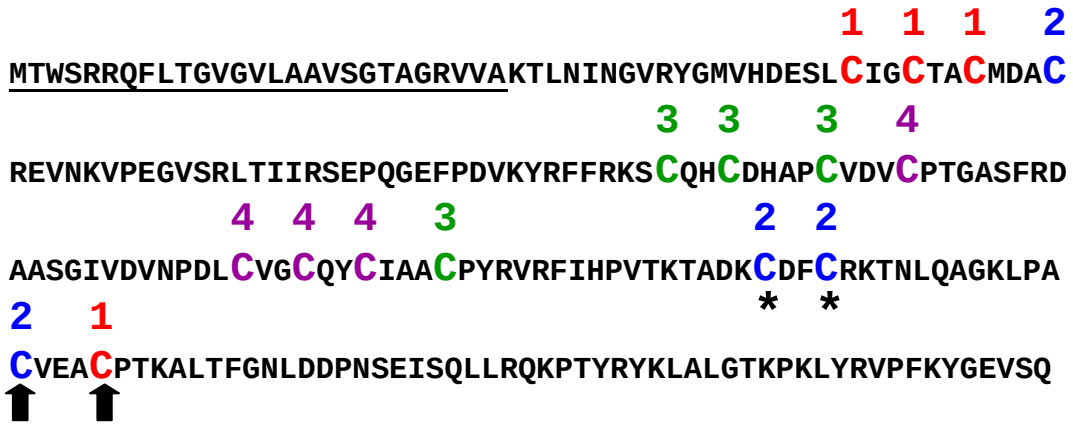


Figure 6.9 Predicted FeS ligands targeted for mutagenesis in NrfC. The primary sequence of *E. coli* pre-NrfC with the signal peptide underlined. The cysteine residues predicted to be involved in the formation of four [4Fe-4S] centres are highlighted, and are numbered with the FeS centres 1–4 that they belong to, based on strong homology to formate dehydrogenase-N subunit β. Residues substituted with alanine in the M1 and M2 NrfC mutants are indicated by asterisks and arrows, respectively. Figure adapted from (Matos *et al.*, 2008).

Two TatA truncations were assessed for their ability to export both wild-type and mutated NrfC. TatA truncated by 40 or 45 amino acids from the C-terminus with a His-6 tag for detection (A-40h, A-45h) were expressed from the pEXT22 vector, which is compatible with pBAD24. Expression of the constructs was confirmed by immunoblots of Δ*AE* cells expressing the truncations and pre-NrfC, Figure 6.10A. To assess the ability of the truncations to export wild-type and mutated NrfC, Δ*AE* cells co-expressing a truncation and an NrfC construct were separated into cytoplasmic, membrane and periplasmic fractions, and immunoblotted, Figure 6.10B. A-40h was able to export the wild-type NrfC, as demonstrated by a mature sized band in the periplasmic lane. As expected A-45h was unable to export wild-type NrfC, with no NrfC detected in the periplasm. This confirms the conclusions in §6.2 concerning the function of these truncations, and demonstrates that the results were not dependent on choice of substrate. None of the truncations were able to export

either of the NrfC mutants, $NrfC_{M1}$ and $NrfC_{M2}$, as demonstrated by the absence of a mature NrfC band in the periplasmic lane. This indicates that the C-terminus of TatA has no direct role in the proofreading of NrfC, as the Tat system is still able to reject these misfolded/misassembled proteins.

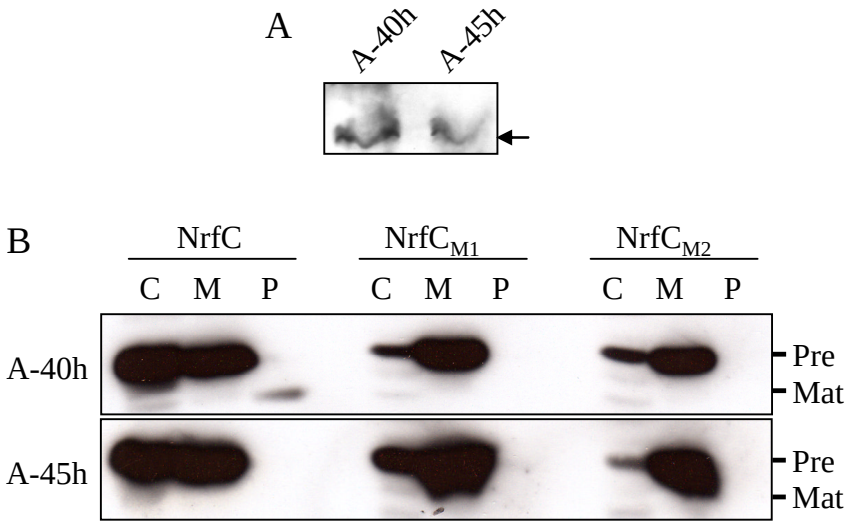


Figure 6.10 Analysis of the role the TatA C-terminus in the proofreading of NrfC. A. Immunoblots of whole Δ*AE* cells expressing TatA truncated by 40 or 45 amino acids with a C-terminal His-6 tag (A-40h, A-45h), using anti-His-6 antibodies. Arrows indicate the TatA constructs. **B.** Cytoplasmic, membrane and periplasmic samples (C, M, P) were prepared from Δ*AE* cells expressing the TatA constructs and wild-type or mutated His-6 tagged NrfC (NrfC, $NrfC_{M1}$, $NrfC_{M2}$). Immunoblots using anti-His-6 antibodies were performed. The precursor and mature forms of the NrfC constructs are indicated (Pre, Mat).

6.7 Discussion

The exact role of the C-terminus of TatA is still unknown, with previous research demonstrating the final 40 residues are not involved in the translocation of TorA (Lee *et al.*, 2002). The work presented here has sought to identify the minimal functional subunit of TatA and important features in the C-terminus, with an aim to determining its role in the Tat pathway.

6.7.1 An acidic motif in TatA is important for function and complex formation

Previous work has highlighted residues Phe39 to Lys49 as being important for function of TatA (Lee *et al.*, 2002). In order to pinpoint the important residues in this region a new C-terminal truncation of TatA by 45 amino acids was prepared. This truncation was unable to export the Tat substrates TorA and NrfC, allowing us to narrow down the important residues to a region just after the APH. Gel filtration analyses also indicated that the identified residues were important for complex formation, with the 45 residue truncation predominantly present in a single peak. Analysis of this region revealed an acidic motif which is highly conserved in TatA homologues from Gram-negative bacteria. This motif generally contains 2 to 4 acidic residues of varying composition, Figure 6.4. In contrast, this motif is not as ubiquitous in Gram-positive bacteria, Appendix A, which may be linked to recent findings that TatAd from *B. subtilis* does not form a variety of different sized complexes and instead forms a smaller, more homogeneous, complex (Barnett *et al.*, 2008).

Single substitutions in the DDE acidic motif of TatA with structurally similar but uncharged residues did not appear to affect function or complex formation. Substituting all three residues with polar, uncharged amino acids (NNQ) led to a reduction in the export of TorA and TorA-GFP by more than 50%. An additional triple substitution with uncharged, hydrophobic residues (LLM) prevented the Tat pathway from functioning. The triple substitutions also affected the formation of ordered TatA complexes. Blue-native gel analyses revealed that both of the triple mutants resulted in fewer, broad bands, when compared to wild-type TatA. Although the predicted pI of the mutants was outside the ideal range for blue-native analyses, the broadening and disappearance of bands points towards destabilisation of the TatA complexes. It is possible that a subset of the NNQ mutant may still form the correct complexes, which allows export but these are masked by the majority of the protein forming the incorrect complexes.

These results suggest that the negative charge of the acidic motif, whilst important for efficient function of the Tat pathway, is not essential. Instead we suggest that the

acidic motif is ideally composed of extremely hydrophilic residues or those with the ability to hydrogen bond.

6.7.2 Mutations of the lysine residues in the APH and the acidic motif cannot reinstate function

It has been previously observed that three lysines in the APH are important for function and complex formation (Barrett *et al.*, 2005; Barrett & Robinson, 2005). Plotting these residues on a helical wheel and creating a three-dimensional model of the APH highlighted their proximity to one another. To test whether this positively charged region was interacting with the negatively charged acidic motif, mutants consisting of the acidic motif triple mutation and substitution of the lysines with glutamines were tested for functional activity. No export of TorA or TorA-GFP by the mutants was detected, demonstrating that the introduction of the additional substitutions was not able to restore function of the Tat pathway.

Although no export was seen, this does not demonstrate conclusively that the APH lysines do not interact with the acidic motif. Simply by mutating the oppositely charged regions to uncharged residues with the ability to hydrogen bond may not be sufficient to restore TatA function. The interactions may be more specific, requiring more than just the ability to hydrogen bond. Additionally, the mutations may cause some conformational changes which prevent the two regions from coming together and contributing to the formation of stable complexes. Other techniques must be used to rule out charge interactions between these two regions.

6.7.3 Truncation of TatA does not affect the affect the proofreading capabilities of the Tat pathway

To determine if the C-terminus of TatA was involved in the proofreading of Tat substrates, TatA truncation mutants were analysed for their ability to export wild-type and mutated NrfC. None of the truncation mutants exported mutated NrfC, demonstrating that the Tat pathway still retained its ability to proofread in the absence of the TatA C-terminus. Although this demonstrates that the C-terminus is not directly responsible for proofreading, it does not rule out the possibility that it is involved in some manner. It is probable that proofreading in the Tat pathway is a

highly regulated system, as export of misfolded/misassembled proteins would be a waste of resources and could have a detrimental affect on the cell. Indeed it has been suggested there are two stages of checks; a transport quality control stage to ensure only correctly folded proteins are exported, and a proofreading stage to confirm correct cofactor insertion (Sargent, 2007). If this is the case, then several factors could be involved in ensuring Tat substrates are acceptable for export. Soluble, substrate-specific chaperones have been identified for molybdoproteins such as TorA and DmsA (Ilbert *et al.*, 2003; Jack *et al.*, 2004; Oresnik *et al.*, 2001). TatA may function cooperatively with other proteins in the Tat pathway or these chaperones, to ensure only correctly assembled proteins are exported. Thus, if the C-terminus of TatA was involved, its deletion alone would not be sufficient to reinstate export of mutated Tat substrates.

The involvement of TatA in the degradation of proteins rejected by the Tat pathway was not investigated here. The involvement of the Tat apparatus in the proofreading and degradation of rejected proteins is a very recent discovery, and more work must be done to determine the exact role of the different components.

In conclusion, the work presented here has led to the identification of an acidic motif after the APH of TatA. These acidic residues are important for both function and complex formation. Initial attempts to identify possible interactions between the motif and an adjacent positively charged region did not demonstrate any charge interactions. More work must be done to identify the exact role of the acidic motif and how its mutation affects the Tat pathway. Truncation of TatA from the C-terminus did not show any effect on the proofreading capabilities of the Tat pathway. However this additional role of TatA is a very recent discovery and much more work must be done to determine how the Tat pathway is able to recognise, reject and degrade misassembled protein.

Chapter 7. Conclusion

The Tat pathway differs from the majority of protein transport systems, since it is able to translocate folded proteins, often with cofactors bound. Substrates that are exported via this pathway can be identified by an invariant twin-arginine motif in their signal peptide, which gives the pathway its name. Since its discovery it has generated much interest due to several factors. The ability to export folded proteins using only the membrane proton motive force (PMF) may enable this pathway to be used in biotechnological protein production of heterologous proteins not compatible with the Sec pathway (Brüser, 2007). Furthermore, the lack of a Tat pathway in animal and fungal cells and the importance of selected Tat substrates in virulence suggest it could be a suitable target for novel antimicrobial compounds (De Buck *et al.*, 2008). Despite being the focus of many pieces of research, our knowledge of the mechanism of the Tat pathway is still behind that of other major transport systems, such as the Sec pathway (Driessen & Nouwen, 2008). Initial research identified three integral membrane proteins – TatA, TatB and TatC – that are essential for function of the Tat pathway in *E. coli*. They form two distinct types of complex; a TatABC complex and separate TatA complexes of varying size, which are believed to form the membrane receptor and translocation pore, respectively. The mechanism by which these complexes recognise and translocate Tat substrates across the membrane is still unknown. The work presented in this thesis sought to gain further understanding of the Tat pathway, using a combination of bioimaging, biochemical and biophysical techniques to answer the following questions:

- Are there differences in the recognition of Tat substrates by the Tat pathways in *Synechocystis* and *E. coli*?
- Does the TM domain of *E. coli* TatA play a role in function and complex formation?
- Is the APH of *E. coli* TatA able to insert into membranes and contribute to complex formation?
- Do parts of the C-terminal domain of *E. coli* TatA contribute to function and complex formation?

This chapter brings together the work undertaken in order to address these questions and summarises how the results of the work has furthered our knowledge of the Tat

pathway. Furthermore, a model for the assembly of TatA complexes suggested by these results is discussed.

7.1 Differences in the *E. coli* and *Synechocystis* Tat pathways

Although components of the Tat pathway have been identified in the cyanobacterium *Synechocystis* sp. PCC6803 (Aldridge *et al.*, 2008; Fulda *et al.*, 2000), it is not known if it differs from the more widely studied *E. coli* and chloroplast thylakoid systems. Since cyanobacteria possess an internal thylakoid membrane system in addition to a plasma membrane, they may contain characteristics of both bacterial and chloroplast Tat pathways. In Chapter 3 work was undertaken to determine if the two bacteria differed in their requirement for a twin-arginine motif in the signal peptide of Tat substrates. Export of TorA-GFP with different substitutions in the twin-arginine motif was examined in both bacteria. Unlike *E. coli*, where a single conservative substitution of arginine with lysine in the twin-arginine motif does not block export, both of the twin-arginines are essential for export in *Synechocystis*. This indicates a closer similarity to the chloroplast thylakoid Tat pathway, where single substitutions of either arginine block translocation (Chaddock *et al.*, 1995).

Bioimaging was used in this study of *Synechocystis* as it presents a non-invasive, *in vivo* technique, which can be used to obtain data about the cellular location of fluorescent molecules. Due to the added complexity of the thylakoid membranes, it is difficult to separate the different cellular compartments of *Synechocystis* and confocal microscopy provides an alternative when determining cellular location. This technique has been successfully used in a recent study to identify the location of native Tat substrates and one of the TatA/B homologues in *Synechocystis* (Aldridge *et al.*, 2008). There are, however, a number of factors that must be considered when analysing *Synechocystis* with confocal microscopy. Due to the photosynthetic nature of the cells they can react to the laser and give off quite intense fluorescence, so images must be taken and analysed with care. It was originally intended to also use imaging techniques such as fluorescence recovery after photobleaching (FRAP) and light inducible GFP to track the movement of TorA-GFP in *Synechocystis* cells. However this proved difficult due to the small size of the cells compared to the laser used for bleaching. Due to these experimental constraints it was decided to change

direction and concentrate on the *E. coli* Tat pathway, by studying the TatA component using biochemical and biophysical techniques.

7.2 TatA function and complex formation in *E. coli*

The *E. coli* TatA presents an interesting subject for study due to its ability to form complexes of varying size (Gohlke *et al.*, 2005; Leake *et al.*, 2008; Oates *et al.*, 2005). It is essential for function of the Tat pathway, and more recently was also identified as playing a role in the targeting of misassembled Tat substrates for degradation (Matos *et al.*, 2008). Despite many studies focusing on TatA and the other components of the Tat pathway, it is still not known how TatA functions and forms complexes of variable diameter. The work presented in Chapters 4, 5 and 6 sought to analyse the role of the different domains of TatA, with a view to identifying the essential regions.

It has been suggested that the TM domain of TatA plays a key role in the formation of complexes (De Leeuw *et al.*, 2001; Greene *et al.*, 2007; Porcelli *et al.*, 2002). The evidence presented for this so far is indirect so we sought to analyse the role of the TM domain directly, using a synthetic peptide and the TOXCAT assay (Russ & Engelman, 1999). Although predicted to be membrane protein there is still some controversy over the cellular localisation of TatA. One recent study claimed it formed cytoplasmic tubes (Berthelmann *et al.*, 2008), whereas another demonstrated it was only present in the cytoplasmic membrane of *E. coli* (Leake *et al.*, 2008). The $TatA_{TM}$ peptide was analysed for its ability to insert into lipid bilayers using a variety of biophysical techniques. In all of the conditions tested the $TatA_{TM}$ peptide was able to insert spontaneously into membrane, giving support to the hypothesis that TatA is a membrane protein.

Interactions between the TatA TM domain were analysed directly *in vitro* using the $TatA_{TM}$ peptide, and in a natural membrane environment using the TOXCAT assay. Both of these approaches demonstrated relatively weak interactions. The weak nature of these interactions was not unexpected, as the need to form complexes of variable size may require promiscuous interactions, which have the ability to change with the oligomeric state. The weakness of these interactions also suggests that there may be factors other than the TM domain which drive complex formation.

Substitution of a TM domain glutamine, Gln8, which had previously been shown to play a role in function (Greene *et al.*, 2007), did not significantly disrupt the helix-helix interactions of the TM domain in the TOXCAT assay, and did not have a visible effect on complex formation, as assessed by blue-native PAGE. The substitution did however have an effect on function, although this was shown to depend on the expression levels of TatA. The ability of the cell to overcome the deleterious effects of the substitution, suggests that although important, the glutamine is not essential and other factors can compensate for its loss at higher expression levels. This demonstrates again that the TM domain may not be the only factor driving complex formation.

Previous research has indicated that TatA may have a dual topology (Chan *et al.*, 2007; Gouffi *et al.*, 2004), which has led to the hypothesis that the TatA APH inserts into the membrane as part of the translocation process in order to form a hydrophilic pore through which proteins can pass. Analysis of $TatA_{APH}$ peptide revealed it was able to associate with lipid bilayers but no insertion was detected. This was also seen when the APH was analysed using the TOXCAT assay. Despite the inability of the APH TOXCAT chimera to insert correctly, they did show strong association with the membrane. Although insertion of the APH into lipid bilayers was not observed here, there may be conditions under which it may insert. The Tat pathway is driven by the PMF, with the electrochemical component playing an important role in *E. coli* (Bageshwar & Musser, 2007). Under the conditions required for transport it might be possible to detect insertion, however at the present time it is not well understood how the components of the PMF drive translocation. The identification of the key components and how they work will be a crucial point to establish in order for the mechanism of the Tat pathway to be determined.

It was also observed that the local environment of the APH has a strong effect on its secondary structure. Analysis of the $TatA_{APH}$ peptide demonstrated it was only α-helical in excess detergent or in the presence of negatively charged lipids. The *E. coli* cytoplasmic membrane contains a significant proportion of negatively charged lipids. These results therefore indicate that the APH must be in the proximity of the cytoplasmic membrane in order to form the correct secondary structure for function. This is in agreement with a recent study of the TatAd APH from the Gram-positive

B. subtilis (Lange *et al.*, 2007), demonstrating some key similarities between the different bacteria.

The C-terminal domain of TatA is unstructured and can be removed without affecting function. Analysis of the important regions of *E. coli* TatA highlighted a stretch of three acidic residues immediately after the APH. Alignments with TatA homologues from other Gram-negative bacteria showed that this region of several negatively charged amino acids was conserved. Removal of the negative charges did not completely abolish function, suggesting that the motif may optimally consist of extremely hydrophilic residues, or those with the ability to hydrogen-bond. Substitution of the acidic residues also affected complex formation, as judged by blue-native PAGE. TatA without the acidic motif was no longer able to form distinct complexes. Previously, a region of positively charged amino acids had been identified (Barrett *et al.*, 2005; Barrett & Robinson, 2005), however preliminary investigations could not determine if there was a charge interaction between these two regions. Since there is currently no three-dimensional structure of TatA, it is difficult to determine which areas of the protein may come into contact in the formation of complexes. This highlights the need for high resolution structures of all the Tat pathway components and the complexes they form in order to fully understand the interactions between the proteins and how they function.

Recently a role of TatA in the proofreading of FeS proteins was established (Matos *et al.*, 2008). Using the same methods, it was demonstrated that removal of the C-terminal domain of TatA could not restore the export of mutated NrfC. The role of the Tat components in proofreading is a very recent discovery and the quality control mechanisms of the Tat pathway are still poorly understood. It is likely that further factors are involved and only upon their identification can the exact role of TatA be determined.

7.3 New model for the assembly of TatA complexes

The results presented here can be summarised in a new model for TatA complex formation, Figure 7.1. In this model, the TM domain of TatA inserts into the membrane, bringing the APH into close contact with the membrane and enabling it to take on its α-helical secondary structure, Figure 7.1A. Weak interactions between

the TM domains enable the TatA monomers to come together to form higher order oligomers, Figure 7.1B. Additional interactions between the acidic motif (identified here) and a previously identified positively charged region, possibly located in the APH, stabilise the formation of these oligomers and the TatA complexes are able to form, Figure 7.1C. The apparently weak, but significant, TM domain interactions could be a factor in the dynamic nature of the TatA, enabling the oligomers to assemble and disassemble with relative ease.

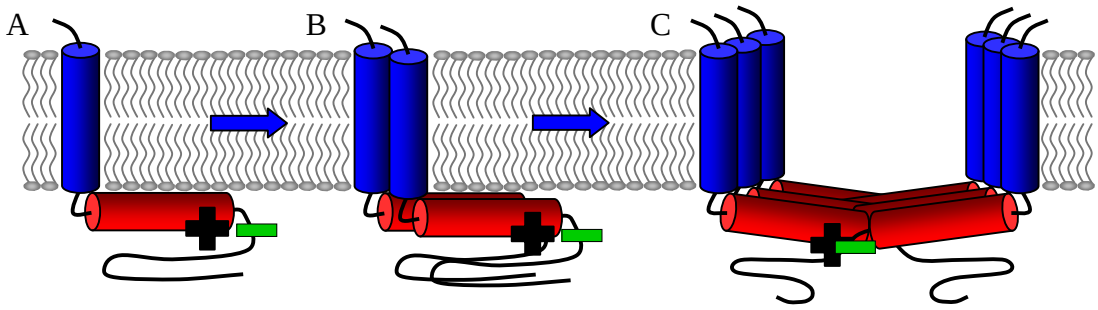


Figure 7.1 Proposed model for TatA complex formation. A. The TatA monomer is inserted into the membrane and association with the membrane stabilises the secondary structure of the APH. **B.** Weak TM domain interactions drive the formation of higher order TatA oligomers. **C.** Charge interactions between the acidic motif in the C-terminal domain and a positively charged region, possibly in the APH, stabilise the formation of TatA oligomers.

7.4 Future directions

Despite the many studies of the Tat pathway there still remains much to be determined about the mechanism of this interesting system. The research presented here has again demonstrated the need for a high resolution structure of the components of the Tat pathway and the complexes they form. High resolution structures of the components of the Sec pathway have proved extremely important in the characterisation of this system. So far attempts to obtain crystals of the components if the Tat pathway have proved unsuccessful, possibly due to the detergents required for purification. Work is currently underway to try and obtain higher resolution structures of the Tat complexes using electron microscopy. It is hoped these structures will provide information on how the components interact with each other, which will enable accurate modelling of the Tat pathway.

Another important future direction is the determination of the factors which drive translocation. The electrochemical gradient has been identified as the major factor in driving translocation (Bageshwar & Musser, 2007). By developing new tools to analyse the Tat pathway *in vitro* it may be possible to determine if the TatA APH inserts into the membrane as part of the translocation process. This would provide essential information on the topology of TatA, which remains somewhat controversial.

The discovery that components of the Tat pathway are involved in the degradation of misassembled Tat substrates is a recent discovery and much more research is needed to fully understand this process (Matos *et al.*, 2008). Work is currently ongoing to try and determine how the Tat pathway identifies, rejects and then directs these misfolded proteins to be degraded. The identification of a role for TatD is an important step towards understanding the proofreading capabilities of the Tat pathway (Matos *et al.*, 2009). It is hoped that by characterising the proofreading process more information will be gained on other aspects of the Tat pathway.

In summary, the work presented here has led to an increased understanding of the Tat pathway in *Synechocystis* and *E. coli*. Differences between the pathway in the two bacteria have been highlighted, and it has been demonstrated that confocal microscopy can be a useful tool in the analysis of the *Synechocystis* Tat pathway. Detailed analysis of the three domains of *E. coli* TatA has identified interactions which contribute to function and the formation of complexes, leading to the proposal of a new model for the formation of TatA complexes. Much work remains to be done in order to fully understand the mechanism of the Tat pathway; however this work provides a platform for future research towards complete characterisation of this important system.

Chapter 8. References

Abramoff, M. D., Magalhaes, P. J., Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International* **11**, 36-42.

Adair, B. D. & Engelman, D. M. (1994). Glycophorin A helical transmembrane domains dimerize in phospholipid bilayers: a resonance energy transfer study. *Biochemistry* **33**, 5539-5544.

Adams, P. D., Arkin, I. T., Engelman, D. M. & Brunger, A. T. (1995). Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat Struct Biol* **2**, 154-162.

Adams, P. D., Engelman, D. M. & Brunger, A. T. (1996). Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. *Proteins* **26**, 257-261.

Akimaru, J., Matsuyama, S., Tokuda, H. & Mizushima, S. (1991). Reconstitution of a protein translocation system containing purified SecY, SecE, and SecA from *Escherichia coli*. *Proc Natl Acad Sci U S A* **88**, 6545-6549.

Alami, M., Trescher, D., Wu, L. F. & Muller, M. (2002). Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli*. *J Biol Chem* **277**, 20499-20503.

Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H. G., Brunner, J. & Muller, M. (2003). Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol Cell* **12**, 937-946.

Alder, N. N. & Theg, S. M. (2003a). Energetics of protein transport across biological membranes. a study of the thylakoid ΔpH-dependent/cpTat pathway. *Cell* **112**, 231-242.

Alder, N. N. & Theg, S. M. (2003b). Energy use by biological protein transport pathways. *Trends Biochem Sci* **28**, 442-451.

Aldridge, C., Spence, E., Kirkilionis, M. A., Frigerio, L. & Robinson, C. (2008). Tat-dependent targeting of Rieske iron-sulphur proteins to both the plasma and thylakoid membranes in the cyanobacterium *Synechocystis* PCC6803. *Mol Microbiol* **70**, 140-150.

Aldridge, C., Cain, P. & Robinson, C. (2009). Protein transport in organelles: Protein transport into and across the thylakoid membrane. *FEBS J* **276**, 1177-1186.

Alfadhli, A., Steel, E., Finlay, L., Bachinger, H. P. & Barklis, E. (2002). Hantavirus nucleocapsid protein coiled-coil domains. *J Biol Chem* **277**, 27103-27108.

Allen, S. C., Barrett, C. M., Ray, N. & Robinson, C. (2002). Essential cytoplasmic domains in the *Escherichia coli* TatC protein. *J Biol Chem* **277**, 10362-10366.

Ames, G. F. (1968). Lipids of *Salmonella typhimurium* and *Escherichia coli*: Structure and Metabolism. *J Bacteriol* **95**, 833-843.

Angelini, S., Deitermann, S. & Koch, H. G. (2005). FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. *EMBO Rep* **6**, 476-481.

Ardhammar, M., Mikati, N. & Norden, B. (1998). Chromophore orientation in liposome membranes probed with flow dichroism. *J Am Chem Soc* **120**, 9957-9958.

Bageshwar, U. K. & Musser, S. M. (2007). Two electrical potential-dependent steps are required for transport by the Escherichia coli Tat machinery. *J Cell Biol* **179**, 87-99.

Barnett, J. P., Eijlander, R. T., Kuipers, O. P. & Robinson, C. (2008). A minimal Tat system from a gram-positive organism: a bifunctional TatA subunit participates in discrete TatAC and TatA complexes. *J Biol Chem* **283**, 2534-2542.

Barrett, C. M., Mathers, J. E. & Robinson, C. (2003a). Identification of key regions within the *Escherichia coli* TatAB subunits. *FEBS Lett* **537**, 42-46.

Barrett, C. M., Ray, N., Thomas, J. D., Robinson, C. & Bolhuis, A. (2003b). Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in *Escherichia coli*. *Biochem Biophys Res Commun* **304**, 279-284.

Barrett, C. M., Mangels, D. & Robinson, C. (2005). Mutations in subunits of the Escherichia coli twin-arginine translocase block function via differing effects on translocation activity or tat complex structure. *J Mol Biol* **347**, 453-463.

Barrett, C. M. & Robinson, C. (2005). Evidence for interactions between domains of TatA and TatB from mutagenesis of the TatABC subunits of the twin-arginine translocase. *FEBS J* **272**, 2261-2275.

Behrendt, J., Standar, K., Lindenstrauß, U. & Brüser, T. (2004). Topological studies on the twin-arginine translocase component TatC. *FEMS Microbiol Lett* **234**, 303-308.

Behrendt, J., Lindenstrauß, U. & Brüser, T. (2007). The TatBC complex formation suppresses a modular TatB-multimerization in *Escherichia coli*. *FEBS Lett* **581**, 4085-4090.

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

Berks, B. C. (1996). A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**, 393-404.

Berks, B. C., Sargent, F. & Palmer, T. (2000). The Tat protein export pathway. *Mol Microbiol* **35**, 260-274.

Berks, B. C., Palmer, T. & Sargent, F. (2003). The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol* **47**, 187-254.

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

Berthelmann, F., Mehner, D., Richter, S., Lindenstrauß, U., Lünsdorf, H., Hause, G. & Brüser, T. (2008). Recombinant expression of *tatABC* and *tatAC* results in the formation of interacting cytoplasmic TatA tubes in *Escherichia coli*. *J Biol Chem* **283**, 25281-25289.

Blaudeck, N., Kreutzenbeck, P., Muller, M., Sprenger, G. A. & Freudl, R. (2005). Isolation and characterization of bifunctional *Escherichia coli* TatA mutant proteins that allow efficient Tat-dependent protein translocation in the absence of TatB. *J Biol Chem* **280**, 3426-3432.

Bogsch, E., Brink, S. & Robinson, C. (1997). Pathway specificity for a ΔpH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein. *EMBO J* **16**, 3851-3859.

Bogsch, E. G., Sargent, F., Stanley, N. R., Berks, B. C., Robinson, C. & Palmer, T. (1998). An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* **273**, 18003-18006.

Bolhuis, A., Bogsch, E. G. & Robinson, C. (2000). Subunit interactions in the twin-arginine translocase complex of Escherichia coli. *FEBS Lett* **472**, 88-92.

Bolhuis, A., Mathers, J. E., Thomas, J. D., Barrett, C. M. & Robinson, C. (2001). TatB and TatC form a functional and structural unit of the twin-arginine translocase from Escherichia coli. *J Biol Chem* **276**, 20213-20219.

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

Bolhuis, A. (2004). The archaeal Sec-dependent protein translocation pathway. *Philos Trans R Soc Lond B Biol Sci* **359**, 919-927.

Brass, V., Bieck, E., Montserret, R., Wolk, B., Hellings, J. A., Blum, H. E., Penin, F. & Moradpour, D. (2002). An amino-terminal amphipathic alpha -helix mediates membrane association of the Hepatitis C virus nonstructural protein 5A. *J Biol Chem* **277**, 8130-8139.

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

Brink, S., Bogsch, E. G., Edwards, W. R., Hynds, P. J. & Robinson, C. (1998). Targeting of thylakoid proteins by the ΔpH-driven twin-arginine translocation pathway requires a specific signal in the hydrophobic domain in conjunction with the twin-arginine motif. *FEBS Lett* **434**, 425-430.

Brown, L.-A. & Baker, A. (2008). Shuttles and cycles: transport of proteins into the peroxisome matrix (Review). *Mol Membr Biol* **25**, 363 - 375.

Brundage, L., Hendrick, J. P., Schiebel, E., Driessen, A. J. & Wickner, W. (1990). The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* **62**, 649-657.

Brüser, T. (2007). The twin-arginine translocation system and its capability for protein secretion in biotechnological protein production. *Appl Microbiol Biotechnol* **76**, 35-45.

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

Buchanan, G., de Leeuw, E., Stanley, N. R., Wexler, M., Berks, B. C., Sargent, F. & Palmer, T. (2002). Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol Microbiol* **43**, 1457-1470.

Burstein, E. A., Vedenkina, N. S. & Ivkova, M. N. (1973). Fluorescence and the location of tryptophan residues in protein molecules. *Photochem Photobiol* **18**, 263-279.

Cabelli, R., Dolan, K., Qian, L. & Oliver, D. (1991). Characterization of membrane-associated and soluble states of SecA protein from wild-type and SecA51(TS) mutant strains of *Escherichia coli*. *J Biol Chem* **266**, 24420-24427.

Casadaban, M. J. & Cohen, S. N. (1979). Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: in vivo probe for transcriptional control sequences. *Proc Natl Acad Sci U S A* **76**, 4530-4533.

Chaddock, A. M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R. G., Klosgen, R. B. & Robinson, C. (1995). A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the ΔpH-dependent thylakoidal protein translocase. *EMBO J* **14**, 2715-2722.

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

Chen, H. & Kendall, D. A. (1995). Artificial transmembrane segments. *J Biol Chem* **270**, 14115-14122.

Christiaens, B., Symoens, S., Vanderheyden, S., Engelborghs, Y., Joliot, A., Prochiantz, A., Vandekerckhove, J., Rosseneu, M. & Vanloo, B. (2002). Tryptophan fluorescence study of the interaction of penetratin peptides with model membranes. *Eur J Biochem* **269**, 2918-2926.

Cline, K., Ettinger, W. & Theg, S. (1992). Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J Biol Chem* **267**, 2688-2696.

Cline, K. & Mori, H. (2001). Thylakoid ΔpH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J Cell Biol* **154**, 719-729.

Cole, J. (1996). Nitrate reduction to ammonia by enteric bacteria: redundancy, or a strategy for survival during oxygen starvation? *FEMS Microbiol Lett* **136**, 1-11.

Cornut, I., Desbat, B., Turlet, J. M. & Dufourcq, J. (1996). In situ study by polarization modulated Fourier transform infrared spectroscopy of the structure and orientation of lipids and amphipathic peptides at the air-water interface. *Biophys J* **70**, 305-312.

Cristobal, S., de Gier, J. W., Nielsen, H. & von Heijne, G. (1999). Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J* **18**, 2982-2990.

Cross, B. C., Sinning, I., Luirink, J. & High, S. (2009). Delivering proteins for export from the cytosol. *Nat Rev Mol Cell Biol* **10**, 255-264.

Dabney-Smith, C., Mori, H. & Cline, K. (2003). Requirement of a Tha4-conserved transmembrane glutamate in thylakoid Tat translocase assembly revealed by biochemical complementation. *J Biol Chem* **278**, 43027-43033.

Dabney-Smith, C., Mori, H. & Cline, K. (2006). Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. *J Biol Chem* **281**, 5476-5483.

Dabney-Smith, C. & Cline, K. (2009). Clustering of C-terminal stromal domains of Tha4 homo-oligomers during translocation by the Tat protein transport system. *Mol Biol Cell* **20**, 2060-2069.

Dafforn, T. R. & Rodger, A. (2004). Linear dichroism of biomolecules: which way is up? *Curr Opin Struct Biol* **14**, 541-546.

De Buck, E., Maes, L., Meyen, E., Van Mellaert, L., Geukens, N., Anne, J. & Lammertyn, E. (2005). *Legionella pneumophila* Philadelphia-1 *tatB* and *tatC* affect intracellular replication and biofilm formation. *Biochem Biophys Res Commun* **331**, 1413-1420.

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

De Buck, E., Lammertyn, E. & Anne, J. (2008). The importance of the twin-arginine translocation pathway for bacterial virulence. *Trends Microbiol* **16**, 442-453.

de Gier, J.-W. L., Scotti, P. A., Sääf, A., Valent, Q. A., Kuhn, A., Luirink, J. & von Heijne, G. (1998). Differential use of the signal recognition particle translocase targeting pathway for inner membrane protein assembly in *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**, 14646-14651.

De Keersmaeker, S., Van Mellaert, L., Lammertyn, E., Vrancken, K., Anné, J. & Geukens, N. (2005). Functional analysis of TatA and TatB in *Streptomyces lividans*. *Biochem Biophys Res Commun* **335**, 973-982.

De Leeuw, E., Porcelli, I., Sargent, F., Palmer, T. & Berks, B. C. (2001). Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway. *FEBS Lett* **506**, 143-148.

DeLisa, M. P., Samuelson, P., Palmer, T. & Georgiou, G. (2002). Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* **277**, 29825-29831.

DeLisa, M. P., Tullman, D. & Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A* **100**, 6115-6120.

Dilks, K., Rose, R. W., Hartmann, E. & Pohlschroder, M. (2003). Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**, 1478-1483.

Ding, F. X., Xie, H. B., Arshava, B., Becker, J. M. & Naider, F. (2001). ATR-FTIR study of the structure and orientation of transmembrane domains of the *Saccharomyces cerevisiae* alpha-mating factor receptor in phospholipids. *Biochemistry* **40**, 8945-8954.

Dixon, A. M., Stanley, B. J., Matthews, E. E., Dawson, J. P. & Engelman, D. M. (2006). Invariant chain transmembrane domain trimerization: A step in MHC class II assembly. *Biochemistry* **45**, 5228-5234.

Driessen, A. J. (1992). Precursor protein translocation by the *Escherichia coli* translocase is directed by the protonmotive force. *EMBO J* **11**, 847-853.

Driessen, A. J. M. & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* **77**, 643-667.

Duong, F. & Wickner, W. (1998). Sec-dependent membrane protein biogenesis: SecYEG, preprotein hydrophobicity and translocation kinetics control the stop-transfer function. *EMBO J* **17**, 696-705.

Dykxhoorn, D. M., St Pierre, R. & Linn, T. (1996). A set of compatible tac promoter expression vectors. *Gene* **177**, 133-136.

Egea, P. F., Stroud, R. M. & Walter, P. (2005). Targeting proteins to membranes: structure of the signal recognition particle. *Curr Opin Struct Biol* **15**, 213-220.

Engelman, D. M., Chen, Y., Chin, C.-N. & other authors (2003). Membrane protein folding: beyond the two stage model. *FEBS Lett* **555**, 122-125.

Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H. & Phillips, G. J. (2000). Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *J Bacteriol* **182**, 4068-4076.

Feldherr, C. M. & Akin, D. (1990). EM visualization of nucleocytoplasmic transport processes. *Electron Microsc Rev* **3**, 73-86.

Finazzi, G., Chasen, C., Wollman, F. A. & de Vitry, C. (2003). Thylakoid targeting of Tat passenger proteins shows no ΔpH dependence *in vivo*. *EMBO J* **22**, 807-815.

Fisher, L. E., Engelman, D. M. & Sturgis, J. N. (1999). Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain. *J Mol Biol* **293**, 639-651.

Fisher, L. E., Engelman, D. M. & Sturgis, J. N. (2003). Effect of detergents on the association of the glycophorin A transmembrane helix. *Biophys J* **85**, 3097-3105.

Freeman-Cook, L. L., Dixon, A. M., Frank, J. B., Xia, Y., Ely, L., Gerstein, M., Engelman, D. M. & DiMaio, D. (2004). Selection and characterization of small random transmembrane proteins that bind and activate the platelet-derived growth factor β receptor. *J Mol Biol* **338**, 907-920.

Fulda, S., Huang, F., Nilsson, F., Hagemann, M. & Norling, B. (2000). Proteomics of *Synechocystis* sp. strain PCC 6803. *Eur J Biochem* **267**, 5900-5907.

Geisse, N. A., Wasle, B., Saslowsky, D. E., Henderson, R. M. & Edwardson, J. M. (2002). Syncollin homo-oligomers associate with lipid bilayers in the form of doughnut-shaped structures. *J Membr Biol* **189**, 83-92.

Gerard, F. & Cline, K. (2006). Efficient twin arginine translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site. *J Biol Chem* **281**, 6130-6135.

Glover, J. R., Andrews, D. W. & Rachubinski, R. A. (1994). *Saccharomyces cerevisiae* peroxisomal thiolase is imported as a dimer. *Proc Natl Acad Sci U S A* **91**, 10541-10545.

Gohlke, U., Pullan, L., McDevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H. R. & Berks, B. C. (2005). The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102**, 10482-10486.

Gouffi, K., Gerard, F., Santini, C. L. & Wu, L. F. (2004). Dual topology of the Escherichia coli TatA protein. *J Biol Chem* **279**, 11608-11615.

Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2001). Polar side chains drive the association of model transmembrane peptides. *Proc Natl Acad Sci U S A* **98**, 880-885.

Greene, N. P., Porcelli, I., Buchanan, G., Hicks, M. G., Schermann, S. M., Palmer, T. & Berks, B. C. (2007). Cysteine scanning mutagenesis and disulfide mapping studies of the TatA component of the bacterial twin-arginine translocase. *J Biol Chem* **282**, 23937-23945.

Guzman, L. M., Barondess, J. J. & Beckwith, J. (1992). FtsL, an essential cytoplasmic membrane protein involved in cell division in Escherichia coli. *J Bacteriol* **174**, 7716-7728.

Halbig, D., Hou, B., Freudl, R., Sprenger, G. A. & Klosgen, R. B. (1999). Bacterial proteins carrying twin-R signal peptides are specifically targeted by the delta pH-dependent transport machinery of the thylakoid membrane system. *FEBS Lett* **447**, 95-98.

Harris, C. R. & Silhavy, T. J. (1999). Mapping an interface of SecY (PrlA) and SecE (PrlG) by using synthetic phenotypes and *in vivo* cross-linking. *J Bacteriol* **181**, 3438-3444.

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

Harwood, L. & Claridge, T. (1997). *Introduction to organic spectroscopy*: Oxford University Press New York.

Hatzixanthis, K., Palmer, T. & Sargent, F. (2003). A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol Microbiol* **49**, 1377-1390.

Hatzixanthis, K., Clarke, T. A., Oubrie, A., Richardson, D. J., Turner, R. J. & Sargent, F. (2005). Signal peptide-chaperone interactions on the twin-arginine protein transport pathway. *Proc Natl Acad Sci U S A* **102**, 8460-8465.

Henry, R., Carrigan, M., McCaffrey, M., Ma, X. & Cline, K. (1997). Targeting determinants and proposed evolutionary basis for the Sec and the ΔpH protein transport systems in chloroplast thylakoid membranes. *J Cell Biol* **136**, 823-832.

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

Hicks, M. G., Lee, P. A., Georgiou, G., Berks, B. C. & Palmer, T. (2005). Positive selection for loss-of-function tat mutations identifies critical residues required for TatA activity. *J Bacteriol* **187**, 2920-2925.

Hicks, M. R., Damianoglou, A., Rodger, A. & Dafforn, T. R. (2008). Folding and membrane insertion of the pore-forming peptide gramicidin occur as a concerted process. *J Mol Biol* **383**, 358-366.

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

Hou, B., Frielingsdorf, S. & Klosgen, R. B. (2006). Unassisted membrane insertion as the initial step in ΔpH/Tat-dependent protein transport. *J Mol Biol* **355**, 957-967.

Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38.

Hynds, P. J., Robinson, D. & Robinson, C. (1998). The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J Biol Chem* **273**, 34868-34874.

Ignatova, Z., Hornle, C., Nurk, A. & Kasche, V. (2002). Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. *Biochem Biophys Res Commun* **291**, 146-149.

Ilbert, M., Mejean, V., Giudici-Orticoni, M.-T., Samama, J.-P. & Iobbi-Nivol, C. (2003). Involvement of a mate chaperone (TorD) in the maturation pathway of molybdoenzyme TorA. *J Biol Chem* **278**, 28787-28792.

Iwamoto, T., You, M., Li, E., Spangler, J., Tomich, J. M. & Hristova, K. (2005). Synthesis and initial characterization of FGFR3 transmembrane domain: consequences of sequence modifications. *Biochim Biophys Acta* **1668**, 240-247.

Ize, B., Gerard, F. & Wu, L. F. (2002). *In vivo* assessment of the Tat signal peptide specificity in *Escherichia coli*. *Arch Microbiol* **178**, 548-553.

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

Jack, R. L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T. & Sargent, F. (2004). Coordinating assembly and export of complex bacterial proteins. *EMBO J* **23**, 3962-3972.

Jenei, Z. A., Borthwick, K., Zammit, V. A. & Dixon, A. M. (2009). Self-association of Transmembrane Domain 2 (TM2), but Not TM1, in Carnitine Palmitoyltransferase 1A: Role of GXXXG(A) motifs. *J Biol Chem* **284**, 6988-6997.

Jeong, K. J., Kawarasaki, Y., Gam, J., Harvey, B. R., Iverson, B. L. & Georgiou, G. (2004). A periplasmic fluorescent reporter protein and its application in high-throughput membrane protein topology analysis. *J Mol Biol* **341**, 901-909.

Johnson, A. E. & van Waes, M. A. (1999). The translocon: a dynamic gateway at the ER membrane. *Annu Rev Cell Dev Biol* **15**, 799-842.

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202.

Jones, J. D., McKnight, C. J. & Gierasch, L. M. (1990). Biophysical studies of signal peptides: implications for signal sequence functions and the involvement of lipid in protein export. *J Bioenerg Biomembr* **22**, 213-232.

Jongbloed, J. D., Antelmann, H., Hecker, M. & other authors (2002). Selective contribution of the twin-arginine translocation pathway to protein secretion in Bacillus subtilis. *J Biol Chem* **277**, 44068-44078.

Jongbloed, J. D., van der Ploeg, R. & van Dijl, J. M. (2006). Bifunctional TatA subunits in minimal Tat protein translocases. *Trends Microbiol* **14**, 2-4.

Jormakka, M., Tornroth, S., Byrne, B. & Iwata, S. (2002). Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. *Science* **295**, 1863-1868.

Kahn, T. & Engelman, D. (1992). Bacteriorhodopsin can be refolded from two independently stable transmembrane helixes and the complementary five-helix fragment. *Biochemistry* **31**, 6144-6151.

Kim, J. Y., Fogarty, E. A., Lu, F. J., Zhu, H., Wheelock, G. D., Henderson, L. A. & DeLisa, M. P. (2005). Twin-arginine translocation of active human tissue plasminogen activator in *Escherichia coli*. *Appl Environ Microbiol* **71**, 8451-8459.

Kipping, M., Lilie, H., Lindenstrauß, U., Andreesen, J. R., Griesinger, C., Carlomagno, T. & Brüser, T. (2003). Structural studies on a twin-arginine signal sequence. *FEBS Lett* **550**, 18-22.

Klosgen, R. B., Brock, I. W., Herrmann, R. G. & Robinson, C. (1992). Proton gradient-driven import of the 16 kDa oxygen-evolving complex protein as the full precursor protein by isolated thylakoids. *Plant Mol Biol* **18**, 1031-1034.

Kostakioti, M., Newman, C. L., Thanassi, D. G. & Stathopoulos, C. (2005). Mechanisms of protein export across the bacterial outer membrane. *J Bacteriol* **187**, 4306-4314.

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

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Laidler, V., Chaddock, A. M., Knott, T. G., Walker, D. & Robinson, C. (1995). A SecY Homolog in *Arabidopsis thaliana*. *J Biol Chem* **270**, 17664-17667.

Lakowicz, J. R. (2006). *Principles of fluorescence spectroscopy*, 3rd edn. New York: Springer.

Lange, C., Muller, S. D., Walther, T. H., Burck, J. & Ulrich, A. S. (2007). Structure analysis of the protein translocating channel TatA in membranes using a multi-construct approach. *Biochim Biophys Acta* **1768**, 2627-2634.

Langosch, D., Brosig, B., Kolmar, H. & Fritz, H.-J. (1996). Dimerisation of the Glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *J Mol Biol* **263**, 525-530.

Lazarova, T., Brewin, K. A., Stoeber, K. & Robinson, C. R. (2004). Characterization of peptides corresponding to the seven transmembrane domains of human adenosine A_2a receptor. *Biochemistry* **43**, 12945-12954.

Leake, M. C., Greene, N. P., Godun, R. M., Granjon, T., Buchanan, G., Chen, S., Berry, R. M., Palmer, T. & Berks, B. C. (2008). Variable stoichiometry of the TatA component of the twin-arginine protein transport system observed by *in vivo* single-molecule imaging. *Proc Natl Acad Sci U S A* **105**, 15376-15381.

Lee, H. C. & Bernstein, H. D. (2001). The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc Natl Acad Sci U S A* **98**, 3471-3476.

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

Lee, P. A., Orriss, G. L., Buchanan, G. & other authors (2006). Cysteine-scanning mutagenesis and disulfide mapping studies of the conserved domain of the Twin-arginine translocase TatB component. *J Biol Chem* **281**, 34072-34085.

Lees, J. G., Miles, A. J., Wien, F. & Wallace, B. A. (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* **22**, 1955-1962.

Lemmon, M., Flanagan, J., Hunt, J., Adair, B., Bormann, B., Dempsey, C. & Engelman, D. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J Biol Chem* **267**, 7683-7689.

Li, R., Gorelik, R., Nanda, V., Law, P. B., Lear, J. D., DeGrado, W. F. & Bennett, J. S. (2004). Dimerization of the transmembrane domain of Integrin α_{IIb} subunit in cell membranes. *J Biol Chem* **279**, 26666-26673.

Liberton, M., Berg, R. H., Heuser, J., Roth, R. & Pakrasi, H. B. (2006). Ultrastructure of the membrane systems in the unicellular cyanobacterium *Synechocystis* sp strain PCC 6803. *Protoplasma* **227**, 129-138.

Lill, R., Dowhan, W. & Wickner, W. (1990). The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. *Cell* **60**, 271-280.

Macek, P., Zecchini, M., Pederzolli, C., Dallaserra, M. & Menestrina, G. (1995). Intrinsic Tryptophan Fluorescence of Equinatoxin-Ii, a Pore-Forming Polypeptide from the Sea-Anemone Actinia-Equina L, Monitors Its Interaction with Lipid-Membranes. *European Journal of Biochemistry* **234**, 329-335.

Marrington, R., Dafforn, T. R., Halsall, D. J. & Rodger, A. (2004). Micro-volume couette flow sample orientation for absorbance and fluorescence linear dichroism. *Biophys J* **87**, 2002-2012.

Marrington, R., Dafforn, T. R., Halsall, D. J., MacDonald, J. I., Hicks, M. & Rodger, A. (2005). Validation of new microvolume Couette flow linear dichroism cells. *Analyst* **130**, 1608-1616.

Matos, C. F., Robinson, C. & Di Cola, A. (2008). The Tat system proofreads FeS protein substrates and directly initiates the disposal of rejected molecules. *EMBO J* **27**, 2055-2063.

Matos, C. F., Di Cola, A. & Robinson, C. (2009). TatD is a central component of a Tat translocon-initiated quality control system for exported FeS proteins in *Escherichia coli*. *EMBO Rep* **In press**.

McDonough, J. A., Hacker, K. E., Flores, A. R., Pavelka, M. S., Jr. & Braunstein, M. (2005). The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial β-lactamases. *J Bacteriol* **187**, 7667-7679.

McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-405.

Melnyk, R. A., Partridge, A. W. & Deber, C. M. (2001). Retention of native-like oligomerization states in transmembrane segment peptides: Application to the *Escherichia coli* aspartate receptor. *Biochemistry* **40**, 11106-11113.

Melnyk, R. A., Partridge, A. W., Yip, J., Wu, Y. Q., Goto, N. K. & Deber, C. M. (2003). Polar residue tagging of transmembrane peptides. *Biopolymers* **71**, 675-685.

Mendel, S., McCarthy, A., Barnett, J. P., Eijlander, R. T., Nenninger, A., Kuipers, O. P. & Robinson, C. (2008). The *Escherichia coli* TatABC system and a *Bacillus subtilis* TatAC-type system recognise three distinct targeting determinants in twin-arginine signal peptides. *J Mol Biol* **375**, 661-672.

Mikhaleva, N. I., Santini, C.-L., Giordano, G., Nesmeyanova, M. A. & Wu, L.-F. (1999). Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in *Escherichia coli*. *FEBS Lett* **463**, 331-335.

Milenkovic, D., Muller, J., Stojanovski, D., Pfanner, N. & Chacinska, A. (2007). Diverse mechanisms and machineries for import of mitochondrial proteins. *Biol Chem* **388**, 891-897.

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C. L., 3rd, Ban, N. & Frank, J. (2005). Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* **438**, 318-324.

Mori, H. & Cline, K. (1998). A signal peptide that directs non-Sec transport in bacteria also directs efficient and exclusive transport on the thylakoid Delta pH pathway. *J Biol Chem* **273**, 11405-11408.

Mori, H. & Cline, K. (2002). A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid ΔpH/Tat translocase. *J Cell Biol* **157**, 205-210.

Mould, R. M. & Robinson, C. (1991). A proton gradient is required for the transport of two lumenal oxygen- evolving proteins across the thylakoid membrane. *J Biol Chem* **266**, 12189-12193.

Mullineaux, C. W. (2004). FRAP analysis of photosynthetic membranes. *J Exp Bot* **55**, 1207-1211.

Mullineaux, C. W., Nenninger, A., Ray, N. & Robinson, C. (2006). Diffusion of green fluorescent protein in three cell environments in *Escherichia coli*. *J Bacteriol* **188**, 3442-3448.

Musser, S. M. & Theg, S. M. (2000a). Proton transfer limits protein translocation rate by the thylakoid ΔpH/Tat machinery. *Biochemistry* **39**, 8228-8233.

Musser, S. M. & Theg, S. M. (2000b). Characterization of the early steps of OE17 precursor transport by the thylakoid ΔpH/Tat machinery. *Eur J Biochem* **267**, 2588-2598.

Nakai, M., Sugita, D., Omata, T. & Endo, T. (1993). Sec-Y protein is localized in both the cytoplasmic and thylakoid membranes in the cyanobacterium *Synechococcus* PCC7942. *Biochem Biophys Res Commun* **193**, 228-234.

Nakai, M., Nohara, T., Sugita, D. & Endo, T. (1994). Identification and characterization of the SecA protein homologue in the cyanobacterium *Synechococcus* PCC7942. *Biochem Biophys Res Commun* **200**, 844-851.

Nevskaya, N. A. & Chirgadze, Y. N. (1976). Infrared-spectra and resonance interactions of amide-I and amide-II vibrations of alpha-helix. *Biopolymers* **15**, 637-648.

Nordh, J., Deinum, J. & Norden, B. (1986). Flow orientation of brain microtubules studies by linear dichroism. *Eur Biophys J* **14**, 113-122.

Norling, B., Zak, E., Andersson, B. & Pakrasi, H. (1998). 2D-isolation of pure plasma and thylakoid membranes from the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett* **436**, 189-192.

Oates, J., Mathers, J., Mangels, D., Kuhlbrandt, W., Robinson, C. & Model, K. (2003). Consensus structural features of purified bacterial TatABC complexes. *J Mol Biol* **330**, 277-286.

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

Oates, J., Hicks, M., Dafforn, T. R., DiMaio, D. & Dixon, A. M. (2008). *In vitro* dimerization of the bovine papillomavirus E5 protein transmembrane domain. *Biochemistry* **47**, 8985-8992.

Oresnik, I. J., Ladner, C. L. & Turner, R. J. (2001). Identification of a twin-arginine leader-binding protein. *Mol Microbiol* **40**, 323-331.

Orriss, G. L., Tarry, M. J., Ize, B., Sargent, F., Lea, S. M., Palmer, T. & Berks, B. C. (2007). TatBC, TatB, and TatC form structurally autonomous units within the

twin arginine protein transport system of *Escherichia coli*. *FEBS Lett* **581**, 4091-4097.

Paetzel, M., Karla, A., Strynadka, N. C. J. & Dalbey, R. E. (2002). Signal Peptidases. *Chem Rev* **102**, 4549-4580.

Palmer, T. & Berks, B. C. (2003). Moving folded proteins across the bacterial cell membrane. *Microbiology* **149**, 547-556.

Papanikolau, Y., Papadovasilaki, M., Ravelli, R. B. G., McCarthy, A. A., Cusack, S., Economou, A. & Petratos, K. (2007). Structure of dimeric SecA, the *Escherichia coli* preprotein translocase motor. *J Mol Biol* **366**, 1545-1557.

Pop, O. I., Westermann, M., Volkmer-Engert, R., Schulz, D., Lemke, C., Schreiber, S., Gerlach, R., Wetzker, R. & Muller, J. P. (2003). Sequence-specific binding of prePhoD to soluble TatAd indicates protein-mediated targeting of the Tat export in *Bacillus subtilis*. *J Biol Chem* **278**, 38428-38436.

Popot, J.-L. & Engelman, D. M. (1990). Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* **29**, 4031-4037.

Popot, J.-L. & Engelman, D. M. (2000). Helical membrane protein folding, stability, and evolution. *Annu Rev Biochem* **69**, 881-922.

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

Porter, R. D. (1988). DNA transformation. *Methods Enzymol* **167**, 703-712.

Potter, L. C. & Cole, J. A. (1999). Essential roles for the products of the napABCD genes, but not napFGH, in periplasmic nitrate reduction by *Escherichia coli* K-12. *Biochem J* **344**, 69-76.

Prodöhl, A., Weber, M., Dreher, C. & Schneider, D. (2007). A mutational study of transmembrane helix-helix interactions. *Biochimie* **89**, 1433-1437.

Pugsley, A. P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* **57**, 50-108.

Punginelli, C., Maldonado, B., Grahl, S., Jack, R., Alami, M., Schroder, J., Berks, B. C. & Palmer, T. (2007). Cysteine scanning mutagenesis and topological mapping of the *Escherichia coli* Twin-arginine Translocase TatC component. *J Bacteriol* **189**, 5482-5494.

Randall, L. L. & Hardy, S. J. (1986). Correlation of competence for export with lack of tertiary structure of the mature species: a study in vivo of maltose-binding protein in E. coli. *Cell* **46**, 921-928.

Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663-669.

Rasband, W. S. (1997–2008). ImageJ. *U S National Institutes of Health, Bethesda, Maryland, USA*, **http://rsb.info.nih.gov/ij/**.

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A* **106**, 1760-1765.

Ray, N., Nenninger, A., Mullineaux, C. W. & Robinson, C. (2005). Location and mobility of twin arginine translocase subunits in the *Escherichia coli* plasma membrane. *J Biol Chem* **280**, 17961-17968.

Richter, S., Lindenstrauss, U., Lucke, C., Bayliss, R. & Brüser, T. (2007). Functional Tat transport of unstructured, small, hydrophilic proteins. *J Biol Chem* **282**, 33257-33264.

Rijkers, D. T. S., Kruijtzer, J. A. W., Killian, J. A. & Liskamp, R. M. J. (2005). A convenient solid phase synthesis of S-palmitoyl transmembrane peptides. *Tetrahedron Letters* **46**, 3341-3345.

Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. & Stanier, R. Y. (1979). Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Journal of General Microbiology* **111**, 1-61.

Rodger, A., Rajendra, J., Marrington, R. & other authors (2002). Flow oriented linear dichroism to probe protein orientation in membrane environments. *Phys Chem Chem Phys* **4**, 4051-4057.

Rodrigue, A., Chanal, A., Beck, K., Muller, M. & Wu, L. F. (1999). Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J Biol Chem* **274**, 13223-13228.

Rose, R. W., Brüser, T., Kissinger, J. C. & Pohlschröder, M. (2002). Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol* **45**, 943-950.

Russ, W. P. & Engelman, D. M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc Natl Acad Sci U S A* **96**, 863-868.

Russ, W. P. & Engelman, D. M. (2000). The GxxxG motif: A framework for transmembrane helix-helix association. *J Mol Biol* **296**, 911-919.

Saint-Joanis, B., Demangel, C., Jackson, M., Brodin, P., Marsollier, L., Boshoff, H. & Cole, S. T. (2006). Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases β-Lactam susceptibility and virulence. *J Bacteriol* **188**, 6669-6679.

Sambrook, J. & Russell, D. W. (2001). *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

San Miguel, M., Marrington, R., Rodger, P. M., Rodger, A. & Robinson, C. (2003). An Escherichia coli twin-arginine signal peptide switches between helical and unstructured conformations depending on the hydrophobicity of the environment. *Eur J Biochem* **270**, 3345-3352.

Santini, C. L., Ize, B., Chanal, A., Muller, M., Giordano, G. & Wu, L. F. (1998). A novel sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J* **17**, 101-112.

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

Sargent, F., Stanley, N. R., Berks, B. C. & Palmer, T. (1999). Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J Biol Chem* **274**, 36073-36082.

Sargent, F., Gohlke, U., De Leeuw, E., Stanley, N. R., Palmer, T., Saibil, H. R. & Berks, B. C. (2001). Purified components of the *Escherichia coli* Tat protein transport system form a double-layered ring structure. *Eur J Biochem* **268**, 3361-3367.

Sargent, F., Berks, B. C. & Palmer, T. (2002). Assembly of membrane-bound respiratory complexes by the Tat protein-transport system. *Arch Microbiol* **178**, 77-84.

Sargent, F. (2007). The twin-arginine transport system: moving folded proteins across membranes. *Biochem Soc Trans* **35**, 835-847.

Schägger, H. & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* **199**, 223-231.

Schägger, H., Cramer, W. A. & von Jagow, G. (1994). Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem* **217**, 220-230.

Schatz, G. & Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science* **271**, 1519-1526.

Schneider, D., Fuhrmann, E., Scholz, I., Hess, W. R. & Graumann, P. L. (2007). Fluorescence staining of live cyanobacterial cells suggest non-stringent chromosome segregation and absence of a connection between cytoplasmic and thylakoid membranes. *BMC Cell Biol* **8**, -.

Schuenemann, D., Amin, P., Hartmann, E. & Hoffman, N. E. (1999). Chloroplast SecY is complexed to SecE and involved in the translocation of the 33-kDa but not the 23-kDa subunit of the Oxygen-evolving complex. *J Biol Chem* **274**, 12177-12182.

Scopes, R. K. (1974). Measurement of protein by spectrophotometry at 205 nm. *Anal Biochem* **59**, 277-282.

Senes, A., Gerstein, M. & Engelman, D. M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J Mol Biol* **296**, 921-936.

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

Shanmugham, A., Wong Fong Sang, H. W., Bollen, Y. J. M. & Lill, H. (2006). Membrane binding of twin arginine preproteins as an early step in translocation. *Biochemistry* **45**, 2243-2249.

Sharma, V., Arockiasamy, A., Ronning, D. R., Savva, C. G., Holzenburg, A., Braunstein, M., Jacobs, W. R. & Sacchettini, J. C. (2003). Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein translocating ATPase. *Proc Natl Acad Sci U S A* **100**, 2243-2248.

Sharpe, S., Barber, K. R., Grant, C. W. M., Goodyear, D. & Morrow, M. R. (2002). Organization of model helical peptides in lipid bilayers: Insight into the behavior of single-span protein transmembrane domains. *Biophys J* **83**, 345-358.

Sheinerman, F. B., Norel, R. & Honig, B. (2000). Electrostatic aspects of protein-protein interactions. *Curr Opin Struct Biol* **10**, 153-159.

Silvestro, A., Pommier, J., Pascal, M. C. & Giordano, G. (1989). The inducible trimethylamine N-oxide reductase of Escherichia coli K12: its localization and inducers. *Biochim Biophys Acta* **999**, 208-216.

Spence, E., Sarcina, M., Ray, N., Moller, S. G., Mullineaux, C. W. & Robinson, C. (2003). Membrane-specific targeting of green fluorescent protein by the Tat pathway in the cyanobacterium Synechocystis PCC6803. *Mol Microbiol* **48**, 1481-1489.

Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252-260.

Srinivas, S., Srinivas, R., Anantharamaiah, G., Segrest, J. & Compans, R. (1992). Membrane interactions of synthetic peptides corresponding to amphipathic helical segments of the human immunodeficiency virus type-1 envelope glycoprotein. *J Biol Chem* **267**, 7121-7127.

Srivastava, R., Pisareva, T. & Norling, B. (2005). Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. *Proteomics* **5**, 4905-4916.

Stanley, N. R., Palmer, T. & Berks, B. C. (2000). The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* **275**, 11591-11596.

Summer, E. J., Mori, H., Settles, A. M. & Cline, K. (2000). The thylakoid ΔpH-dependent pathway machinery facilitates RR-independent N-tail protein integration. *J Biol Chem* **275**, 23483-23490.

Switzer, R. C., 3rd, Merril, C. R. & Shifrin, S. (1979). A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal Biochem* **98**, 231-237.

Tatulian, S. A. (2003). Attenuated total reflection Fourier transform infrared spectroscopy: A method of choice for studying membrane proteins and lipids. *Biochemistry* **42**, 11898-11907.

Taylor, P. D., Toseland, C. P., Attwood, T. K. & Flower, D. R. (2006). TATPred: a Bayesian method for the identification of twin arginine translocation pathway signal sequences. *Bioinformation* **1**, 184-187.

Therien, A. G., Grant, F. E. M. & Deber, C. M. (2001). Interhelical hydrogen bonds in the CFTR membrane domain. *Nat Struct Biol* **8**, 597-601.

Thomas, J. D., Daniel, R. A., Errington, J. & Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli. *Mol Microbiol* **39**, 47-53.

Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354.

Treptow, N. A. & Shuman, H. A. (1985). Genetic evidence for substrate and periplasmic-binding-protein recognition by the MalF and MalG proteins, cytoplasmic

membrane components of the Escherichia coli maltose transport system. *J Bacteriol* **163**, 654-660.

Ullers, R. S., Luirink, J., Harms, N., Schwager, F. o., Georgopoulos, C. & Genevaux, P. (2004). SecB is a bona fide generalized chaperone in *Escherichia coli*. *Proc Natl Acad Sci U S A* **101**, 7583-7588.

van Dalen, A., Killian, A. & de Kruijff, B. (1999). $\Delta\psi$ stimulates membrane translocation of the C-terminal part of a signal sequence. *J Biol Chem* **274**, 19913-19918.

van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. & Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. *Nature* **427**, 36-44.

Vassylyev, D. G., Mori, H., Vassylyeva, M. N., Tsukazaki, T., Kimura, Y., Tahirov, T. H. & Ito, K. (2006). Crystal structure of the translocation ATPase SecA from *Thermus thermophilus* reveals a parallel, head-to-head dimer. *J Mol Biol* **364**, 248-258.

Voelker, R. & Barkan, A. (1995). Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J* **14**, 3905-3914.

von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane-proteins correlates with the trans-membrane topology. *EMBO J* **5**, 3021-3027.

von Heijne, G. (1990). The signal peptide. *J Membr Biol* **115**, 195-201.

Voulhoux, R., Ball, G., Ize, B., Vasil, M. L., Lazdunski, A., Wu, L. F. & Filloux, A. (2001). Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* **20**, 6735-6741.

Walton, P., Hill, P. & Subramani, S. (1995). Import of stably folded proteins into peroxisomes. *Mol Biol Cell* **6**, 675-683.

Weiner, J. H., Bilous, P. T., Shaw, G. M., Lubitz, S. P., Frost, L., Thomas, G. H., Cole, J. A. & Turner, R. J. (1998). A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93**, 93-101.

Wexler, M., Bogsch, E. G., Klosgen, R. B., Palmer, T., Robinson, C. & Berks, B. C. (1998). Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the delta pH pathway. *FEBS Lett* **431**, 339-342.

Wexler, M., Sargent, F., Jack, R. L., Stanley, N. R., Bogsch, E. G., Robinson, C., Berks, B. C. & Palmer, T. (2000). TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export. *J Biol Chem* **275**, 16717-16722.

White, S. H. & von Heijne, G. (2008). How translocons select transmembrane helices. *Annu Rev Biophys* **37**, 23-42.

Whitmore, L. & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392-400.

Xie, K., Hessa, T., Seppala, S., Rapp, M., von Heijne, G. & Dalbey, R. E. (2007). Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. *Biochemistry* **46**, 15153-15161.

Xie, K. & Dalbey, R. E. (2008). Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. *Nat Rev Micro* **6**, 234-244.

Yahr, T. L. & Wickner, W. T. (2001). Functional reconstitution of bacterial Tat translocation in vitro. *EMBO J* **20**, 2472-2479.

Yen, M. R., Tseng, Y. H., Nguyen, E. H., Wu, L. F. & Saier, M. H., Jr. (2002). Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol* **177**, 441-450.

Yuan, J., Henry, R., McCaffery, M. & Cline, K. (1994). SecA homolog in protein transport within chloroplasts: evidence for endosymbiont-derived sorting. *Science* **266**, 796-798.

Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. (2001). Polar residues drive association of polyleucine transmembrane helices. *Proc Natl Acad Sci U S A* **98**, 2250-2255.

Zimmer, J., Li, W. & Rapoport, T. A. (2006). A novel dimer interface and conformational changes revealed by an X-ray structure of *B. subtilis* SecA. *J Mol Biol* **364**, 259-265.

Appendix A. TatA alignments

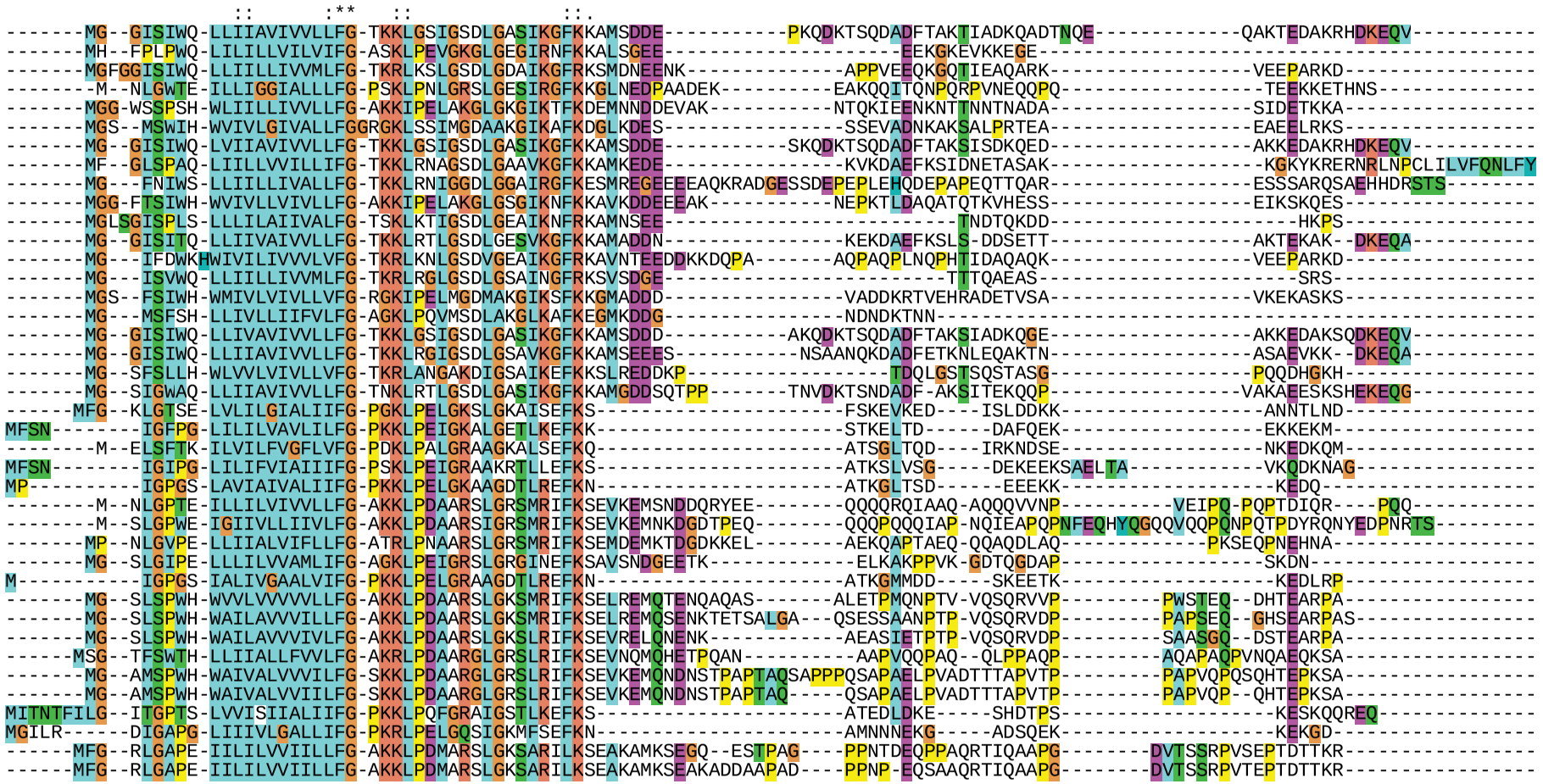
Number	Gram-negative Bacteria	Accession number
1	*Escherichia coli* (strain K12)	P69428
2	*Aquifex aeolicus*	O66477
3	*Azotobacter chroococcum*	A1EHP9
4	*Bdelliovibrio bacteriovorus*	Q6MK41
5	*Campylobacter jejuni*	Q9PNB9
6	*Caulobacter crescentus*	Q9A6T0
7	*Enterobacter cloacae*	A6N9K1
8	*Haemophilus influenzae*	P57046
9	*Halorhodospira halophila*	A1WW03
10	*Helicobacter pylori* J99	Q9ZMB8
11	*Legionella pneumophila*	Q5WSQ5
12	*Pasteurella multicoda*	Q9CKD3
13	*Pseudomonas aeruginosa*	Q9HUB5
14	*Pseudomonas stutzeri*	P95557
15	*Rhizobium loti*	Q98LC5
16	*Rickettsie prowazekii*	Q9ZCJ1
17	*Salmonella typhimurium*	P0A2H3
18	*Vibrio cholerae*	P57051
19	*Xyella fastidiosa*	Q9PFU3
20	*Yersina pestia*	Q8ZAM2

Number	Gram-positive bacteria	Accession number
21	*Alkaliphilus oremlandii*	A8MGV3
22	*Bacillus anthracis*	Q81R17
23	*Bacillus subtilis* TatAc	O31804
24	*Bacillus subtilis* TatAd	O31467
25	*Bacillus subtilis* TatAy	O05522
26	*Corynebacterium diphtheriae*	Q6NH98
27	*Corynebacterium glutamicum*	Q8NQE4
28	*Corynebacterium jeikeium*	Q4JVQ1
29	*Heliobacterium modesticaldum*	B0TFT7

30	Listeria monocytogenes	Q8YA04
31	*Mycobacterium leprae*	P54079
32	*Mycobacterium paratuberculosis*	Q73YX3
33	*Mycobacterium tuberculosis*	P66889
34	*Nocardia farcinica*	Q5YUY0
35	*Rhodococcus erythropolis*	P72267
36	*Rhodococcus* sp. (strain RHA1)	Q0SIG6
37	*Staphylococcus aureus*	Q5DVX3
38	*Streptococcus gordonii*	A8AXD8
39	*Streptomyces avermitilis*	Q828H7
40	*Streptomyces lividans*	Q7B115

Alignment of TatA homologues from Gram-negative bacteria
Alignment of TatA homologues from Gram-positive bacteria

Alignment of TatA homologues from Gram-negative and Gram-positive bacteria



Appendix B. Journal article

doi:10.1016/j.jmb.2009.02.060 J. Mol. Biol. (2009) 388, 122–132

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Contributions of the Transmembrane Domain and a Key Acidic Motif to Assembly and Function of the TatA Complex

Gemma Warren[1], Joanne Oates[2], Colin Robinson[3] and Ann M. Dixon[2*]

[1]*MOAC Doctoral Training Centre, University of Warwick, Coventry CV4 7AL, UK*

[2]*Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK*

[3]*Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK*

Received 28 October 2008; received in revised form 9 February 2009; accepted 25 February 2009
Available online 3 March 2009

Edited by J. Bowie

The twin-arginine translocase (Tat) pathway transports folded proteins across bacterial and thylakoid membranes. In *Escherichia coli*, a membrane-bound TatA complex, which oligomerizes to form complexes of less than 100 to more than 500 kDa, is considered essential for translocation. We have studied the contributions of various TatA domains to the assembly and function of this heterogeneous TatA complex. The TOXCAT assay was used to analyze the potential contribution of the TatA transmembrane (TM) domain. We observed relatively weak interactions between TatA TM domains, suggesting that the TM domain is not the sole driving force behind oligomerization. A potential hydrogen-bonding role for a TM domain glutamine was also investigated, and it was found that mutation blocks transport at low expression levels, while assembly is unaffected at higher expression levels. Analysis of truncated TatA proteins instead highlighted an acidic motif directly following the TatA amphipathic helix. Mutating these negatively charged residues to apolar uncharged residues completely blocks activity, even at high levels of TatA, and appears to disrupt ordered complex formation.

© 2009 Elsevier Ltd. All rights reserved.

Keywords: protein transport; twin-arginine translocase; TatA; transmembrane domain; TOXCAT

Introduction

In bacteria, two primary pathways are used for the export of proteins across the cytoplasmic membrane: the general secretory (Sec) pathway, and the twin-arginine translocase (Tat) pathway.[1,2] Unlike the Sec pathway, where proteins are threaded through the Sec translocase in an unfolded state, the Tat pathway is able to export folded proteins. The pathway derives its name from the twin-arginine motif in the N-terminal domain of the signal peptide of Tat substrates.[3,4]

The minimal components of the Tat pathway in *Escherichia coli* are TatA, TatB and TatC.[5–7] TatA and TatB are sequence-related proteins and are predicted to have the same secondary structure: an N-terminal α-helical transmembrane (TM) domain, followed by an amphipathic α-helix (APH) and a largely unstructured C-terminal region.[8–10] Despite their similarity, TatA and TatB have distinct roles in the Tat pathway,[10] and TatA is expressed at much higher levels than TatB and TatC.

Two distinct types of complex have been characterized in *E. coli* cytoplasmic membranes: a TatABC complex and separate TatA complexes. The TatABC complex has been purified and shown to contain equal amounts of TatB and TatC and variable amounts of TatA.[11] The majority of TatA is found in separate homo-oligomeric complexes that range in size from less than 100 kDa to well over 500 kDa.[8,9] These complexes have variable diameters and have been proposed to form the translocation pore through which proteins are exported.[8] It has been suggested that TatA size heterogeneity may be linked to the need to transport a variety of substrates, with the translocation pore perhaps assembled from appropriately sized Tat complexes. Counter to this idea, it has been shown that TatA heterogeneity is unaffected in export-defective TatC mutants.[12]

*Corresponding author. E-mail address: ann.dixon@warwick.ac.uk.

Abbreviations used: Tat, twin-arginine translocase; TM, transmembrane; Sec, secretory; APH, amphipathic α-helix; GpA, glycophorin A; MBP, maltose binding protein; CAT, chloramphenicol acetyl transferase; TMAO, trimethylamine *N*-oxide; TMAO reductase, TorA; BN-PAGE, blue-native PAGE; GFP, green fluorescent protein.

0022-2836/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.

It is highly unusual for single-span membrane proteins to assemble into such a heterogeneous range of complexes, and the mechanism of TatA complex formation is still unknown. A recent study has shown that, although TatB and TatC are required for the correct formation of large TatA complexes, smaller complexes containing an average of about four TatA molecules still form in the absence of TatB and TatC.[13] It has been hypothesized that the TatA TM domain is an important factor in TatA oligomerization.[14–16] When TatA is expressed without its TM domain it becomes a monomeric, soluble protein and loses some of its secondary structure.[16] Furthermore, a recent cysteine-scanning study highlighted a key glutamine residue in the TM domain that was important for Tat function.[15] Polar residues in TM domains have been implicated in the oligomerization of other membrane proteins,[17,18] although it has not been determined whether substitution of the glutamine in the TatA TM domain affects the structure of the TatA complex.[15]

It is also clear that regions other than the TM domain play a role in functional assembly of TatA. While the extreme TatA C-terminal domain can be removed without loss of function,[19] the APH is essential for activity.[20,21] Moreover, the TatA F39A mutation (in the middle of the APH) was shown to completely block Tat function[21] and severely disrupt TatA oligomerization.[12]

In this report we have set out to explore the roles of the TM domain and other regions of the protein in TatA assembly and function. We have used the TOXCAT assay[22] in conjunction with study of a synthetic peptide to directly probe helix–helix interactions between TatA TM domains, and we demonstrate relatively weak interactions between TM domains [compared to glycophorin A (GpA)]. Additionally, a new C-terminal truncation of TatA has identified the minimal functional subunit of TatA, which extends to a patch of negatively charged residues immediately after the APH. Through mutational analyses of this region, we show this acidic motif is important for TatA complex formation and function. These results have led us to discuss a new model for TatA complex formation that includes this acidic motif as well as the TM domain.

Results

The primary sequence of the 89-residue TatA polypeptide is shown in Fig. 1a. The putative structure of TatA is shown in Fig. 1b together with the various constructs used in this work. The TM domain stretches from (approximately) residues 6–20, followed by a predicted "hinge" region, an APH and a more unstructured C-terminal domain. The aim of this study was to analyze the role of these various domains in TatA assembly and function and to better understand the relationship between TatA heterogeneity and function.

Homo-oligomerization of the TatA TM domain

In recent studies, it has been suggested that oligomerization of TatA is mediated through TM

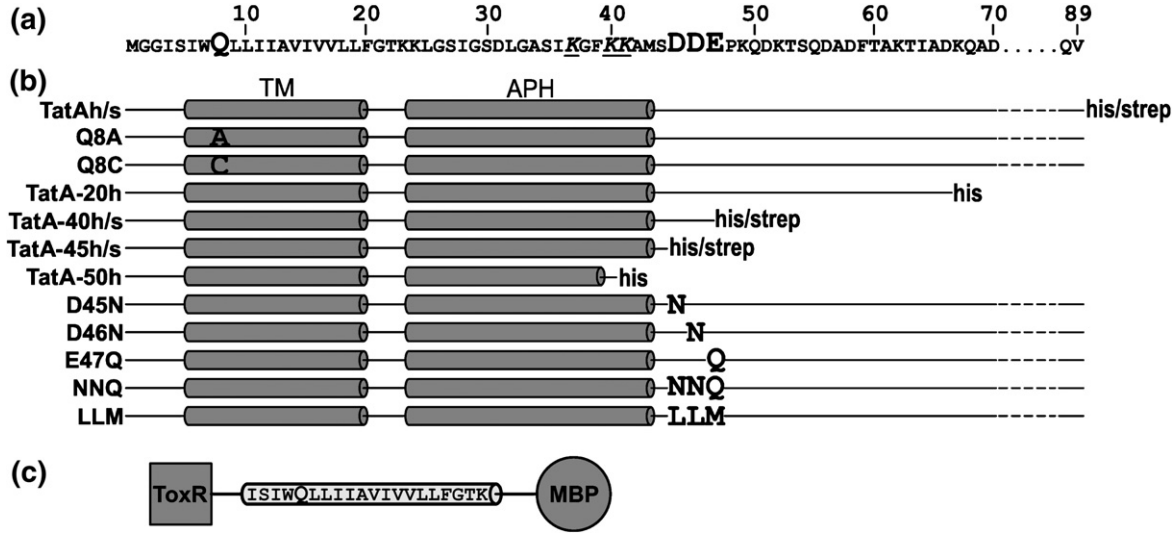


Fig. 1. TatA sequence and constructs used in this study. (a) The primary sequence of *E. coli* TatA. The glutamine in the TM domain and an acidic motif identified in this study are highlighted, and three lysines in the APH that have been shown to be important for function and complex formation are underlined.[12] (b) TatA constructs used in this study, with the TM domain and the APH labelled. Full-length TatA and C-terminal truncations by 20, 40, 45 and 50 amino acids were tagged with His-6 (his) or Strep II™ (strep) tags (TatAh/s, TatA-20h, TatA-40h/s, TatA-45h/s, TatA-50h). Mutations were introduced in the TM domain (Q8A, Q8C) and in the acidic motif after the APH (D45N, D46N, E47Q, NNQ, LLM). (c) The TOXCAT chimera used in this study. The TatA TM domain was inserted between the N-terminal dimerization-dependent DNA binding domain of ToxR (■) and the monomeric periplasmic anchor MBP (●). The glutamine that was mutated to both alanine and cysteine is highlighted.

domain interactions,[14–16] but the actual TatA assembly process is still poorly understood. To test this hypothesis *in vitro*, we analyzed a synthetic peptide corresponding to the TatA TM domain ($TatA_{TM}$; residues G2–K24). A lysine was added to the C-terminus to aid solubility of this very hydrophobic peptide, an approach that has been shown in the past not to affect the oligomerization of strongly interacting TM domains.[23,24] SDS-PAGE was used to analyze the oligomeric state of the $TatA_{TM}$ peptide at different concentrations (Fig. 2a). At low concentrations (10–50 μM), the peptide runs as a single band of less than 4 kDa; however, at higher concentrations (≥100 μM) a second band can be detected at about 6 kDa. The positions of the peptide bands do not correspond exactly to the predicted monomer and dimer masses (2.65 and 5.30 kDa, respectively). This aberrant behavior has been seen in previous work on membrane proteins and is thought to be caused by conformational effects.[25] Despite the difference in size, the bands were assigned to monomer and dimer, with the majority of the peptide present as a monomer. The tendency of the TM domain peptide to remain primarily as a monomer indicates that although a small fraction of the TM domain does oligomerize, these interactions are largely disrupted by SDS.

To analyze the TM domain interactions in a natural membrane environment, as opposed to detergent micelles that can be highly denaturing,[26] we used the TOXCAT assay[22] (see Materials and Methods). TOXCAT measures the ability of a given TM domain, in isolation from the rest of the protein, to self-associate in a membrane bilayer. Although TOXCAT data cannot be used to determine the precise oligomeric state, the data give a good indication of the strength of TM helix–helix interactions and, thus, of whether the TM domain is significantly involved in homo-oligomerization. A TOXCAT chimera containing the predicted TatA TM domain (Fig. 1c) was constructed and its topology in the *E. coli* inner membrane was confirmed with the *malE* complementation assay[22] (Fig. 2b). Briefly, in this assay, NT326 cells [which lack endogenous maltose binding protein (MBP)] are grown on plates where the sole carbon source is maltose. Growth in this medium is supported only for cells expressing TOXCAT chimera with the correct topology (i.e., MBP in the periplasm). Combined with the response to oligomerization [chloramphenicol acetyl transferase (CAT) activity], which confirms the cytoplasmic location of the ToxR domain, both the topology and the correct membrane insertion are confirmed. The resulting [CAT] activity of the chimera containing the TatA TM domain is shown in Fig. 2c. The TatA signal is normalized to a positive control, namely, the TM domain of GpA, which is known to strongly self-associate into dimers. Also shown are data from a point mutant of GpA, G83I, which greatly reduces the population of dimers[27] and is used as a negative control for the TOXCAT assay. Compared to the strongly and weakly dimerizing controls (GpA and G83I, respectively), the TatA TM yielded intermediate levels of CAT in the assay ([CAT] ~38% of GpA), suggesting that there is a significant level of interaction between TatA TM domains, but that the strength of this interaction is weaker than that observed for GpA.

A recent cysteine-scanning study of the TM domain and APH of TatA highlighted a glutamine residue, Gln8, in the TM domain that was important for function (Fig. 1a).[15] Mutation of this residue to

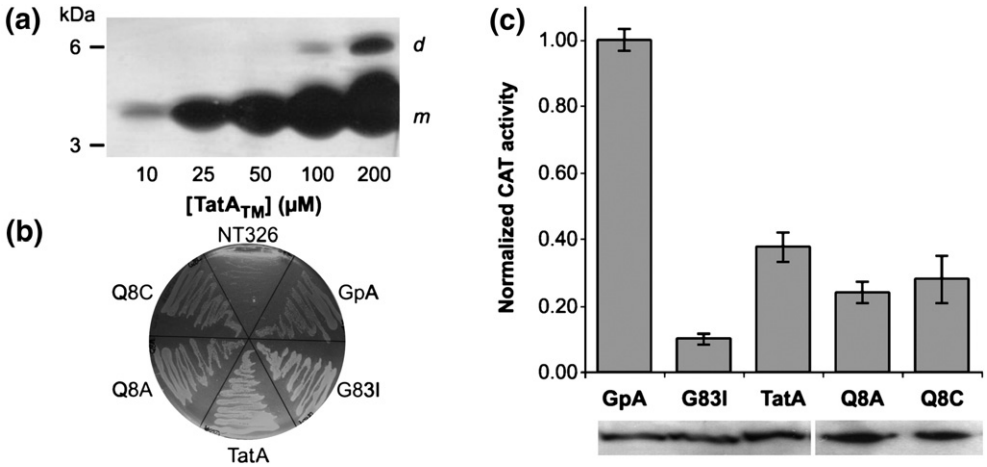


Fig. 2. TatA TM domains demonstrate a small but significant ability to self-associate as measured using SDS-PAGE gels and the TOXCAT assay. (a) SDS-PAGE analysis of the synthetic $TatA_{TM}$ peptide was carried out over a range of concentrations (stated below each lane). The majority of the peptide was present as monomer (*m*), with dimer (*d*) visible at the highest concentrations. (b) *malE*-deficient *E. coli* cells (NT326) and NT326 cells expressing TOXCAT chimera containing the TM domain of GpA (positive control), a mutant of GpA (G83I, negative control), the TatA TM domain (TatA) and the TatA TM domain with Gln8 mutated to alanine or cysteine (Q8A, Q8C), were grown on M9 agar containing 0.4% maltose to test for the topology of the TOXCAT chimera. (c) Levels of CAT activity in NT326 cells expressing the TatA TM TOXCAT chimera (TatA, Q8A, Q8C) were assayed (as detailed in Materials and Methods) along with GpA and G83I as positive and negative controls, respectively. CAT activity was normalized to GpA. Error bars show the standard deviation of three separate measurements. Anti-MBP immunoblots of each construct are shown below.

cysteine or alanine severely affected the export of TMAO (trimethylamine *N*-oxide) reductase (TorA) to the periplasm. The presence of hydrophilic residues in TM domains, and specifically Gln residues, has been shown to drive TM helix–helix association via electrostatic interactions and formation of interhelical hydrogen bonds.[28,29] Therefore, we studied the role of Gln8 in TatA TM domain oligomerization by mutagenesis of this residue and subsequent measurement with TOXCAT. Gln8 was substituted with cysteine (Q8C) and alanine (Q8A), and the resulting TOXCAT signals are shown in Fig. 2c. These mutations led to a noticeable decrease in [CAT] activity compared to wild-type TatA TM levels ($[CAT]_{Q8A}=24\%$ and $[CAT]_{Q8C}=28\%$); however, the effect is only significant within the error for the Q8A mutant.

Since these mutations did not dramatically affect the strength of TatA TM interactions as measured with the TOXCAT assay, we decided to analyze these mutations in the full-length protein, using a different expression system from the previous study mentioned above.[15] Mutants of TatA were prepared in the arabinose-inducible plasmid pBAD-ABC, and the effects of each mutation were analyzed in terms of both Tat export efficiency and TatA complex formation. Δ*ABCDE* cells containing pBAD-ABC or mutated pBAD-ABC were grown in the presence or in the absence of arabinose. The mutants were analyzed at both low and high expression levels in order to assess more rigorously any concentration-dependent effects of mutation. Previous work[11] has shown that the pBAD-ABC plasmid provides a near-wild-type level of Tat components when induced with very low (5 μM) concentrations of arabinose, and TatABC levels are slightly lower than wild-type levels in the absence of arabinose. We observe an approximately 25-fold increase over wild-type levels when induced with arabinose at 50 μM concentrations: similar expression levels have been reported in other studies of plasmid-borne Tat proteins.[30]

We first tested the activity of the Gln8 mutations using our expression system. Cells were separated into periplasmic and cytoplasmic fractions and assayed for the presence of TorA activity with the use of a native gel assay (Fig. 3a). SDS-PAGE and immunoblotting were also performed on membrane fractions to monitor the expression of TatA and TatB (Fig. 3b). In the control samples (Fig. 3a), TorA activity is clearly apparent in the periplasm (P) of wild-type MC4100 cells and no export is evident in the *tat* deletion strain (Δ*ABCDE* panel), where a slowly migrating form of TorA (TorA*) is observed in the cytoplasm (C).[20,31] When arabinose was omitted from the growth medium, the Tat components were expressed at very low levels (Fig. 3b). At these low expression levels, wild-type TatA (Δ*ABCDE*+ABC) is able to support export of TorA, but no export is detectable with the mutated forms (Q8A, Q8C). This result is in agreement with previous reports[15] studied in a different expression system. Interestingly, when arabinose was used to induce the expression of the Tat components, the TatA Q8A and Q8C mutants were both able to support export of TorA to the periplasm, as indicated by white bands in the periplasmic lanes of Fig. 3a. However, it is clear that these mutants are not exporting at wild-type levels and therefore the mutation is still having some effect even at higher concentrations. This result suggests that Gln8 is important, but not absolutely essential, for functional activity. At higher levels of expression, the increased concentration of TatA (~25-fold increase over wild type) can compensate for the deleterious effects of the mutations, especially given that the Q8 mutants still oligomerize (albeit more weakly) when analyzed using TOXCAT. A previous cysteine-scanning study[15] showed that the same mutations

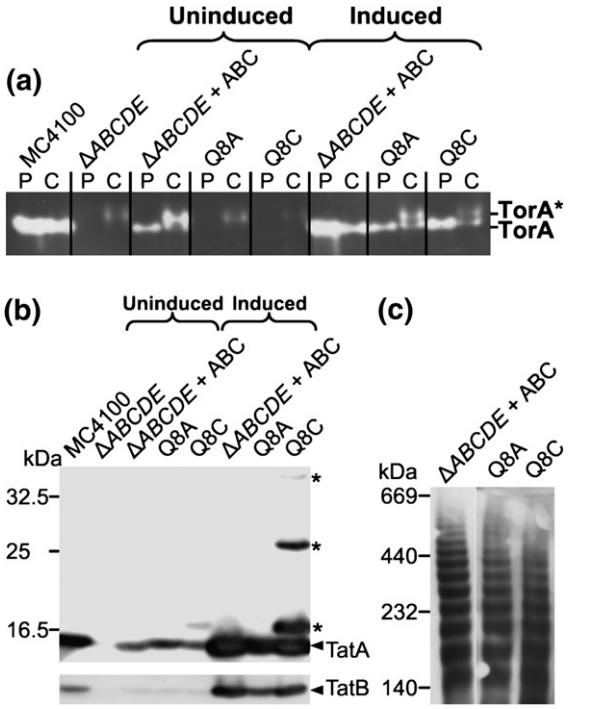


Fig. 3. The Gln8 mutation affects TatA function at low expression levels but does not affect complex formation. Periplasmic (P) and cytoplasmic (C) samples were prepared from wild-type *E. coli* cells (MC4100), from Δ*ABCDE* cells, and from Δ*ABCDE* cells expressing wild-type TatABC from plasmid pBAD-ABC (Δ*ABCDE*+ABC), or the same vector with TatA Gln8 mutated to alanine or cysteine (Q8A, Q8C), grown in the absence or presence of arabinose (uninduced, induced). (a) Native polyacrylamide gels of periplasmic and cytoplasmic samples were stained for TorA activity. White bands indicate the presence of TorA; TorA* indicates a slower migrating cytoplasmic form of TorA. (b) Immunoblots of membrane samples using TatA and TatB antibodies. Arrows indicate TatA and TatB. Asterisks denote unidentified TatA adducts caused by the Gln8→Cys mutation. (c) Membranes solubilized in digitonin were subjected to blue-native gel analyses (as described in Materials and Methods) and immunoblotted using TatA antibodies. The mobility of molecular mass markers (in kilodaltons) are indicated on the left.

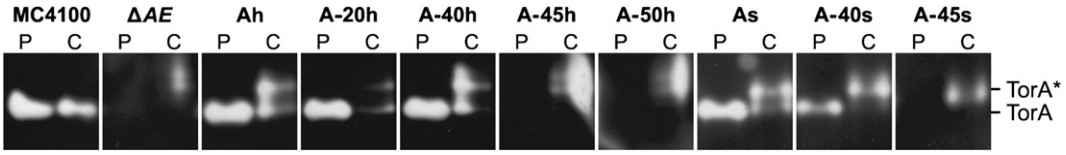


Fig. 4. C-terminal truncation of TatA identifies an acidic motif that is essential for function. Periplasmic (P) and cytoplasmic (C) samples were prepared from wild-type *E. coli* cells (MC4100), Δ*AE* cells (Δ*AE*), and Δ*AE* expressing full-length TatA, TatA truncated by 20, 40, 45 or 50 amino acids from the C-terminus with a C-terminal His-6 tag (Ah, A-20h, A-40h, A-45h, A-50h), and full-length TatA, TatA truncated by 40 or 45 amino acids from the C-terminus with a Strep II™ tag (As, A-40s, A-45s) from the pBAD24 plasmid. A native polyacrylamide gel was stained for TorA activity. White bands indicate the presence of TorA; TorA* indicates a slower migrating cytoplasmic form of TorA.

blocked export completely, but it is possible that the mutants were expressed at very low concentrations in that study.

By moderately overexpressing TatA, the higher concentration of protein also allows us to study the oligomeric state(s) of TatA mutants using blue-native gels.[9] Currently there is no method sensitive enough to visualize TatA oligomers at wild-type levels. High protein concentrations are necessary for characterization of protein oligomeric state, and this approach has been taken in several studies of TatA.[8,9,32] Although data must be interpreted with care, valuable information on assembly of this and many other protein complexes has been gained at higher concentration.

Using blue-native PAGE (BN-PAGE), we analyzed overexpressed wild-type TatA as well as the Q8A and Q8C mutants to determine how these mutations affect complex formation. Membranes isolated from Δ*ABCDE E. coli* cells expressing the wild-type and mutated pBAD-ABC plasmid were run on blue-native gels and immunoblotted with TatA antibodies to detect the range of TatA complexes (Fig. 3c). These analyses were carried out at protein concentrations similar to those used for the induced TorA assay, where we saw reduced export compared to those at wild-type levels. However, both TatA mutants were able to form oligomers identical to those formed by wild-type TatA (spanning from 100 to 500 kDa). Taken together, the above data suggest that Gln8 and the TM domain are only moderately involved in oligomer formation at these concentrations.

An acidic motif after the APH is important for TatA function

TOXCAT, SDS-PAGE, TorA export and blue-native analyses suggest that, while the TatA TM domain demonstrates a small but significant degree of self-association, there are most likely other regions in the protein that also contribute to the functional assembly of TatA. Previous work has shown that truncation of up to 40 amino acids from the C-terminus of TatA (A-40) still resulted in fully functional protein; however, truncation by 50 amino acids (A-50) abolished function, suggesting that the stretch of amino acids from residues 39 to 49 were crucial for function.[19] To further delineate important residues in this region, a construct corresponding to TatA truncated by 45 amino acids from the C-terminus was prepared, containing either a His-6 or Strep II™ C-terminal tag (Fig. 1b). The tags were necessary for detection in the absence of a suitable TatA antibody, and previous work has shown that a C-terminal His-6 tag on TatA does not affect its functional activity.[33] Truncations of TatA by 20, 40 and 50 residues with C-terminal His-6 or Strep II™ tags (Fig. 1b) were also prepared to allow comparison of our expression system with the previous study in a different expression system.[19] The activities of the His-6- and Strep II™-tagged truncation mutants were mea-

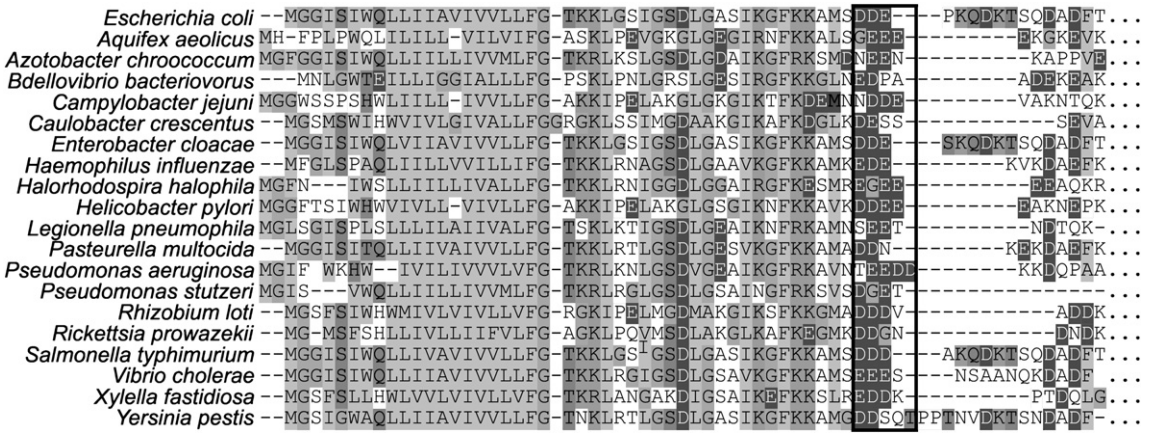


Fig. 5. The TatA acidic motif is conserved in TatA homologs from other Gram-negative bacteria. Sequences of TatA homologs from Gram-negative bacteria were identified and aligned using ClustalX[34]. The alignment shows the TM domain, the APH and the region immediately following the APH. An acidic region after the predicted APH is boxed.

sured with the TorA export assay (Fig. 4). In the control samples, TorA activity is clearly apparent in the periplasm of wild-type MC4100 cells and no export is evident in cells lacking TatA and TatE (Δ*AE* panel). Analysis of the tagged truncations show that A-20 and A-40 mutants (A-20h, A-40h and A-40s panels) support export of TorA, whereas the A-50 truncation mutant (A-50h panel) does not. These results agree with the data reported by Lee *et al.*[19] for plasmid-borne, overexpressed TatA, confirming that our tagged proteins were behaving in the same manner as previously studied untagged versions. In this study, an additional TatA mutant was prepared, A-45, in which 45 amino acids were deleted from the C-terminus. This mutation also abolished export via the Tat pathway (Fig. 4, A-45h and A-45s panels), identifying a stretch of five residues (Asp45–Lys49) that is important for function.

Analysis of these five residues revealed three consecutive acidic residues immediately following the predicted APH (DDE sequence highlighted in Fig. 1a). Alignment of TatA sequences from other Gram-negative bacteria shows that this represents a highly conserved acidic "motif" (Fig. 5). We use the term motif not to suggest the specific conservation of the sequence DDE, but instead to indicate a region of two or three acidic residues almost always present after the APH of many TatA proteins. The structural and functional relevance of these residues was analyzed by mutagenesis of TatA. Each of the three residues was mutated individually to a structurally similar but uncharged residue (D45N, D46N and E47Q), and a triple mutant was also prepared (NNQ) (Fig. 1b). Because asparagine and glutamine are able to hydrogen-bond and thus participate in strong non-covalent interactions, an additional triple mutant, which would be unable to hydrogen-bond with other residues, was prepared (LLM, Fig. 1b). The effect of each mutation on TatA function was analyzed with the TorA export assay in which TatA expression was induced with arabinose. The single mutations in the DDE sequence have no apparent effect on the functioning of the Tat pathway (Fig. 6a), with most of the TorA activity found in the periplasm in each case. However, the triple NNQ mutation significantly inhibited the export of TorA, with the majority of TorA found in the cytoplasm. When arabinose was omitted from the growth medium, resulting in almost wild-type expression levels, no TorA activity was seen in the periplasm (data not shown). The other triple mutation (LLM mutant) has a more drastic effect, completely blocking TorA export (Fig. 6a) even at high protein concentrations.

Since the TorA assay is a rather qualitative assay, the triple mutants were also examined using the more quantitative TorA–GFP (green fluorescent protein) export assay (Fig. 6b) to assess more accurately the efficiency of export. This assay involves expression of a chimeric construct composed of the TorA signal peptide linked to GFP. Export efficiencies can be assessed by immunoblotting.[35] The data show that TorA–GFP is efficiently exported

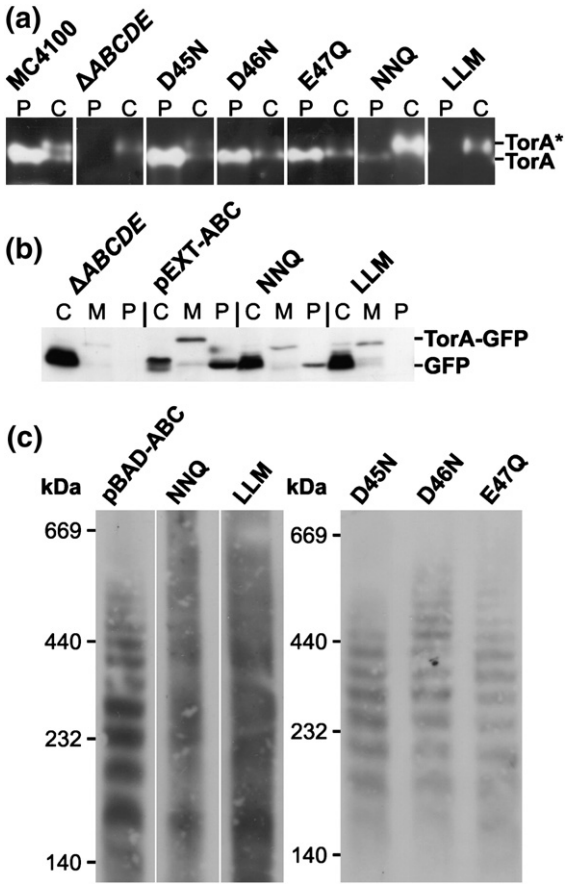


Fig. 6. The acidic motif is critical for function, and potentially for complex formation, but uncharged polar residues can support translocation activity. The effects of five different mutations of the acidic motif of TatA (Asp45→Asn, D45N; Asp46→Asn, D46N; Glu47→Gln, E47Q; Asp46→Asn, Asp46→Asn and Glu47→Gln, NNQ; Asp45→Leu, Asp46→Leu and Glu47→Met, LLM) were analyzed by a TorA activity assay, a TorA–GFP export assay and blue-native gel analysis. (a) Native polyacrylamide gel stained for TorA activity. Periplasmic (P), cytoplasmic (C) and membrane (M) samples were prepared from wild-type *E. coli* cells (MC4100), from Δ*ABCDE* cells, and from Δ*ABCDE* cells expressing TatABC with mutated TatA (D45N; D46N; E47Q; NNQ; LLM). White bands indicate the presence of TorA; TorA* indicates a slower migrating cytoplasmic form of TorA. (b) Immunoblots using a GFP antibody of C, M and P samples from Δ*ABCDE* cells expressing TorA–GFP from the pJDT1 plasmid (Δ*ABCDE*), Δ*ABCDE* cells expressing TorA–GFP and wild-type TatABC from the pEXT-ABC plasmid (pEXT-ABC), and Δ*ABCDE* cells expressing TorA–GFP and mutated TatABC from the pEXT-ABC plasmid (NNQ, LLM). (c) Membranes from Δ*ABCDE* cells expressing wild-type or mutated TatA from the pBAD-ABC plasmid solubilized in digitonin were subjected to blue-native gel analyses and immunoblotted using TatA antibodies.

in Δ*ABCDE* cells expressing either TatABC (from plasmid pEXT-ABC) or the same construct carrying the TatA NNQ mutation. In each case, most of the GFP is found as mature-sized protein in the periplasm. However, the triple mutation LLM leads to a complete block in export (Fig. 6b), as observed

above with TorA. The accumulation of mature-sized GFP in the cytoplasm is due to proteolysis of TorA–GFP when export is blocked.[35,36] These data confirm the importance of this motif of acidic amino acids and show that polar residues such as Gln and Asn can at least partially rescue function.

Oligomerization of TatA mutants carrying substitutions in the acidic motif

In order to investigate the link between correct TatA oligomer formation and function, blue-native gel analyses were performed on each of the mutants described above (Fig. 6c). The single mutants D45N, D46N, and E47Q, all of which were functional in the TorA assay, yield a banding pattern identical to that of wild-type TatA, indicating that these mutations have not affected the ability of TatA to assemble into oligomers. The triple mutations, however, result in substantial destabilization of the TatA complexes as judged by this approach (Fig. 6c, panels NNQ and LLM). Although there is some indication of oligomer formation for these mutants, many fewer bands are observed, and these bands are much broader and appear at different molecular mass values. This effect could be accounted for by the high theoretical p*I* of the triple mutants (8.96) compared with the ideal running conditions for BN-PAGE (p$I \leq 8.6$).[37] Proteins with p*I*s of 9.3 or above typically display broad bands in BN-PAGE, and their slow migration down the gel reflects an inaccurately high molecular mass.[37] However, the NNQ and LLM mutants are only just outside the range for BN-PAGE and, although we do see broadening of the bands, we also see many fewer bands. Taken with the observation that we see no such effect when any one of the acidic residues is removed independently, these data suggest that there may be a link between TatA structure and function, with most of the mutants either active and correctly structured (the single mutants), or inactive and disordered (LLM triple mutant). However, this will need to be tested further with complementary methods.

Discussion

TatA TM domains interact relatively weakly with one another

TatA is proposed to form homo-oligomeric complexes of variable diameters that act as the translocation pore. The mechanism by which assembly occurs is not known, but the TM domain of TatA has been strongly implicated in the formation of homo-oligomers. Previous work has shown that TatA expressed without its TM domain purifies as a soluble monomer,[16] leading to the conclusion that the TM domain was involved in oligomerization.[14–16] This conclusion was supported by a study that found that a glutamine residue in the TatA TM domain (Q8) was important for functional activity of the Tat pathway.[15] In this work, using a synthetic peptide and the TOXCAT assay, respectively, we have analyzed the TatA TM domain *in vitro* and *in vivo* in order to assess its ability to self-associate. We have also analyzed the role of Gln8 in this process, because this Gln could potentially influence TM helix–helix interactions through its ability to hydrogen-bond. Perhaps surprisingly, our data indicate the TM domains interact with each other relatively weakly. Measurements in a natural bilayer using the TOXCAT assay, an assay that has been highly effective for the study and quantification of inter-TM association,[18,22] suggests that TatA TM domains do not have a strong intrinsic ability to oligomerize, but instead demonstrate intermediate self-association. In addition, we find that these TM domain interactions are only mildly disrupted by the mutation of Gln8. The data from TOXCAT measurements are further supported by our results from blue-native gel analyses, where mutation of Gln8 produced no effect on the oligomerization of full-length TatA.

In light of the previous studies of TatA, where formation of variable-sized complexes has been observed, these results are less surprising. If one expects promiscuous interactions for TatA (i.e., variable-sized complexes), one would not expect an extremely strong interaction. These data also suggest that the TM domain is, at most, only partially responsible for the formation of TatA complexes, and other regions of the protein almost certainly play a role in the oligomerization observed for wild-type TatA.

An acidic motif after the APH is important for TatA function

Previous work has shown that the APH region is important for both assembly and function of TatA;[20,21] the apparently disordered C-terminal domain appears to be less important and can be deleted without blocking activity.[19] We have studied a new C-terminal truncation of TatA in an attempt to pinpoint specific areas within this region critical for function and have discovered a key region—a motif of consecutive acidic residues lying just outside the APH. This acidic motif is highly conserved within TatA subunits from Gram-negative bacteria, which generally contain two to three acidic residues in this region (see Fig. 5). Interestingly, the motif is much less prevalent in Gram-positive bacteria (not shown), and we speculate that this may be correlated with the recent observation that the TatA subunit from the *Bacillus subtilis* does not oligomerize in the same manner, instead forming a much smaller and more homogeneous homo-oligomeric complex.[38]

Single mutations in the acidic motif (to uncharged amino acids) did not apparently affect TatA function or complex formation, and triple mutation to polar but uncharged residues (NNQ) was still able to support export of TorA and TorA–GFP to the periplasm, although at a much reduced efficiency. However, substitution by uncharged and hydrophobic residues (LLM) leads to a complete loss in activity. These data indicate that negative charges

are not an absolute requirement for translocation activity, and we suggest that the motif may optimally consist of extremely hydrophilic residues, or perhaps those with an ability to hydrogen-bond.

How might the acidic motif be involved in TatA function? Previous studies have highlighted the importance of the three positively charged amino acids in the APH of TatA: lysine 37, lysine 40 and lysine 41.[12,20] Mutating these three lysine residues to glutamines, thus removing their charge, disrupted both functional activity of the Tat pathway and complex formation. Plotting these residues on a helical wheel shows that they are located on the same face of the APH (data not shown). Identification of an important positively charged region in the APH and a nearby negatively charged region leads to the hypothesis that there is a charge (electrostatic) interaction between these two regions, and the disruption of this interaction affects complex formation and functional activity.

These results have led us to propose the following method of TatA complex formation. In this model, the TatA TM domain allows the protein to fold correctly and stabilizes interactions between the APH and the membrane. Weak interactions between the TM domains also bring the proteins together in the correct orientation, allowing the negative region from one monomer to interact with positive residues from another monomer, possibly located in the APH. While further work is required to confirm the mechanism of association, these charge interactions could stabilize the formation of complexes and allow TatA to oligomerize correctly and perform its role in the export of proteins.

Materials and Methods

Bacterial strains and growth conditions

Bacterial strains used in this study are listed in Table 1. For TorA activity assays, TorA–GFP export assays and blue-native analyses, *E. coli* strain MC4100[39] was the parental strain. Δ*AE* and Δ*ABCDE* have been described before,[6,40] and arabinose-resistant derivatives were used as described previously.[33] *E. coli* was grown aerobically in Luria broth (LB) at 37 °C and anaerobically in LB supplemented with 0.5% glycerol, 0.4% TMAO and 1 μM sodium molybdate. Media supplements were used at the following concentrations: ampicillin, 100 μg/ml; kanamycin, 50 μg/ml; arabinose, 50 μM; and IPTG, 1 mM unless otherwise stated. For TOXCAT assays, *E. coli* NT326 cells[41] were grown aerobically in LB supplemented with ampicillin (200 μg/ml) as appropriate.

Plasmid construction

Plasmids used in this study are given in Table 1, and TatA constructs made in the study are shown in Fig. 1b. For the His-6- and Strep II™-tagged truncations, PCR was used to amplify *tatA* truncated by the relevant number of codons. Primers were designed to introduce a C-terminal tag and NcoI/PstI sites during amplification. The NcoI/PstI-digested PCR-amplified fragments were ligated into the NcoI/PstI site of pBAD24.

For the TOXCAT assay, residues corresponding to the TM domain of TatA (residues 4 to 23) were cloned into the NheI/BamHI site of pccKAN to make plasmid pccTatA. The TOXCAT chimera is shown in Fig. 1c.

Site-specific mutagenesis was used to mutate *tatA* in pBAD-ABC, pEXT-ABC and pccTatA, using the Quikchange® mutagenesis kit (Stratagene) according to the manufacturer's instructions. All primers used in this study are available on request.

TOXCAT assay

The TOXCAT assay was used to investigate the TM domain of TatA in a natural membrane as described previously.[22] In the TOXCAT system, a chimeric protein is used, which consists of the α-helical TM domain of interest inserted between the N-terminal dimerization-dependent DNA binding domain of ToxR, and the monomeric periplasmic anchor MBP. The chimera is expressed in *malE*-deficient *E. coli* NT326 cells along with a chloramphenicol acetyltransferase (CAT) reporter gene under the control of the ToxR-responsive *ctx* promoter. If the TM domains interact, the ToxR domain is able to dimerize and activate the *ctx* promoter, leading to the expression of

Table 1. Plasmids and strains

	Relevant properties	Reference
E. coli strains		
MC4100	F⁻ Δ*lacU*169 *araD*139 *rpsL*150 *relA*1 *ptsF rbs flbB*5301	39
MC4100 Δ*ABCDE*	tat deletion strain	40
MC4100 Δ*AE*	*tatA* and *tatE* deletion strain	6
NT326	*F-(argF-lac)U169, rpsL150, relA1, rbsR, flbB5301, ptsF25, thi-1, deoC1,* Δ*malE444, recA, srlA*⁺	41
Plasmids		
pBAD24	Cloning vector *araC, Amp*ʳ	42
pBAD-ABC	pBAD24 derivative containing the *E. coli tatABC* operon with a Strep II™ tag on the C-terminus of TatC; Ampʳ	11
pEXT-ABC	pEXT22 derivative containing the *E. coli tatABC operon*; Kanʳ	35
pBAD-Ah	pBAD24 derivative containing the *E. coli tatA* gene with a C-terminal His-6 tag; Ampʳ	33
pccKAN	Pkk232-8 (Stratagene) derivative containing the ctx::CAT reporter construct and the gene encoding the ToxR′-(kanamycin resistance cassette)-MBP chimera; Ampʳ and Kanʳ	22
pccGpA	pccKAN derivative containing the GpA TM domain; Ampʳ	22
pccG83I	pccGpA containing the G83I mutation (Gly83→Ile); Ampʳ	22

CAT. The level of CAT expression in the cells expressing the chimera is proportional to the strength of oligomerization. The TM domain of GpA, which is known to strongly dimerize, and its dimerization-defective mutant (G83I) were used as positive and negative controls, respectively. Expression levels of the chimera were checked by SDS-PAGE and Western blotting using anti-MBP antibodies as described below. Correct insertion and orientation of chimera in the membrane were confirmed using sodium hydroxide washes[43], protease sensitivity in a spheroplast assay,[22] and the *malE* complementation assay[22] where cells were grown on M9 agar plates containing 0.4% maltose. CAT assays were performed using the *FAST* CAT® Green (deoxy) Chloramphenicol Acetyltransferase assay kit (Invitrogen) according to the manufacturer's instructions.

Peptide sequence and purification

A peptide corresponding to the TM domain of TatA ($TatA_{TM}$) was synthesized by the Keck Facility at Yale University using F-moc chemistry. The sequence of the $TatA_{TM}$ peptide was GGISIWQLLIIAVIVVLLFGTKKK, containing residues G2–K24 from TatA as well as a non-native lysine residue to aid solubility and end caps on the C- and N-termini. The peptide was purified by reversed-phase HPLC using a linear acetonitrile gradient including 0.1% trifluoroacetic acid on a Phenomenex C4 column. The purity of pooled fractions was confirmed by mass spectrometry before subsequent lyophilization.

SDS-PAGE and Western blotting

SDS-PAGE of the synthetic peptide was carried out using NuPAGE 12% Bis–Tris gels (Invitrogen, Paisley, UK). Peptide concentrations were estimated from the absorbance at 280 nm using a molar extinction coefficient (ε_{280}) of 5500 mol^{-1} cm^{-1} for $TatA_{TM}$. Peptide visualization was achieved by staining with silver nitrate.

Proteins were separated by SDS-PAGE and immunoblotted using antibodies to TatA, TatB, GFP (Invitrogen), MBP (Sigma) or His-6 tag (Invitrogen). TatA, TatB and GFP were detected using horseradish-peroxidase-conjugated anti-rabbit IgG secondary antibody (Promega), and His-6 tag and MBP were detected using horseradish-peroxidase-conjugated anti-mouse IgG secondary antibody (Promega). The Strep II™ tag was detected directly using a Streptactin–horseradish peroxidase conjugate (Institut für Bioanalytik). An ECL detection kit (Amersham Biosciences) was used to visualize the proteins.

TorA activity and TorA–GFP export assay

TorA activity assays were performed as previously described.[11,44] *E. coli* cells were grown anaerobically for 3.5 h with appropriate amounts of arabinose before being separated into periplasmic, cytoplasmic and membrane fractions as described in Ref. 31. Fractions were run on a 10% native gel that was assayed for TorA activity as described previously. TorA–GFP export assays were performed as previously described.[35] *E. coli* cells containing pJDT1 and mutated pEXT-ABC were grown for 2 h in the presence of 200 μM arabinose to induce the expression of TorA–GFP. The cells were then washed and grown in the presence of 1 mM IPTG for 3 h before being separated into periplasmic, cytoplasmic and membrane fractions. The presence of TorA–GFP was determined by SDS-PAGE and immunoblotting with GFP antibodies as described above.

Blue-native gel electrophoresis

Δ*ABCDE* cells containing pBAD-ABC were grown anaerobically for 3.5 h with 50 μM arabinose. Membranes were prepared as described above and solubilized in 50 mM Bis–Tris, pH 7.0, 750 mM 6-aminocaproic acid and 2% (w/v) digitonin. Solubilized membranes were separated on a 5–13% gradient polyacrylamide gel. Blue native gel analysis was performed as described,[9] based on the method described by Schägger *et al.*[37,45] Proteins were detected by immunoblotting as described above.

Acknowledgements

We thank J. Crawford for peptide synthesis. This work was supported by an Engineering and Physical Sciences Research Council studentship to G.W., by a Medical Research Council Grant to A.M. D., and by Biotechnology and Biological Sciences Research Council grants to C.R.

References

1. Muller, M. (2005). Twin-arginine-specific protein export in *Escherichia coli*. *Res. Microbiol.* **156**, 131–136.
2. Robinson, C. & Bolhuis, A. (2004). Tat-dependent protein targeting in prokaryotes and chloroplasts. *Biochim. Biophys. Acta*, **1694**, 135–147.
3. Chaddock, A. M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R. G., Klosgen, R. B. & Robinson, C. (1995). A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase. *EMBO J.* **14**, 2715–2722.
4. Stanley, N. R., Palmer, T. & Berks, B. C. (2000). The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J. Biol. Chem.* **275**, 11591–11596.
5. Bogsch, E. G., Sargent, F., Stanley, N. R., Berks, B. C., Robinson, C. & Palmer, T. (1998). An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem.* **273**, 18003–18006.
6. Sargent, F., Bogsch, E. G., Stanley, N. R., Wexler, M., Robinson, C., Berks, B. C. & Palmer, T. (1998). Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J.* **17**, 3640–3650.
7. Weiner, J. H., Bilous, P. T., Shaw, G. M., Lubitz, S. P., Frost, L., Thomas, G. H. *et al.* (1998). A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell*, **93**, 93–101.
8. Gohlke, U., Pullan, L., McDevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T. *et al.* (2005). The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc. Natl Acad. Sci. USA*, **102**, 10482–10486.
9. Oates, J., Barrett, C. M., Barnett, J. P., Byrne, K. G., Bolhuis, A. & Robinson, C. (2005). The *Escherichia coli* twin-arginine translocation apparatus incorporates a

distinct form of TatABC complex, spectrum of modular TatA complexes and minor TatAB complex. *J. Mol. Biol.* **346**, 295–305.
10. Sargent, F., Stanley, N. R., Berks, B. C. & Palmer, T. (1999). Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J. Biol. Chem.* **274**, 36073–36082.
11. Bolhuis, A., Mathers, J. E., Thomas, J. D., Barrett, C.M. & Robinson, C. (2001). TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *J. Biol. Chem.* **276**, 20213–20219.
12. Barrett, C. M., Mangels, D. & Robinson, C. (2005). Mutations in subunits of the *Escherichia coli* twin-arginine translocase block function via differing effects on translocation activity or tat complex structure. *J. Mol. Biol.* **347**, 453–463.
13. Leake, M. C., Greene, N. P., Godun, R. M., Granjon, T., Buchanan, G., Chen, S. *et al.* (2008). Variable stoichiometry of the TatA component of the twin-arginine protein transport system observed by in vivo single-molecule imaging. *Proc. Natl Acad. Sci. USA*, **105**, 15376–15381.
14. De Leeuw, E., Porcelli, I., Sargent, F., Palmer, T. & Berks, B. C. (2001). Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway. *FEBS Lett.* **506**, 143–148.
15. Greene, N. P., Porcelli, I., Buchanan, G., Hicks, M. G., Schermann, S. M., Palmer, T. & Berks, B. C. (2007). Cysteine scanning mutagenesis and disulfide mapping studies of the TatA component of the bacterial twin arginine translocase. *J. Biol. Chem.* **282**, 23937–23945.
16. Porcelli, I., de Leeuw, E., Wallis, R., van den Brink-van der Laan, E., de Kruijff, B., Wallace, B. A. *et al.* (2002). Characterization and membrane assembly of the TatA component of the *Escherichia coli* twin-arginine protein transport system. *Biochemistry*, **41**, 13690–13697.
17. Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2001). Polar side chains drive the association of model transmembrane peptides. *Proc. Natl Acad. Sci. USA*, **98**, 880–885.
18. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. (2001). Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. USA*, **98**, 2250–2255.
19. Lee, P. A., Buchanan, G., Stanley, N. R., Berks, B. C. & Palmer, T. (2002). Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation. *J. Bacteriol.* **184**, 5871–5879.
20. Barrett, C. M. & Robinson, C. (2005). Evidence for interactions between domains of TatA and TatB from mutagenesis of the TatABC subunits of the twin-arginine translocase. *FEBS J.* **272**, 2261–2275.
21. Hicks, M. G., Lee, P. A., Georgiou, G., Berks, B. C. & Palmer, T. (2005). Positive selection for loss-of-function tat mutations identifies critical residues required for TatA activity. *J. Bacteriol.* **187**, 2920–2925.
22. Russ, W. P. & Engelman, D. M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl Acad. Sci. USA*, **96**, 863–868.
23. Ding, F. X., Xie, H., Arshava, B., Becker, J. M. & Naider, F. (2001). ATR–FTIR study of the structure and orientation of transmembrane domains of the *Saccharomyces cerevisiae* alpha-mating factor receptor in phospholipids. *Biochemistry*, **40**, 8945–8954.
24. Melnyk, R. A., Partridge, A. W., Yip, J., Wu, Y., Goto, N. K. & Deber, C. M. (2003). Polar residue tagging of transmembrane peptides. *Biopolymers*, **71**, 675–685.
25. Therien, A. G., Grant, F. E. & Deber, C. M. (2001). Interhelical hydrogen bonds in the CFTR membrane domain. *Nat. Struct. Biol.* **8**, 597–601.
26. Fisher, L. E., Engelman, D. M. & Sturgis, J. N. (2003). Effect of detergents on the association of the glycophorin A transmembrane helix. *Biophys. J.* **85**, 3097–3105.
27. Doura, A. K., Kobus, F. J., Dubrovsky, L., Hibbard, E. & Fleming, K. G. (2004). Sequence context modulates the stability of a GxxxG-mediated transmembrane helix–helix dimer. *J. Mol. Biol.* **341**, 991–998.
28. Dixon, A. M., Stanley, B. J., Matthews, E. E., Dawson, J. P. & Engelman, D. M. (2006). Invariant chain transmembrane domain trimerization: a step in MHC class II assembly. *Biochemistry*, **45**, 5228–5234.
29. Freeman-Cook, L. L., Dixon, A. M., Frank, J. B., Xia, Y., Ely, L., Gerstein, M. *et al.* (2004). Selection and characterization of small random transmembrane proteins that bind and activate the platelet-derived growth factor beta receptor. *J. Mol. Biol.* **338**, 907–920.
30. Buchanan, G., de Leeuw, E., Stanley, N. R., Wexler, M., Berks, B. C., Sargent, F. & Palmer, T. (2002). Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol. Microbiol.* **43**, 1457–1470.
31. Barrett, C. M., Mathers, J. E. & Robinson, C. (2003). Identification of key regions within the *Escherichia coli* TatAB subunits. *FEBS Lett.* **537**, 42–46.
32. Berthelmann, F., Mehner, D., Richter, S., Lindenstrauss, U., Lunsdorf, H., Hause, G. & Bruser, T. (2008). Recombinant expression of tatABC and tatAC results in the formation of interacting cytoplasmic TatA tubes in *Escherichia coli*. *J. Biol. Chem.* **283**, 25281–25289.
33. Bolhuis, A., Bogsch, E. G. & Robinson, C. (2000). Subunit interactions in the twin-arginine translocase complex of *Escherichia coli*. *FEBS Lett.* **472**, 88–92.
34. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H. *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
35. Barrett, C. M., Ray, N., Thomas, J. D., Robinson, C. & Bolhuis, A. (2003). Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **304**, 279–284.
36. Thomas, J. D., Daniel, R. A., Errington, J. & Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol. Microbiol.* **39**, 47–53.
37. Schagger, H., Cramer, W. A. & von Jagow, G. (1994). Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* **217**, 220–230.
38. Barnett, J. P., Eijlander, R. T., Kuipers, O. P. & Robinson, C. (2008). A minimal Tat system from a gram-positive organism: a bifunctional TatA subunit participates in discrete TatAC and TatA complexes. *J. Biol. Chem.* **283**, 2534–2542.
39. Casadaban, M. J. & Cohen, S. N. (1979). Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: in vivo probe for transcriptional control sequences. *Proc. Natl Acad. Sci. USA*, **76**, 4530–4533.

40. Wexler, M., Sargent, F., Jack, R. L., Stanley, N. R., Bogsch, E. G., Robinson, C. *et al.* (2000). TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export.. *J. Biol. Chem.* **275**, 16717–16722.
41. Treptow, N. A. & Shuman, H. A. (1985). Genetic evidence for substrate and periplasmic-binding-protein recognition by the MalF and MalG proteins, cytoplasmic membrane components of the *Escherichia coli* maltose transport system. *J. Bacteriol.* **163**, 654–660.
42. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130.
43. Chen, H. & Kendall, D. A. (1995). Artificial transmembrane segments. *J. Biol. Chem.* **270**, 14115–14122.
44. Silvestro, A., Pommier, J., Pascal, M. C. & Giordano, G. (1989). The inducible trimethylamine *N*-oxide reductase of *Escherichia coli* K12: its localization and inducers. *Biochim. Biophys. Acta*, **999**, 208–216.
45. Schagger, H. & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223–231.