Twin Arginine Translocase (Tat): structural and functional insight

Roshani Patel (MSc)

A thesis submitted for the degree of Doctor of Philosophy

School of Life Sciences

June 2014

Table of Contents

Lists of Tables

Lists of Figures

i. Acknowledgements

I am dearly grateful to my supervisor Colin Robinson for giving the opportunity to work with him and helping me grow as a researcher. I would also like to thank my current supervisor Corinne Smith for always having an open door and being extremely supportive.

I am also thankful to my secondary supervisor Dan Mitchell for all his help and advice with the SPR work. I would acknowledge my collaborators Cvetelin Vasilev for his work with AFM and Carmine Monteferrante for providing the AyCy samples.

Although it seems like a long time ago, I am so thankful to the old members of CR group, who were like family and made the labs so enjoyable. In particular, my lab mum Nishi Vasisht, for just about everything and holding everything together. Thank you Jacopo, Cristina, Chao, Anna and Dan.

Sarah Smith and Charlotte Carroll, you guys are amazing friends always pushing me to do my best. Thank you for caring and the uncontrollable laughs.

Finally and most importantly, I am indebted to my family, especially my parents for providing such loving home to comeback to everyday. Suraj, Chandni and Ashish, for your help with various things. It would not have been possible without your love and support.

Thank you

ii. Declaration

The work presented in this thesis is original, and was conducted by the author, unless otherwise stated, under the supervision of Professor Colin Robinson, Dr Corinne J Smith and Dr. Daniel Mitchell. None of the work presented in this thesis has been submitted previously for another degree.

This research was funded by the Medical Research Council (MRC).

The Cryo-EM imaging work was carried out by Ian Hands-portman.

Collaborative work has been carried out contributing towards material presented in this thesis. The extent of the collaborative work is indicated in the relevant chapters (Chapter 4). In summary, collaborative work was carried out with the following:

Cvetelin Vasilev, (Department of Molecular Biology and Biotechnology, University of Sheffield) with the AFM study.

The bacterial membranes from *B. subtilis* were provided by Carmine Monteferrante (Department of Medical Microbiology, University Medical Center Groningen)

iii. Summary

The twin arginine tranlocation (Tat) pathway is responsible for the transport of folded protein across the membrane. In bacteria, this occurs at the cytoplasmic membrane. In Gram-negative bacteria, Tat forms a 3-component machinery named TatABC. The current hypothesised mechanism is compiled from the model organism *Escherichia coli*.

In Gram-positive bacteria the Tat machinery lack the TatB component and so raises the question on the validity of the mechanism assumed from the TatABC system. To date there is limited research available on the TatAC system with in Gram-postive bacteria. Recent characterisation has been focused on the *Bacillus subtilis* system, which contain two TatAC systems; the TatAdCd and TatAyCy.

This thesis focuses on the structural characterisation of the TatAC system of *B. subtilis* and their similarity to the TatABC system in *E. coli.* The TatAdCd complex was studies by electron microscopy (EM) to show structure similarity to the TatBC complex. Mutation within the N-terminus region of the TatAy protein showed functional involved in complex assembly. The TatAyCy was also analysed by EM show conserved round complexes similar to the TatAdCd and TatBC. The similarity in structure may suggest the Gram-negatives and the Gram-Positives share a similar mechanism of transport despite difference in the components.

iv. Abbreviations

AFM	Atomic force microscopy
APH	Amphipathic Helix
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BN-PAGE	Blue Native PAGE
BSA	Bovine serum albumin
C11	Undecyclamine
C-	Carboxy terminus
C	Cytoplasmic fraction
C12E9	Nonaoxyethylene n- dodecyl ether
CMC	Critical Micelle Concentration
CpTat	Chloroplast Tat
Δ Delta	Gene deletion
DDM	*n*-Dodecyl-*β*-D-maltoside
DHFR	Dihydrofolate reductase
DMSO	Dimethyl sulphoxide
E	Elution fraction
ECL	Enhanced chemifluoresence
EDTA	Ethylenediaminetetraacetic acid

EM	Electron microscopy
FT	Flowthrough
g	Grams
g	Gravity
GF	Gel-Filtration
GFP	Green fluorescent protein
GLC	general amine coupling
His-tag	Hexa-histidine tag
h	Hour
HRP	Horseradish peroxidise
kDa	Kilodaltons
LB	Luria Bertani medium
M	Membrane fraction
min	Minutes
nM	Nanometres
OD	Optical density
P	Periplasmic fraction
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PMF	Proton motive force
PVDF	polyvinylidene fluoride membrane

rpm	revolutions per minute
s	seconds
Sec	General secretory pathway
SPR	Surface plasmon resonance
Strep-Tag	Strep II tag
Tat	Twin arginine translocation
TEM	Transmission electron microscopy
TorA	Trimethylamine-*N*-oxide reductase
V_o	Void Volume
v/v	Volume per volume
W	Wash fraction
WT	Wild type
w/v	Weight per volume
YFP	Yellow fluorescent protein

Chapter 1

Introduction

1 Introduction

1.1 Protein transport in bacteria

All cell processes are dependent on the transport of proteins, both by intracellular and extracellular transport. Even the simplest organisms are compartmentalised to produce microenvironments for the cell to carry out different biological processes simultaneously. The compartmentalisation of the cell is achieved through a semi-permeable membrane, which allows regulated transport of proteins and ions.

The survival of a bacterium is highly dependent on the interaction with the surrounding environment. The interaction involves export of proteins and ions across the semi permeable membrane. Bacteria are divided into two groups; Gram-positive and Gram-negative depending on the structure of the cell envelope. A typical Gram-negative bacterium has two membranes, a cytoplasmic membrane, followed by a single layer of peptidoglycan in the periplasmic space and finished with an outer membrane. The outer membrane in Gram-negative bacteria, is composed of lipopolysaccharides (LPS), porins and proteins involved in bacterial pathogenicity, membrane integrity and cell adhesion. The cell wall of the Gram-positive bacteria consists of a multiple layers of peptidoglycan with teichoic acid embedded in the layers. The difference in cell wall structure permits the Gram stain dye to be retained in cell carrying a thick peptidoglycan layer hence are classed as Gram-positive, while Gram-negative bacteria do not stain due to their outer membrane, which prevents the stain from interacting. Figure 1.1 illustrates difference in outer leaflet of Gram-positive and Gram-negative bacteria.

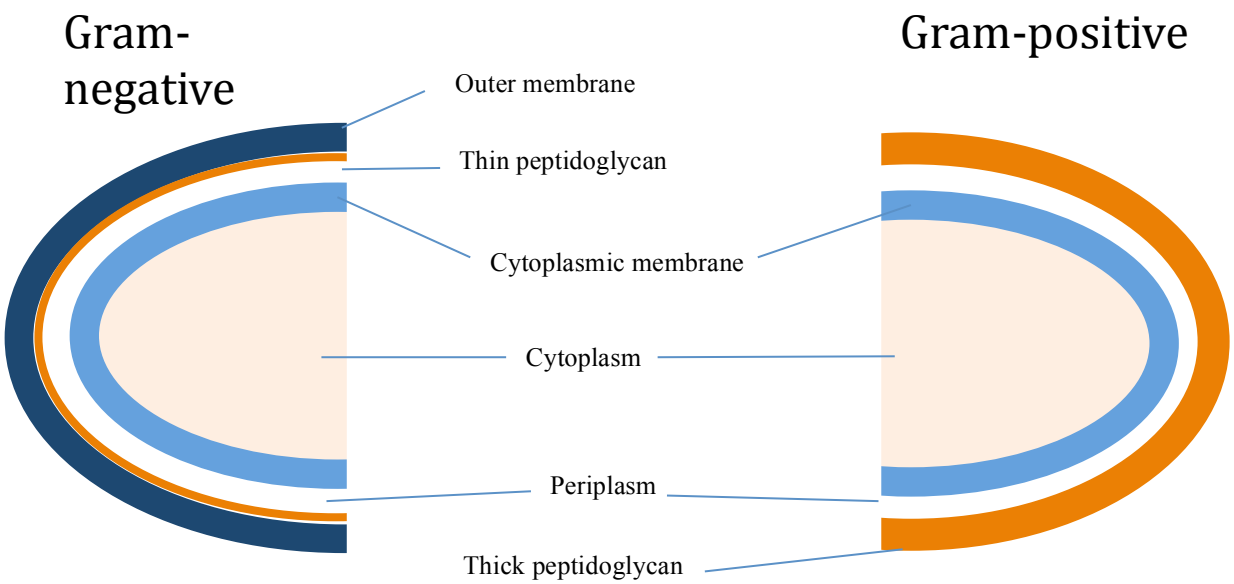


Figure 1.1: The bacterial cell wall of a Gram-negative and Gram-positive bacteria

The Gram-negative cell shows two membranes; the cytoplasmic membrane also called the cytoplasmic membrane and an outer membrane, with lipopolysaccharides on the outer leaflet. The space between the two membranes is known as the periplasmic space and within this space, resides a thin layer of peptidoglycan. In a Gram-positive cell the cytoplasmic membrane is followed by a periplasmic space with a thick layer of peptidoglycan exposed to the external environments.

In bacteria, 20-30% of the proteins that are synthesised are required to be transported out of the cell or incorporated into the membrane (Tjalsma *et al.*, 2000; Holland, 2010). Therefore the cell has to have appropriate machinery to carry out this function. The majority of the proteins that are destined for transport are transported by a well characterised secretory pathway (Sec pathway) (Pugsley, 1993). However other proteins that require more sophisticated machinery use the Twin Arginine Translocation pathway (Tat pathway) (Cline, 1992; Klösgen *et al.*, 1992). The Sec and Tat machineries are the two main protein-transporting machineries in the cytoplasmic membrane of prokaryotes. In Gram-negative bacteria, both systems are capable of transporting proteins either into the membrane or across the cytoplasmic membrane.

1.2 The Sec pathway

The sec pathway is the most common route of protein transport across the membrane or inserting the proteins into the bilayer. The Sec pathway is present in both prokaryotes and eukaryotes. In the latter, it is localised in the cytoplasmic membrane of yeast, bacterial and mammalian cells, the ER membrane and the thylakoid membrane of chloroplasts. The Tat pathway works alongside the Sec pathway and so understanding the Sec pathway is crucial to understand the partnership and similarity between the two pathways.

The Sec translocase is composed of a heterotrimeric complex; made up of SecY, SecE and SecG (Pugsley *et al.*, 2004). The SecY protein is the largest component (48kDa) of the translocase with 10 TMH that arranges into a concave structure (van den Berg *et al.*, 2004; Akiyama & Ito, 1987). The structure of SecY is separated into two parts; TMH1-5 makes up one half and TMH6-10 makes the other half of the structure (Kuhn *et al.*, 2011). It is the cytoplasmic loop of SecY that forms direct interactions with the ribosome, SecA and the signal recognition particle (SRP) (Kuhn *et al.*, 2011). The second component SecE is 14kDa large and is located on back of the SecY structure, which clamps the two halves together. There have been some

implication of the SecE protein in the opening and closing the channel however, other have found the presence of SecE has no influence on this occurrence. SecG is a 12kDa protein with 2 TM helices and associated on the N-terminus of SecY (Satoh *et al.*, 2003). The function of SecG remains to be eluded however its direct relationship to SecA-mediated transport suggests it may be functional in a more specific targeting (Koch & Muller, 2000; Zimmer *et al.*, 2008). The SecYEG complex is fully functional with a single copy of each subunit however larger complexes have also been observed through crosslinking and BN-PAGE gels (Boy & Koch, 2009; Deville *et al.*, 2011; Veenendaal *et al.*, 2001) in E. coli the SecYEG translocase assembles into a dimer and forms a structure of 20-25Å on the periphery and the centre of the channel constricts to ~4Å; giving a general hourglass structure. It has been shown that the pore of the translocase is constricted in the centre and likely to open laterally (Higy *et al.*, 2005).

The post-translational targeting mechanism by the Sec pathway also involves other cytoplasmic components; SecA and SecB. There are multiple routes by which a protein is targeted to SecYEG. One of these is the SecA/B mechanism where the SecB binds to the preprotein and targets it to the membrane bound SecA. The preprotein is then transferred to to SecYEG via SecA where it is translocated (Fekkes *et al.*, 1998). The SecB protein is a chaperone protein that forms a tetrameric complex. The SecB was thought to function in the targeting the preprotein to the membrane (Hartl *et al.*, 1990); however some proteins are shown to be targeted independently of SecB. In such instances, the cytosolic SecA can co-translationally binds to a translating ribosome to form a ribosome-associated nascent protein complex (RNC) and targets it to the SecYEG translocase directly (Huber *et al.*, 2011). SecA is a motor protein; driving the translocation of the preprotein upon ATP hydrolysis. The conformation change as a result of ATP hydrolysis is what drives the translocation.

Chaperone trigger factors are thought to be the first to bind the emerging preprotein through the ribosome channel; consequently allowing other chaperone such as GroEl and DnaK to bind (Eisner *et al.*, 2003; Lakshmipathy *et al.*, 2007). The principle role of these chaperones is to maintain the polypeptide in an unfolded conformation.

In cases where preproteins are co-translationally targeted to SecYEG, the translation process remains on hold and bound by SRP, which target the complex to the FtsY protein; associated to the SecYEG translocon. The SRP is a ribonucleo-protein complex; commonly used for membrane protein insertion (Kuhn *et al.*, 2011). The SRP specifically recognises a largely hydrophobic signal anchor (SA) sequence lacking a peptidase cleavage site (Luirink *et al.*, 2005). The SA acts as a mode of anchoring the membrane protein in the lipid bilayer. The SRP is thought to scan all ribosomes for an emerging SA sequence and also has a higher affinity to displace SecA that may be pre-bound to the ribosome (Holtkamp *et al.*, 2012; Huber *et al.*, 2011).

In general once the preprotein is in contact with the SecYEG translocase, the protein is then threaded through the channel either post-translationally or co-translationally. This threading mechanism is dependent on ATP hydrolysis; a single ATP powers the translocation in small increments (30-40 residues); hence a repeated cycle of ATP hydrolysis is required (Schiebel *et al.*, 1991; Tomkiewicz *et al.*, 2006). In addition, a proton motive force is also required for unidirectional insertion of the signal peptide inside the channel (Tani *et al.*, 1989). There have been different model proposed for the SecA driven translocation including; examples include power stroke model, Brownian ratchet model and the piston model (Whitehouse *et al.*, 2012; Mitra & Frank, 2006; Tomkiewicz *et al.*, 2006). Despite much research on the Sec pathway many mechanistic understanding still remain unclear and more specifically the cross talk between Sec and other pathways; more significant the Tat pathway.

1.3 The Tat secretory pathway

The Twin Arginine Translocation pathway was discovered after the Sec pathway in the early 1990s. Earlier the pathway was named the "Δ-pH pathway", on the basis that the protein's transport was not dependent on ATP hydrolysis as with the Sec system, but rather used a proton motive force (PMF) (Cline, 1992; Mould & Robinson, 1991). Later, the pathway was shown to transport proteins in a folded state (Creighton *et al.*, 1995). The analysis of the prefolding protein in bacterial cells showed a conserved "twin Arginine" motif in the signal peptide in both bacteria and chloroplast, hence leading to the generic term Twin Arginine Translocation (Tat)

pathway (Sargent *et al.*, 1998; Berks, 1996). The Tat pathway in chloroplast is known as the cpTat pathway and is composed of 3 minimal components; Tha4, Hcf106 and cpTatC (Settles *et al.*, 1997; Cline & Mori, 2001).

The characterisation of the Tat pathway in bacteria, found homologs of cpTat components Tha4, Hcf106 and cpTatC, which were named TatA, TatB and TatC respectively (Sargent *et al.*, 1998; Sargent *et al.*, 1999). Earlier Tat studies focused on the model Gram negative bacterium *E. coli,* however the characterisation of the Tat pathway in Gram positive bacteria showed only a 2 component system is in place. *Streptomyces* are an exception to this finding which also has a 3 component system (Yen *et al.*, 2002). The Tat pathway in Gram positive and Gram-negative bacteria differs and so the subsection below describes the pathway individually.

The utilisation of the Tat pathway within bacteria is varied from 6% of secreted proteins in *E. coli* to >90% in *Haloferax volcanii* (*H. volcanii*) (Dilks *et al.*, 2003; Rose *et al.*, 2002). In cases where the Tat pathway is not heavily used, the deletion of this pathway is not detrimental to the growth of the cell, such as in *E. coli* and *B. subtilis* (Dilks *et al.*, 2005). However, with *H. volcanii*, deletion of the Tat machinery is detrimental to the cell (Dilks *et al.*, 2005).

The current research on the Tat pathway has been heavily focused on understanding the functioning of the pathway, particularly through structural characterisation. Other implications of the pathway, such as in pathogenicity and its application in biotechnology have developed a further need to understand the pathway more closely.

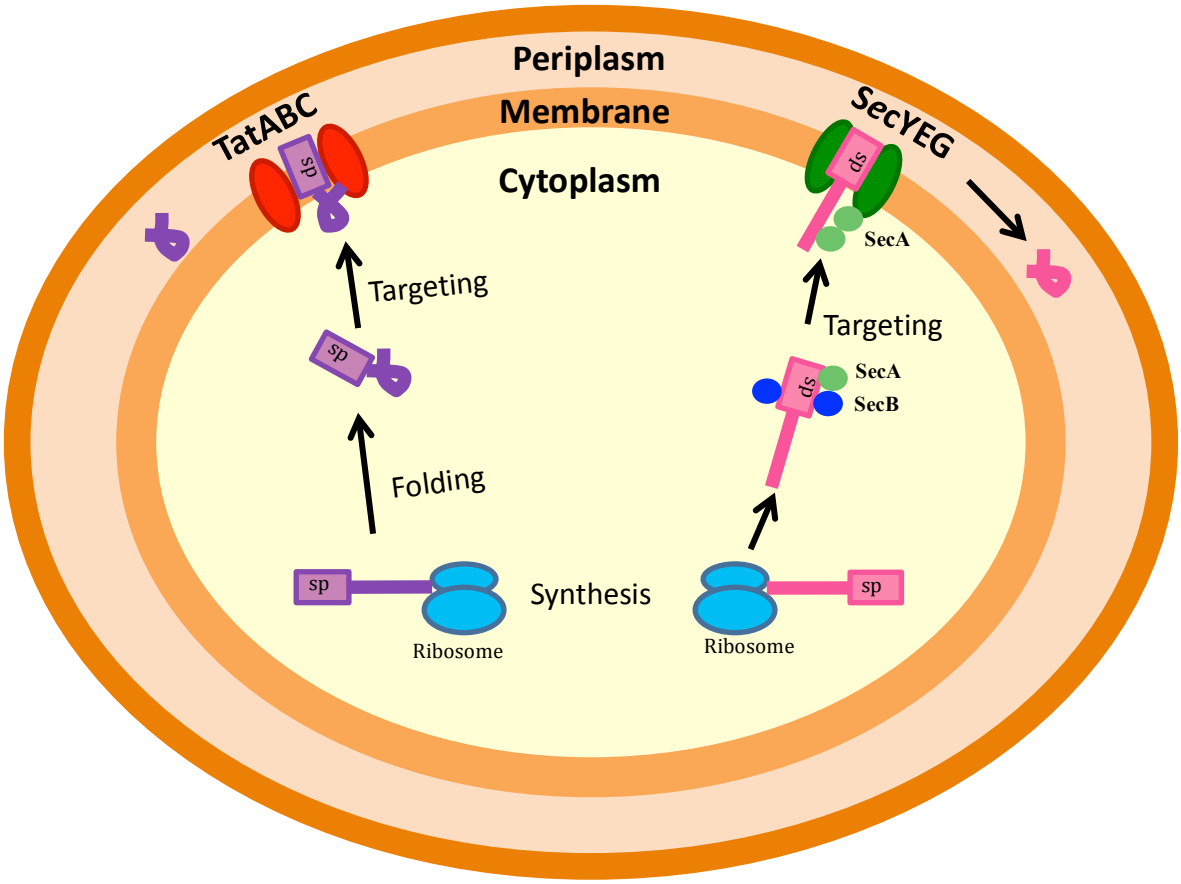


Figure 1.2: The two secretory pathways in the cytoplasmic membrane of an *E. coli*

The *E. coli* cell has two pathways for cytoplasmic protein transport, one dedicated to unfolded protein (Sec pathway) and one for folded proteins (Tat pathway). All proteins have a signal peptide coding the target sequence to a specific pathway (represented by the rectangle on the N- terminus of the protein, when it is known as a preprotein. In the Sec pathway, the preprotein is synthesised and bound by chaperone to prevent folding. The chaperone bound preprotein is targeted to the SecYEG machinery and transported through in a threading like manner. In contrast, the TatABC machinery transports preproteins that fold within the cytoplasm prior to transport. The signal peptide is cleaved off in the periplasm to produce a mature protein.

1.4 Tat components in Gram negative bacteria

In Gram-negative bacteria such as in *E. coli*, the Tat system is encoded by four *tat* genes *tatABCD* in a single operon, with a monocistronic *tatE* gene found elsewhere in the genome. In *E. coli,* the Tat machinery is minimally composed of the TatA, TatB and TatC genes which are translated in to membrane proteins (Wexler *et al.*, 2000). The *tatD* gene encodes a protein for DNase activity, and was later found to be involved in the proof reading of the Tat substrates (Wexler *et al.*, 2000; DeLisa *et al.*, 2003; Matos *et al.*, 2009). Earlier expression studies have reported that TatA:TatB:TatC:TatD is expressed in a relative ratio of 25:1:0.5:0.1 (Jack *et al.*, 2001), where the TatA is produced in the highest abundance. The *tatE* gene is thought to have resulted from a gene duplication of the *tatA* gene. Complementation assays have shown that replacing TatA with TatE can maintain the functionality of the Tat translocase activity (Sargent *et al.*, 2002). Although *tatE* shares 50% sequence homology with *tatA*, it is expressed in lower abundance (Sargent *et al.*, 2002).

1.5 Tat components in Gram positive bacteria

In the Gram-positive bacteria, the Tat system is composed of 2 components machinery, TatA and TatC. The sequence homology shared between TatA and TatB allows the function of TatB to be carried out by TatA, making the TatA in Gram-positive bacteria bifunctional (Jongbloed *et al.*, 2002; Barnett *et al.*, 2008). More specifically, the *B. subtilis* bacterium has two copies of the TatAC machinery; termed TatAdCd and TatAyCy (Jongbloed *et al.*, 2000) (Jongbloed *et al.*, 2002). The two homologous systems are differentiated by their substrate specificity; the TatAdCd system is specific for the substrate PhoD. The *phoD* gene is upstream of the *tatAy* and *tatCy* genes and expressed within a single operon (Jongbloed *et al.*, 2004; Pop *et al.*, 2002). The *tatAy* and *tatCy* genes are expressed on a separate operon with its specific substrate *ywbN* expressed elsewhere is the genome. An additional *tatAc* gene is expressed elsewhere in the genome; with no functional role identified (Jongbloed *et al.*, 2004).

1.6 Tat Signal Peptide

Proteins that are targeted to the Sec and Tat pathways have a short amino acid sequence on the N-terminus, which is cleaved after transport known as a signal peptide. This signal peptide is generally entails a tripartite structure with a basic N-region, a hydrophobic centre and a polar C-region (Figure 1.3). The C-region of the signal peptide has a cleavage site, which is recognised by either type I or type II signal peptidase. Type I signal peptidases are responsible for cleavage of the precursor protein with an AXA motif preceding the cleavage site (Tuteja, 2005). While the type II signal peptidase is responsible for cleavage of lipoproteins and recognises the conserved motif L(A/S/T)(G/A) (Tjalsma *et al.*, 2002). The C-terminus cleavage recognition site is similar for both Tat and Sec substrates, using Type I signal peptidase. The substrate recognition sites in located in the N-region are different to allow specific targeting to either the Sec or Tat machineries. The Sec substrates have a generic positively charged motif in the N-region that is recognised by the Sec machinery while the Tat substrates have a more specific conserved motif SRRXFLK, as shown in Figure 1.3 (x is any polar residue) (Cristóbal *et al.*, 1999; Stanley *et al.*, 2000; Berks, 1996). The twin arginine residues are essential for Tat pathway targeting and hence gives the pathway its name (Twin Arginine Translocation). In plants, a single arginine mutation completely abolishes the transport of the protein, while in bacteria a single arginine substitution is not detrimental, however, when both arginine residues are substituted, the translocation is blocked (Chaddock *et al.*, 1995; Buchanan *et al.*, 2002; DeLisa *et al.*, 2002; Mendel *et al.*, 2008). In a few rare cases in bacteria the RR residues are not conserved and a single arginine residue is sufficient. The substrate TtrB (tetrathionate reductase) has only a single arginine residue in *Salmonella enterica* and penicillin G amidase (PGA) in *E. coli* has the motif RNR with a residue dividing the twin arginine residues (Hinsley *et al.*, 2001; Ignatova *et al.*, 2002).

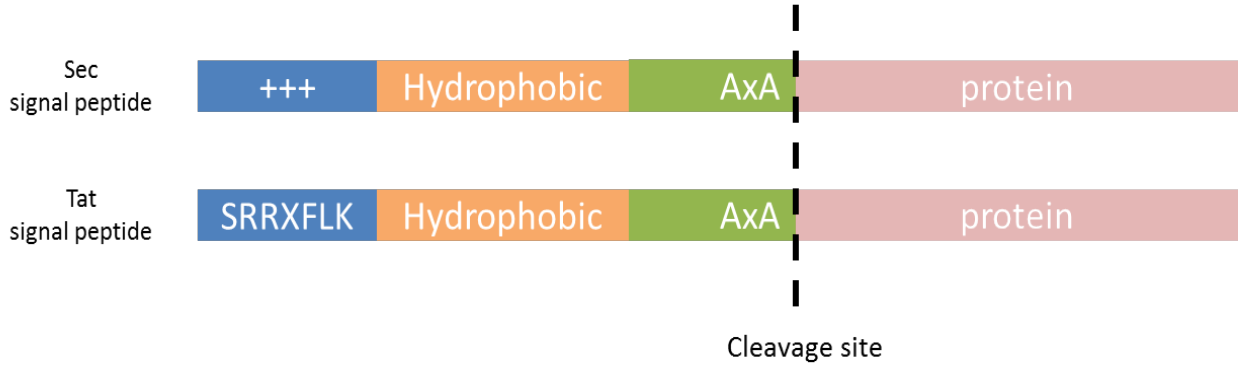


Figure 1.3: The architecture of a signal peptide for protein transport.

The signal peptides for both the Tat and Sec targeted proteins have a similar structure characteristic with a charged N-region, a hydrophobic centre and a polar C-region. Within the C- region is the AxA signal peptidase recognition site. The Sec targeted signal peptides have residues in the N-terminus region that are not specific and positively charged. The Tat signal peptide is more specific and contains a conserved motif SRRXFLK where x is any polar residue.

Although the RR is the significant motif in the Tat signal peptide, the adjacent residues to the RR sequence have also been conserved, implicating the significance of the whole motif rather than just the RR residues. As mention above, the SRRXFLK motif (X is any polar residue) as a whole is responsible for the Tat pathway targeting. In *E. coli,* the mutation of the serine residue resulted in reduced translocation activity of two substates; TorA and DmsA (Mendel *et al.*, 2008). However the same mutation in SufI does not influence the translocation activity (Stanley *et al.*, 2000). In addition, when the leucine of the SRRXFLK motif is substituted by aspartate, the translocation of TorA-GFP is blocked but when substituted for either Serine or Alanine, the translocation activity remains functional (Mendel *et al.*, 2008). In general, the Tat system in plants is more rigorous in signal peptide recognition compared to that in the well characterised *E. coli*, where a single substitution of the arginine residue does not completely block the translocation activity (Stanley *et al.*, 2000).

In addition to the SRRXFLK motif, various other regions are shown to be involved in substrate recognition process. For example the basic residues in the C-region have been named the "Sec avoidence motif" to prevent mis-targetting of the substrate in chloroplasts (Bogsch *et al.*, 1997). Another example includes the proline residue at the end of the H-region of the signal peptide and the overall reduced hydrophobicity in the Tat signal peptide compared to the Sec signal peptide (Ize *et al.*, 2002; Cristóbal *et al.*, 1999).

The signal peptide is also responsible for chaperone binding during and post translation. For example the model *E. coli* Tat substrate TorA contains lucine-rich residues L27, L31, L32 involved in recognising the chaperone TorD (Genest *et al.*, 2006; Buchanan *et al.*, 2008). The N-region of the Tat signal peptide is thought to have a regulatory function by controlling the transport activity of the substrate (Bowman *et al.*, 2013).

1.7 Tat substrates

Proteins exported by the Tat machinery are predicted from the signal peptide of the proteins. In *E. coli*, there are 29 putative Tat substrates, which represent 6% of secreted proteins (Dilks *et al.*, 2005). Many of these proteins require insertion of co-factor such as dimethyl sulfoxide reductase DmsA, which contains the organic cofactor molybdopterin (Palmer *et al.*, 2005). There are other proteins carrying inorganic cofactors such as Fe+ ions, Fe-S clusters, Cu^{2+} ions (Palmer *et al.*, 2005). A list of the substrates that require cofactor insertion is shown in Table 1.1. Most of the cofactor-containing proteins are enzymes involved in anaerobic respiration. Therefore the Tat machinery is significant for bacterial survival and metabolism. The requirement of these proteins to fold in the cytoplasm to allow cofactor insertion is what makes these proteins dependent on the Tat pathway.

There are also other substrates that do not require a cofactor but are still targeted to the Tat machinery such as AmiA, AmiC, FhuD, and SufI but are thought have rapid folding kinetics and therefore are folded within the cytoplasm, making them incompatible with the Sec pathway (Palmer *et al.*, 2005; Ize *et al.*, 2003). However it can also be argued that the SRP pathway is capable of co-translationally transporting rapid folding proteins DsbA (Shimohata *et al.*, 2005). Another possibility is that the external environment of the bacteria may not be ideal for protein folding, as disulphide bond may not be able to form. An example of this is in *Halophilic archea,* where the ionic environment would lead to aggregation of unfolded proteins (Rose *et al.*, 2002).

Another mode of transport by the Tat machinery was termed “hitch-hiking” where proteins without the Tat specific signal peptide interacts with another protein that does carry the signal peptide and it can be transported together (Rodrigue *et al.*, 1999). The Tat mechanism is somewhat tolerant as it allows some proteins to “piggy back” across the membrane. Which is a surprising quality as it has a stringent mechanism in place which can identify misfolded of Tat substrates (DeLisa *et al.*, 2003). Example of such proteins include DmsA/DmsB, YnfF/YnfH, HyaA/HyaB are among the 9 proteins predicted to piggy-back onto a Tat substrate (Rodrigue *et al.*, 1999; Palmer *et al.*, 2005).

Table 1.1: List of Tat substrates in *E. coli* that require cofactor insertion

Adapted from (Palmer *et al.*, 2005) showing predicted Tat substrates in *E. coli* K12 with the role of the protein and its associated cofactors. The molybdopterin guanine dinucleotide (MGD), or molydopterin (MPT) are organic cofactors. The remaining substrates have an inorganic cofactor as either ions or Fe-S clusters.

Substrate	**Function**	**Cofactors**
HyaA	Hydrogen oxidation	3 x Fe-S clusters
HybO	Hydrogen oxidation	3 x Fe-S clusters
HybA	Hydrogen oxidation	4 x Fe-S clusters
NapG	Nitrate reduction	4 x Fe-S clusters
NrfC	Nitrite reduction	4 x Fe-S clusters
YagT	unknown	2 x Fe-S clusters
YdhX	unknown	4 x Fe-S clusters
TorA	TMAO reduction	MGD
TorZ	TMAO reduction	MGD
NapA	Nitrate reduction	MGD, 1 x Fe-S cluster
DmsA	DMSO reduction	MGD, 1 x Fe-S cluster
YnfE	DMSO reduction	MGD, 1 x Fe-S cluster
YnfF	DMSO reduction	MGD, 1 x Fe-S cluster
FdnG	Formate oxidation	MGD, 1 x Fe-S cluster
FdoG	Formate oxidation	MGD, 1 x Fe-S cluster
YedY	TMAO/DMSO reduction	MPT
CueO	Copper homeostasis	4 x Cu ions
YahJ	unknown	1 x Fe ion

1.8 Topology of TatABC proteins

TatA is an 89 amino acids membrane protein with the N' terminus protruding on the periplasmic side while the C' terminus extends to the cytoplasmic side. This protein forms a hydrophobic helix followed by a smaller amphipathic helix (Figure 1.4).

Until recently the topology of the TatA unit was thought to be a helical structure spanning through the membrane with a amphipathic helix aligning to the cytoplasmic face of the membrane (Figure 1.4) (Lange *et al.*, 2007; White *et al.*, 2010). However Gohlke *et al.* (2005) explored the possibility of the amphipathic helix switching between membrane resting and membrane spanning topologies (Gohlke *et al.*, 2005; Gouffi *et al.*, 2004; Chan *et al.*, 2007). In some unusual cases, soluble TatA has also been identified in bacteria and chloroplasts (Pop *et al.*, 2002; De Keersmaeker *et al.*, 2005; Westermann *et al.*, 2006; Berthelmann *et al.*, 2008b; Barnett *et al.*, 2008)

This observation of cytoplasmic TatA has spurred some debate on the cytoplasmic role of TatA in substrate targeting, hence its similarity to the Sec pathway.

The TatB protein (171 amino acids) is larger than TatA, with a single hydrophobic helix and two amphipathic helices on the cytoplasmic face. The hydrophobic helices in TatA and TatB seem to form the membrane-spanning domain and share 20% sequence homology while the amphipathic helices adhere to the periphery of the cytoplasmic membrane (Hicks *et al.*, 2003). Although there is homology between TatA and TatB in terms of sequence and topology, the role of TatB is not completely understood (Figure 1.4). There are some inferences of its involvement in stability of the complex (Sargent *et al.*, 1999).

TatC is the largest and most conserved subunit of the TatABC translocon, with 6 transmembrane helices. The total length of TatC is 259 amino acids and both the N- and C-termini are facing the cytoplasmic side as shown in Figure 1.4 (Behrendt *et al.*, 2004). The functional region of TatC is embedded in the cytoplasmic loop regions, where the signal peptide of the substrate is thought to bind (Buchanan *et al.*, 2002; Allen *et al.*, 2002). Recent crystallisation studies of TatC from *Aquifex aeolicus* have shown the transmembrane helices to be tilted up to 20° within the membrane (Ramasamy *et al.*, 2013; Rollauer *et al.*, 2012). The transmembrane helices (TMH) 5

and 6 are relatively shorter and do not completely span the membrane (Figure 1.4). The overall topology of TMHs forms a concave structure (baseball glove like structure) to give a central cavity (Ramasamy *et al.*, 2013; Rollauer *et al.*, 2012). This recent structural characterisation has developed the understanding of the Tat mechanism that will be later described.

1.9 The topology of the Tat components in B. subtilis

The *B. subtilis* genome codes for 3 types of TatA (TatAc, TatAy and TatAd) and two types of TatC (TatCy and TatCd) (Jongbloed *et al.*, 2004). There are two homologous systems working simultaneously, the TatAdCd and the TatAyCy system but the role of TatAc is not yet known, and it is also not capable of compensating for the function of either TatAd or TatAy in *B. subtilis*. Nevertheless, when expressed in *E. coli* the TatAc could form complexes with TatAc and TatAd to form functional complexes (Monteferrante *et al.*, 2012). The predicted topology for TatAd and TatAy suggests they both form 1 TMH and 1 APH as with TatA but the APH region is shorter compared to TatA (Barnett *et al.*, 2011). Structural characterisation by NMR studies confirms the predicted topology of TatAd (Walther *et al.*, 2010). The TatCd and TatCy proteins are also predicted to form 6 transmembrane helices in a similar fashion to TatC. Figure 1.4 shows the Tat proteins of the two working systems in *B. subtilis*.

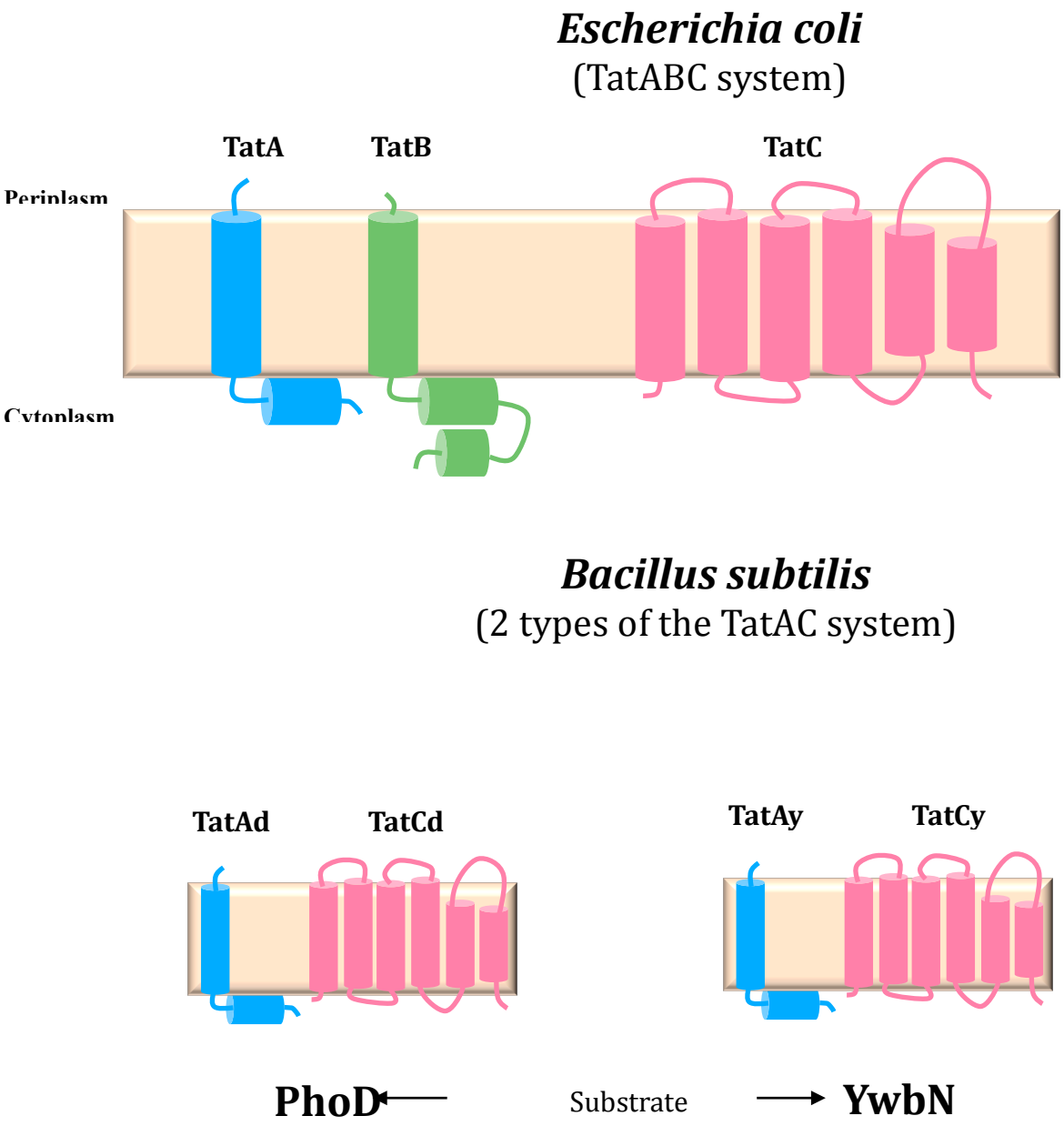


Figure 1.4: The topology of Tat components in *E. coli* and *B. subtilis*.

TatABC proteins of *E. coli* are shown above with the topology of the individual membrane protein. TatA and TatB proteins contain one transmembrane helix (TMH) and one amphipathic helix (APH). The TatB protein is slightly larger with an unstructured region towards the C-terminus. The TatC protein forms 6 TMH with both the N- and C-termini in the cytoplasmic side. The figure below shows the topology of the Tat proteins in *B. subtilis*. Where the TatAd and Ay both have 1 TMH and 1 APH as with TatA and TatB. Similarly TatCd and TatCy has the same topology at TatC. The difference in the TatAC systems in *B. subtilis* is the specific targeting of the PhoD to TatAdCd and of YwbN to TatAyCy system.

1.10 The Tat complexes

Initially, TatABC was thought to form a single functional translocon like its cousin the Sec translocon, which also is composed of 3 subunits. As further studies developed, it soon became apparent that TatABC functioned in two halves, with a separate pore-forming TatA complex and its counterpart TatBC complex; the substrate recognising unit. The two complexes only come together to form an active translocon upon substrate binding and in presence of PMF. There have been sighting of TatAB complexes, however these are thought to result from a mid-stage translocation event such as assembly or disassembly of the TatABC complex (Sargent *et al.*, 2001; Oates *et al.*, 2005). Earlier studies in chloroplasts suggest a clear distinction between the TatBC complex and the TatA complexes. In *E. coli* TatA is also shown to associate to the TatBC complex to form a Tat(A)BC with varying amount of TatA (Oates *et al.*, 2005; Cline & Mori, 2001; de Leeuw *et al.*, 2002).

1.1.1 The function of the TatA complex

TatA is found at a higher abundance than other Tat proteins, more specifically in 20-fold excess of TatB (Jack *et al.*, 2001; Sargent *et al.*, 2001). TatA forms a homo-oligomeric complex as well as a Tat(A)BC hetero-oligomeric complex in *E. coli*. This high abundance of TatA and its ability to form complexes ranging size from 50-600kDa prompted the hypothesis, that TatA is responsible for forming the pore of the machinery and the variable range in the complex sizes was accommodate different size substrates (Jack *et al.*, 2001; Sargent *et al.*, 2001; Oates *et al.*, 2005). An individual unit of TatA has an average molecular mass of 17kDa, suggesting the complex is minimally composed of 3 subunits. This trimeric complex was also observed from cross-linking experiment with TatA over-expression (De Leeuw *et al.*, 2001). Blue-native gel has shown TatA to form complexes ranging from 50-600kDa with ~40kDa increments (Oates *et al.*, 2005). This increasing increment is similar to the trimeric complexes, with an estimated weight of 51kDa, implying these complexes associate together to form the larger complexes (Oates *et al.*, 2005; Gohlke *et al.*, 2005). The TatA complex has been allocated the role as the pore-

forming complex, however recent mechanistic views suggest that a pore is not formed but rather TatA is used in destabilisation of the membrane (Brüser & Sanders, 2003; Walther *et al.*, 2010). Therefore the role of the TatA in pore formation is debatable.

1.1.2 Structure of the TatA complex

TatA complexes form a ring structure, with a lid covering one side of the complex (Gohlke *et al.*, 2005). The calculated pore size ranges from 30-70Å in diameter, which is large enough to translocate one of the largest Tat substrates, TorA as shown in Figure 1.5 (Sargent *et al.*, 2002). The large size range of the complex suggests a tightly controlled channel must be in place to accommodate all size of substrates. The membrane spanning ring structure has an estimated height of 45-55Å and width of ring around 25-30Å (Figure 1.5) (Tarry *et al.*, 2009). The height of the ring is large enough to span the lipid bilayer while the width of ring has the potential to accommodate two helices. Gohlke *et al.* (2005) suggests two modes of assembly for the ring width: either each hydrophobic helix forms the channel while the amphipathic helices align on the membrane as in

Figure 1.4. On the other hand, every second TatA unit has a hairpin structure where the amphipathic helix folds into the membrane following a dual topology mechanism (Gohlke *et al.*, 2005; Gouffi *et al.*, 2004). From the molecular weight estimation of the complex the latter seems plausible.

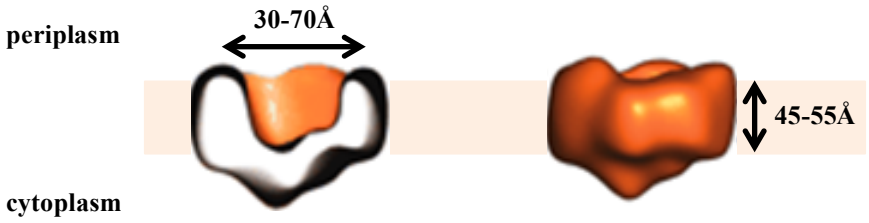


Figure 1.5: 3D model of the TatA complex by single particle reconstruction

Figure adapted from Gohlke *et al.* (2005) showing the side view of the TatA complex with the lid covering the channel on the cytoplasmic side, with cross sectional view on the right of the complex. The height of the complex is approximately 50Åwhich fit in with the plane of the bilayer.

1.1.3 The function of the TatBC complex

While some authors believe that TatA has some role in stabilising the TatBC complex others would suggest that these interactions are not present under native conditions and TatA has no role in the stability of TatBC (Mangels *et al.*, 2005; McDevitt *et al.*, 2006; Orriss *et al.*, 2007). The TatBC complex (Hcf106-cpTatC in chloroplast) was first shown to be the substrate binding complex in chloroplasts but is now accepted as the "substrate binding complex" universally for all Tat systems (Mori & Cline, 2002). TatB and TatC are found in an equimolar ratio of 1:1 in the TatBC complex (Bolhuis *et al.*, 2001; Mori & Cline, 2002). The substrate binding and recognition role of the TatBC complex remains functional when only the TatB and TatC units are overexpressed; TatA is not required during the early stages of substrate binding in both *E. coli* and chloroplasts (Tarry *et al.*, 2009; Alami *et al.*, 2003; Mori & Cline, 2002). Size estimation of the TatBC complex from Blue-Native protein electrophoresis (BN-gel) was estimated at 500kDa (Tarry *et al.*, 2009). Earlier size estimation from BN-gels suggested the TatBC-His complex is 440kDa with a fainter band of 580kDa, resulting from in vivo substrate binding (Richter & Brüser, 2005). However later evaluation of the size estimation showed the His-tag

was influencing the mobility of the protein on a BN-gel and when removed, the complex was formed at 500kDa (Tarry *et al.*, 2009). There appears to be a general agreement that the substrate binds initially to TatBC (Hcf106-cpTatC in thylakoids), and this binding is independent of other Tat components. Once the substrate is bound to the TatBC complex, the TatA complex associates with the TatBC complex in the presence of a ΔpH (Cline & Mori, 2001).

1.1.4 Structure of the TatBC complex

The $TatBC_{His}$ complex runs on a BN gel at 443kDa producing a single band, however from the electron microscopy (EM) data, two distinct groups of $TatBC_{His}$ were identified. Both the structures produced a hemispherical structure, however the density of the complexes differed with 420kDa and 330kDa complexes. The cross-section view of these $TatBC_{His}$ complex shows a central cavity that is too small to accommodate the substrate itself, however there have been inferences of cavity forming a potential site for signal peptide insertion (Maurer *et al.*, 2010; Fröbel *et al.*, 2012).

Figure 1.6: 3D model of the $TatBC_{His}$ structure generated from single particle reconstruction

The structure shows the complex ranging in size of 9-11nm. The structure has central cavity, which increases in the larger $TatBC_{his}$ complexes. The smaller TatBC complex forms a hemispherical structure with height in the range of 4.5nm, which sit in the plane of the bilayer (grey rectangle). The figure have been edited from Tarry *et al.* (2009)

1.11 The Tat mechanism

The overall mechanism can be divided into 3 major steps, substrate binding, active translocon assembly and translocation. All three steps will be described in detail in the following subsections and summarised in Figure 1.7.

1.1.5 Substrate binding

The Tat pathway is responsible for the transport of folded pre-protein across the cytoplasmic membrane in both bacteria and chloroplasts. The Tat substrate contains a signal peptide on the N-terminus, with a conserved SRRXFLK motif in majority of these substrates. The specific targeting of the substrate is thought to occur via the twin arginine residues. The signal peptide is synthesised first and possibly bound by a chaperone to maintain the signal peptide in an unfolded conformation (Brüser & Sanders, 2003).

The TatC is thought to recognise the twin arginine signal peptide and primarily bind the signal peptide (Frobel *et al.*, 2012). The lack of TatB in *E. coli* prevents the transport of endogenous substrates suggesting it has some functional role in substrate binding (Blaudeck *et al.*, 2005; Ize *et al.*, 2002; Fröbel *et al.*, 2012). Both TatB and TatC proteins are thought to make contacts with the signal peptide, more specifically the H-region of the signal peptide crosslinks to the TatB (Alami *et al.*, 2003; Panahandeh *et al.*, 2008; Gerard & Cline, 2006). The TatC component is thought to interact specifically with the RR residues of the signal peptide (Alami *et al.*, 2003; Gerard & Cline, 2006). It has also been shown the interaction between TatBC and the signal peptide does not require a PMF of any kind (Alami *et al.*, 2003; Whitaker *et al.*, 2012). Even the labouring process of signal peptide insertion into the bilayer with a hair-pin fold of the signal peptide does not seem to require a PMF.

Recent analysis of signal peptide binding suggests the signal peptide is inserted into the lipid bilayer via a groove formed by the TatC structure. From the cross linking studies, the N-terminus of the signal peptide has been shown to protrude on the cytoplasmic side forming a “hairpin” like structure within the TatBC complex (Frobel *et al.*, 2012). This type of insertion also leads to prior exposure of signal

peptide cleavage site to the periplasmic side (Rollauer *et al.*, 2012). The actual binding event is not limited by an energy-producing step and is an autonomous step in the pathway in thylakoids (Alcock *et al.*, 2013; Fröbel *et al.*, 2011).

1.1.6 Active translocon assembly

Upon completion of the substrate binding/recognition via the signal peptide, the complexes must assemble into an active translocon. This second step is highly dependent on the PMF to drive rearrangement of the components (Mould & Robinson, 1991). There are two components to the PMF; the pH gradient (Δ-pH) and electrical gradient (ΔΨ), which can be exchangeable when required (Bageshwar & Musser, 2007). The TatA is recruited by the TatBC complex in the presence of PMF, where the TatA forms interactions directly with the TatB and the periplasmic loops of TatC (Barrett & Robinson, 2005; Ramasamy *et al.*, 2013). The requirement of the PMF is also seen after the early assembly stages where the signal peptide makes contact with TatA (Froebel *et al.*, 2011; Alami *et al.*, 2003; Pal *et al.*, 2013). Live cell imaging shows TatA-YFP is only recruited in the presence of the TatBC-substrate complex and PMF. TatA only dissociates from the TatBC complex after translocation, when the PMF is absent (Alcock *et al.*, 2013). This interaction between TatA and the signal peptide is shown to be PMF dependent, where the signal peptide is passed on from the TatBC complex to the TatA complex (Bageshwar & Musser, 2007). The first PMF requirement is during the recruitment of the TatA complex forming a functional TatABC complex in *E. coli* (Dabney-Smith *et al.*, 2006; Leake *et al.*, 2008; Mori & Cline, 2002). The general mechanism of the Tat pathway is shown in Figure 1.7.

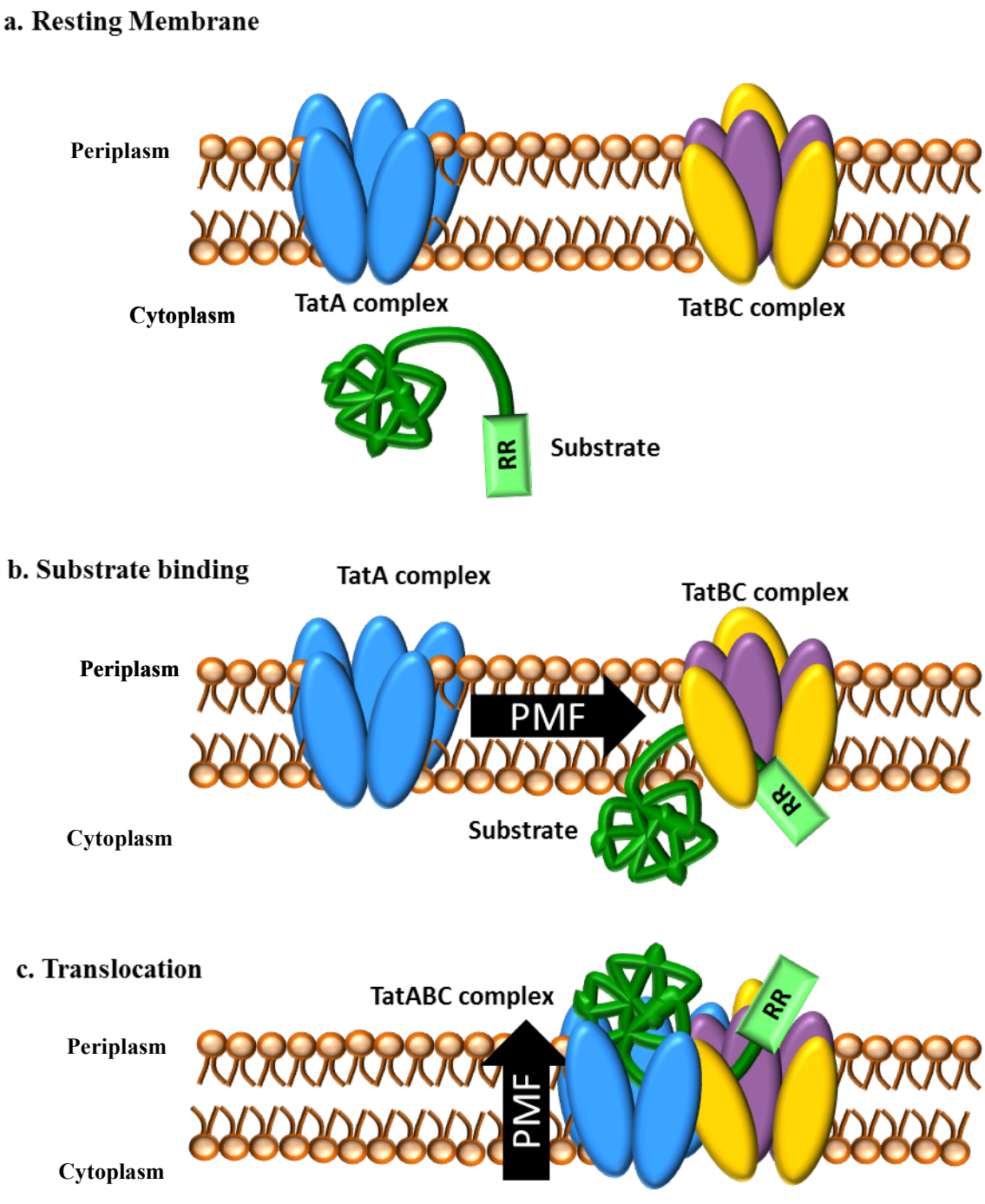


Figure 1.7: The hypothesised mechanism for the TatABC machinery

The mechanism of the Tat pathway begins with a resting membrane with two complexes; TatA complex and the TatBC complex (a). Once the substrate is folded in the cytoplasm, its RR motif targets the substrate to the TatBC recognition complex (b). The substrate binding triggers the recruitment of the TatA pore complex, driven by PMF. The substrate is then translocated using the PMF across the membrane (c).

1.1.7 Translocation

The final event is the translocation of the substrate, which is dependent on a second round of PMF. Recent structural characterisation has led to two plausible modes of Tat translocation (Fröbel *et al.*, 2012; Maurer *et al.*, 2010). The first to be proposed was the pore model, where TatA is thought assemble into a structured pore through which the protein is translocated. The second favours the new emerging idea of membrane destabilisation and thereby increasing the permeability of the membrane. Both model are dependent on the homo-oligomerisation of TatA forming variable sizes of TatA complexes (Gohlke *et al.*, 2005; Oates *et al.*, 2005). However the functional analysis of the oligomerisation has led two opposing theories of pore formation and membrane destabilisation, which will be described in detail.

a. Pore-forming model

The pore-forming model was first proposed from evidence of the ability of TatA to oligomerise into variable sized complexes. The later structural characterisation and topology studies proposed that the role of the amphipathic helix could be involved as a trapdoor action, where it gates the pore initially and then flips into the membrane to allow the substrate to pass through (Gouffi *et al.*, 2004; Oates *et al.*, 2005). The pore-forming model is illustrated in Figure 1.8 a. In *E. coli*, TatA is a highly abundant protein relative to TatB and TatC and its ability to form complexes of variable size suggested a possible role as translocation pore (De Leeuw *et al.*, 2001; Oates *et al.*, 2005; Gohlke *et al.*, 2005; Beck *et al.*, 2013).

In addition, Gohlke *et al.* (2005) proposed the ring of the pore was large enough to accommodate the TMH of TatA and was later supported by further NMR studies (White *et al.*, 2010). The pore of the TatA was clearly visible with the range between 20-70Å to accommodate all the Tat substrates (Berks *et al.*, 2000; Gohlke *et al.*, 2005). The TatAd complexes again showed similar pore like structure however in

this case the pore is not large enough to accommodate the substrates (Beck *et al.*, 2013). Other theoretical work suggest the APH of TatA forms a "hair pin" fold at the C- terminus which then flips into the membrane forming a gate to the pore, which agree with structure of a single-sided lid on the cytoplasmic side (Walther *et al.*, 2013; Gohlke *et al.*, 2005). The interaction between TatA and the signal peptide observed suggests the TatA is directly involved during the translocation event (Fröbel *et al.*, 2011; Maurer *et al.*, 2010; Pal *et al.*, 2013). The current evidence suggests TatA makes contact with signal peptide and then leads to a seeding like recruitment of monomeric TatA to form a form in presence of PMF (Fröbel *et al.*, 2012; Rodriguez *et al.*, 2013).

b. Membrane destabilisation

This model suggests the lipid bilayer becomes weak with large number of TatA localising at a particular site (Figure 1.8 b).The substrate is then able to take advantage of the weakened membrane and pass through the membrane. This model would suggest why multiple sized complexes can be translocated. As previously mentioned the structural organisation of the Tat complexes in bacteria and chloroplasts is highly variable. The pore model that was suggested with evidence to the excessive TatA in *E. coli*, with variable complex sizes supports a pore analogy. However in chloroplasts, the Tha4 (TatA ortholog) is not excessively expressed, suggesting that the pore model is more unlikely because Tha4 is not as readily available.

The membrane destabilisation model can be considered plausible with the recent finding of the TatA topology showing the APH domain can incorporate into the lipid bilayer partially, causing a disruption of the lipid organisation (Walther *et al.*, 2010). The TatA protein is shown to have dual topology where the APH domain can alternate position from the cytoplasmic face to wedged into the bilayer (Walther *et al.*, 2010). In addition, a recent study on TatAd suggests the assumed pore may not be large enough to accommodate substrate, and rather the accumulation of TatAd serves to disrupt the lipid bilayer (Beck *et al.*, 2013; Froebel *et al.*, 2011).

Finally the involvement of a protein that regulates membrane destabilisation has been shown to be associated with the Tat machinery. In particular the phage shock protein PspA in *E. coli*, which works in complement to the Tat system during translocation by directly interacting with the TatA protein (Lo & Theg, 2012; DeLisa *et al.*, 2004; Mehner *et al.*, 2012). A similar function is carried out in the thylakoid by a VIPPI protein. The involvement of a membrane stabilising protein counteracting the compromised permeability of the lipid bilayer has led to the second theory of membrane destabilisation (Kudva *et al.*, 2013).

Currently the evidence for the two proposed mechanism for translocation are limited and with two opposing theories within the field, understanding the mechanistic details of the Tat translocation has increased significantly.

Aim

The functioning of the Tat pathway has proved extremely difficult to understand to date. The lack of structural characterisation of the Tat complexes has limited progress of resolving many ambiguity of the Tat mechanism. The main steps in the mechanism of the Tat pathway that is poorly understood are the initial membrane targeting and assembly of the active translocon. The chapters within this thesis aims to gain insight different stages of the mechanism from signal peptide binding using SPR, to the function of the TatA in *B. subtillis* using Transmission Electron Microscopy (TEM). The structure of the TatAC complexes of *B. subtilis* has not been previously studied, and so visualising these complexes would allow the similarity between the TatAC complex and the TatABC complexes from *E. coli* need to be evaluated to form conclusions about the similarity in the Tat pathway between Gram-positive and Gram-negative bacteria. The overall aim of this thesis is to gain structural and functional insight into the TatAC complex of *B. Subtilis* by structural characterisation. In addition the much debated mechanism of signal peptide targeting was investigated further to compare its similarity to the Sec signal peptide targeting, and, its implication on the cross talk between Tat and Sec during the early cytoplasmic targeting step.

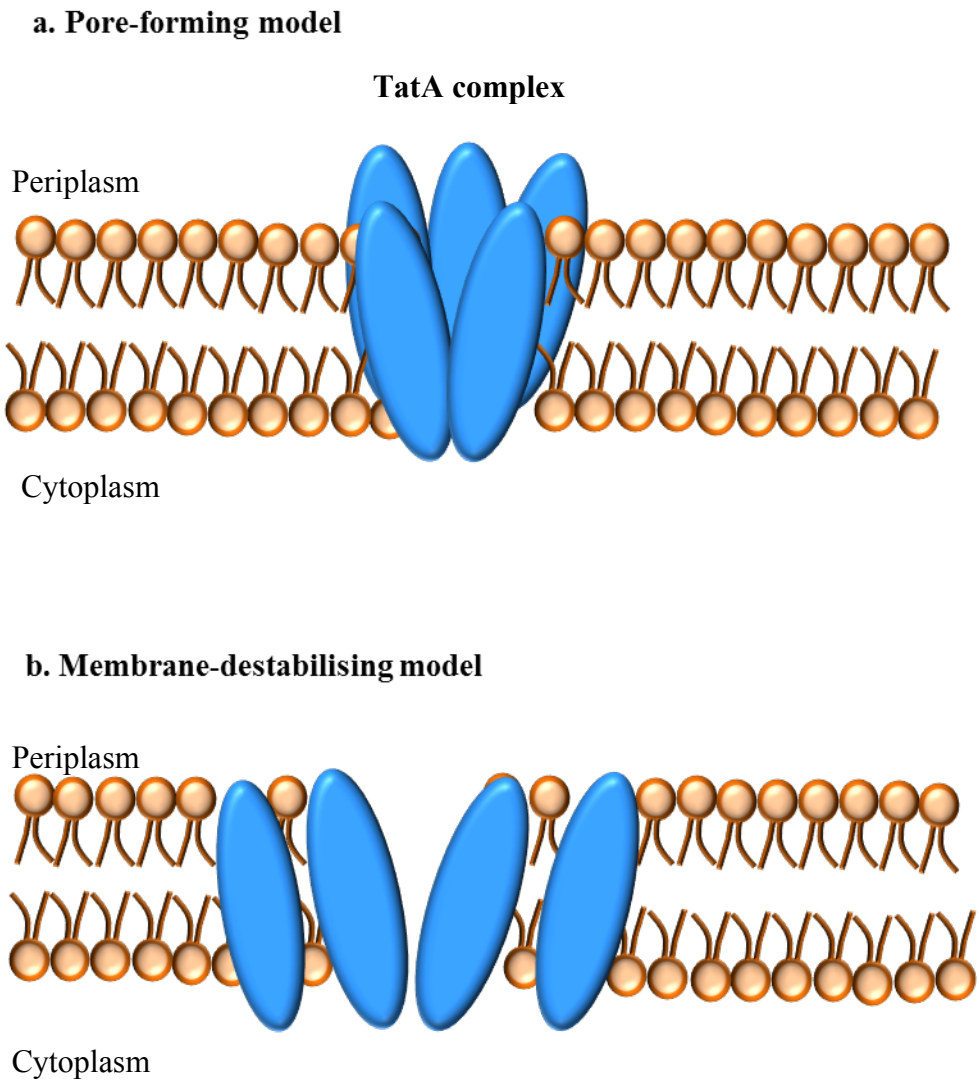


Figure 1.8: The two mechanisms for substrate transport by the TatA complex

The mechanism by which the substrate is translocated, following its binding to the TatBC complex is depicted in the figure above. The first and most supported hypothesis of the pore forming complex, in which the substrate passes through, is shown in (a). The second is a recent mechanism of lipid destabilisation, where the substrate passes through the bilayer that becomes disrupted by TatA accumulation (b).

Chapter 2

Methods and Material

2 Methods and material

2.1 Chemicals, Reagents, Material

Agar Scientific (UK): formvar/carbon coated copper grids (300 mesh).

Avanti lipids (USA): Mini extruder and *E.coli* Lipids

BioRad: Bio-bead^{TM} SM-2 adsorbent, GLC chips, SPR reagents (EDC, NHS, acetate buffer pH 4.5/5, 1 M Methanolamine, 4mM CHAPs, C11 reagent)

GE healthcare (UK): Superdex 200 10/300 GL pre-packed gel filtration column; ECL™ detection reagents; Native protein markers for gel electrophoresis and gel filtration chromatography; Hybond TM-P PVDF membrane.

Calbiochem (Germany): *n*-Dodecyl-*β*-D-maltoside (DDM)

Expedeon (UK): Instant Blue™ ready-to-use Coomassie® stain.

Fermentas (UK): Bradford Reagent ready-to-use

Fisher Scientific (UK): Acetic acid (glacial); Acetone; Ethanol; Formaldehyde; Glycerol; Glycine; Hydrochloric acid; Methanol; Sodium chloride; sodium dodecyl

Sulphate (SDS); Sucrose, and Tris.

Fuji (Japan): Super RX film.

Hamilton (Switzerland) 1ml syringes (for extruder)

IBA (Germany): Buffer E; Strep-tactin™ HRP-conjugate, and Strep-tactin™ affinity chromatography resin, Biotin blocking buffer.

Invitrogen (USA): Anti-His (C-terminal) antibody.

Melford (UK): Isopropyl-*ß*-D-thiogalactoside.

New England Biolabs (USA): Pre-stained broad range protein markers.

Premier International Foods (UK): Marvel milk powder.

Promega (UK) Anti-rabbit-HRP and anti-mouse-HRP conjugates

Roche applied science (UK): Complete™ protease inhibitor cocktail tablets.

Sigma-Aldrich (UK): Ampicillin; L-arabinose; Avidin; Bis-Tris; *ß*-mercaptoethanol; Bovine serum albumin; C12E9; Kanamycin; Lysozyme; Protein A-Sepharose™; Silver nitrate; Sodium carbonate; Sodium dithionite; Sodium thiosulphate; Sodium molybdate; TEMED; Tricine; Triton X-100, Trizma, anti-GFP

SPI-Chem™ : Uranyl Acetate

VWR (UK): 40% Acrylamide solution; Ammonium persulphate; Bromophenol blue; Calcium chloride; Disodium hydrogen phosphate; EDTA; Glucose; Magnesium chloride; Potassium chloride; Sodium dihydrogen phosphate; Sodium hydroxide pellets; Trichloroacetic acid and Tween 20.

Whatmann (UK): 3mm chromatography paper and membranes (0.1nm, 0.2nm, 0.6nm, 0.2um)

2.2 Growth and storage of E.coli cells

2.2.1 Bacterial Strains used

Table 2.1: Strains of *E. coli* used in this thesis

Strain	Properties	Reference
MC4100	F- Δ*lacU169, araD139, rpsL150, relA1, ptsF, rbs, flbBS301*	(Casadaban & Cohen, 1980)
ΔtatABCDE	MC4100 Δ*tatABCDE, Ara*^{r}	(Sargent *et al.*, 1998)
ΔtatA/E	MC4100 ΔtatA/E, Ara^r	(Sargent *et al.*, 1998)
DH5α	*supE44,* Δ*lacU169 (80lacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1*	(Sambrook *et al.*, 1989)

Table 2.2: Bacterial plasmids used in this thesis

Constructs	Properties	Reference	Source
TatABC_S	pBAD24 + *E. coli* TatABC-Strep in ΔtatABCDE	(Bolhuis *et al.*, 2001)	Robyn Eijlander (University of Groningen)
TatAdCd_S	pBAD24 + *B. subtilis* TatAdCd-Strep in ΔtatABCDE	(Barnett *et al.*, 2008)	Robyn Eijlander (University of Groningen)
TatAyCy_S	pBAD24 + *B. subtilis* TatAyCy-Strep in ΔtatABCDE	(Barnett *et al.*, 2009)	Robyn Eijlander (University of Groningen)
S_spDmsA^{RR}_YFP	pBAD24 + Strep-Signal peptide of *DmsA^{RR}_YFP* in ΔtatABCDE	(Mendel *et al.*, 2008)	Sharon Mendel (University of Warwick)
S_spDmsA^{KK}_YFP	pBAD24 + Strep-Signal peptide of *DmsA^{KK}_YFP* in ΔtatABCDE	(Mendel *et al.*, 2008)	Sharon Mendel (University of Warwick)

2.2.2 LB supplements

Luria Broth (LB) media (10 g/L NaCl, 10 g/L bactotryptone and 5 g/L yeast extract) was supplemented with Amipicillin (100 µg/µl) with pBAD24.

2.2.3 Competent Tat Cell

E. coli strain, MC4100 and its knockout mutant listed above were inoculated in 5 ml culture and grown overnight at 37°C. The culture was then diluted 1/10 into fresh LB and grown for 3 h. The cells harvested by centrifuging at maximum 3000 *g* for 10 min. The cells were resuspended in 10 ml of 0.1 M $MgCl_2$ and incubated for before harvesting cells as before. The second round of resuspension was carried out in1ml of 0.1 M $CaCl_2$ and left on ice for a minimum of 30 min. Once competent the cells remain viable for 1 day.

2.2.4 Transformation

100 µl of competent cell and ~100 ng of DNA were equilibrated on ice for 30 min before mixing. Then a further 15 min equilibration period was introduced. The competent cell were then subjected to heat-shock at 42°C for 1 min and immediately placed back on ice for 2 min.

2.3 Fractionation of the *E. coli* cells

2.3.1 Protein expression in *E. coli*

A 5 ml pre-culture was set up with *E. coli* cells containing the appropriate plasmids was cultures overnight at 37°C with the appropriate antibiotics. The pre-culture was diluted 1/100 and grown to an OD_{600} of 0.1-0.2 before induction. For cells containing the pBAD24 plasmid, 200 µM of arabinose was used to induce the cells until the cells reach stationary phase (OD 0.6-0.8).

2.3.2 Fractionation of the *E. coli* cell

Cell were expressed in 10 ml culture by induction with 1 mM arabinose and then grown for 3 h. The cell pellet was the isolated by centrifugation at 4000 *g* and resuspended in 1 ml of disruption buffer (100 mM Tris-acetate pH 8.2, 500 mM

sucrose and 5 mM EDTA). The periplasm was collected by the lysis of the outer membrane, using 20 μg/ml of lysozyme and kept on ice. Osmotic pressure was introduced by the addition of 1ml of ice-cold dH_2O to burst the outer membrane open. Finally 40 μl of 1 M $MgSO_4$ was added to the suspension to inhibit the enzyme activity. The periplasm was collect by centrifugation at 4000g for 1m. The pellet was then resuspended in 0.5 ml of wash buffer (50 mM Tris-acetate pH 8.2, 2.5 mM EDTA) and centrifuged at 4000 *g* for 1 min again to clean up the pellet. The pellet was then resuspended in 1 ml of sonication buffer (50 mM Tris-acetate pH 8.2 and 2.5 mM EDTA) and sonicated at 8 micron for 10 s with 30 s intervals and repeated for 3 cycles. The cytoplasm fraction was then isolated from the membrane fraction by ultracentrifugation at 70000 rpm for 30 min (Beckman TLA 100.3). The cytoplasm is collected as the supernatant and the membrane is isolated by resuspending the pellet in solubilisation buffer (20 mM Tris-HCl pH 8.0, 10% Gylcerol, 50 mM NaCl and 1% C12E9) for a minimum of 1 h at 4°C.

2.3.3 Membrane isolation and solubilisation for Tat purification

For a standard 500 ml culture the cell were centrifuged at 6000 *g*, for 15 min at 4°C to harvest the *E. coli* cells. To a prep of 500ml, 24ml of pre-chilled disruption buffer (100 mM Tris-acetate pH 8.2, 500 mM sucrose and 5 mM EDTA) with 1 protease inhibitor tablet (Roche) was used to resuspend the pellet. The outer membrane was disrupted using 1 mg/ml of lysozyme and followed by immediate addition of 24 ml of dH_2O to release the periplasm as result of osmotic pressure. Finally 1 ml of 1 M $MgSO_4$ was added to the suspension to inactivate the further disruption by the lysozymes. The spheroplasts were then pelleted by centrifugation at 4000 *g* for 15 min at 4°C and resuspended in 10 ml of sonication buffer (50 mM Tris-acetate pH 8.2 and 2.5 mM EDTA). The inner membrane was broken by sonication for 30 s at 10 microns for 4 cycles, with 1 min intervals between each pulse. The membrane was pelleted from the cytoplasm by ultracentrifugation at 70 000 rpm for 35 min using Beckman TL80.6 rotor. The pelleted membranes were solubilised in 10 ml solubilisation buffer (20 mM Tris-HCl pH 8.0, 10% Gylcerol, 50 mM NaCl and 2% DDM or 1% C12E9 unless stated otherwise) overnight at 4°C.

2.4 Protein Chromatography

2.1.1 Purification using anion exchange Q-sepharose chromatography

Ion exchange was used prior to affinity chromatography in some cases to reduce amount of contaminating proteins loaded on the affinity column. The solubilised membranes from a 500 ml culture were applied to a 5ml Q-sepharose column (GE healthcare), which had been equilibrated with 10 ml of equilibration buffer (20 mM Tris-HCl pH 8.0, 50mM NaCl and 10% Glycerol). Once the flow-through was collected, the column was subjected to 5 cycles of 5 ml washes using the wash buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% Gylcerol and 0.1% C12E9 or 0.02% DDM). The protein was then eluted with 5X 2ml elution buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% Gylcerol and 0.1% C12E9 or 0.02% DDM. The elutions were directly into 2 ml of no salt buffer (20 mM Tris-HCl pH 8.0, and 0.1% C12E9 or 0.02% DDM to immediately restore the salt concentration to 150mM.

2.4.1 Purification using Streptactin^{TM}-Sepharose affinity chromatography

Any solubilised membranes or Q-Sepharose elutions that were to be applied onto a streptactin column were incubated with 1/1000 (*v/v*) of avidin at 1 mg/ml for a minimum of 30 min prior to loading. The streptactin column was equilibrated with 2 column volumes (CVs) equilibration/wash buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.1% C12E9 or 0.02% DDM) before the sample was applied. Proteins were then applied to 4ml streptactin column, and then washed 5X with 4 ml of equilibration/wash buffer. Finally the strep-tagged proteins were eluted in 10X 1ml elution fraction using the elution buffer (20 mM Tris-HCl, 150 mM NaCl, 2.5 mM desthiobiotin and 0.1% C12E9 or 0.02% DDM).

2.4.2 Purification using Talon^{TM} affinity chromatography

Firstly, 10 ml of Talon resin was equilibrated with 1 CVs of buffer 1 (20 mM Tris-HCl pH 8.0, 400 mM NaCl, 5 mM imidazole and 0.02% DDM) and left rotating overnight at 4°C. The resin was pelleted by centrifugation at 2000 rpm for 5 min to

remove buffer. The solubilized membrane or pooled Q-sepharose elutions were mixed with the equilibrated Talon resin and left rotating for a minimum of 4 h at 4°C. The slurry was then applied to a column and allowed to settle before collecting the flow-though. The column was washed in 6 CVs of buffer 1 and then eluted with 2.5 CV of buffer 2 (20 mM Tris-HCl pH 8.0, 400 mM NaCl, 150 mM imidazole and 0.02% DDM). The elution fractions were collected in 0.5 ml fractions, which were then run on a SDS-PAGE gel and visualized by western blot. Peak fractions were concentrated using 30kDa VivaSpin concentrator for TatAyCy, TatAdCd and TatABC complexes.

2.4.3 Gel-Filtration Chromatography

Complexes destined for Electron Microscopy (EM) were further resolved by gel-filtration microscopy to polish the purification and also get a overall size estimation of complexes that may have formed using molecular markers. A Superdex200 10/300 GL column was used in conjuction with Äkta purifier FPLC system (GE healthcare), operated by the UNICORN® v.4.00 software. The superdex200 10/300 was used to resolve protein in range of 10 – 600kDa. The column was equilibrated with 2 CVs of the selected GF buffer (20 mM Trizma-HCl pH 7, 150 mM NaCl) with respective detergent supplement. Concentrated sample was centrifuged at 10000 rpm for 10 min before application to the column. A sample volume of 250 µl was applied onto a 0.5 ml superloop and then the separation was performed at 0.5 ml/min. The elution was collected in 0.5 ml fractions. Between sample runs the column was rinsed with 2 CVs buffer and the 0.5 ml loading loop was manually flushed with the same buffer. The following cleaning protocol was used routinely after weekly usage or change in buffer conditions to maintain column performance and resin purity: Wash with 2 CVs dH2O, then 2 CVs 0.5 M NaOH, then 2 CVs dH_2O. Eluted fraction were collected and westernblotting for detection of protein complexes.

2.4.4 Gel filtration calibration

Calibration of the Superdex200 was performed using protein standards of known molecular mass. Blue Dextran was used to estimate the column void volume (Vo);

the volume at which all matter too large to enter the resin matrix is eluted. The protein standards and their elution volumes (Ve) are shown below in

Table 2.3.

Table 2.3: The elution volumes for protein standard on Superdex200

Mol. Wt. (kDa)	**Mol. Wt. (Da)**	**Protein standard**	**Ve**	**Ve/Vo**
669	669000	Thyroglobulin	9.30	1.13
440	440000	Ferritin	10.92	1.32
158	158000	Aldolase	12.87	1.56
66	66000	Albumin	14.03	1.70
			Vo	
2000	2000000	Blue Dextran	8.25	

2.5 Protein Detection techniques

2.5.1 SDS-PAGE

Proteins were seperated on a 17.5% SDS-page resolving gel for proteins sepeartion in the range of 5kDa-175kDa. For protein lkarger than 10kDa a 15% SDS-page resolving gel was used. The stacking gel was made to 5%. Table below shows the component of the gel are listed on table below. The gels were run using C.B.S vertical gel system as per manufacturers instructions. Samples to be run on the gel

were mixed in 2X loading sample buffer (Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.02% bromophenol blue and 5% β-mecaptoethanol)

Table 2.4: Chemical used to make SDS-PAGE stacking and resolving gel

Reagent	17.5% resolving gel	15% resolving gel	6% stacking gel
Acrylogel	17.5%	15%	6%
Tris-HCl pH 8.8	375 mM	375 mM	-
Tris-HCl pH 6.2	-	-	125 mM
SDS	0.1%	0.1%	0.1%
APS	0.05%	0.05%	0.05%
TEMED	0.01%	0.01%	0.01%

2.5.2 Coomassie Stain

SDS-PAGE gels were incubated with 50 ml of Instant Blue Coomassie stain for a minimum of 1 h before washing with dH_2O until background is clear.

2.5.3 Silver stain

Once run, the SDS-PAGE gels were incubated in fixer solution (50% acetone, 1.25% TCA and 0.015% formaldehyde) for 15 min. The fixer was then washed off with dH_2O, 3 times in 5 min. The gel was then washed with 50% acetone for a further 5 min. The gel was then immidiately incubated in enhancer solution (0.02% (*w/v*)

sodium thiosulphate) before a 5 min wash in dH_2O. The gel was then incubated for 8 min in staining solution (0.25% (*w/v*) silver nitrate and 0.4% formaldyhyde (*v/v*)) and developing the gel until bands become visible in the developing solution (0.2 mM Sodium carbonate, 0.004% Sodium thiosulphate (*w/v*) and 0.015% formaldyhyde). Once the band were visible the developing was stopped using 1% acetic acid. Finally the gel was rinsed in water before transferring to a plastic cover for scanning.

2.5.4 Western blotting

Proteins were separated on a SDS-page gel and transferred to a PVDF membrane using semi-dry western blotting system (Sigma, UK). The membrane was then blocked in either 5% milk or 3% BSA for a minimum of 1 h. Membranes blocked with BSA an addition 10 min incubation with 20 µl of biotin blocking buffer in 10ml PBS/T before antibody addition. Western blots were performed using anti-His IgG and horseradish peroxidase (HRP) conjugated anti-Mouse IgG. The blot was visualised using chemillucent EZ-ECL detection kit.

2.5.5 Protein concentration determination

The protein concentration was determined using the PierceTM BCA Protein Assay Kit as per manufacturer instructions. The kit used is suitable with detergent and can be used to determine the protein concentration of membrane proteins.

2.5.6 Densitometry

2D densitometry was performed on Western blot bands with the AIDA (Advanced Image Data Analyzer) software v. 3.28.001.

2.6 Trasmission Electron Microscopy

2.6.1 Grid preparation

The peak elutions of the Tat complexes from gel filtration were diluted to ~ 0.1 mg/ml in buffer A before fixing to the grid. The grid were prepared as previously described by Beck *et al.* (2013) in brief a carbon coated grid (200 mesh, Agar scientific) was negatively glow discharged for 1 min. Immediately 4 µl of protein sample was applied for 1 min and washed twice with buffer A minus detergent. The grid was then stained twice for 20 s using 2% Uranyl acetate and then air-dried for a minimum of 5 min.

2.6.2 Transmission electron microscopy

The grid was imaged under a 200kV JEOL 2010 field emission gun (FEG) TEM with a 4k Gatan Ultrascan CCD camera. The images were taken at 40000X magnification at -1.5DV defocus.

2.6.3 Density profiling

The DM3 file was used on the Digital Micrograph software and used the density profile tool to measure the size of the complexes.

2.7 Atomic Force Microscopy (AFM)

All AFM work was carried out by Cvetelin Vasilev (University of Sheffield)

2.7.1 Sample preparation

The purified protein was diluted 10 or 50 folds in imaging buffer (10 mM Tris pH 7.4, 120 mM KCl, 0.03% DDM) and 40 µl of the protein solution was deposited onto freshly cleaved mica surface attached to a magnetic puck (AFM sample support). After leaving the protein to adsorb for approximately 15 min, the sample was rinsed with imaging buffer (3 times) and imaged with the AFM.

2.7.2 AFM measurements

The AFM data was collected in PeakForce QNM mode using a Multimode 8 instrument equipped with a 15 μm scanner (E-scanner) and coupled to a NanoScope Vcontroller (Bruker). NanoScope software (v8.15, Bruker) was used for data collection and Gwyddion (v2.30, open source software covered by GNU general public license, www.gwyddion.net) and OriginPro (v8.5.1,OriginLab Corp.) software was used for data processing and analysis. All measurements were performed in nearly-physiological conditions in imaging buffer (10 mM Tris pH 7.4, 120 mM KCl, 0.03% 13-DDM) at room temperature using BL-AC40TS probes (Olympus). The Z-modulation amplitude was adjusted to values in the range 20 – 24 nm, while the Z-modulation frequency was 2 kHz and the contact tip-sample force was kept in the range 60 – 80 pN.

2.8 Protein reconstitution in liposomes

2.8.1 Liposomes preparation

The *E. coli* polar lipids mixture dissolved in chloroform was sourced from Avanti. 1ml of the lipid mixture (25 mg) was dry in a 50 ml round bottom flask using gentle nitrogen airflow in 60°C waterbath. The thin yellow lipid film was then left under a vacuum overnight to remove any residue solvents. The lipid film was then hydrated in nitrogen saturated liposome buffer (20 mM Tris-HCl pH 8 and 150 mM NaCl) to give a 2 mg/ml lipid solution. The buffer is incubated with the lipid film for 4-6 h with gentle stirring, forming multilamellar vesicles. Once the lipid layer was fully hydrated, liposomes were subjected to 5 cycles of freeze-thaw. The liposomes were then extruded through a 600nm (3X) and 200nm membrane (20X) to form uniform 200nm size liposomes.

2.8.2 Protein Reconstitution

After extrusion, the pre-formed liposomes (2 mg) were subjected to 0.12% TritonX100 disruption. Upon addition of TritonX100, the liposomes were rotated for

15 min. Then, 20 µg of protein was added to the 2 mg of liposomes and further incubated for 45 min at room temperature for protein to be incorporated. The tritonX100 detergent was then removed using SM2 Biobeads. 30 mg of wet Biobeads were gradually added to the 1 ml liposomes mixture and incubated at room temperature for 30 min. Further 60 mg of beads and incubate for 1 h. The final batch of 60 mg Biobeads were added and rotated overnight at 4°C. The liposome solution was extracted from the Biobeads and concentrated by ultracentrifugation at 55000 rpm for 1.5 h (Beckman TLA 100.3 rotor).

2.8.3 Preparation of Wet Bio-beads SM-2

Biobeads SM-2 are adsorbent beads that can capture multiple detergent molecules in a single bead. Bio-beads (1 mg) were washed in methanol by rotating for 15 min and then discarded. Another round of 16 ml of methanol was used to wash the beads for 30m and then discarded as previously explained. The methanol was then washed out by multiple washes in 30 ml dH_2O for 30 min. Finally, the beads were left to equilibrate with the liposome buffer (20 mM Tris/HCl pH 8 and 150 mM NaCl) until ready to use.

2.8.4 Sucrose density gradient centrifugation (SDGC)

The sucrose density gradient is used to separate protein and liposomes based on the density of the sample. Liposomes have lower density than protein so the liposomes can be purified out from the excess protein that is not incorporated into the liposomes by this method. Firstly solutions contained either 60% and 20% sucrose are made up with liposomes buffer (20 mM Tris/HCl pH 8 and 150 mM NaCl) and chilled. First the 60% sucrose is diluted down to 40% by addition of the liposomes solution (400 µl 60% sucrose and 200 µl liposomes). This sucrose-liposomes mix forms the bottom layer of the gradient. This layer is then topped up 2 ml of 20% sucrose and final layer of the 400 µl liposome buffer is used to top up the gradient. The gradient is then subjected to ultracentrifugation at 55000 rpm (Beckman TLA 100.3) for 1.5 h at 4°C. the low density of the liposomes pulls them to the top middle layer while any proteins not incorporated protein remain in the bottom, high sucrose density layer.

2.9 Surface Plasmon Resonance SPR

2.9.1 Activating GLC chip

The GLC chip was activated with 40 mM EDC and 10 mM NHS reagent mix which coats the surface of the chip. The EDC-NHS mix allows proteins/compounds to be immobilised to the chip by amine coupling. Once the compound of interest is immobilised onto the chip, the surface is then blocked with ethanolamine () to prevent further amine coupling reactions. The running buffer (20 mM Tris/HCl pH 8 and 150 mM NaCl) used with the ProteOn XPR36 system was supplemented with 0.005% tween to prevent non-specific binding. Variable concentration of liposomes and substrate were diluted in the running buffer to measure.

2.9.2 Substrate Immobilisation to the GLC chip

The anti-GFP monoclonal antibody was immobilised to the activated GLC chip with the running buffer PBS/Tween at 0.005%. The antibody was diluted to 0.3 mg/ml with acetate buffer. The buffer was then changed (20 mM Tris/HCl pH 8 and 150 mM NaCl) for substrate immobilisation (S_spDmsA^{RR}_YFP or S_spDmsA^{KK}_YFP) and measuring analyte binding. The concentration of DmsA-YFP used, varied from 1 mg/ml to 0.25 mg/ml. The TatABC and TatAdCd liposomes were then flown over the immobilised substrates at a rate of 25 µl/ml to measure the binding for 480 s. All reagents and protein samples used on the ProteOn XPR36 were prepared in micro-plates (BioRad) and made up to 300 µl for each step.

2.9.3 Liposome immobilisation to the GLC chip

The C11 reagent (Undecylamine in MES buffer and 10% DMSO) was used after the EDC-NHS activation. C11 forms lipophilic interaction with the liposomes to immobilise them onto the GLC chip. Variable concentrations of liposomes were used (0.125-1 mg/ml) to immobilise to the C11 surface. The running buffer (20 mM Tris/HCl pH 8 and 150 mM NaCl) used for the immobilisation of liposomes was free of detergent to prevent liposomes disruption.

2.9.4 Measuring Binding of analyte

The analyte was flowed over the ligand of interest at a rate of approximately 30 µl/min and so the binding was measured over 300 s. The dissociation was measured by flowing the running buffer over the chip from 600-1200 s and remove unbound proteins. The dissociation of the protein was in the running buffer (20 mM Tris/HCl pH 8 and 150 mM NaCl).

Chapter 3

Expression and purification of the TatAdCd_S complexes

3 Expression and purification of the TatAdCd_S complexes

3.1 Introduction

The *E. coli* cell contains two types of translocons, capable of transporting protein across the cytoplasmic membrane; the Sec translocon and Tat translocon. The Tat machinery in *E. coli* is minimally composed of 3 membrane proteins, TatA (9kDa), TatB (18kDa) and TatC (28kDa) (Oates *et al.*, 2003). In a resting membrane, the Tat machinery exists as two separate complexes; the hetero-oligomeric Tat(A)BC complex and homo-oligomeric TatA complex. Upon substrate binding, a conformational change induces the recruitment of the TatA complex by the TatABC complex, forming an active translocon of 600kDa (Oates *et al.*, 2005). It has proved extremely difficult to isolate this ultimate complex due to its requirement for a proton motive force, and its transient nature.

In comparison, the Tat machinery in most Gram-positive bacteria consists of only two components, TatA and TatC. More specifically, *B. subtilis* has two homologous systems working in parallel known as the TatAyCy and TatAdCd_S (Jongbloed *et al.*, 2004). The specificity for the substrates is what differentiates the two systems. TatAdCd_S has been estimated to form complexes of 230 kDa (Barnett *et al.*, 2008). The stringency of specificity of the TatAC systems is compromised when heterologously expressed in *E. coli,* with substrates exceeding the usual size (Barnett *et al.*, 2011; van der Ploeg *et al.*, 2011). The differences presented between the Gram-positive and Gram-negative systems suggest the mechanisms between the two systems may not be as comparable, as first perceived.

The active Tat complex has proved difficult to isolate and characterise, and so far the focus has been on studying the two complexes individually, to understand the mechanism. As the Tat complexes are membrane embedded, detergent must be used

to stabilise these complexes. In addition, for structural characterisation, these complexes they must be isolated with minimum detergent concentration to reduce interference. This chapter focuses on the optimisation of the expression conditions to increase the yield of Tat protein expression. Once expressed, the Tat complexes were isolated in the minimum amount of detergent to stabilise the complex and reduce interference for biophysical studies.

Detergents molecules contain a hydrophilic head and hydrophobic tail, similar to a phospholipids found in the membranes. These individual molecules can dislocate the membrane, to disrupt the phospholipid bilayer and then accumulate together to form a micelle. There are many types of detergent with different properties; those commonly used for membrane protein isolation include Digitonin, n-dodecyl-β-D-maltopyranoside (DDM) and Nonaethylene-glycol-dodecyl-ether (C12E9). These detergents are all non-ionic and resemble the *in vivo* lipid membrane environment. Depending on the size and aggregation number of the detergent, the detergent has a minimum concentration that is required to form a micelle, known as the critical micelle concentration (CMC). Different detergents form different sizes of micelle, where large micelles are commonly used for larger protein or protein complexes. In some cases, using larger detergent micelles can prevent protein-protein interactions and would not be favourable in a complex assembly. Even small micelle-forming detergents have their disadvantages such as undesirable protein aggregation as result of lack of detergent.

For this study, C12E9 (Nonaethylene-glycol-dodecyl-ether) was chosen for protein isolation and purification because of it low CMC and large micelle size (70 kDa). This chapter aimed to optimise the expression of Tat complex for higher yield purification, using minimal detergent concentration during purification.

3.2 Time-dependent expression of TatAdCd_S

TatAdCd_S was overexpressed from a pBAD24 plasmid using 1mM arabinose at multiple time points. The current lab protocol for protein expression was based on immediate induction of the cells, however using this method the yield of purified protein remains minimal, around 100 µg per 1 L culture. In an attempt to increase protein expression levels, the cells were allowed to reach a stable growth condition (stationary) before induction of the plasmid. The *E. coli* cells were either induced immediately (exponential phase), at 1.5 h (end of exponential phase) or at 3 h (stationary phase). The optical density of the cells was measured at 600 nm to record cell growth. The result of this preliminary study showed that when *E. coli* cells carrying the pBAD_TatAdCd_S plasmid were not induced, the cells require approximately 2 h to reach stationary phase Figure 3.1. However, when cells were induced with arabinose irrespective of the induction time, the optimal density (OD_{600}) is drastically reduced. This suggests the overexpression of pBAD_TatAdCd_S may induce a toxic effect on cell growth.

With the aim to reduce this toxic effect, the cells were either grown at a lower temperature (data not shown) or with a lower concentration of arabinose, as shown in Figure 3.2. It is evident that, as the time of induction increases, cell density also increases; therefore number of cell expressing pBAD_TatAdCd_S also increases. During the first induction time point (exponential phase), inducing applies stress to the cells that are already under stress. So although the initial OD may be similar at 0 h of the growth phase, induction causes cells to become stresses and therefore reproduce slowly compared to the non-induced cells (Figure 3.1).

The cell growth curves for cells that are induced with arabinose reach the stationary phase within the first 2 h irrespective of induction time point. This may suggest why the earlier protocol used immediate induction to achieve maximum time of expression before the cells reach the stationary phase. However when the cells are grown in absence of the inducer, the optical density of the stationary phase is higher Figure 3.1, showing greater density of cell are viable in the stationary phase as opposed when induced earlier (0 h). Therefore in presence of an inducer the cell stress levels are increased and so cell growth is limited. As the cells reach stationary

phase within 2 h of induction, and Tat requires a minimum of 2 h induction of express and localise the proteins, another approach to reduce cell stress/toxicity was also explored. As the expression of membrane proteins is generally toxic to the cell, reducing the rate of the expression by varying the concentration of arabinose was investigated.

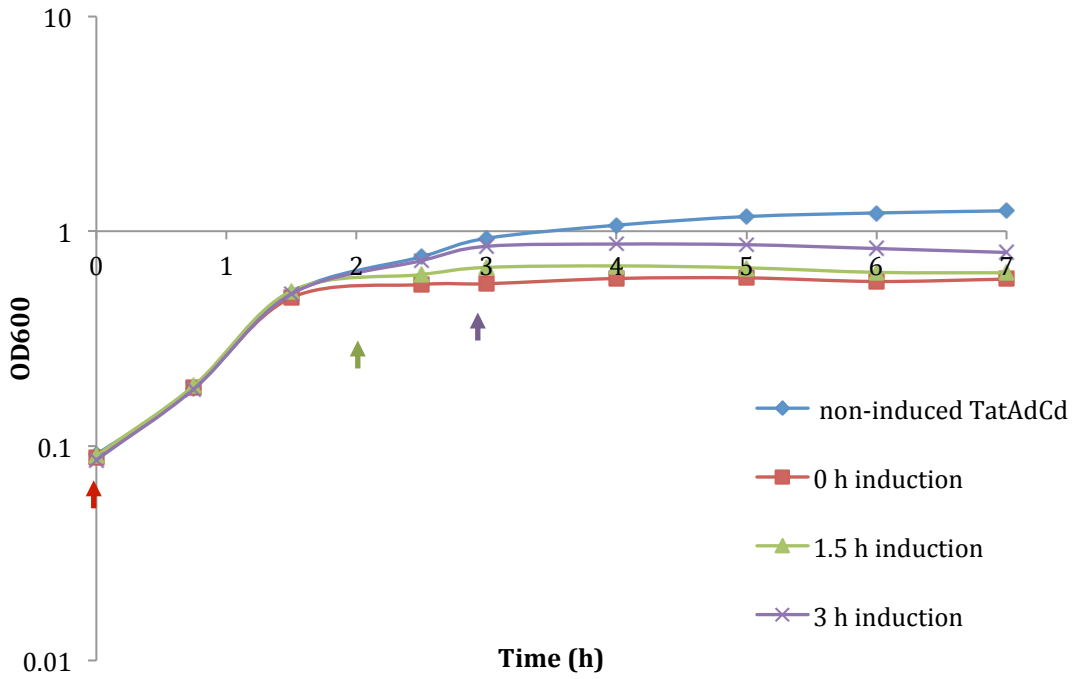


Figure 3.1: Growth curve for *E. coli* cells expressing TatAdCd_S_S at varied time points

pBAD_TatAdCd_S_S was overexpressed in ΔTatABCDE *E.coli* cell with 1 mM arabinose in a 500ml culture. The cell was cultured at 37°C and the optical density of the cell was measured at 600 nm every 30 min. The time of induction was varied from 0 h (red), late-exponential at 1.5 h (green) and stationary at 3 h (purple). A control growth was measured without induction (blue). The time of induction is marked by the coloured arrows.

3.3 Concentration-dependent expression of TatAdCd_S

The concentration of arabinose for induction is also thought to influence cell growth and may also influence the toxicity to the cell. The experiment was designed to vary the concentration of arabinose used when the cells are induced at time point 0 h. The standard protocol states 1 mM arabinose is required for induction, so the concentration was varied in a 2-fold range above and below 1 mM.

The control cell were grown without induction, and as a result, the cells are capable of growing to an OD_{600} of 1.2 before they reach a stationary phase Figure 3.2. However, it is clear that when pBAD_AdCd_S is induced the yield of cells at stationary phase is reduced. As shown earlier in Figure 3.1, cells that are induced reach the stationary phase before the non-induced cells because of increased stress and consequently reduced growth. The effect of varying the arabinose concentration has no significant effect on cell growth as maximum OD_{600} obtained, which does not exceed the OD600 of 0.9, irrespective of the arabinose concentration. The induced cells reach the stationary phase within 2.5 h while without induction; the stationary phase is delayed 3 h.

It is evident that the concentration of arabinose (inducer) does not influence cell growth but the protein expression activity may still be influenced by the manipulation of the inducer concentration. This is observed when cells reach higher optical density; the increased concentration of arabinose allows increased number of cells to be induced. Therefore an increase in TatAdCd_S expression is a result of cell density at time of induction, rather than the concentration of arabinose.

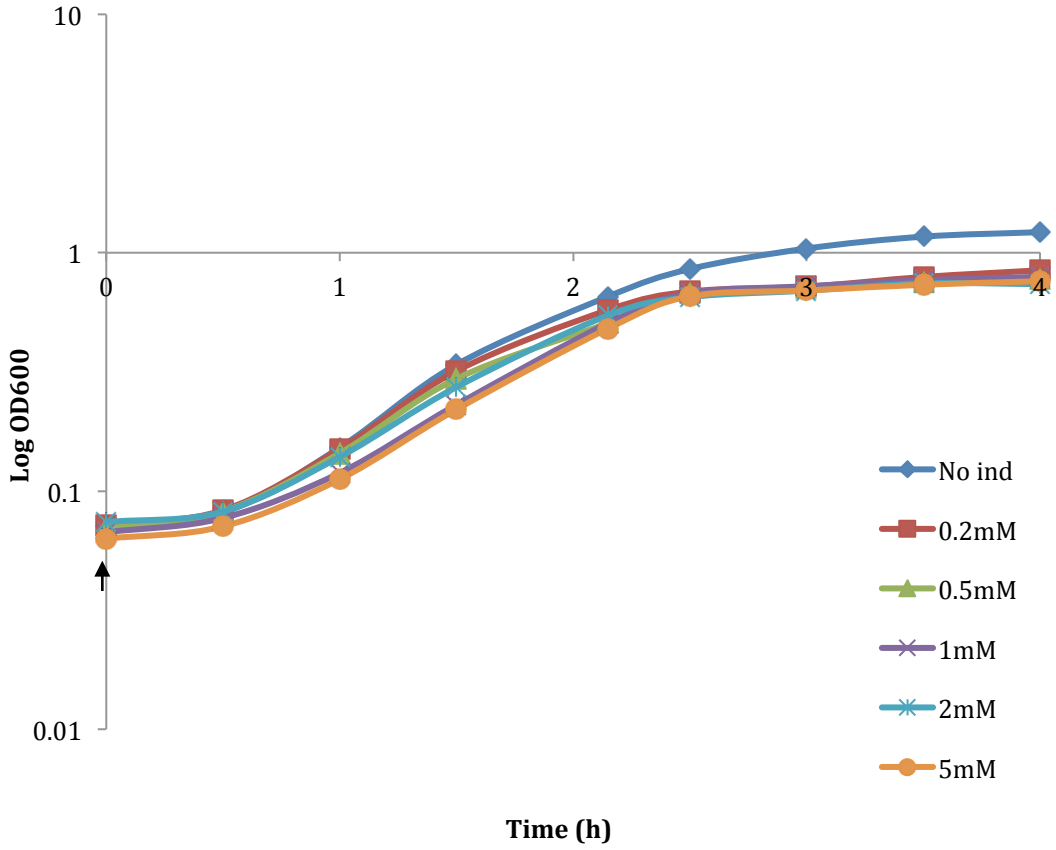


Figure 3.2: Growth curve for *E. coli* cells expressing TatAdCd_S with varied concentration of arabinose

pBAD_TatAdCd_S-Strep was expressed formed in ΔTatABCDE. *E.coli* cell were induced with a range of concentration of arabinose (0-5 mM) in a 500 ml culture. The cells were induced at 0 h as with the standard protocol. The control (blue line) was not induced during the growth. The cells were cultured at 37°C and the optical density was measured at 600 nm at 30 min intervals. The black arrow marks the induction time point.

3.3.1 Growth curve of *E. coli* strains without induction

Complementary to the expression studies described in section 3.2/3.3, a control experiment was performed to ensure that the growth limitations observed are due to overexpression rather than as a consequence of the strain used. Growth curve for both the MC4100 strain (WT Tat) and the Tat knockout strain (Δ Tat) with the deletion of the Tat genes (ΔTatABCDE) show no difference in growth (Figure 3.3). This suggests that deletion of that Tat machinery in *E. coli* does not have a detrimental effect on the cell growth under these conditions. However, when the native *B. subtilis* Tat machinery (TatAdCd_S) is incorporated into the ΔTat *E. coli* strain, slight stress is applied to the cell. This also suggests that with increasing manipulation of the *E. coli* strain, the robustness of the cells prior to overexpression is compromised. The maximum OD_{600} for WT Tat and Δ-Tat strain plateau just above to 2. With the TatAdCd_S strain, the optimal density does not exceed 2 (Figure 3.3), suggesting TatAdCd_S strain used for the expression studies are predisposed to stress before induction. It is also evident that the cells reach the stationary phase within 4hrs and are at mid exponential phase by 1 h. Ideally, for the expression of the complete Tat machinery and its incorporation into the membrane requires a minimum of 2hrs. This presents a valid reason for immediate induction, allowing sufficient time for expression of Tat before the cells reach the stationary phase in the previous protocol.

The previous protocol of immediate induction allowed protein to be expressed efficiently however the yield produced was lower. For this project, the yield needed to be optimised to allow further characterisation of TatAdCd_S complex for biophysical studies. Therefore the effect of expressing TatAdCd_S protein on the complex during the late-exponential phase was further investigated.

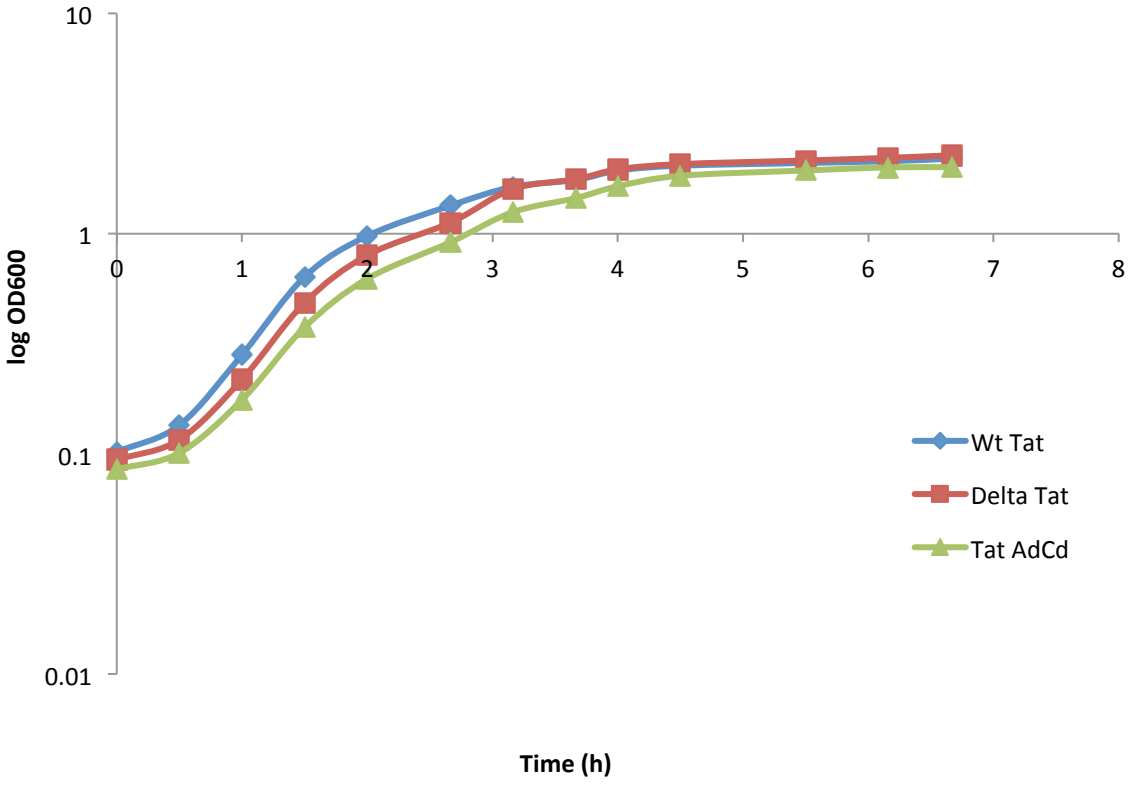


Figure 3.3: Growth curve for native expression of TatABC and TatAdCd_S in *E. coli*

The three strains were grown at 37°C without induction of the cells in a 500 ml culture. The growth curve for TatABC (WT-Tat) and Tat knock out (Δ-Tat) and TatAdCd_S (Δ-Tat) was measured for this experiment. The OD of the cells was measured at 30 min interval over 7 hours. The cells reached a stationary phase at 4 h of growth and reach the mid-exponential phase within 1.5 h.

3.4 Time-dependent and concentration-dependent expression of TatAdCd_S

It was shown that varying the concentration of inducer for immediate induction did not alter the growth phase of the cells significantly. From this experiment it was shown that the toxicity of TatAdCd_S expression could not be regulated by the concentration of arabinose in this case.

3.4.1 Growth curve for cells expressing TatAdCd_S under different condition

In this experiment three different conditions were tested against the control cell growth to determine the ideal conditions for the expression of TatAdCd_S. The standard protocol represented by the green line in Figure 3.4, shows the cells induced at 0 h has a short exponential phase and the growth rate remains low. In addition, the concentration of arabinose at induction was reduced to 0.2 mM to see if this extends the exponential phase in Figure 3.4 (shown in red line). The result again showed that the cells reached the stationary phase at the same time however, the OD_{600} maxima is increased with lower concentration of arabinose. This result is contradictory to that observed previously which suggested that reducing the concentration of arabinose had no effect on OD_{600} maxima at the 3 h stationary phase (Figure 3.2). When the cells were induced at time-point 0 h with a lower concentration of arabinose (red line), the cell growth was stopped at 3 h after induction. This effect shows that reducing the concentration may have an effect on the early stages of cell growth. For the control (blue line) and the 3 h induction (purple line, Figure 3.2) cell growth reached an OD_{600} of ~0.5 before induction and was allowed to growth for further 3 h. Once induced at 3 h the growth rate is reduced, however the drop in growth does not restrict the yield. Therefore the cells are able to reach maxima of OD_{600} = 1 as with the control cells (Figure 3.4). For this experiment the cell were all induced for 3hr and then harvested for purification so the growth curve is not complete for 0 h immediate induction.

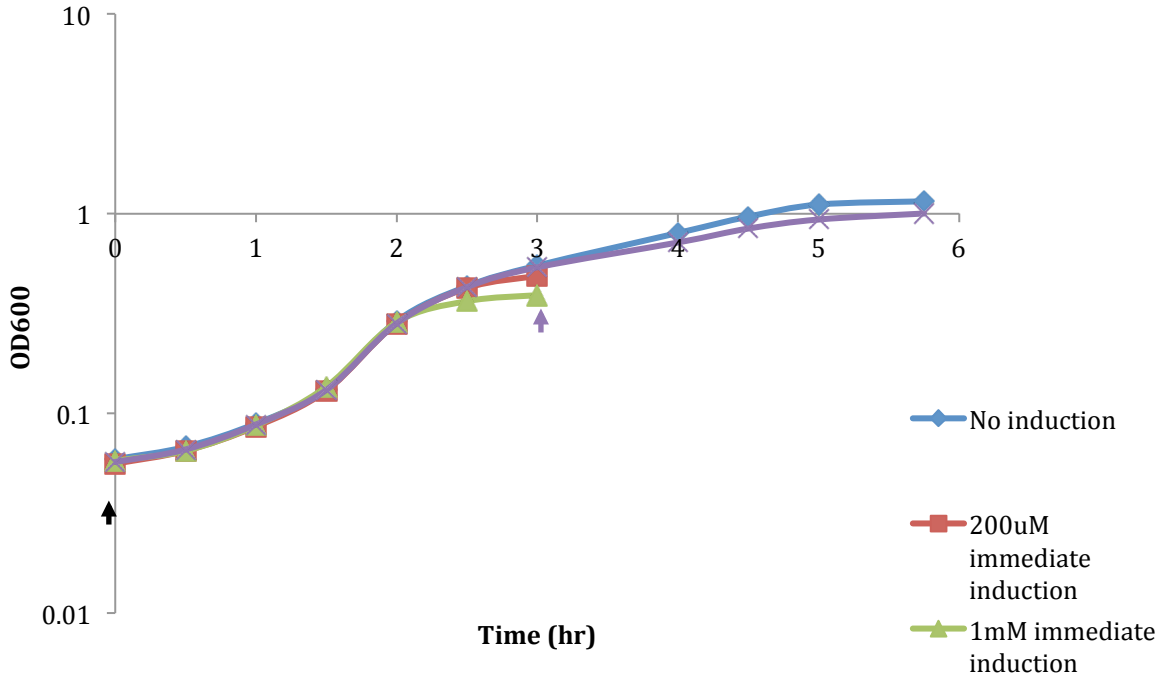


Figure 3.4: Growth curve for TatAdCd_S expression in *E. coli* under varied conditions.

TatAdCd_S_S was expressed from a pBAD24-based vector transformed in ΔTatABCDE. *E.coli* cells were induced with 1 mM arabinose in 500 ml culture. The time of induction was varied with concentration of arabinose (200 µM and 1 µM) at 0 h or with 1 mM arabinose at 3 h induction. The cell was cultured at 37°C and the optical density of the cell was measured at 600 nm every 30 min. the arrows mark the induction time point at 0 h or 3h.

3.4.2 Purification of cell expressed under different conditions

To further investigate the effect of varying the time point and inducer concentration on the expression levels of TatAdCd_S directly, the cell were pelleted and fractionated to isolate the membranes. The membrane fraction was then solubilised in 1% C12E9 overnight before using ion exchange chromatography (Q-Sepharose), followed by StrepTactin affinity chromatography at 4°C. The elution fractions collected from the StrepTactin column is shown in Figure 3.5.

The western blot of the StrepTactin purification suggests, the amount of protein expressed is relatively higher when cells are induced at 3 h, as the yield of cells expressing TatAdCd_S, is also higher. With the standard protocol, cells are stressed prematurely when induced at 0 h with 1mM arabinose. This demonstrated in the results observed in the western blot, where the cell growth is restricted and the expression of TatAdCd_S remains low (Figure 3.5). Similarly, although reducing the concentration of arabinose to 0.2 mM results in a small increase in cell growth, the overall expression and purification yield is lower (Figure 3.5). The expression of TatAdCd_S appears to be optimal for high yield once the cells are in a stationary phase. Therefore, even with early induction and limiting the concentrations of arabinose, the cells are capable of expressing the protein as efficiently in the early stages of the growth cycle. When the cells are grown to the mid-exponential phase, they express proteins more efficiently. In addition, the increased optical density at the time of induction results in a higher yield of protein expression. From this study, it can be concluded that cells must be induced in the stationary phase still express proteins and leads to increased protein yield. The overall time of induction must be kept at 3 h to provide sufficient time for protein expression and localisation.

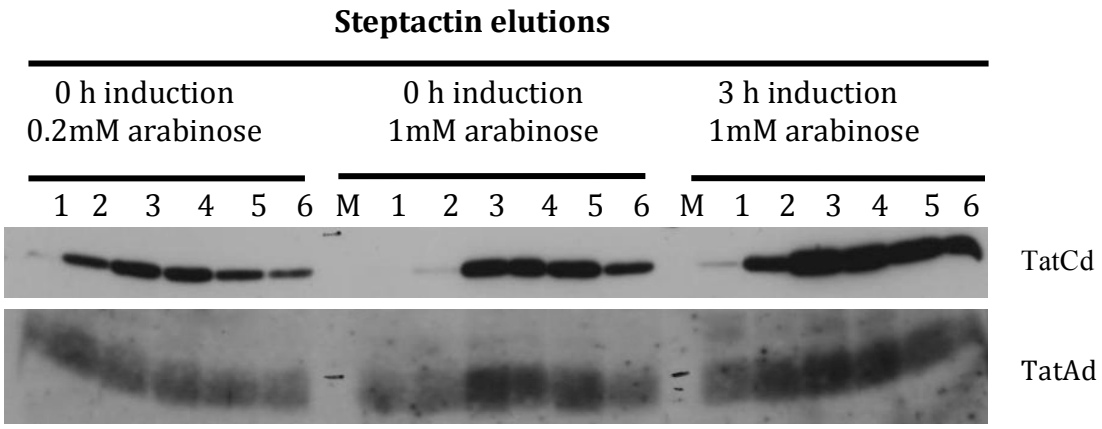


Figure 3.5: Western blot of TatAdCd_S expressed under varied conditions

TatAdCd_S was expressed with variable arabinose concentration (0.2 mM and 1 mM) at 0 h and at 3 h with 1mM. The TatAdCd_S was isolated from the membrane from 250 ml cultures and purified on a 5 ml StrepTactin column. Elutions from the StrepTactin column are shown in the western blots.

3.4.3 Gel filtration of TatAdCd_S complexes expressed under different conditions

The purification of TatAdCd_S shows the highest yield is obtained when the cells are induced in the later stage of the growth cycle, such that the cells reach the stationary phase following 3hrs induction. Therefore as the cell growth rate decreases, it is important to check expression is not compromised. If compromised, this could lead to the formation of inclusion bodies and aggregation of proteins, due to misfolding. As part of the final purification step, gel filtration chromatography was used to assess the quality of protein expression (formation aggregation or inclusion bodies) by size estimation. If the proteins were forming aggregates then the aggregation can be visualised as a peak. Gel filtration chromatography allows protein of different size to be resolved based on the size of the protein complexes. The absorbance of the protein was measured at 220nm and 280nm to show the peak protein elution. The absorbance at 280nm was used to detect the protein from its aromatic groups, which absorb in this range; these include tryptophan and tyrosine. However the actual peptide bonds of the proteins are detected at 220nm and also the aromatic group phenylalanine. The 220nm absorbance resulted in a higher RU spectrum and for this reason this absorbance was used for a clearer interpretation of the data.

Using the standard protocol (blue line-Figure 3.6) the TatAdCd_S complexes are eluted at 11.9ml and by using the calibration marker, the size of these complexes was estimated to ~220kDa. When the cell was induced with 0.2 mM arabinose (red line-Figure 3.6) the peak position remained constant, showing the size of the complexes is not altered by reduced expression. However, when the cells are induced at 3 h with 1 mM arabinose, the peak (green line-Figure 3.6) shifts to 11.6 ml which increases the average size of the complexes to ~260kDa. The change in complex size is not significantly varied but this either suggests that there may be some unordered organisation of the complexes that may be occurring. In addition, the increase in the average size of the complex can also be a result of increased expression level of the proteins, and so the cell can produce larger oligomers depending on the protein expression levels.

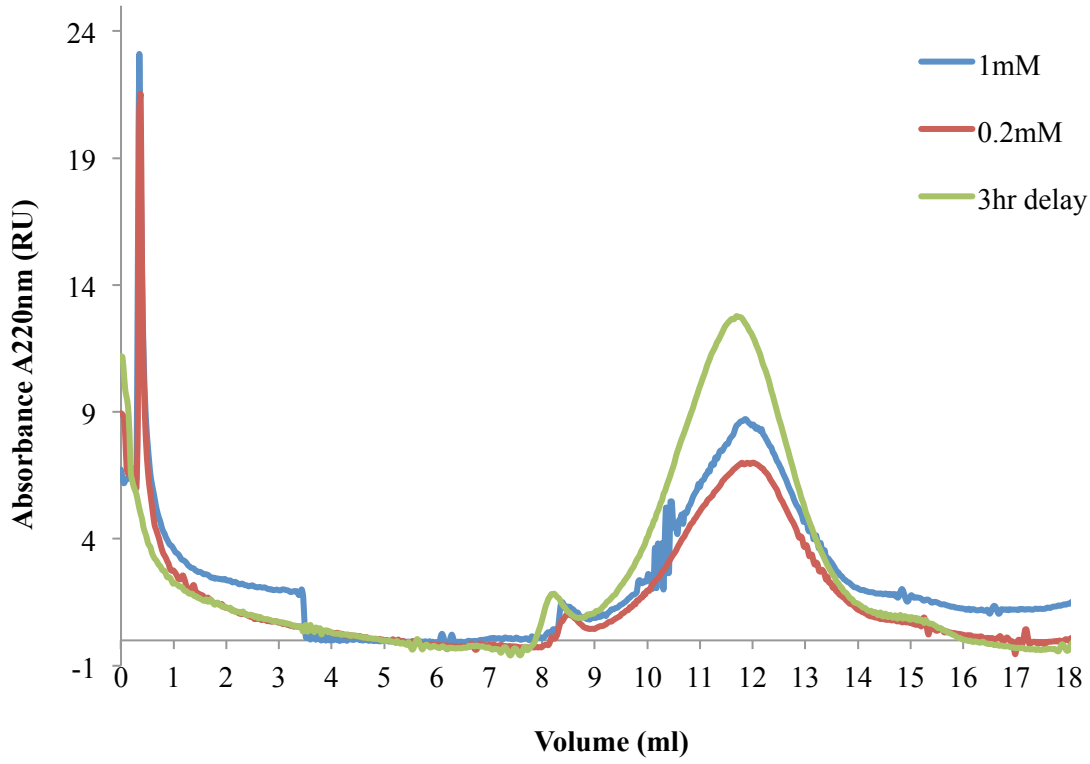


Figure 3.6: Gel filtration curve for TatAdCd_S complex expressed under different conditions

TatAdCd_S strain was grown at 37°C and expressed either immediately with either 0.2mM (red line) or 1mM (blue line) of arabinose. In addition, during the exponential phase (3 h after inoculation) the cell were induced with 1mM arabinose (green line). The peak elutions from Figure 3.5 were concentrated to 500 µl and 250 µl was applied to the superdex200 column. The size of the TatAdCd_S complexes is not altered when the arabinose concentration is altered; only the concentration of protein expressed and purified is reduced. However when the induction is delayed to the exponential phase, the complex size is increased by ~50 kDa.

Upon collection of the peak fractions, the eluents were run on a silver-stained SDS PAGE gel, which confirmed the presence of the TatAdCd_S protein and validated that the peak is attributed to the TatAdCd_S complex. The TatAd protein is 7 kDa and TatCd is 25 kDa. Unfortunately with a protein as small as 7 kDa, the resolution of the band is compromised, as the ratio of TatAd is relatively low, compared to TatCd in the TatAdCd_S complex. For this reason the TatAd is not visible in Figure 3.7 and therefore it cannot be concluded that the peak observed is of the entire TatAdCd_S complex. However for this purpose, it can be concluded that the elution pattern for the TatAdCd_S complex is largely similar in all three conditions. In particular, when the concentration of arabinose is changed from 0.2 mM to 1 mM the elution pattern is identical with the majority of the complexes eluting within 9-11ml.

The broad elution range suggests that there may be multiple sizes of complexes forming. What is also clear from the western blot but skewed by the gel filtration peak, is the elution pattern for delayed induction, is similar to the elution pattern in other conditions. It can be concluded that after purification the yield of TatAdCd_S can be increased by delaying the induction stage even after the cells reach the stationary phase. In addition, the western blot (Figure 3.7) also confirms that increasing expression levels does not alter the size of complexes formed, as portrayed by the gel filtration graph in Figure 3.6.

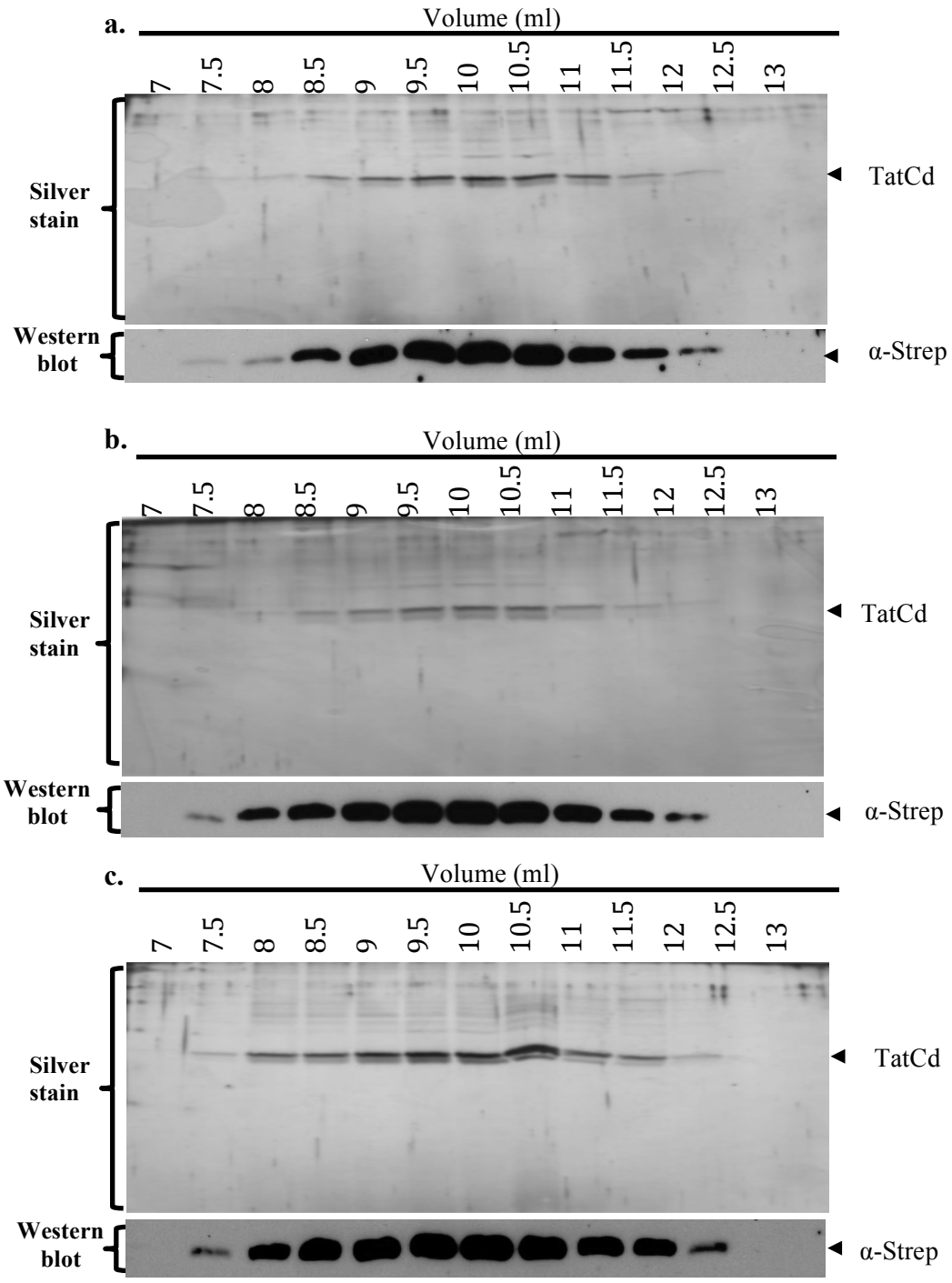


Figure 3.7: Silver stained gel and Western blot of the peak gel filtration fractions of TatAdCd_S

The peak fractions containing TatAdCd_S after gel filtration were analysed by SDS-PAGE and visualised with Silver stain. The elution for the variable condition with 0.2 mM arabinose at 0 h induction (a.), 1 mM arabinose at 0 h induction (b.) and 3 h induction (c.) is shown. Western blot was also performed using α-Strep to bind the Strep-tag on TatCd with high proportion of TatAdCd_S eluting in 9-11 ml for all three conditions.

3.5 Optimisation of purification by varying detergent concentration

The TatAdCd translocon is composed of membrane proteins and must be extracted from the membrane using a detergent. Detergents acts to destabilises the membrane by disrupting the lipid bilayer and extracting the complex. As Tat forms heterologous complexes, it is important that the concentration of the detergent is sufficient to completely isolate the complex from the membrane without interrupting complex assembly. In addition, where biophysical studies such as Electron Microscopy (EM) and Surface Plasmon Resonance (SPR) are involved, the presence of detergent can increase background noise extensively. The aim of the following experiments was therefore to optimise the detergent concentration to improve the quality of data obtained from the biophysical studies, which are carried out in the following chapters.

The detergent used for the purification of Tat complexes is C12E9, a non-ionic detergent with a low Critical Micelle Concentration (CMC). This detergent is commonly used in membrane solubilisation, as it forms larger micelles to extract larger complexes such as Tat. In addition, C12E9 also proves economical when the scale of purification is increased for higher yield. The CMC of a detergent is the minimum concentration of detergent required to form a micelle, and for C12E9 it is 0.003%. The previous protocol used 1% C12E9 for the initial solubilisation of the membrane and 0.1% C12E9 was supplemented all the purification buffers to maintain the stability of the micelles. In the solubilisation process, 1% C12E9 was used in excess to completely solubilise the membrane. With the aim to minimise the detergent level at the purification stage, a series of purification trials using a range of lower detergent concentrations were carried out for the purification of TatAdCd_S and then characterised by EM.

3.5.1 Purification of TatAdCd_S using varied concentration of C12E9

A 500 ml culture expressing TatAdCd_S in *E. coli* was grown to an OD_{600} of 0.4-0.6 and then induced with 1mM arabinose for 3 h before harvesting by centrifugation. The cells were then fractionated as described in section 2.2.3 and the membrane fraction was then solubilised in 1% C12E9 (>300X CMC) overnight.

After solubilisation, the Tat complex was purified by 3-step purification; ion exchange chromatography, affinity chromatography and gel filtration. For ion exchange, the negative charge of the Tat complex is used to bind the protein to the Q-Sepharose resin. During this process the C12E9 concentration was reduced to 0.5% in the running buffers. Q-Sepharose purification is a crude method of purification to separate protein based on the charge of the protein. The elutions containing the TatAdCd_S complex were identified using Bradford's reagents to measure the protein concentration and hence where Tat complex is eluted.

The eluted proteins were incubated with avidin before it was applied to the StrepTactin column. Avidin can bind other biotinylated proteins that could compete to bind the column and hence reduce non-specific binding. The Strep-tag on TatCd_S binds to the column, with any TatAd that may be associated in form of a complex. The protein was then eluted by addition of desthiobiotin, which exhibit higher affinity for the StrepTactin resin. The buffers used for affinity chromatography were supplemented with 0.05% C12E9. From the point of solubilisation to purification to the end of affinity chromatography, the detergent concentration of C12E9 was gradually reduced down to from 1% to 0.05% C12E9 (16X CMC). The final polishing purification step was used to test the effect of reducing detergent concentrations with varied concentration of C12E9 ranging from 0.05% (~16X CMC) to 0.0005% (~6X below CMC).

3.5.1.1 Purification: Ion Exchange Chromatography-Q shepharose and StrepTactin

TatAdCd_S was purified from the solubilised membrane fraction, by an initial crude method using Q-Sepharose. The Tat protein complex has an overall negative charge and this is used to purify the complex from other proteins that may be positively charged. As shown in Figure 3.8, from the elution from the Q-Sepharose column (QE3) there is purification of both the TatAd and TatCd_S proteins, however, from a crude purification, it is difficult to say whether these proteins are still forming complexes. This is because the proteins may have similar charge, which results in co-elution. The high salt concentration during elution may also interfere with protein-protein interactions, which may prevent the complexes from forming.

The QE3 fraction was then applied to the StrepTactin column, which would bind the Strep-tag on TatCd_S and any TatAd that may be associated in the complex (Figure 3.8). As expected, the TatCd_S purified from the StrepTactin column (SE4-SE8) co-elutes with TatAd, confirming the presence a TatAdCd_S complex. The Silver-stain gel shown above suggests the complex is largely composed of TatCd_S with very small amount of TatAd associated with it, as the band for TatAd is hardly visible on the gel at 7kDa. The lack of band is likely to be a result of its low molecular weight but the west blot confirms co-elution of TatAd. In addition, it should also be noted that the total TatAd amount is higher in the starting membrane fraction and during the Q-Sepharose purification step, large amount of this TatAd is dissociated from the complex or is not forming and complex with TatCd_S. With these purification steps, we confirm that the TatAdCd_S complex is formed and maintained when purified with C12E9. To continue to the final stage of purification, the peak elutions (SE5-8) were concentrated to 0.5 ml and then applied to a gel-filtration column to test the effect of detergent concentration on the complex assembly.

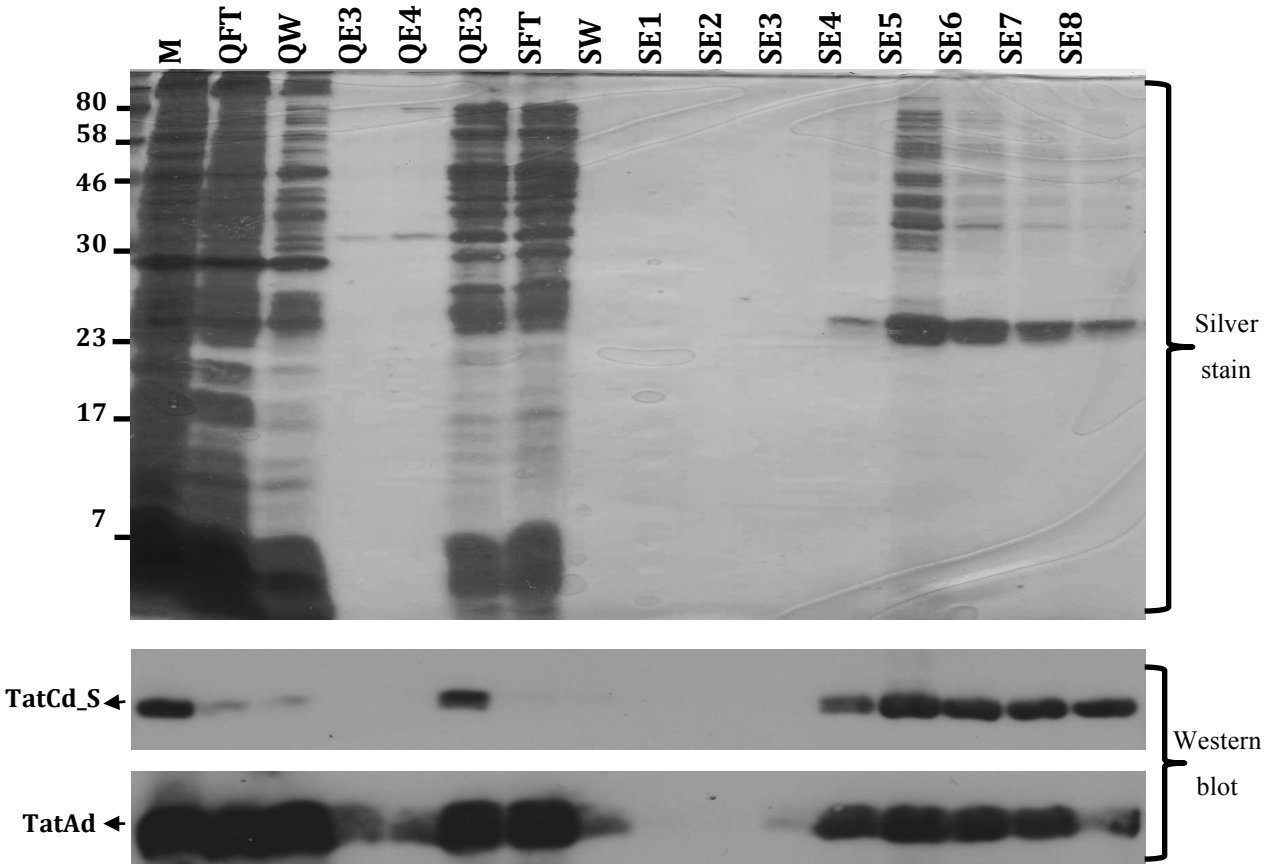


Figure 3.8: Silver Stained Gel and Western Blots showing purification of TatAdCd_S on a Q-Sepharose and StrepTactin Column.

Purification of TatAdCd_S after solubilisation of the membrane is shown by the silver stain gel and western blots. The solubilised membrane was applied to a Q-Sepharose column (QFT1) and washed with 5 column volumes (QW) the protein was then eluted off the column with high salt buffer (QE1-3). QE3 was then applied to a StrepTactin column (SFT) and again washed with 5 column volumes (SW). The proteins were then eluted in 1ml fractions (SE1-8). TatAdCd_S are co-eluted from the StrepTactin column however some contaminants remain in elution. TatCd-Strep is resolved at 25 kDa and TatAd is resolved at 7 kDa. The low level of TatAd means it is not clearly visualised on the gel however a targeted antibody against TatAd confirms the presence of TatAd.

3.5.1.2 Purification: effect of reducing detergent using Gel filtration chromatography

Once the protein complex was purified and the peak elution fractions are concentrated and applied to a gel filtration column superdex200. The running containing was supplemented with different concentration of detergent to assess the stability of the protein with decreasing amount of C12E9. The concentration was reduced from ~16X CMC (0.05% C12E9 *v/v*) to 6X below CMC (0.0005% *v/v*). The gel filtration chromatography step used Superdex200 to resolve the proteins in the range of 600-10 kDa. Where proteins that are forming complexes greater than 600 kDa, pass straight through the column and as a result are eluted in the void volume (Vo) at 8.5 ml. The smaller protein/complexes endures a delayed journey through the column due to the porous beads that capsulate the smaller proteins.

The most unfavourable detergent condition is when the C12E9 concentration is dropped below its CMC (blue line, Figure 3.9), which resulted in the TatAdCd_S to elute in the void volume. This suggests that TatAdCd_S aggregates when the concentration of C12E9 is dropped below its CMC (0.0005% v/v) preventing the formation of protein-detergent complexes. With the lack of C12E9 in conditions below the CMC, the hydrophobic transmembrane regions of the Tat protein are exposed to solution leading to an unordered aggregation of the proteins/complex. When the running buffer was supplemented with detergent at CMC (red line, Figure 3.9) a double peak was observed; one at the void volume and other at 9.7 ml. This suggests roughly half of the protein complexes are still aggregating and producing a peak in the void volume, however the other half seems to form large TatAdCd_S micelles, with an estimated molecular weight of 600 kDa. As shown in Figure 3.9, not all protein complexes are incorporated into the micelle, confirming that protein complexes are not stable even at CMC.

The next gel filtration run was carried out with 3X CMC C12E9 (0.01% *v/v*) and with this the peak of protein complex elution was shifted to 10.9 ml, which is calculated to be 370 kDa from the standard markers. These results are contradictory to the result presented from previous gel filtration estimates showed complex size in the range of 230 kDa (Barnett *et al.*, 2008). The larger size of the complexes may

suggest larger complexes forming when detergent is kept minimal but without structural analysis, it is difficult to conclude whether these larger complexes are in fact TatAdCd_S complexes or an outcome of detergent aggregation.

It appears that once the detergent is in excess of ~8X CMC (0.025% *v/v*) C12E9, the protein forms complexes of 235 kDa, as those previously observed with the detergent Digitonin. It can also be noted that the absolute minimum concentration of C12E9 for the purification of TatAdCd_S (~50 μg/ml) is 0.025% C12E9. The concentration of detergent used for higher concentration of protein complexes would therefore need to be adjusted equivalently. In addition, increasing the detergent concentration from 8X CMC to 12X and then 16X, no further significant effect on complex size is observed, as their average size of complex remains at 235 kDa (Figure 3.9). The evidence from this suggests the complexes are stable under these conditions. The standard protocol uses ~33X CMC, which is approximately double the amount of the highest condition tested in this study. Although at this stage it cannot be concluded with certainty, it appears from the graph that TatAdCd_S complexes are formed in the range of 8-16X CMC of detergent. To confirm this, further structural characterisation must be performed.

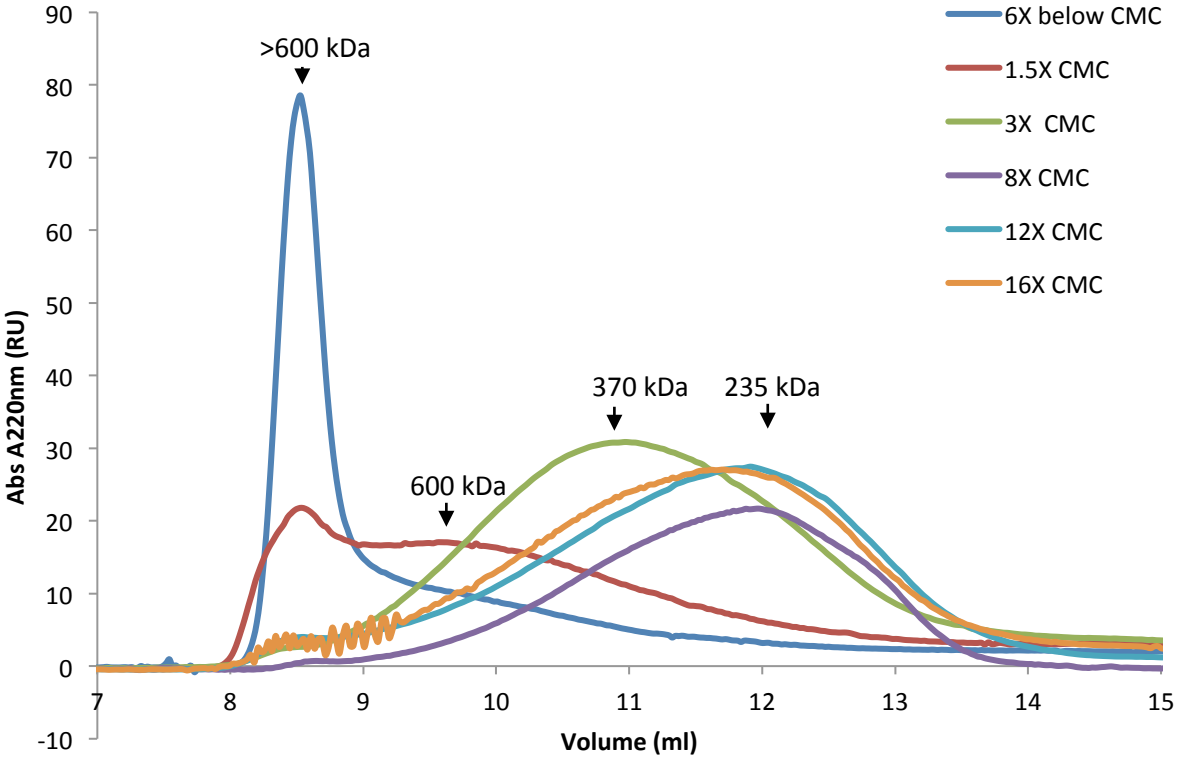


Figure 3.9: Gel-filtration of TatAdCd_S complexes purified using varied detergent conditions

The TatAdCd_S complexes were expressed in *E. coli* and isolated from the membrane using 1% C12E9 detergent. During the next purification steps the detergent concentration was varied in term of the CMC of C12E9 (0.003%). The conditions tested were: below CMC (blue line), 1.5X CMC (red), 3X CMC (green), 8X CMC (purple), 12X CMC (cyan) and 16X CMC (orange). Elution of protein at 8.5ml is the cut-off of the column, which shows complexes greater than 600kDa. At 3X using the standard protocol produces complexes of 370 kDa and when the CMC is increased above 8X, complexes of 235 kDa are formed.

3.6 Structural analysis of varied detergent concentration

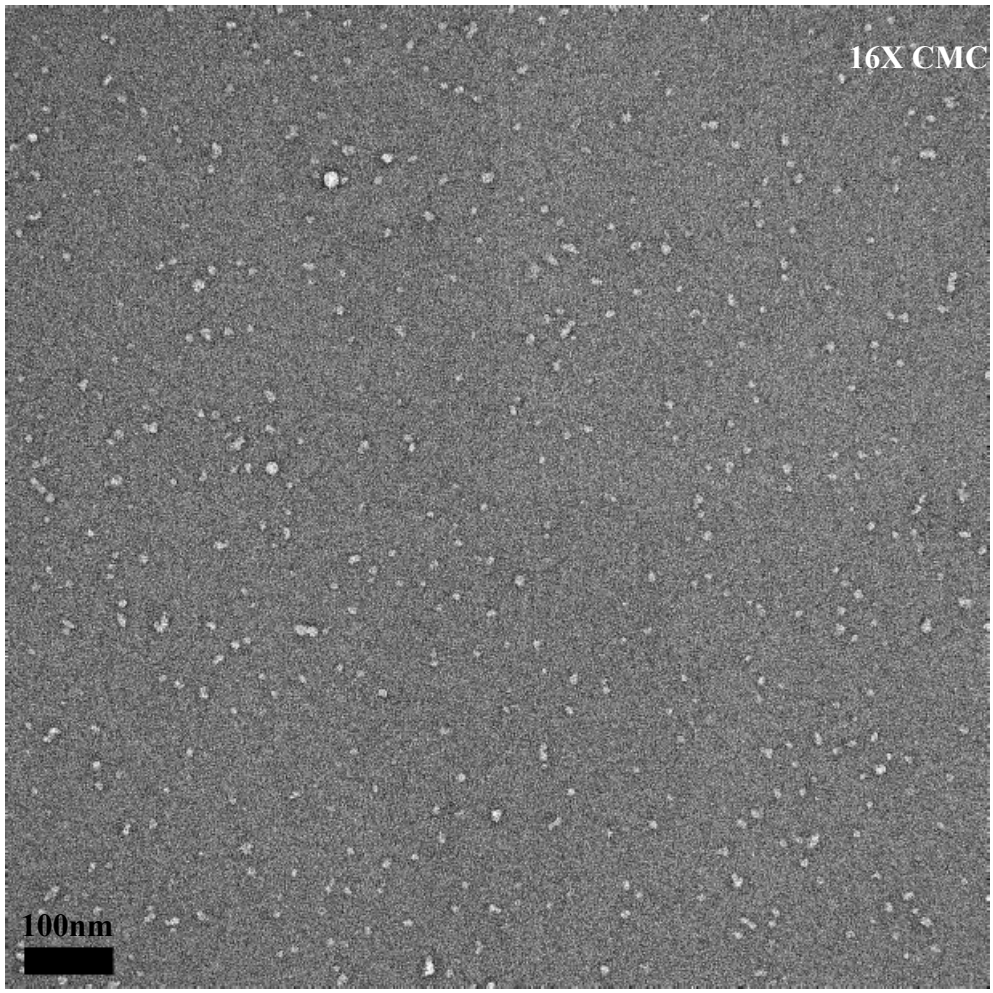
TatAdCd_S forms protein-detergent complexes of ~235kDa provided sufficient detergent is supplemented to maintain the stability. There were 3 detergent conditions that were identified which formed stable complexes; 8X CMC, 12X CMC and 16X CMC (Figure 3.9). As a secondary approach, it is important to characterise the structure of these complexes to evaluate the effect of minimising the detergent on the structure of the complex.

The peak elution fraction(s) containing the TatAdCd_S complexes were identified by running the peak fraction on a SDS-PAGE gel and run a western blot using TatAd (α-Ad) and TatCd (α-Strep) antibodies. The fractions containing the highest concentration of TatAdCd_S were observed by Transmission Electron Microscopy (TEM). TEM is an ideal method for assessing size distribution of particle and visualising the structural effect inflicted by detergent concentrations. Each peak fraction was roughly diluted to 10 μg/ml to obtain an even distribution of the complexes on the grid. The samples were applied to a carbon-coated grid and stained with the heavy metal uranyl acetate (2%) to give contrast against the carbon surface. The images were then captured using a CCD camera at 40000X magnification using the digital micrograph software.

TatAdCd_S complexes of ~235kDa form fairly homogenous particles averaging around 9-11nm. With detergent concentration as high as 16X CMC (0.05% *v/v*), there are some large round complexes observed with a solid white colour. These are empty detergent micelle so have a uniform electron density. The detergent micelles are above 20 kDa, so these complexes can be easily differentiated from the TatAdCd_S protein-detergent complexes. In addition, any ambiguity in identification of the TatAdCd_S particles was settled by a density profile. The Tat complexes have been structurally characterised previously, elucidating a central cavity/pore (Beck *et al.*, 2013; Tarry *et al.*, 2009; Gohlke *et al.*, 2005). Therefore by using the density profile tool, particles exhibiting a pore-like feature can be identified. The size range of TatAdCd_S complexes with 16X CMC C12E9 is between 7-14 kDa with majority of the complexes observed at 9 nm (Figure 3.10).

As the concentration of C12E9 was reduced, larger complexes of detergent-protein complexes were observed. The size distribution of TatAdCd ranges more toward the larger complex size. This is observed with 8X CMC, where size distribution becomes broader in range of 7-14nm (Figure 3.12). These large structures were a result of aggregation as there was no order observed in the assembly. In addition, the larger structures are not of a consistent size and shape, suggesting they form random aggregates when the detergent concentration is reduced. This is observed with 3X CMC, with the size range of the complexes increased between 7-19nm (Figure 3.13). The complexes were most favourable with 0.035% C12E9 which corresponds to ~12X CMC where the complexes are most homogeneous (7-11nm) and no larger aggregates are present (Figure 3.11). At 1.5X CMC most TatAdCd_S aggregates due to lack of detergent, to stabilise the complexes (Figure 3.14). When the detergent concentration is dropped below CMC, even detergent micelles are unable to form so both the detergent molecules and TatAdCd_S cluster together in large structure up to 0.5μm in size (Figure 3.14). The overall effect of varying detergent concentration leads to aggregation of TatAdCd_S complexes, when there is a lack of detergent. More specifically, the C12E9 detergent is shown to be more compatible with the TatAdCd_S complexes, when the CMC is maintained above 8X CMC.

a.



b.

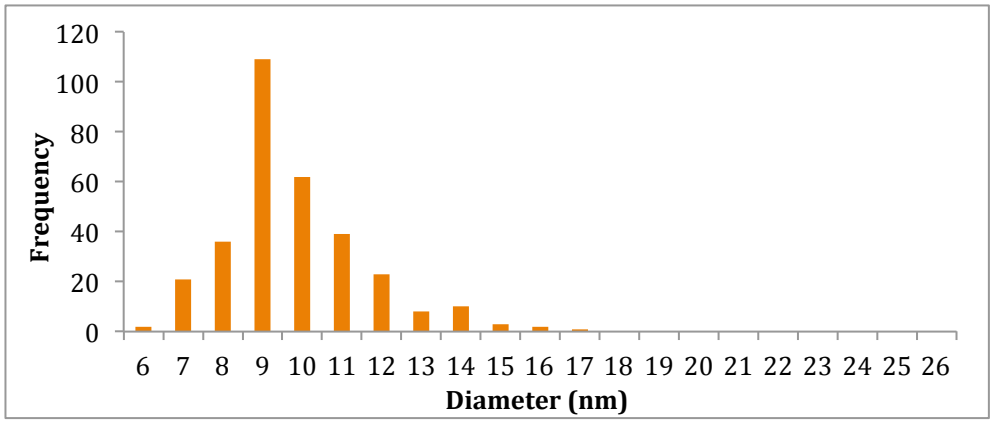
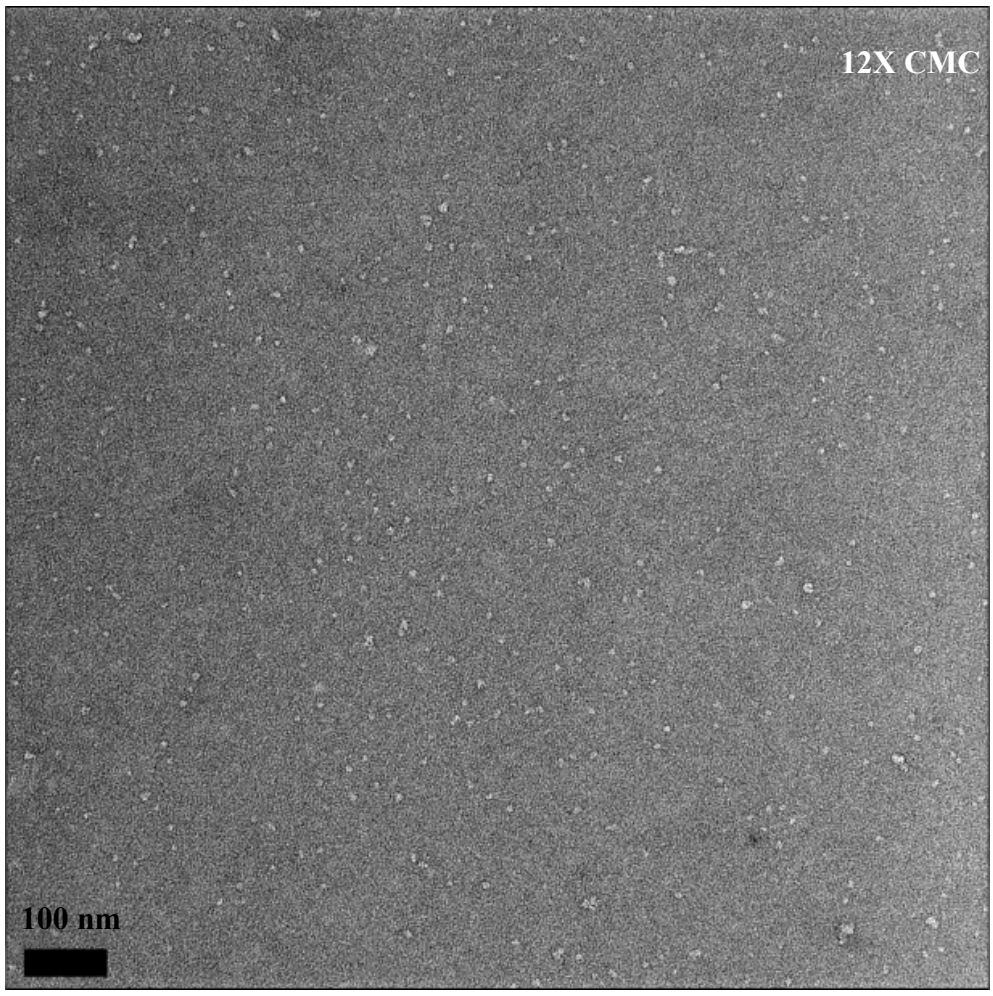


Figure 3.10: TEM micrograph of TatAdCd_S purified in 16X CMC with size distribution of the complexes

TatAdCd_S was purified in 0.05% C12E9 (*v/v*), (16X CMC) on the gel filtration column and the peak elution sample was negatively stained and visualised by TEM (a). The complexes from the micrograph were measure in diameter on ImageJ to show the sixe distribution of the complexes (b). The micrograph is taken at 40000X magnification with a scale-bar of 100nm.

a.



b.

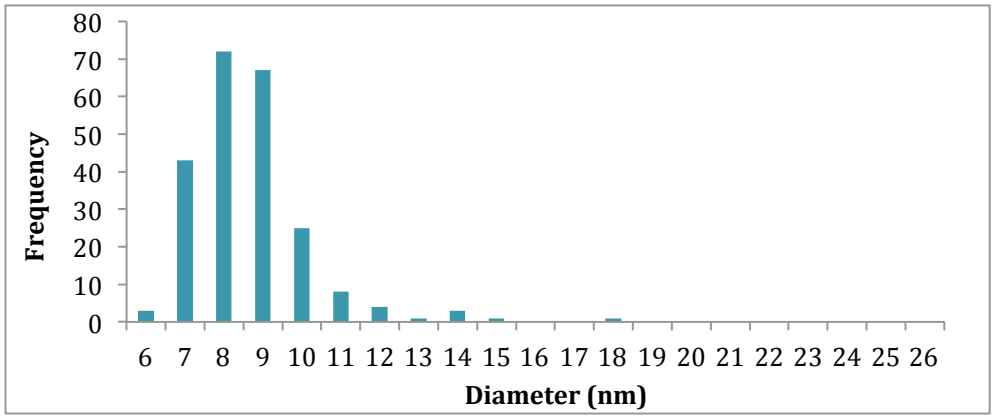
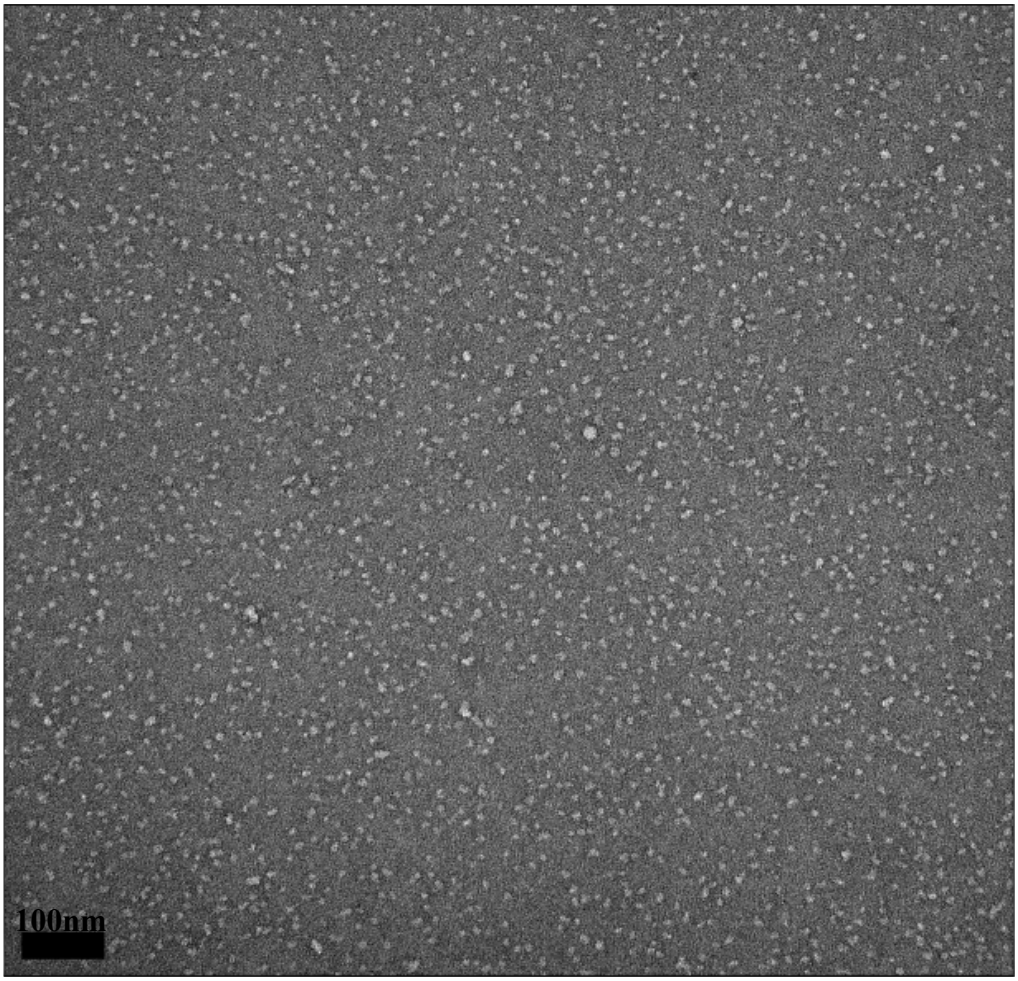


Figure 3.11: TEM micrograph of TatAdCd_S purified in 12X CMC with size distribution of the complexes

TatAdCd_S was purified in 0.035% C12E9 (*v/v*), (12X CMC) on the gel filtration column and the peak elution sample was negatively stained and visualised by TEM (a). The complexes from the micrograph were measure in diameter on ImageJ to show the sixe distribution of the complexes (b). The micrograph is taken at 40000X magnification with a scale-bar of 100nm.

a.



a.

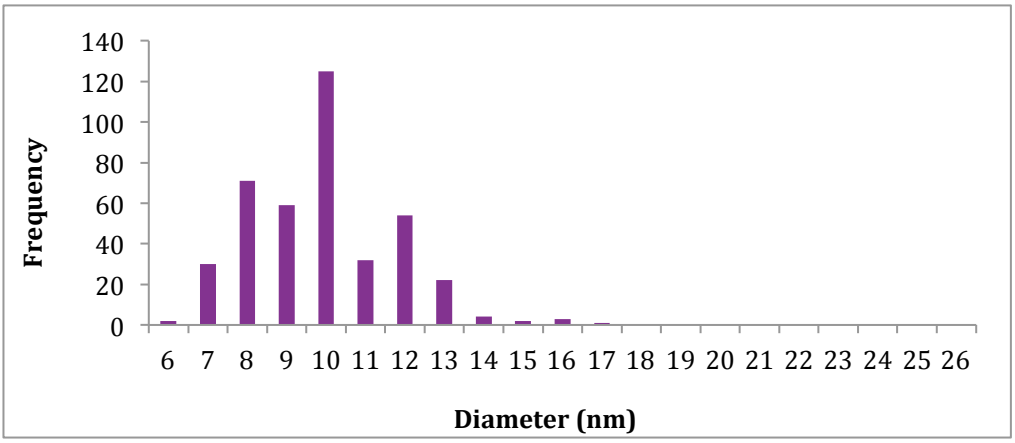
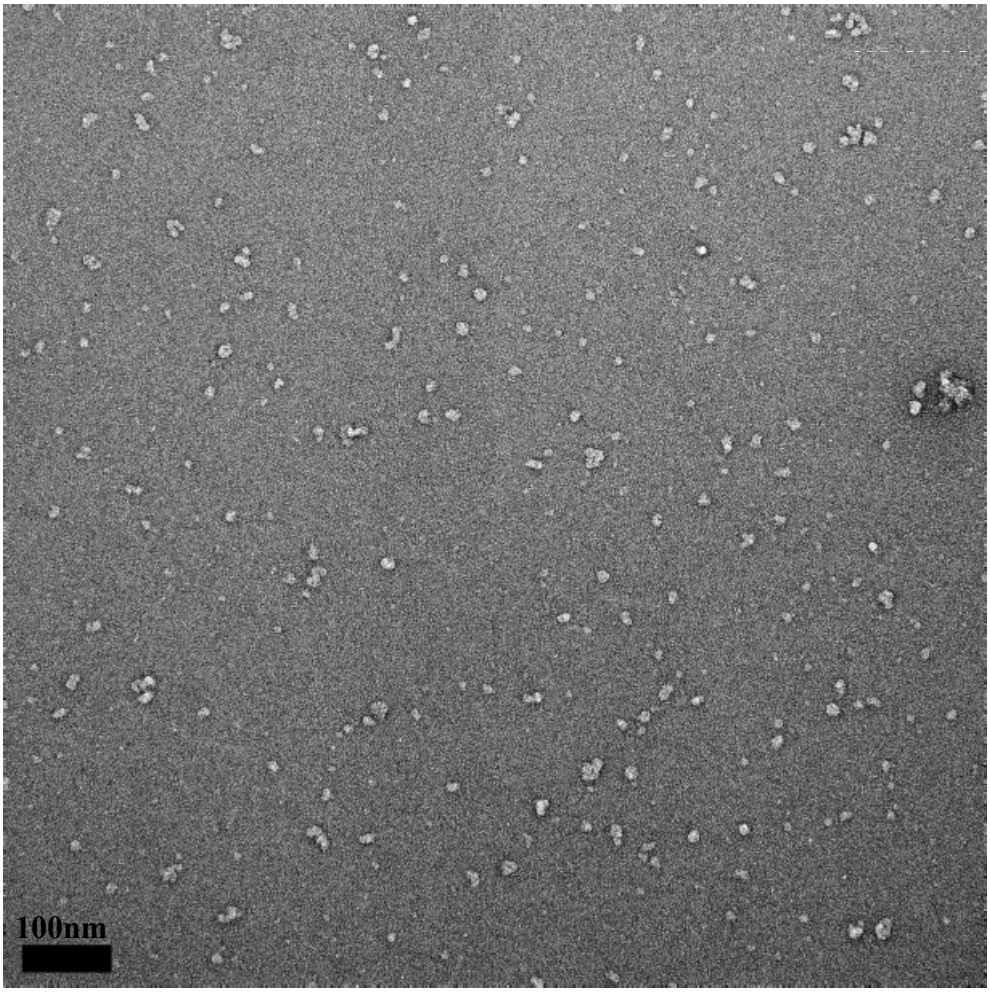


Figure 3.12: TEM micrograph of TatAdCd_S purified in 8X CMC with size distribution of the complexes

TatAdCd_S was purified in 0.025% C12E9 (*v/v*), (8X CMC) on the gel filtration column and the peak elution sample was negatively stained and visualised by TEM (a). The complexes from the micrograph were measure in diameter on ImageJ to show the sixe distribution of the complexes (b). The micrograph is taken at 40000X magnification with a scalebar of 100nm.

a.



b.

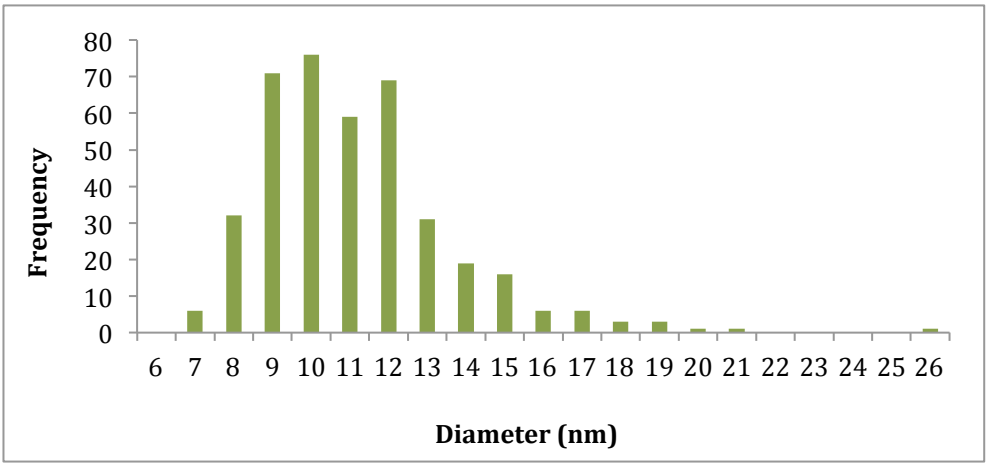


Figure 3.13: TEM micrograph of TatAdCd_S purified in 16X CMC with size distribution of the complexes

TatAdCd_S was purified in 0.01% C12E9 (*v/v*), (3X CMC) on the gel filtration column and the peak elution sample was negatively stained and visualised by TEM (a). The complexes from the micrograph were measure in diameter on ImageJ to show the sixe distribution of the complexes (b). The micrograph is taken at 40000X magnification with a scale-bar of 100nm.

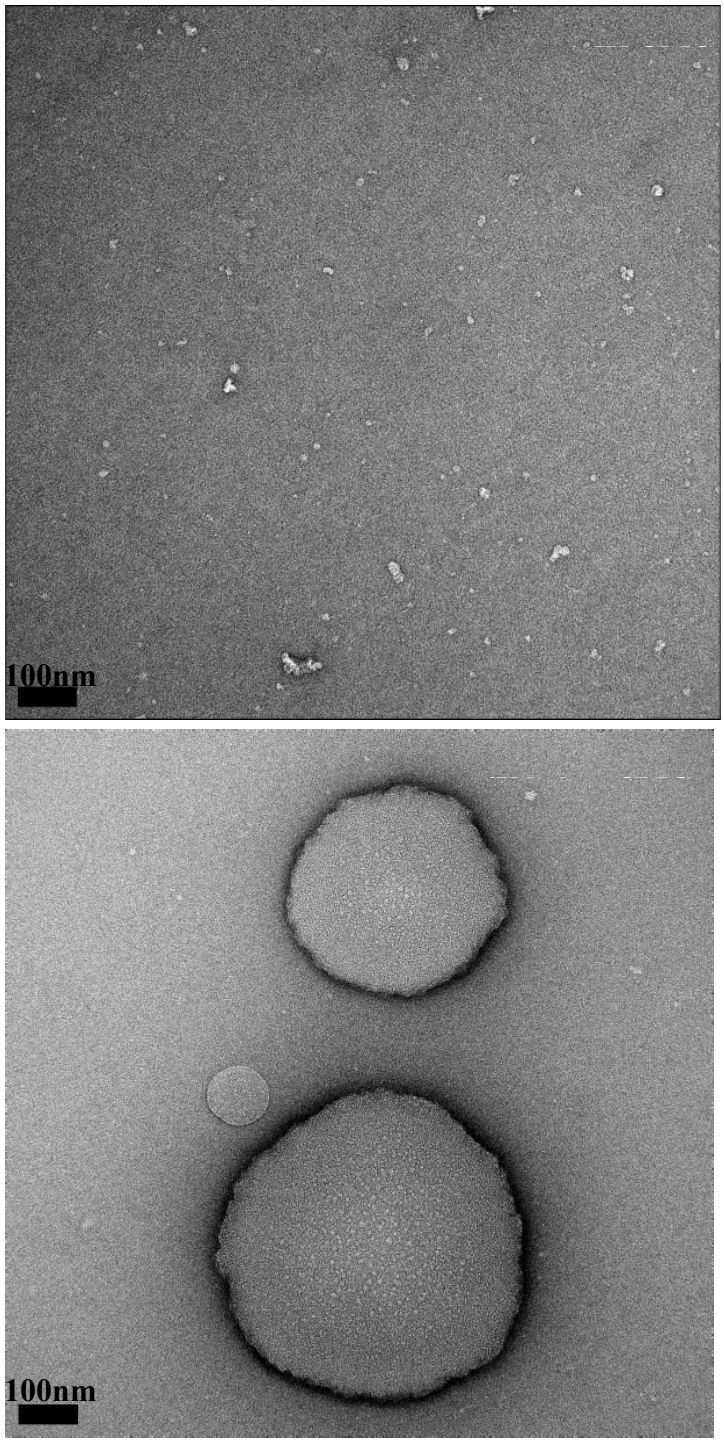


Figure 3.14: Negatively stained EM images of TatAdCd_S with reducing concentration of C12E9

TatAdCd_S complexes were purified using ion exchange, affinity chromatography and gel filtration chromatography. The peak elution fraction was used for visualising complexes by electron microscopy. The images were taken at 40000X magnification after gridding onto a carbon mesh grid and stained with 2% uranyl acetate. With decreasing C12E9 concentration (1.5X CMC and 6X below CMC) there is an increase in formation of aggregates of the TatAdCd_S complexes.

3.7 Discussion

The *B. subtilis* TatAdCd complex is similar to that of *E. coli* TatABC complex, yet the mechanism remains ambiguous. Although some mutational studies have provided some insight into the function of the mechanism, the structural characterisation of the TatAC complexes is key to understand the mechanism in detail (Barnett *et al.*, 2009; van der Ploeg *et al.*, 2011). This chapter begins to identify the optimal condition for the structural characterisation of the TatAdCd complexes using transmission electron microscopy (TEM). Earlier research has focused on the understanding of TatABC complex, as it is native within the model Gram-negative bacterium *E. coli*. However, the heterogeneity in TatABC has made hypothesising difficult due to the variability in results. It is hoped that understanding a simpler and more a specific system, such as TatAdCd, can find clearer answers.

This chapter has focused on the expression and purification of the TatAdCd_S complexes, in conditions suitable for further structural studies. TatAdCd_S was expressed in *E. coli* and results show that heterologous expression applies cell stress prior to induction (3.3.1). It was discovered that the protein was required to be expressed during the exponential phase of the cell growth, rather than immediate induction as the standard protocol suggests. This may be a result of gradual manipulation of the MC4100 strain, which has caused the cells to become less robust.

Cell growth of the MC4100 *E. coli* (WT-Tat) strain showed cell growth lasts up to 2-3 h before it reached stationary phase, as with TatAdCd_S (Δ-Tat) strain. Therefore the standard protocol for expressing either TatABC_S or TatAdCd_S used immediate induction to allow a minimum of 3 h for the protein to be expressed during the exponential phase. The disadvantage of this protocol was the general low yield of protein produced.

This chapter has used various time and concentration dependent inductions to show the protein yield can be increased, by inducing at a later stage of cell growth (late-exponential or stationary). The significant difference in protein expression was observed when cells were induced after 3 h growth with the optical density reaching

0.6. The TatAdCd_S complexes were purified as normal and despite late induction there were no aggregation observed. The concentration of arabinose did not affect the growth of the cells, thereby demonstrating that it is not toxic to the cell and higher concentrations can be used in the future. The EM images of TatAdCd_S complexes expressed at 3 h induction, show spherical complexes similar to those previously observed in TatABC complexes (Barrett & Robinson, 2005). These images confirm no structural defects at low resolution as a result of expressing TatAdCd_S in the stationary phase.

Once a high yield protocol was established, the purification of these complexes with optimal concentration of detergent was to be determined. C12E9 is a non-ionic detergent with CMC of 0.003% and, was found to be ideal in the purification of TatA (de Leeuw *et al.*, 2002). The ideal concentration for the purification of TatAdCd_S complexes was between 8X-16X CMC during the gel filtration step. Complexes isolated in these conditions were estimated to be of 235 kDa, which is similar to the 230 kDa previously observed (Barnett *et al.*, 2008).

Structural characterisation of the complexes by EM begins to show that with increasing CMC, there may be a risk of introducing detergent aggregates. Additionally, by unnecessarily increasing the detergent concentration, it proves more difficult to differentiate between detergent artefact and the structure of the complexes. This is ability of differentiate between detergent artefact is very significant for the single particle analysis. It can be concluded that the optimal C12E9 concentration for protein purification would be 10X CMC (0.03%). The purification of the TatAdCd_S complex from a 500 ml culture, produced on average concentration of 0.05 mg/ml of TatAdCd_S in the peak elution fraction. Therefore the ideal protein: detergent ratio for TatAdCd_S: C12E9 is calculated to 1.5.

The structural analysis of TatAdCd_S obtained from the 8X CMC purification was used to study the TatAdCd_S complexes in detail. The size range of the TatAdCd_S complexes ranged from 7-13 nm. This TatBC complex in *E. coli* had an average size range of 9-11nm, which is comparable to the size observed for TatAdCd_S (Tarry *et al.*, 2009).

The density profile of the TatAdCd_S complex displays a low density in the centre of the complex and this implies one of two things. Firstly, the TatAdCd_S complex has a small cavity in the centre of the complex; this may not be large enough to accommodate the substrate, similar to that seen with TatBC (Tarry *et al.*, 2009). Current characterisation of TatC structure shows this cavity may be the site of signal peptide insertion (Rollauer *et al.*, 2012; Frobel *et al.*, 2012; Ramasamy *et al.*, 2013). The second possibility is that the presence of TatAd may form a premature pore that expands when further TatAd is recruited. This hypothesis is based on the topology of TatAd that assembles into a pore-like structure (Gohlke *et al.*, 2005). Structural characterisation of TatBC suggests the former to be the likely event (Tarry *et al.*, 2009).

The structural analysis of the TatAdCd_S complexes resembles the TatABC complex, with regards to the size and shape of the complexes. This may be a weak evidence for the possibility of the TatAdCd system sharing a similarity to the hypothesised TatABC mechanism. To further assess this hypothesis, a detailed analysis using single particle reconstruction would be more likely to show clearer differences. If the mechanism between TatAdCd and TatABC is similar then bifunctional TatA protein is likely to be a more evolved system using a similar mechanism.

Chapter 4

A mutation in TatAy leads to excessive complex assembly

4 TatAyCy mutation leads to excessive recruitment of TatAy complexes

4.1 Introduction

The Tat machinery has been heavily studied in the model organism in *Escherichia coli* (*E. coli*), but the identification of a simpler, 2-component TatAC system within Gram-positive bacteria has now become the focus within the Tat field. In *Bacillus subtilis* (*B. subtilis*), there are two Tat systems co-existing; the TatAdCd and the TatAyCy, that differ by the substrates they translocate. The reason for the presence two Tat system homologs in *B. subtilis* is not clearly understood but it seems that the specificity for the substrate is more stringent compared that to *E. coli*. TatAyCy is responsible for translocation of YwbN (now named EfeB), while TatAdCd translocates PhoD (Jongbloed *et al.*, 2004; Jongbloed *et al.*, 2000; Goosens, 2013). It was shown that when the endogenous TatABC system of *E. coli* was substituted by the TatAdCd system, the specificity of the system becomes loosely regulated. Therefore allowing other Tat substrates such as AmiA-GFP, DmsA-GFP and MdoD-GFP to be transported (Barnett *et al.*, 2008).

TatAdCd has now become a popular system to study due to its simplicity, compared to the TatABC system in *E. coli*. The TatAd protein shares ~20% sequence homology to its ortholog TatA, with conservation of residues within the cytoplasmic loop (Robinson *et al.*, 2011). The homologous TatAyCy complex has not been heavily characterised to date. Although previous export assays were able to identify the effect of mutation in the N-terminus and amphipathic helix on the translocation ability, the functional significance the residues remain unresolved (Barnett *et al.*, 2009; van der Ploeg *et al.*, 2011). It was clear from the export assays that many residues within the periplasmic N-terminus region of TatAd were crucial for the function as depicted in Figure 4.1 (Barnett *et al.*, 2009). A complementary study for TatAyCy found a similar occurrence, where mutation in the periplasmic N-terminus of TatAy blocked Tat function (van der Ploeg *et al.*, 2011). The N-terminus domain

in TatAy is shorter than TatAd with only one significant residue identified in this region.

van der Ploeg *et al.* (2011) showed the proline-2 residue in the N-terminus of TatAy has a significant role in the translocation. This is particularly interesting because it is part of a small N-terminus (5 residue) causing a significant effect on the function. In addition, the residue is not conserved with other TatA homologs. This chapter aims to explore the structure of the $TatAy^{P2A}$ mutant in the assembly of the $TatAy^{P2A}Cy$ complex and understand its functional role.

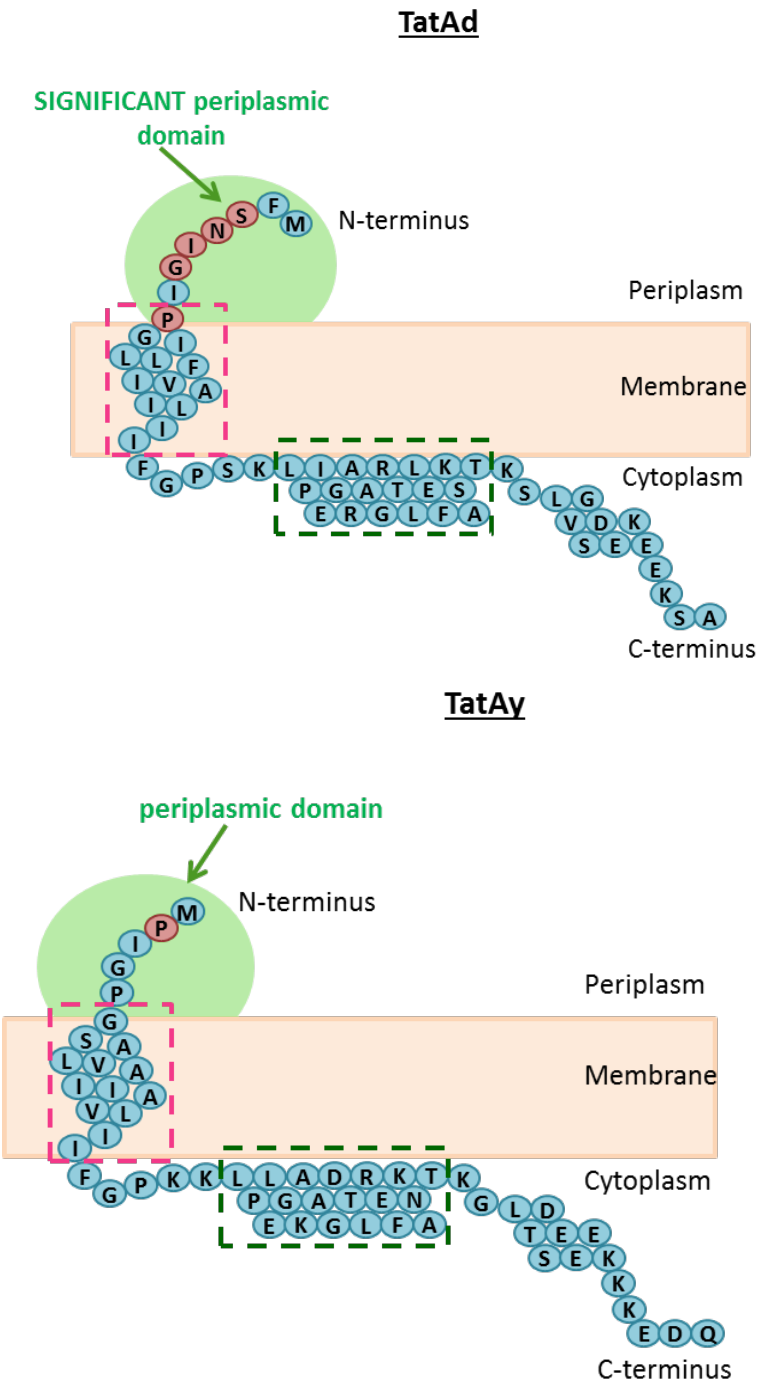


Figure 4.1: Predictive topologies of the TatAd and TatAy proteins from *Bacillus subtilis*

Adapted from b(Barnett *et al.*, 2009) the predicted topologies of TatAd and TatAy is shown. The figure shows TatAy and TatAd share sequence homology with similar secondary structure. The significant residues identified in the periplasmic region are highlighted in red. The pink-dashed box represents the transmembrane helix and the green-dashed box on the cytoplasmic face represents the amphipathic helix. The residues shaded in red represent mutation that prevents translocation.

4.2 TatAy accumulates in the TatAyCy complex when mutated in the N-terminus

In this study, the effect of mutating the P2 residue to alanine on the functional and complex assembly of TatAyCy was examined. The TatAy^{P2A}Cy mutant was expressed within its native host, *Bacillus subtilis,* to investigate the effect on complex formation. Both, TatAyCy_H and TatAy^{P2A}Cy_H were previously engineered with a His-tag on the C-terminus of the TatCy (van der Ploeg *et al.*, 2011). This would allow the purification of TatCy with any TatAy that co-elutes with it. The TatAyCy_H and TatAy^{P2A}Cy_H constructs will be referred as TatAyCy^{WT} and TatAy^{P2A}Cy respectively, for the simplicity.

The isolated membranes expressing TatAyCy and P2A TatAyCy were provided by Carmine Monteferrante (University Medical Center Groningen).

Once TatAyCy^{WT} was expressed, the membrane fraction was isolated and solubilised in DDM (0.3%). From the SDS PAGE gel, stained with Coomassie blue, the membrane fraction showed no major difference in the ratio of TatAy and TatCy_H expression between WT and P2A suggesting that both are expressed equally (Figure 4.2). It is also clear that the majority of the TatAy is associating with the TatCy, and hence no dissociation of TatAy is observed in the washes (Figure 4.2).

The elutions of the TatAyCy^{WT} complex shows the TatCy_H (25kDa) elutes across fractions 7-12 on the western blots, whereas the TatAy (7kDa) only peaks across elution fractions 8-11 (Figure 4.2a). This ratio of TatCy and TatAy is as expected from a resting Tat translocase, where recruitment of TatAy is relatively low. During the elution of TatAy^{P2A}Cy complex, the relative ratio of TatAy is drastically increased and the elution pattern across both proteins is almost identical. This unusual behaviour of TatAy suggests either that the mutation leads to structural aggregation or an increased assembly of TatAy^{P2A}Cy.

The increase in TatAy^{P2A} association to TatCy raises a few questions about the structural organisation of complex in terms of stoichiometry and aggregation. in both, TatAyCy^{WT} and TatAy^{P2A}Cy the proteins are shown to co-elute showing complex assembly between the two proteins. However there is a significant

difference is the stoichiometry of these TatAyCy complexes, which is altered by the P2A mutation. In addition, an unexpected band of 26kDa is also present in the elution fractions 8-12 of both WT and P2A TatAyCy. The 26kDa band is not detected by the α-His antibody and this observation implies that either self-cleavage of the His-tag occurs or that there is another protein bound to the TatAyCy complex. From a previous study using a Strep-Tag this double band is absent when the His-tag is replaced therefore this result is likely to be due to His-tag cleavage (van der Ploeg *et al.*, 2011).

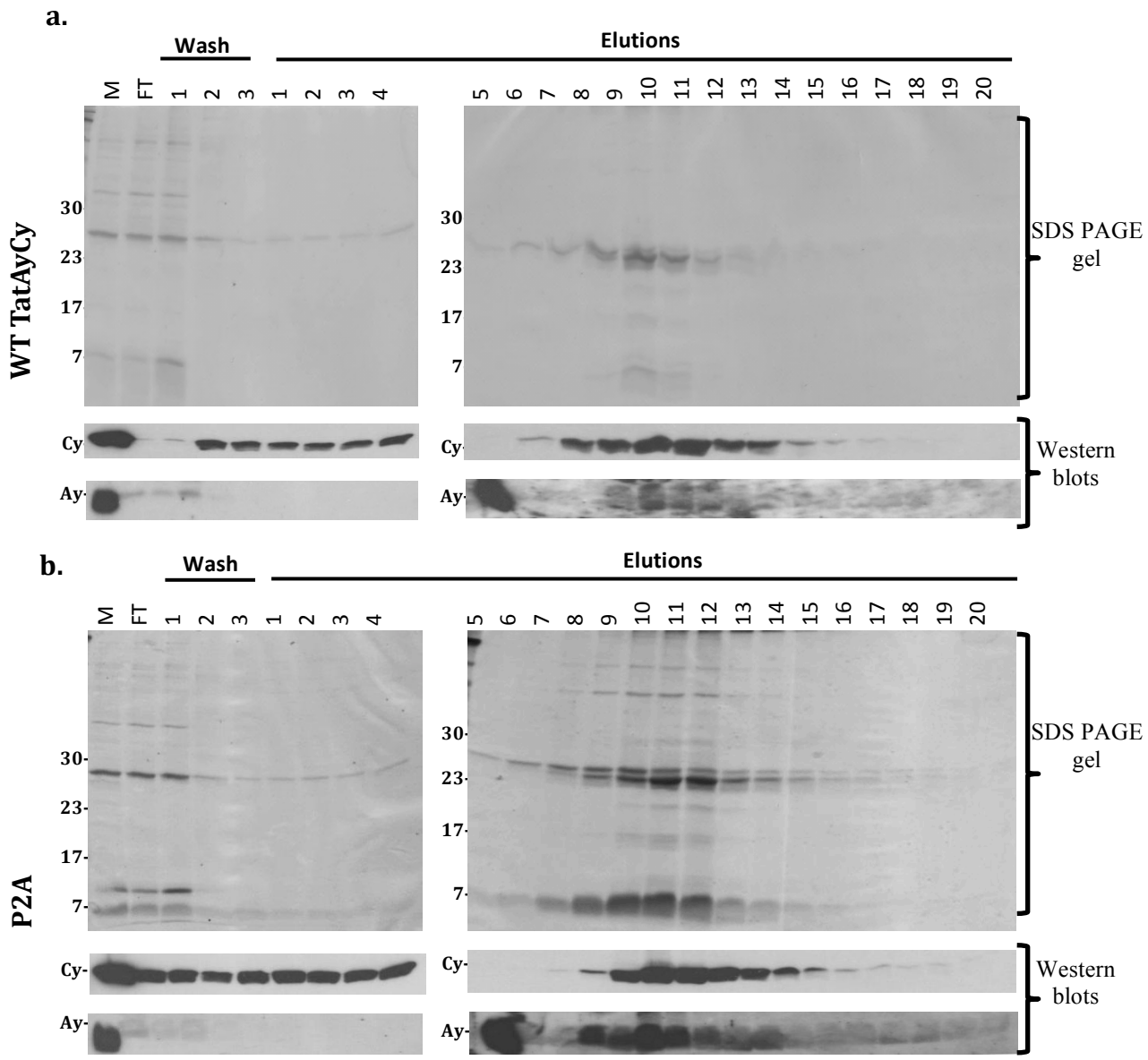


Figure 4.2: A Coomassie stained SDS-PAGE gel of TatAyCy^{WT} and TatAy^{P2A}Cy purification by Talon affinity chromatography.

Membranes were prepared from ΔTatAyCy *B. subtilis* expressing either TatAyCy^{WT} or TatAy^{P2A}Cy. The membrane was solubilised in DDM and applied to a Talon affinity column. The fractions containing whole membrane (M), flow-through (FT), wash and elutions are indicated on Coomassie stained gel of purified TatAyCy complexes. The corresponding western blots using antibody against the TatCy were performed with anti-His and against TatAy with anti-Ay (shown below the gel). TatAy (7kDa) co-elutes with the TatCy-his (25kDa) for both WT and P2A mutant.

4.3 TatAy^{P2A}Cy forms complexes larger than the TatAyCy^{WT} complexes

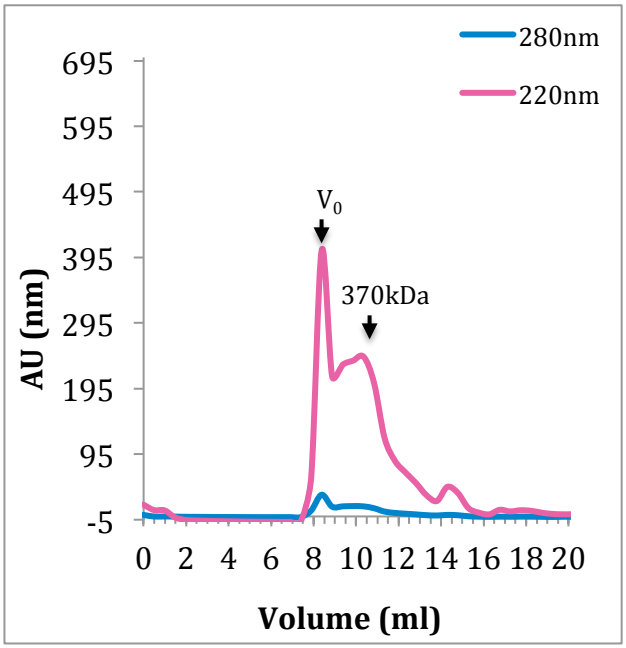
To investigate the effect of the P2A mutation on the size of these complexes, the peak elutions from the Talon-affinity column (described above) were concentrated and applied onto a Superdex200 gel filtration column. Using gel filtration chromatography, the size of the TatAyCy^{WT} complexes was estimated in range of 350-600 kDa. The peak of the elution is at approximately 370 kDa suggesting majority of the TatAyCy population forms complexes within this range. The broad range of elution across a large volume suggests that the TatAyCy^{WT} complex is heterogeneous in size. As with other Tat systems, the size variation is due to the dynamic stoichiometry of the complex. In comparison, a previous study estimated TatAyCy to form a 230kDa complex using gel filtration (Barnett *et al.*, 2009). However the TatAyCy complex presented in this study was isolated from *B. subtilis* as opposed to *E. coli*, which could cause the differences in complex assembly.

When the TatAy^{P2A}Cy mutant was run on the gel filtration column, a large peak was observed in the void volume (V_0). The V_0 is the maximum cut-off range of the column, where complexes greater than 600 kDa are deposited. From this result it can be assumed that the TatAy^{P2A}Cy is forming complexes $\geq$ 600 kDa (Figure 4.3b). Such result can be expected if the mutation is structurally significant or causing aggregation-type clustering. The SDS-PAGE gel from the gel filtration fractions shows that the TatAy and TatCy_H proteins are co-eluting in both TatAyCy^{WT} and the TatAy^{P2A}Cy. This observation confirms that TatAy and TatCy are forming a complex and are not affected by the mutation, although the complex size is altered. In addition, a smaller shoulder peak similar to that observed in TatAyCy^{WT} (Figure 4.3a) is also present at 10 mls. This may suggest some native size 370 kDa complexes are also present amongst the population of larger complexes.

The western blot showing the presence of TatCy and the TatAy confirms that the TatAyCy complexes eluted from the gel filtration column are in form of a complex. The purity of the TatAyCy complexes after gel filtration chromatography allows other contaminating proteins to be removed (Comparing Figure 4.2 to Figure 4.4). It should also be noted that the cleavage of the TatCy_H is still prominent in both

$TatAyCy^{WT}$ and $TatAy^{P2A}Cy$ complexes. The cleaving is observed as a double band for TatCy, showing that the lower band (24 kDa) elutes with a similar pattern to TatAy suggesting these proteins (24 kDa and 7 kDa) are likely to be co-eluting. The higher band at 26 kDa seem to elute as two separate peaks, first near the void volume around fractions 14-16 and then again around fraction 26-30 (Figure 4.4). Fractions 26-30 are observed as peak at 13-15 ml on the filtration graph (Figure 4.4). At this position on the gel filtration profile, a small peak is also observed. Also a western blot using anti-His does not pick up this higher band pattern suggesting that this higher band represents the TatCy without the His-tag.

a.



b.

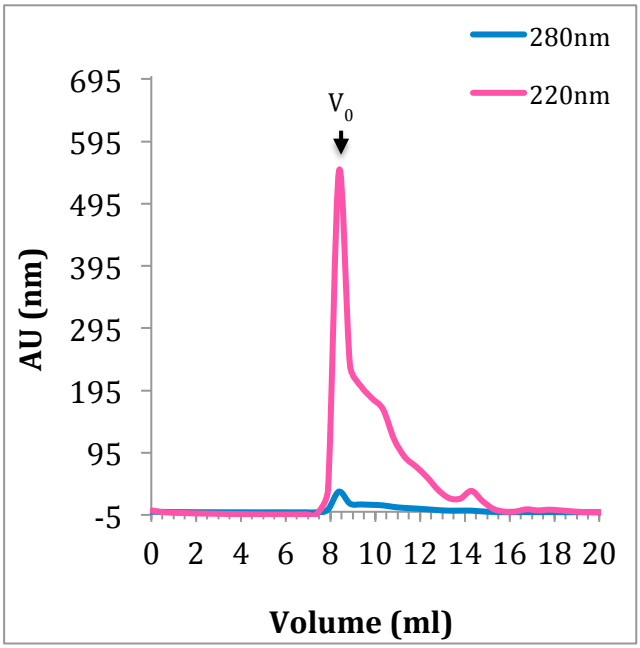


Figure 4.3: absorbance of protein elution for TatAyCy^{WT} and TatAy^{P2A}Cy resolved by a gel filtration chromatography

After concentration of the peak Talon elution the concentrated TatAyCy sample for TatAyCy^{WT} and TatAy^{P2A}Cy were resolved on a gel filtration column. The WT TatAyCy complexes resolved at 370kDa (a) and the P2A TatAyCy complexes resolved in Vo (b).

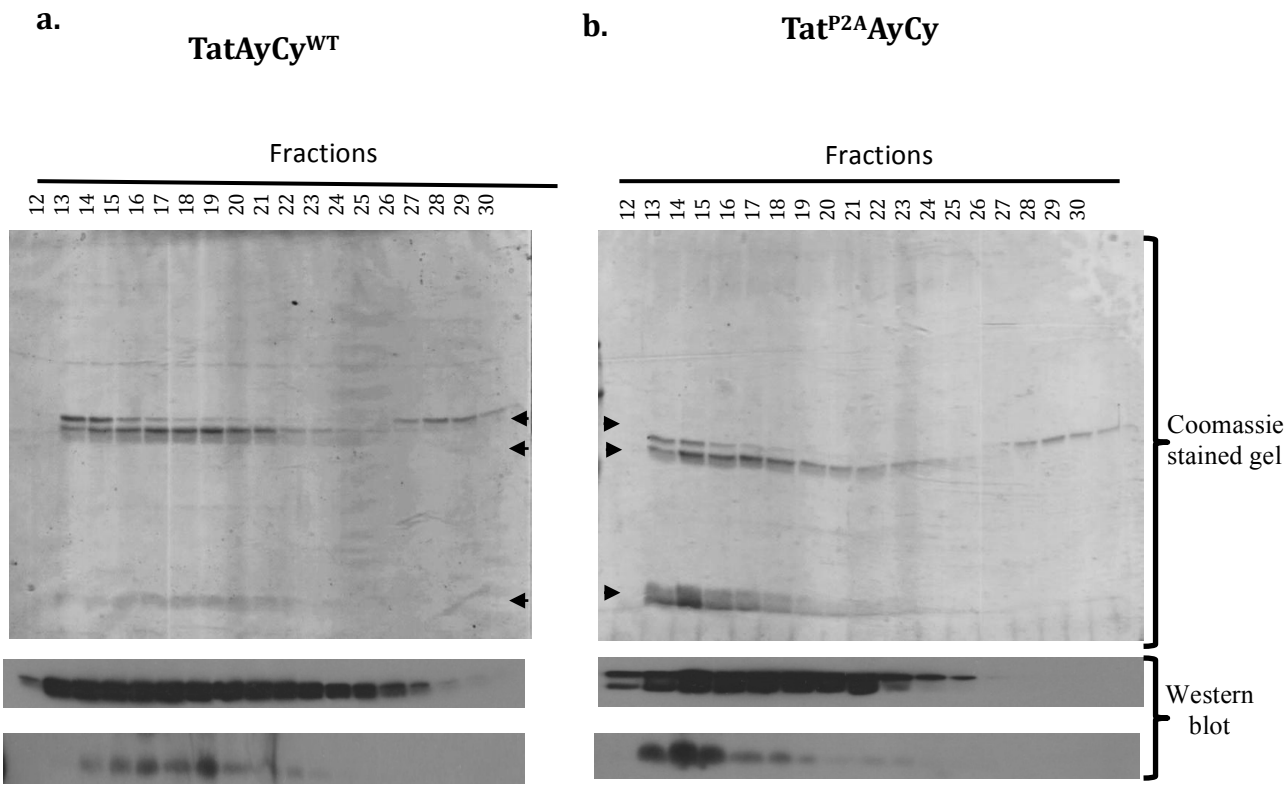


Figure 4.4: Elution fraction from the gel filtration chromatography for TatAyCy^{WT} and TatAy^{P2A}Cy.

After concentration of the peak Talon elution the concentrated TatAyCy sample for TatAyCy^{WT} (a.) and TatAy^{P2A}Cy (b.) was resolved on a gel filtration column. The Coomassie stained gel for fractions 12-30 containing the TatAyCy complexes after resolution is shown above. A western blot against the TatCy_H was performed with anti-His and against TatAy with anti-TatAy.

4.4 The Ratio of TatAy:Cy is increased in the P2A TatAyCy mutant

In a resting membrane, the TatABC machinery has relatively low abundance of TatA associated in the complex, while the TatB and TatC are present in higher and equimolar ratio (Oates *et al.*, 2005). Within the TatAdCd or the TatAyCy system, the TatA equivalent has a bifunctional activity of both TatA and TatB (Barnett *et al.*, 2008). With a crude method of densitometry, the ratio of TatAyCy can be quantitated to understand the effect the mutation on the complex assembly of the TatAyCy complex. The Coomassie stained SDS-PAGE gel from the Streptactin-affinity column and gel filtration column was analysed using densitometry to quantify the ratio of TatAy and TatCy in the peak elution fractions (Figure 4.5). The densitometry of the $TatAyCy^{WT}$ suggests there are 2 TatAy proteins per 3 TatCy in terms of density. Despite this the stoichiometry cannot be determined quantitatively as the size of the proteins is not directly comparable. The density ratio of TatAy (red) to TatCy (blue) in the $TatAyCy^{WT}$ complexes was 1:2 with the ratio of TatAy increasing by 2-fold in the $TatAy^{P2A}Cy$ complex (Figure 4.5b and c). In the later stage of the purification during the gel filtration, the ratio of TatAy increased to 3-fold (Figure 4.5d). This varying ratio of TatAy:TatCy suggests there are some transient changes occurring in the stoichiometry of $TatAy^{P2A}Cy$ complex, as a result of the mutation. In comparison the ratio of proteins in the $TatAyCy^{WT}$ complex remains the same in each purifications steps (Figure 4.5). Although using this crude method, the stoichiometry of the complex cannot be determined; the relative increase in TatAy protein has been significantly quantitated to increase by a minimum of 2-fold.

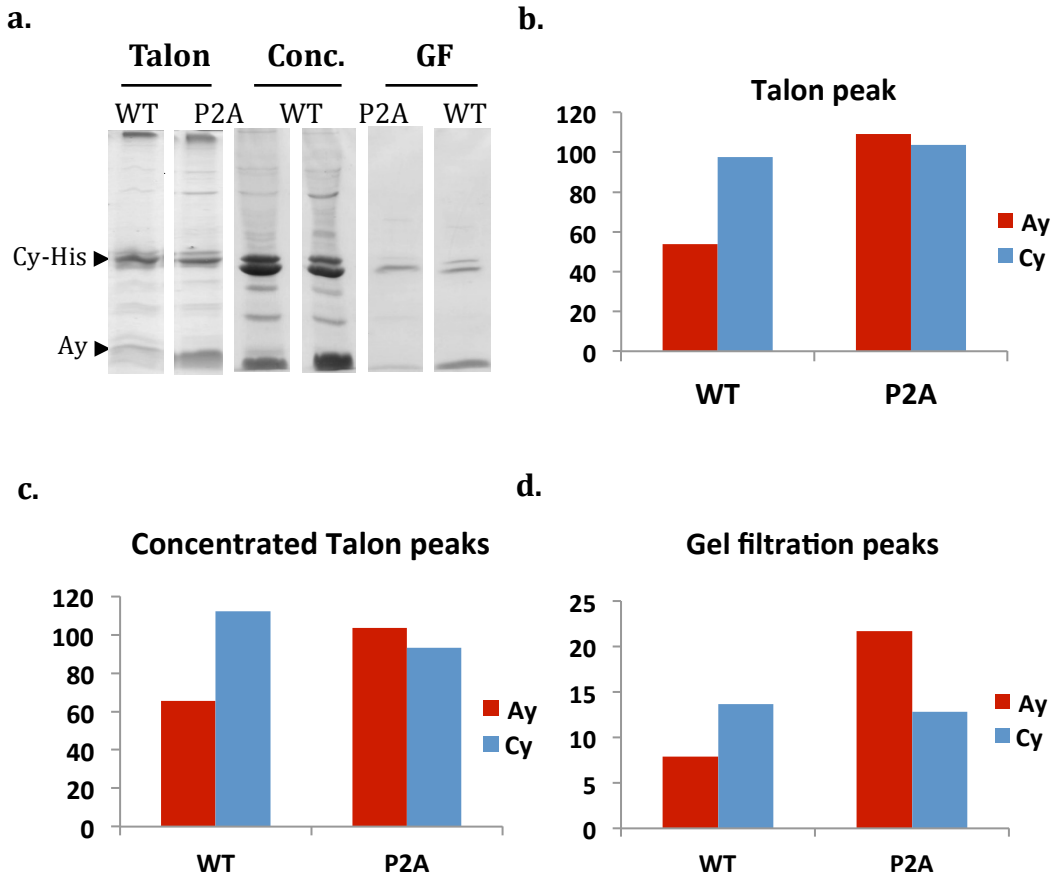


Figure 4.5: Stoichiometry of the TatAyCy^{WT} and TatAy^{P2A}Cy complexes observed from the SDS-PAGE gel with densitometry of the proteins

SDS-PAGE gels of the peak elution from the Talon purification, after concentration of the Talon elutions and peak elution from the gel filtration (a.) The arrows identify the TatCy_H (26 kDa) and TatAy (7 kDa) proteins on the gel. The pixel density of the protein bands in figure: a. was measured using the Aida software. The density from a single band is shown as a bar graph. Graph of the density comparing the ratio of TatAy:TatCy_H in TatAyCy^{WT} and TatAy^{P2A}Cy from the talon affinity column (b.). The ratio was also measured after concentration of elution fractions (c.) and on the peak of the gel filtration elution (d.).

4.5 Electron Microscopy of the TatAyCy^{WT} complex

It is evident from the purification of TatAy^{P2A}Cy that the mutation leads to accumulation of excessive amount of TatAy^{P2A} in the complexes. These results were both visualised on a SDS-PAGE gel and quantitated by densitometry. This sort of accumulation suggests there may be some structural irregularities occurring, possibly in the form of aggregation. With the aim to visualise the effect of exccessive TatAy^{P2A} recruitment on the structure of the complex, transmission electron microscopy was employed. The TatAyCy^{WT} sample from the peak gel filtration elution was analysed (fraction 19 Figure 4.4) at 40000X magnification after negatively staining the in uranyl acetate.

A large population of the particles observed in this fraction had a circular complexes forming, ranging in diameter between 9-16nm (Figure 4.6). The particles thought to be TatAyCy^{WT} complexes, are magnified (red boxes Figure 4.6) for a clearer observation of the structure. It seems that in some orientations, these round complexes appear to have a darker patch in the centre of the complexes. This is a characteristic trait of a hollow domain, suggesting either, a central cavity or pore within the complex. The TatAyCy complex is assumed to function as the substrate-recognition complex, and so, as with TatABC, the TatC equivalent may also be forming a cavity like domain within the TatAyCy complex. In addition, some larger irregular complexes of 20-50 nm were also observed (blue-boxed in Figure 4.6). However these structures were not consistent in shape or size. This irregularity in these structures is a characteristic trait of complex/detergent micelles aggregation.

The structure of TatAdCd_S was shown to be on average 8-11nm with majority of the particles constituting of approximately 9 nm (chapter 3). However, the TatAyCy^{WT} is forming larger complexes ranging from 9-16 nm. This size range observed within the two homologous systems is likely to be related to the size of the substrates the systems translocate. Unlike TatABC, where the size of TatA complexes is variable, the TatAC system in *B. subtilis* forms more homogeneous complexes as they only have a single substrate (Gohlke *et al.*, 2005; Oates *et al.*, 2005). Therefore the size of TatAdCd or TatAyCy is dependent on the substrate

.

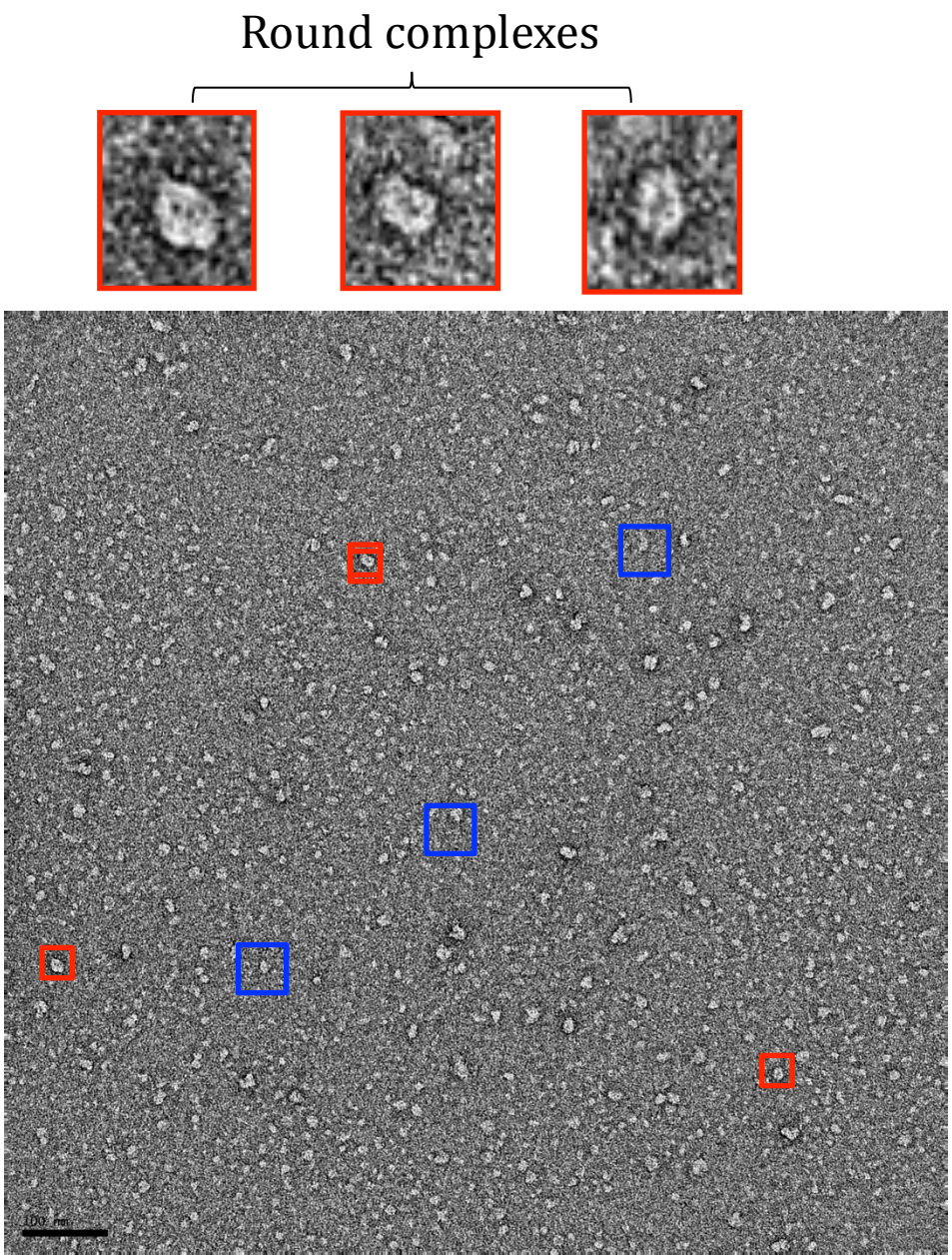


Figure 4.6: Negatively stained EM micrograph of the TatAyCy^{WT} complexes

Purified TatAyCy^{WT} complexes eluted from the gel filtration column were negatively stained with uranyl-acetate and fixed on a carbon mesh grid. The micrograph is of TatAyCy^{WT} by Transmission Electron Microscopy at 40000X magnification. Three of these round TatAyCy^{WT} (red boxes) complexes have been magnified on top of the image. Larger aggregates of detergent/protein complexes are highlighted in the blue boxes. The scale bar is set at 100nm for this image and all particles observed in this image are below 30nm.

4.6 Transmission Electron Microscopy of the TatAy^{P2A}Cy complex

Next, the TatAy^{P2A}Cy complex was visualised on the 200kV JEOL 2010 field emission gun (FEG) TEM. With TatAy^{P2A}Cy, two distinctive groups of particles were identified; one forming the round complexes, previously observed in only TatAyCy^{WT} sample and the other a long fibril structures with a helical-type arrangement (cyan box in Figure 4.7). The size of the round complexes matches what was previously observed with WT TatAyCy (Figure 4.6), showing the complex size ranging from 9-14 nm (Figure 4.7).

The fibril structure ranges in length from 100nm to above 500nm. It is also evident from the figure that the segments making up the fibril are more consistent in size compared to the round complexes. As the size of the segment and round complex are not are not comparable it is likely that the fibril may not be assembled in the same stoichiometry as the round complexes. It is also clear that the P2A mutation leads to increasing accumulation of the TatAy^{P2A} proteins. Therefore it can be hypothesised that multiple TatAy^{P2A} complexes accumulate to form the fibril. In complement the purification suggests that as the fibril are purified by affinity chromatography, TatCy_H must also be a constituent of these fibrils.

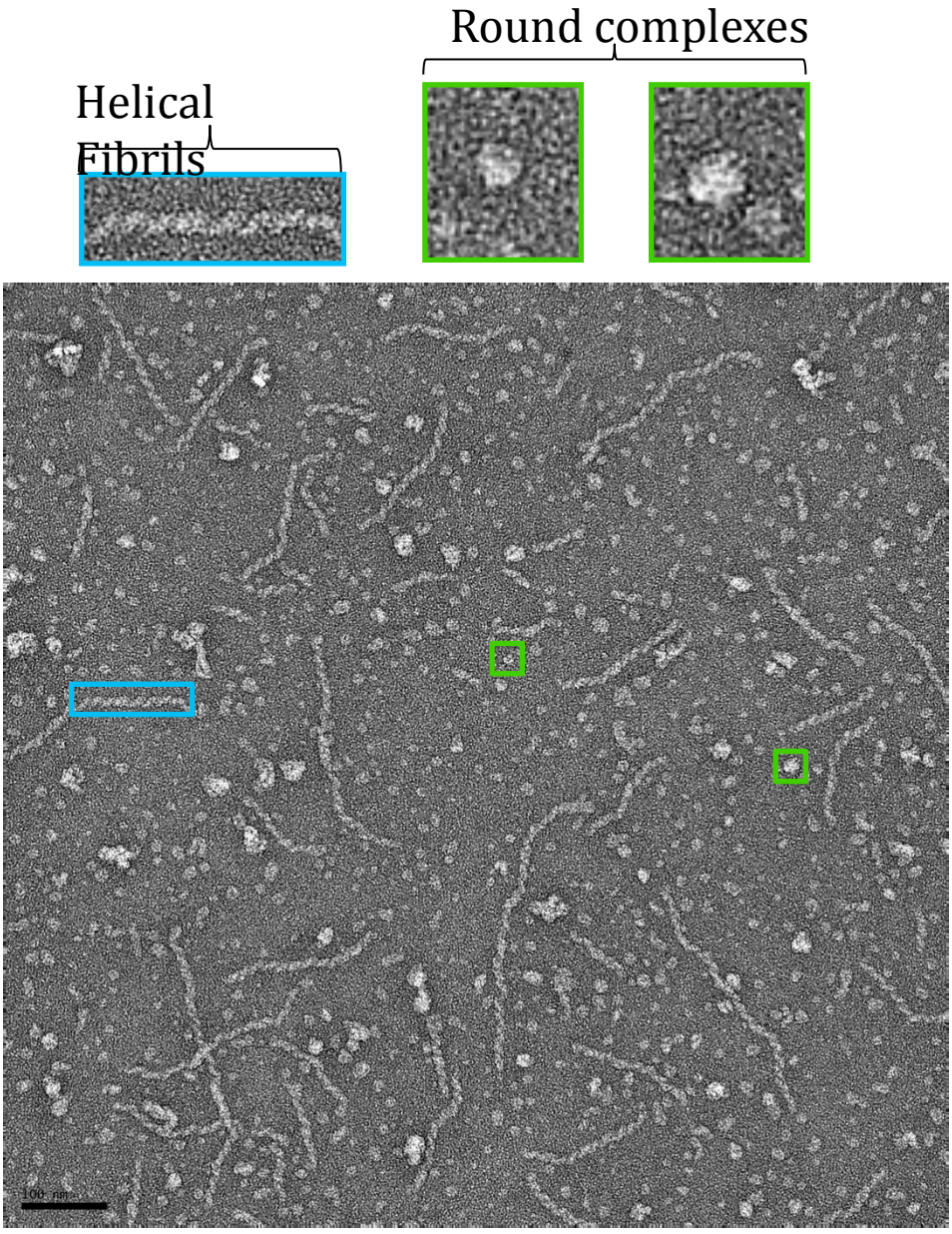


Figure 4.7: Micrograph of TatAy^{P2A}Cy complexes after purification by electron microscopy

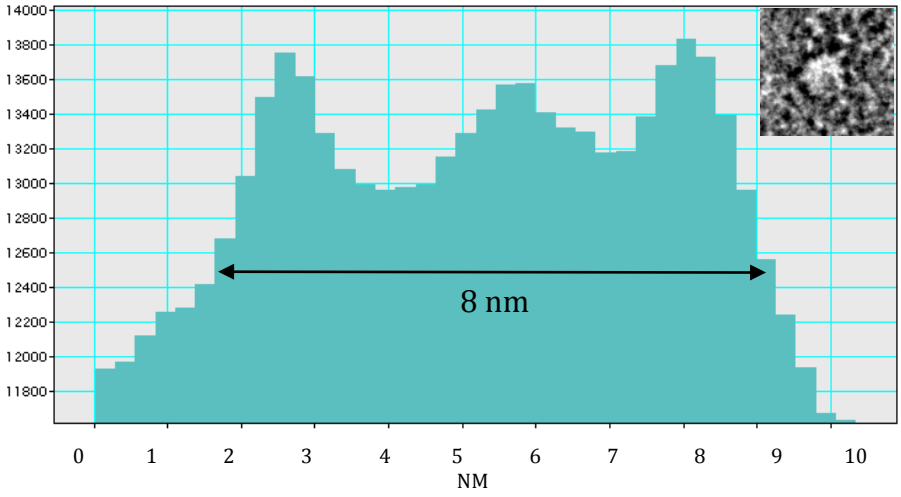
The TatAy^{P2A}Cy was purified using a 3-step purification method: Ion exchange chromatography, affinity chromatography and size exclusion chromatography. The His-tagged TatAy^{P2A}Cy complex was stained after purification using 2% uranyl acetate and visualised by TEM. The image taken showed two groups of particles, the round complexes and the helical fibrils ranging 100-500nm in length. Larger unordered structures are also present in a range of 16-40 nm. The scale bar for this image was set to 100 mn and the image was taken at 40000X magnification and -1.5 DV in defocus.

4.7 Density profile of the round TatAyCy^{WT} complexes

Further analysis of the round TatAyCy^{WT} complexes using the density profile tool allows both the characterisation of the complexes based on their density and size estimation of the complexes. For the TatAyCy^{WT} complexes, the diameter range of approximately 8-13 nm was observed (Figure 4.8). The range in size is similar to that observed for TatAdCd_S in Chapter 3. However the population of the TatAyCy seems to have a higher percentage of larger complexes.

Another interesting characteristic of these complexes was the conservation of the central density pattern in all the TatAyCy^{WT} complexes. The density across the complex shows a "W-shaped" pattern, which suggests, all the complexes have a central cavity. The peak in the centre of the complex not only shows that the TatAyCy^{WT} complexes exhibit some density in the centre, but also acts a means for differentiating from other detergent micelles which would have a hollow centre. The detergent micelles are expected to produce a "U-shaped" density profile, as the centre is completely hollow. The central peak in the density of TatAyCy^{WT} complexes could result a one-side lid like that observed in TatA complexes (Gohlke *et al.*, 2005).

a.



b.

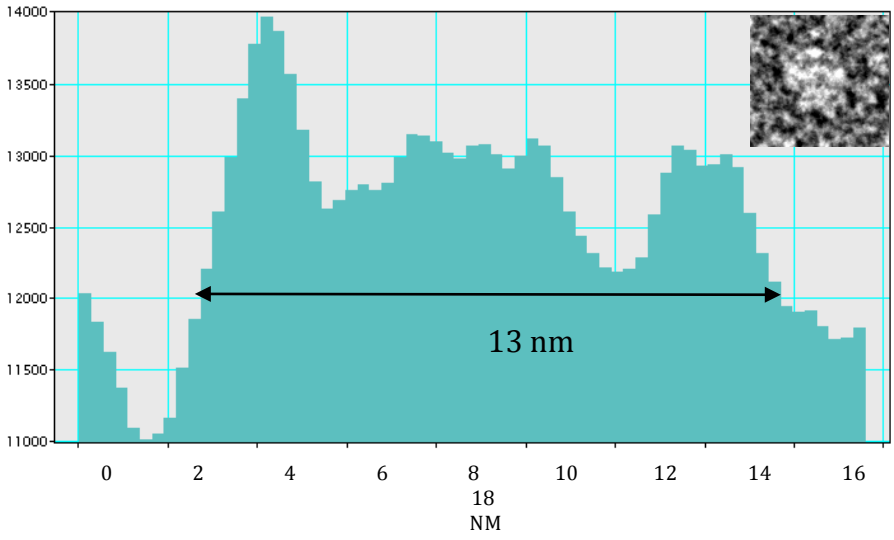


Figure 4.8: The density profile of the TatAyCy^{WT} round complexes with the cropped micrograph of the particle

TatAyCy^{WT} complexes were purified from the 5 ml talon column and run on superdex200 column. The peak elution of the complexes (a) was visualised by TEM at 40000X magnification. The density profile across a small and large complex was performed using the Digital micrograph software to measure the density of the selected complexes. TatAyCy^{WT} complexes ranged in from 8nm for small complex (a) to 13nm (b). The density of the TatAyCy^{WT} complexes shows a "W-shape" profile.

4.8 Density profile of the round TatAy^{P2A}Cy complexes

Similarly, a density profile across 4 complexes in the TatAy^{P2A}Cy complexes was produced to confirm the round complexes within the mutant sample are similar to what is observed in TatAyCy^{WT}. The average size range in the TatAy^{P2A}Cy complexes ranged largely between 10-14 nm. Larger complexes in the range of 14 nm and large aggregates from 20 nm were also observed. The larger 20 nm particles do not exhibit the characteristic "W-shaped" density pattern and so more likely to be detergent aggregates. However the smaller and more homogenous particles (10-14 nm) did portray this conserved "W-shaped" pattern. The round complexes in the TatAy^{P2A}Cy complexes are not exactly as observed in TatAyCy^{WT} but the signature density profile is still conserved. The decrease in the size range of the round complexes may inherently be linked to the formation of the fibrils. It may also be possible that the smaller round complexes that are less prevalent in the TatAy^{P2A}Cy sample maybe used in the assembly of the helical fibrils.

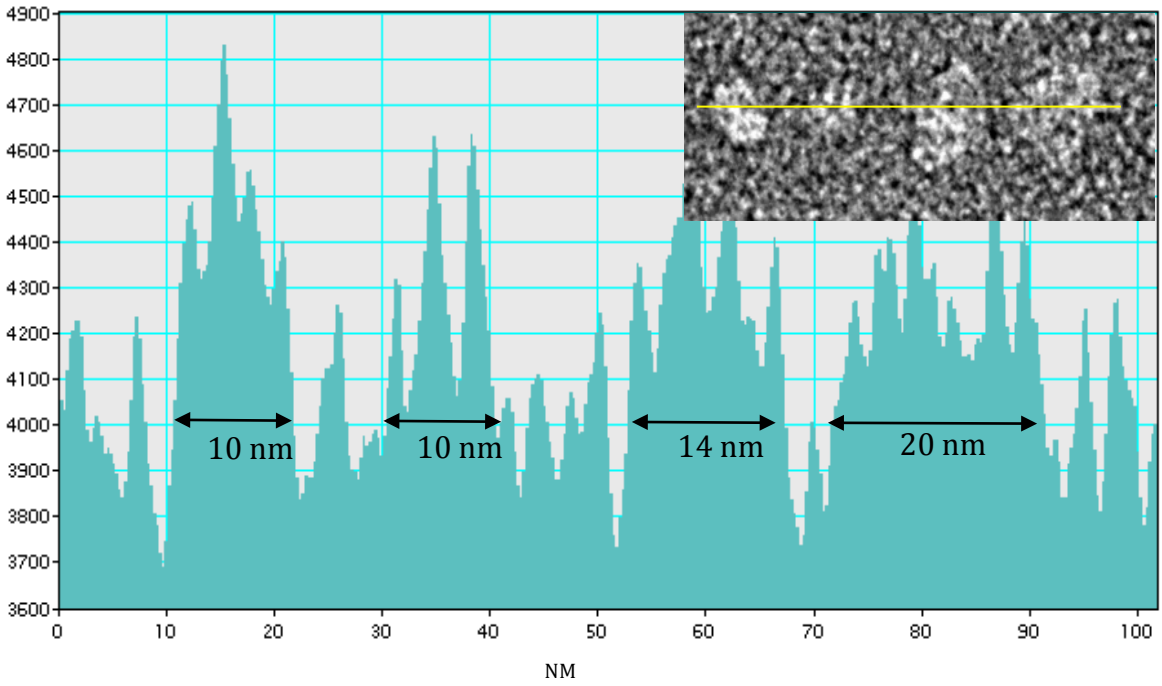


Figure 4.9: Density profile of round TatAy^{P2A}Cy complexes with a TEM micrograph of complexes analysed

The round TatAy^{P2A}Cy complexes were purified from a Talon column and run on superdex200 column. The peak elution of the complexes (Figure 4.4b) was visualised by TEM at 40000X magnification. The density profile across a range of complexes was performed using the Digital micrograph software. The complex size was estimated by using density above 4000 RU. Most of the complexes show a "W-shape" characteristic for complex in the range of 10-14 nm. The 20 nm complex shows disordered pattern suggesting aggregation.

4.9 Density profile of the helical TatAy^{P2A}Cy fibrils

To gain further insight into what is making up the helical fibrils, a series of density profiles were carried out in the fibril itself. The first was a density map across what looked like two helical turns, as shown in Figure 4.10. Within the profile there were 5 distinctive peaks observed with a uniform density pattern. The individual complexes range from 4-7 nm in size, which suggests they may be too small to be the complexes previously observed on the TatAyCy^{WT} complexes (Figure 4.10). It may also be possible, that the complexes have an altered orientation and so the width of the complex would no longer be comparable to the width the fibril. In the 3D model generated for TatA and TatBC (Tarry *et al.*, 2009; Gohlke *et al.*, 2005), the height of the complexes is an average of 5 nm, comparable to that observed with the fibrils (Figure 4.11a). Therefore suggesting the complexes may stack sideways and give a "U-shaped" profile instead.

In addition, the width of the fibril was also measure to understand the possible orientation of the individual complexes that may be producing the fibril structure. Firstly the width along one turn of the helix was analysed as shown in Figure 4.11a, where the diameter of the turn was 12 nm. The density profile across the fibril was used to measure the overall width of the fibril as shown in Figure 4.11a, showing the fibril maintained an average width of 12 nm. This width of the fibril matches the average size of the round complexes observed with both the TatAyCy^{WT} and TatAy^{P2A}Cy. As the cross-sectional density profile along the fibril (Figure 4.10) shows smaller sized particles, the conformation and the constituents of the fibril are difficult to predict at this level of analysis.

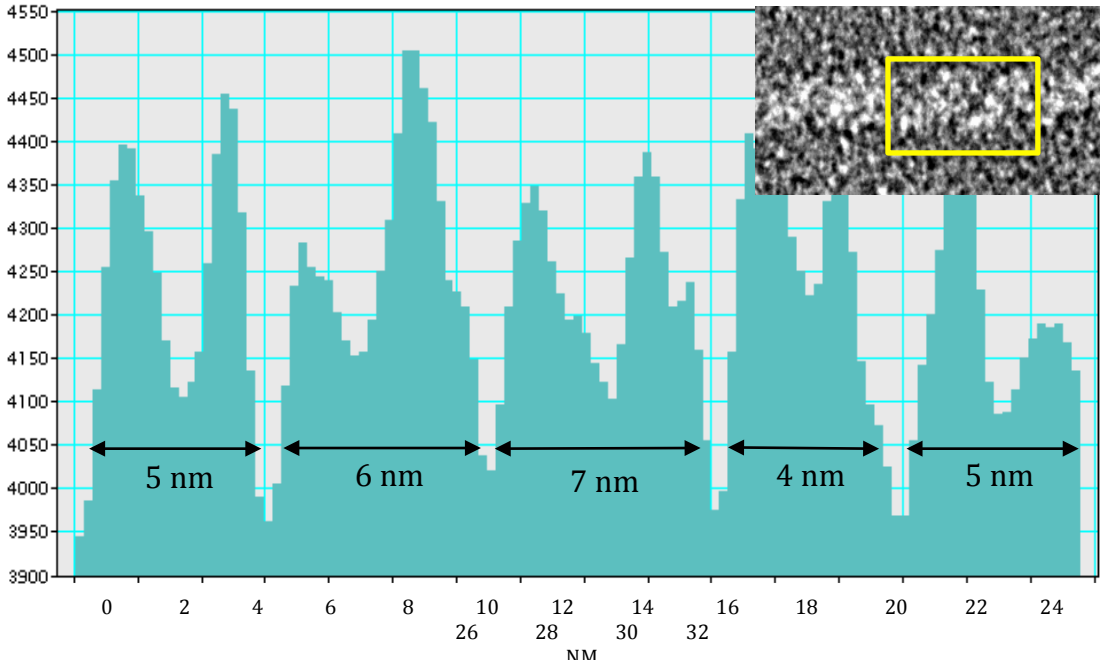
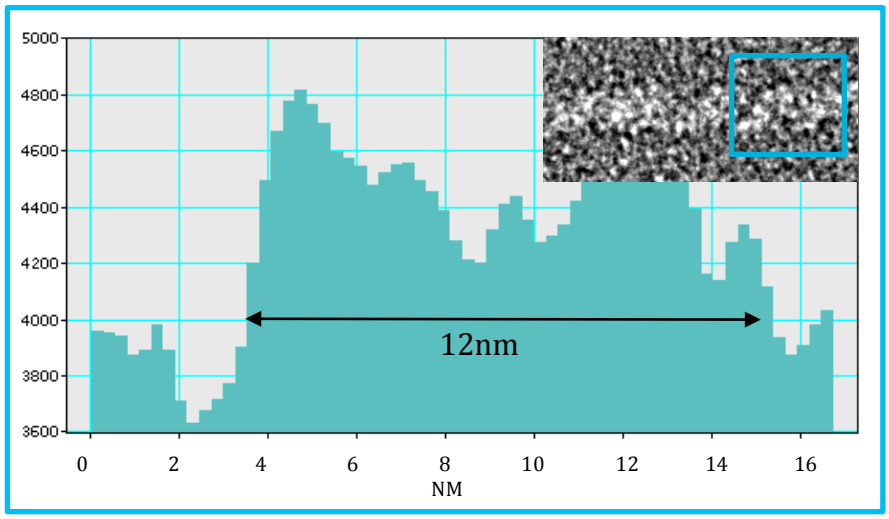


Figure 4.10: Density profile across 2 helical turns on the TatAy^{P2A}Cy fibrils with TEM micrograph of the fibril

TatAy^{P2A}Cy was negatively stained and observed by TEM at 40000X magnification. The density profile along the fibril formed TatAy^{P2A}Cy was taken along the fibril. The yellow box shows the density of the fibril taken from left to right of the boxed image. The fibril that appears as two helical turn shows 5 distinctive arrangements with a "U-shaped" profile. The size of the 5 groups ranges from 4-7 nm when density if measured from 4000 RU.

.

a.



b.

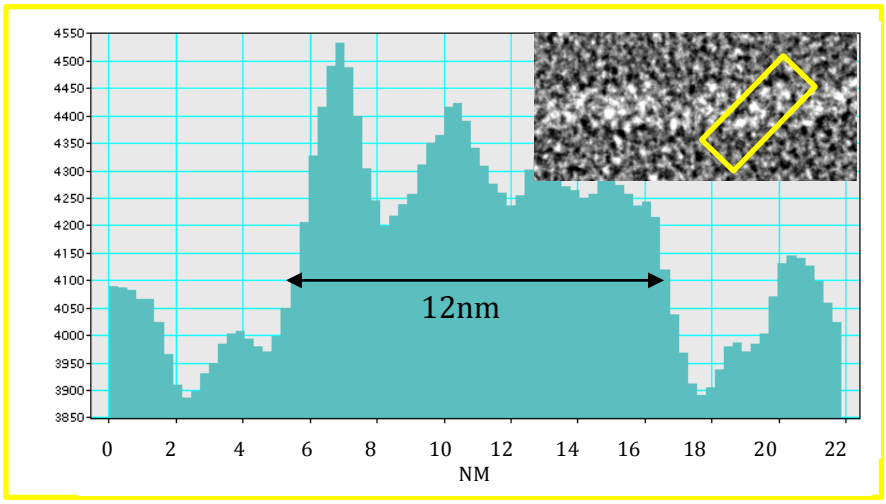


Figure 4.11: Density profile of the TatAy^{P2A}Cy fibril with the TEM micrograph of the fibril

The fibrils formed by the TatAy^{P2A}Cy complexes were visualised by TEM at 40000X magnification. The density profile was taken across the fibril (a) and along one helical turn of the fibril (b) using the Digital micrograph software. In both density profiles, the diameter across two helical turns was 12 nm (a) and the diameter along one turn was 12 nm (b). The size of complex was measured for complex density above the background density at 4000 RU.

4.10 Atomic force Microscopy of TatAyCy^{WT} and TatAy^{P2A}Cy complexes

Cvetilin Vasilev (University of Sheffield) produced the AFM images presented in this chapter. The AFM work was used to complement the EM data shown earlier to gain addition structural insight.

The AFM has the advantage of being able to obtain nm-scale resolution in physiological (or nearly-physiological) environment (in liquid, at the right pH and temperature). The main disadvantage is that the sample has to be 'fixed' (immobilised) onto a flat support in order to allow the AFM probe to be scanned over it. In this case, a mica substrate was used, which provided a very clean and flat (atomically flat over hundreds of μm) surface. The principle of AFM uses a probe to scan over the mica surface and as the probe moves along the surface, a pixel is generated in correlation with the Z-axis. It uses a serial acquisition technique, so the image pixels are recorded in sequence along the scan lines forming a topological image. Therefore this technique is better suited for imaging samples, which remain static over time.

The topography images show partial coverage of the mica surface with fibrils as well as with some large aggregates and smaller isolated TatAyCy complexes. Again the conformation of the fibrils appears to be dynamic. The TatAyCy^{WT} complexes formed small round complexes, with additional larger round complexes ranging from 20-50nm in size (Figure 4.12a). This was not observed with EM images in Figure 4.7 and these complexes form during sample preparation of the mica surface (section 2.7.1). As with EM, the AFM of the TatAy^{P2A}Cy mutant also shows two distinctive groups of complexes, the round complexes and the helical fibrils (Figure 4.12b).

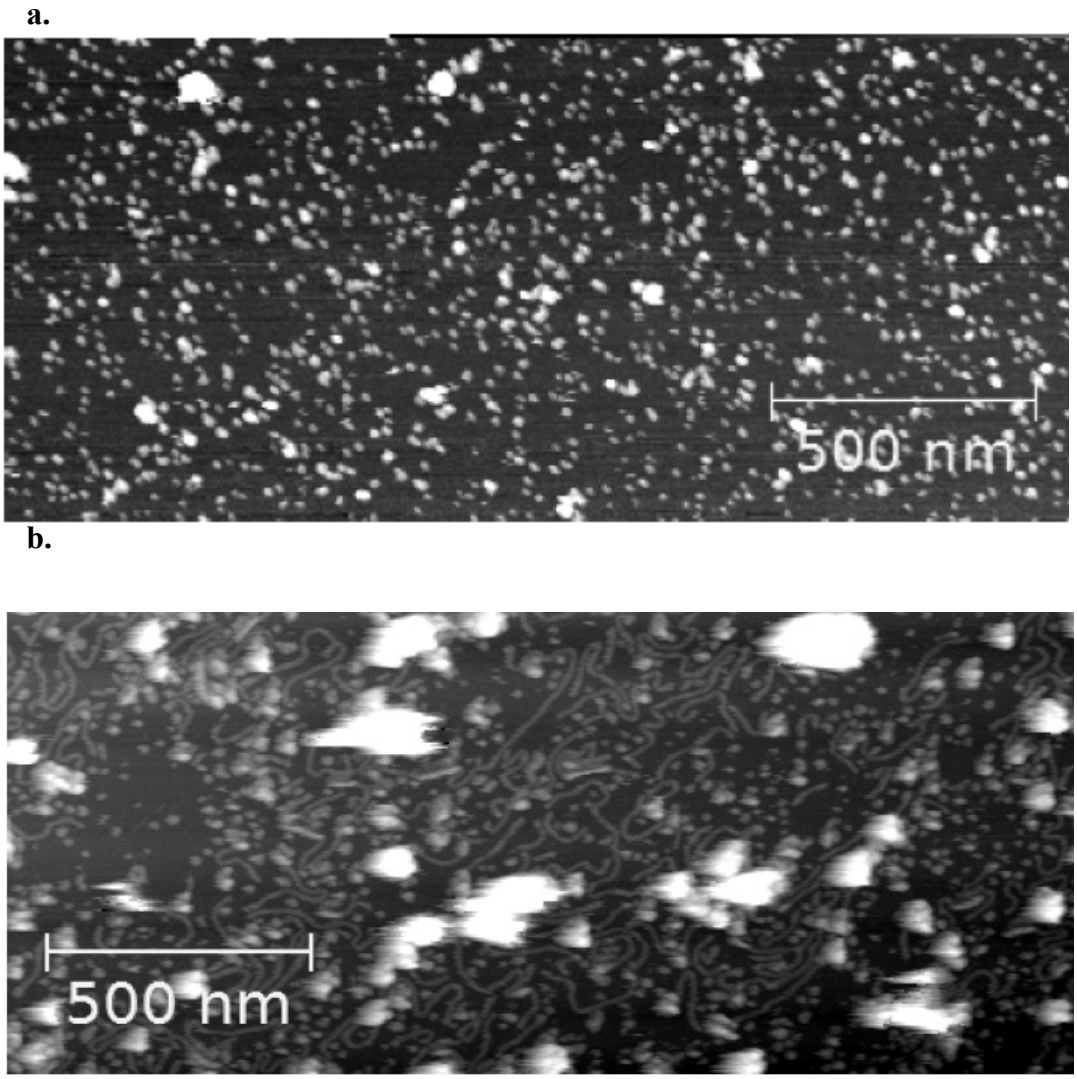


Figure 4.12: Topograph of TatAyCy^{WT} and TatAy^{P2A}Cy mutant using AFM

The purified samples of WT/P2A TatAyCy were applied to a mica surface and suspended in Tris-buffer. AFM probe was scanned over the sample to obtain topological data showing round complex for TatAyCy^{WT} (a) and the TatAy^{P2A}Cy mutant (b). The fibril like structure was observed only with TatAy^{P2A}Cy (b). The length of the fibrils ranges in size from 200-800 nm.

With the aim to increase the topography image resolution for the $TatAy^{P2A}Cy$, the sample was imaged by AFM using higher concentration of protein. Using a higher density sample would reduce the free space on the mica surface during the probe scan. This consequently restricts the movement of the fibril, resulting in a higher resolution data.

Upon analysis of the concentrated sample, it was evident that increasing the sample density on the mica surface leads to a tighter packing of the complexes as shown in Figure 4.12. Interestingly, at a higher concentration the fibril conformation is transformed to a straight and tight tubule like structure (Figure 4.12b). This occurrence is likely to have resulted from packing problems presented with a concentrated sample. In this case, the fibrils cannot rest over each other on the mica surface so this conformation is induced to accommodate the maximum number of fibrils. As the samples are more static, the resolution of the topography image is enhanced to show a helical-like structure within these fibrils (Figure 4.12b). The long-range alignment of the tubules may be imposed by the chirality of the fibrils, resulting in an ordered alignment. The regular spacing of the fibrils could result from long-range electrostatic interactions provided by the charged residues on the periphery of the $TatAy^{P2A}Cy$ complex. Majority of the charged residues are found largely on the cytoplasmic loops of TatC (positive) and on the C-terminus of TatA (negative), as this is the site of substrate recognition (Barnett *et al.*, 2011; Allen *et al.*, 2002). However, in the absence of the substrate, these charged residues may be causing electrostatic interaction when brought in close proximity. The direction of the ordering of the fibrils can be either flow-induced during the sample adsorption/washing on the mica surface. It is also possible that couple of fibrils can line up with the edge of the mica surface and then the order is preserved the next fibrils across the surface.

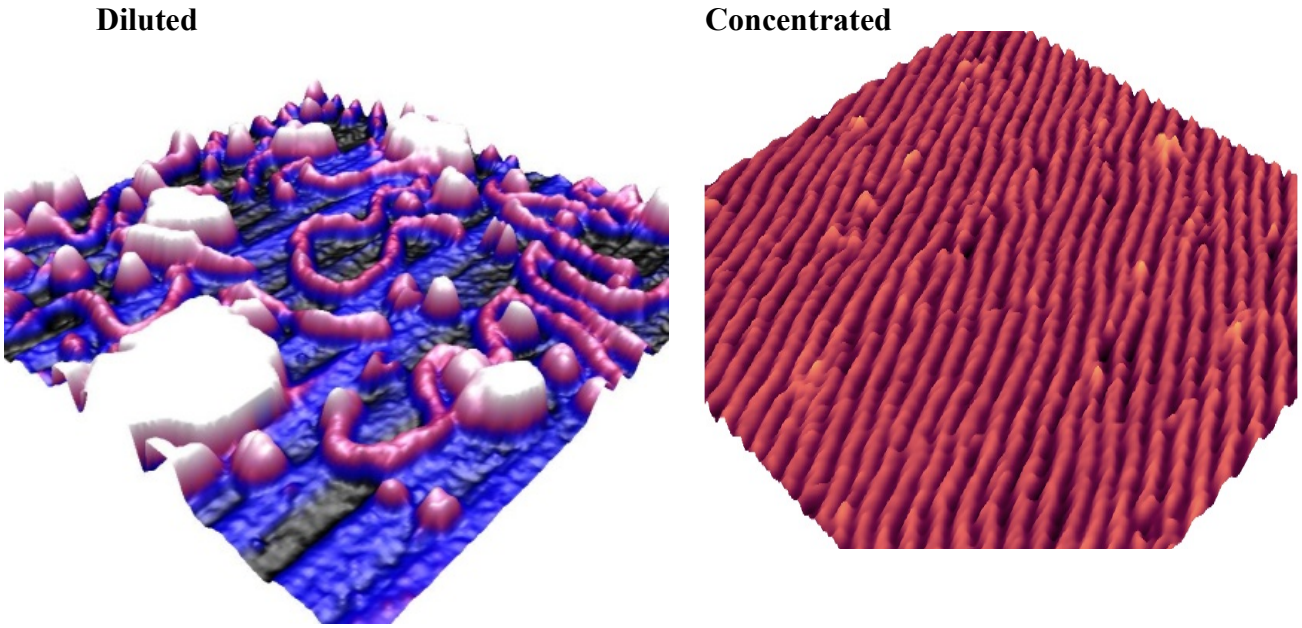


Figure 4.13: 3D model of the TatAy^{P2A}Cy with diluted and concentrated sample adhered to the mica surface.

The TatAy^{P2A}Cy mutant was analysed by AFM under two conditions: diluted and concentrated samples. The height and width dimension of the scan can be used to form a 3D model as shown above. In the diluted sample, there was more free space on the surface so the conformation of the fibrils remains random. However when a concentrated sample is applied to the surface the fibrils are induced into a straight aligning structure to accommodate other fibrils. It was also apparent that a more concentrated sample produced higher density and longer fibrils compared to round complexes.

4.11 Surface profile of the aligned fibrils of TatAy^{P2A}Cy

A final surface profile for the concentrated fibrils was carried out to understand the dimensions of the fibrils. The size range along the fibril was 8-9 nm between the peaks along the fibrils. The surface profile across a set of fibrils shows that the width of the fibrils varies from 9 to 14 nm (Figure 4.14). These results are similar to those observed with the surface profile from the diluted sample. The length of the fibrils is in the range 30 – 850 nm with the majority of the fibrils ranging in length between 300 – 400 nm. The length of the fibrils in the concentrated sample was found to be higher, possibly suggesting that with increased dilution and more free space the fibrils are prone to disassembly.

The height of the fibrils is difficult to determine for the concentrated sample, as the probe does not reach the base of the mica surface. Therefore although the surface profile measures the height of fibrils at 2.5 nm, it is not an accurate evaluation (Figure 4.14). This is because with concentrated sample the space between fibrils is limited and so the probe cannot estimate the height accurately in this sample. Therefore the actual height estimation from the diluted sample was used to determine the height of the fibril, which is on average 5 nm. The curvature of the probe is likely to produce error in the measurement of the width of complexes but the height (Z-axis) is high resolution irrespective of the size or curvature of the probe tip. As the height of the complexes is maintained at 5 nm it suggests that fibril may also have a preferred orientation as seen with individual round complexes.

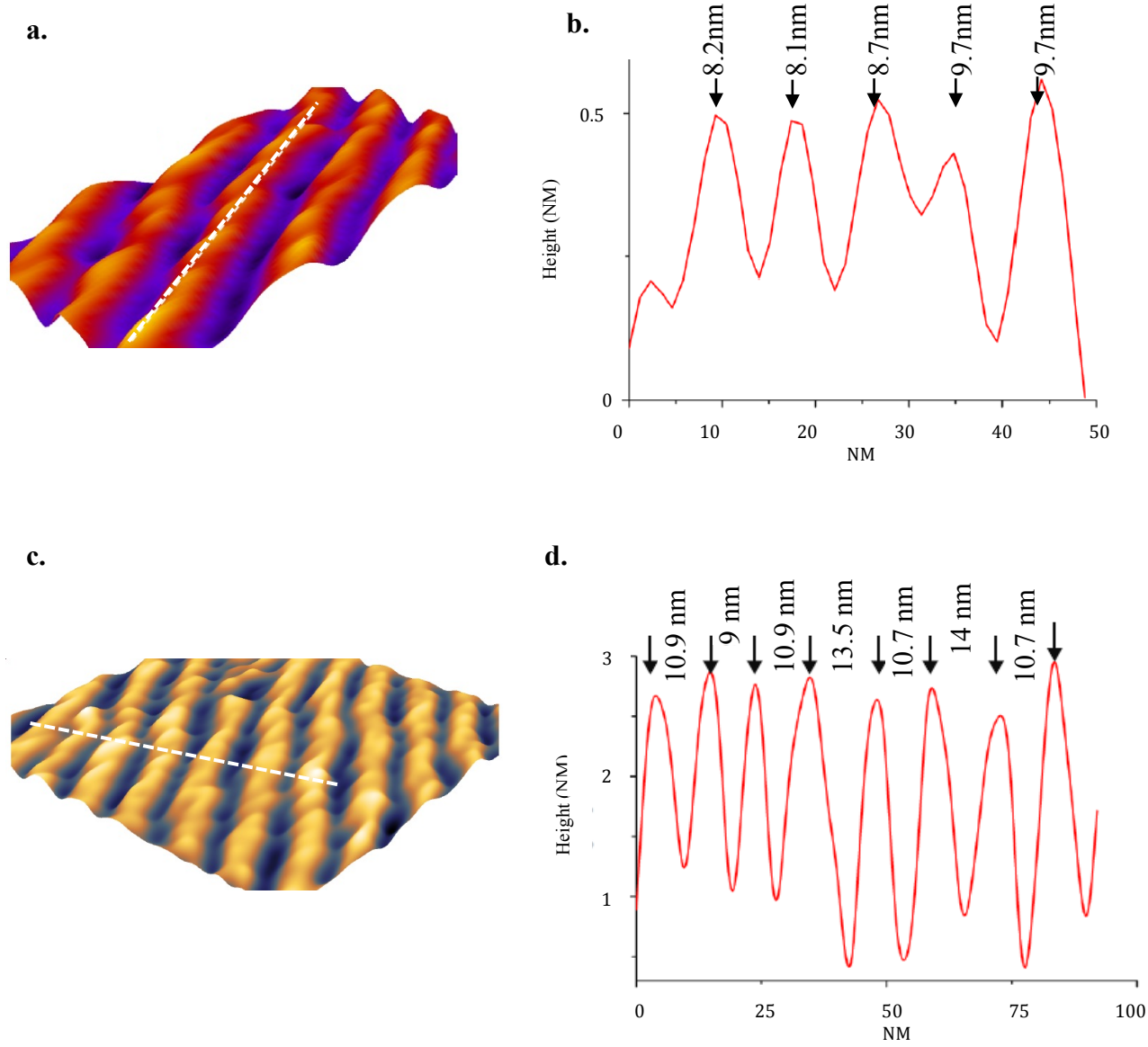


Figure 4.14: magnified topograph of TatAy^{P2A}Cy fibril with it corresponding surface profile plot

Magnification of the fibrils was produced to generate surface profile across (c.) and along the TatAy^{P2A}Cy fibril in a concentrated sample. High-resolution AFM topography (a) and 3D representation (b.) acquired over the area marked with the dashed square in. The helical structure of the fibrils can be seen. Cross-section across 8 fibrils with the period (peak-to-peak) was taken for the fibril (d).

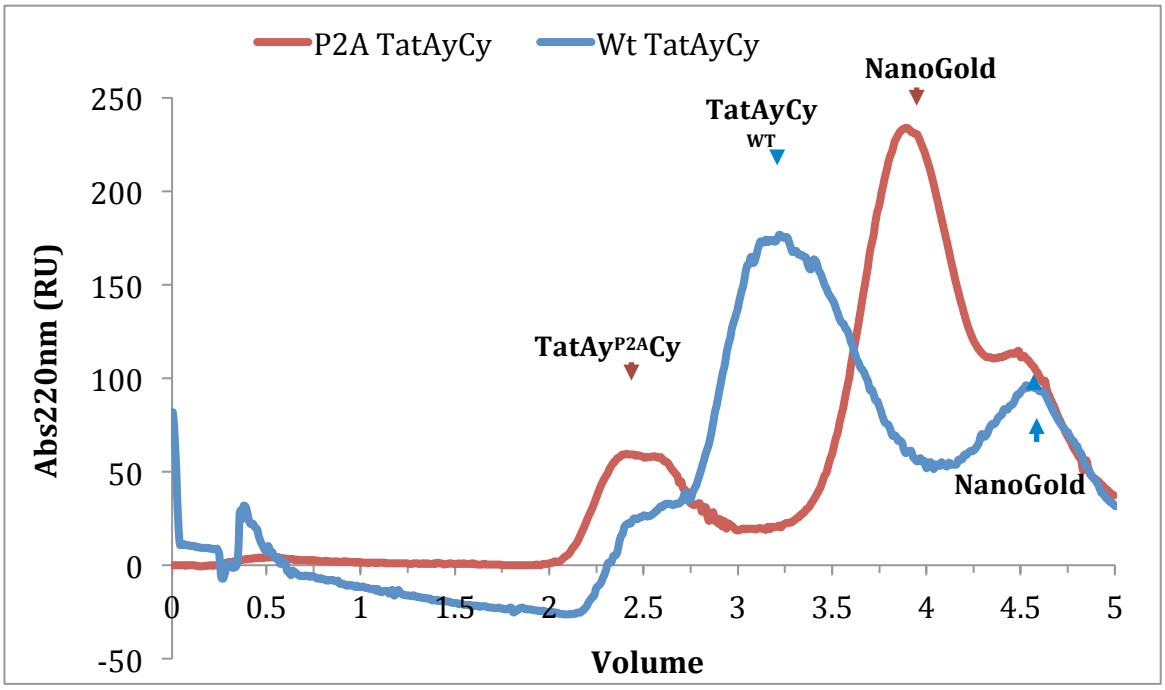
4.12 Nano-Gold binds specifically to the round complexes TatAyCy

From the structural analysis, two significant findings lead to the hypothesis that the two populations of the round complexes and the fibrils are actually composed of two different types of complexes. The homogeneity and conserved size (8-11nm) within the fibrils suggested that these might be $TatAy^{P2A}$ complexes that are present as a result of the mutation and hence seen in higher ration to TatCy during purification. With the round complexes from both $TatAyCy^{WT}$ and $TatAy^{P2A}Cy$ mutant, the sizes of the complexes are much larger and range between 8-14 nm. As the round complexes are natively observed $TatAyCy^{WT}$ with both the TatAy and TatCy_H co-purified, these complexes are TatAyCy complexes. Similarly, both $TatAy^{P2A}$ and TatCy_H should also form the round complexes observed with $TatAy^{P2A}Cy$. In brief, it is hypothesised that the fibril formation is initiated with the $TatAy^{P2A}Cy$ complex (round complex), which then leads to assembly of fibrils from only $TatAy^{P2A}$ complexes (smaller homogeneous complexes).

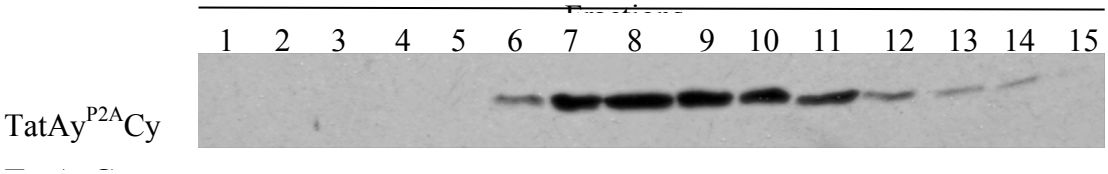
4.12.1 Purification of Nano-Gold labelled TatAyCy

To test this hypothesis, anti-His Nano-Gold particles were used to label TatCy_H protein, to determine if the fibrils consisted of any TatCy_H, other that the initiator complex. The peak elution from the gel filtration column was incubated with anti-His Nano-Gold overnight at 4°C with a ratio of 2:1 (Nano-Gold: protein). The Nano-Gold labelled TatAyCy complexes were then isolated from free anti-His Nano-Gold using a 5ml superdex200 gel filtration chromatography. The 5ml column provided enough separation between the large $TatAy^{P2A}Cy$ complexes (fibrils) that were produced as results from the mutation. Free Nano-Gold antibody monomers are smaller in size so the unlabelled Nano-Gold could was eluted off later. The gel filtration shows two peaks that were observed; in addition Nano-Gold also absorbs in the range of 220nm, similar to that of the residues present in the protein (Figure 4.15 a.). There are two peaks observed from the gel filtration graph, where the peak that represents TatAyCy complexes labelled with Nano-Gold and the peak at represent free Nano-Gold, where no TatAyCy complexes are eluted.

a.



b.



c.

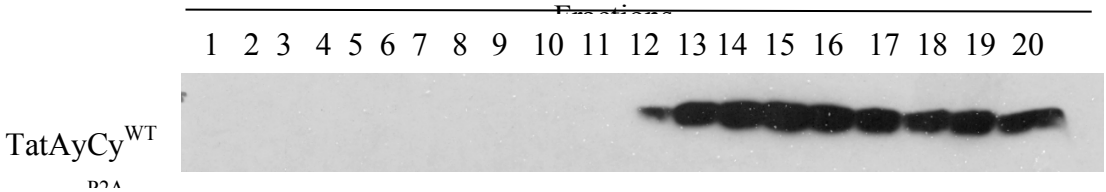


Figure 4.15: The purification of Nano-Gold labelled TatAyCy^{WT} and TatAy^{P2A}Cy complexes on superdex200 with western blot of the elution fractions.

The purified TatAyCy complexes were incubated with anti-His Nano-gold overnight. The TatAyCy^{WT} and TatAy^{P2A}Cy complexes were then isolated using a 5ml Superdex200 beads. The protein elution was measured at 220nm to visualise the anti-His Nano-Gold elution too (a). The presence of the TatAyCy complexes in the gel filtration fraction was confirmed by a Western blot using anti-His antibody (b.and c).

4.12.2 Electron microscopy of H_Nano-Gold labelled TatAyCy^{WT}

The TatAyCy^{WT} sample was similarly incubated with a 2:1 ratio of anti His-tag specific Nano-Gold (H_Nano-Gold) for 2 h and purified out a 5ml superdex200 gel filtration column. As shown in Figure 4.15, the purification of the TatAyCy^{WT} complexes is not as efficient as the TatAy^{P2A}Cy because the complex size (370 kDa) is smaller and the overlap between unbound Nano-Gold and TatAyCy^{WT} complex suggests there may be some free H_Nano-Gold contamination. However, upon visualisation of these complexes by TEM, there were a very low number of unbound H_Nano-Gold particles. In Figure 4.16, the first magnified box (right) seems to represent two free H_Nano-Gold particles while the remaining H_Nano-Gold was bound to a TatAyCyWT complex (box 2-3, Figure 4.16).

The H_Nano-Gold old is bound to the round complexes and in particular the larger complexes (boxes 2-4 on the right, Figure 4.16). From this incubation it was apparent that a very low amount of H_Nano-Gold particles were bound after an incubation of 2hrs and using a H_Nano-Gold to protein ratio of 2:1. The size of the Nano-Gold particles is in range between 3-5 nm, which is in the smaller range and so difficult to determine if the Nano-Gold particle would be able to access the cytoplasmic C-terminus of TatCy_H when assembled into the TatAyCy^{WT} complex. The negative glow discharge of the sample grid prior to sample application allows all complexes to sit in a preferred orientation; so all complexes are observed in the same orientation. This means that if the H_Nano-Gold is accessible to one complex, is should be accessible to all complexes, therefore the limiting factor in the labelling must be the concentration of H_Nano-Gold used during the labelling incubation.

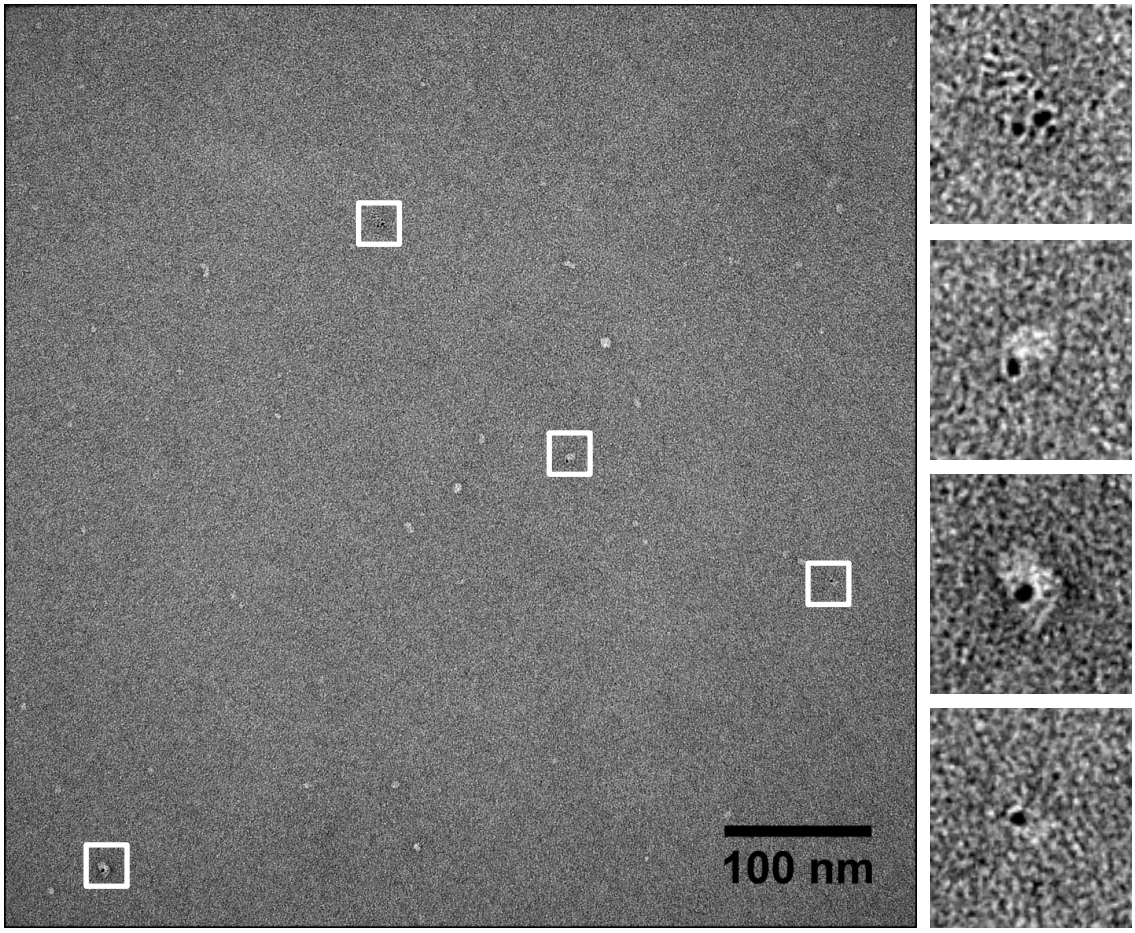


Figure 4.16: TatAyCy^{WT} labelled with anti-His Nano-Gold and visualised by TEM

Purified TatAyCy^{WT} was incubated with H_Nano-Gold at ratio of 2:1 overnight and visualised by transmission electron microscopy (TEM). The sample was stained with 1% uranyl acetate for better contrast with H_Nano-Gold. The images were taken at 40000X magnification and a defocus of -1.5DV was applied. 10% of the round TatAyCy complexes have of particles conjugated to the H_Nano-Gold.

4.12.3 Electron microscopy of H_Nano-Gold labelled TatAy^{P2A}Cy

As it was shown from the TatAyCy^{WT}, the labelling protocol was not sufficient and so during the labelling of TatAy^{P2A}Cy with H_Nano-Gold, the ratio of was increased 10-fold to that of the protein (*v/v*) and left to incubate overnight. With this optimisation step the labelling was significantly improved. The peak elution fraction after labelling with H_Nano-Gold and purification was visualised using TEM. The coverage of the H_Nano-Gold was increased to 50%, so on average, half of the complexes were labelled with H_Nano-Gold (Figure 4.16).

The labelling shows that majority of the H_Nano-Gold is restricted to the round complexes in TatAy^{P2A}Cy. In comparison to the earlier images of TatAy^{P2A}Cy, the image below shows a relatively low abundance of fibrils present and is also shorter in length. This is a result of the dilution with the initial 3-fold dilution during the labelling of the antibody (2:1 ratio). Further 20-fold dilution was subjected during the superdex200 separation step (removing unbound H_Nano-Gold). Therefore multiple images were taken of the grid to locate particles with the H_Nano-Gold specifically bound to the round complexes (Figure 4.17). As observed earlier for the AFM sample, after dilution the population of fibrils is greatly reduced. Therefore when the sample is diluted to this extent during, it seems likely, that the fibrils also start to disassemble. Therefore the ratio of round complexes to fibril is altered when the sample is diluted. In spite of this occurrence, it was still possible label the TatAy^{P2A}Cy complexes.

The H_Nano-Gold labelling of TatAy^{P2A}Cy is specific to the larger round complex as seen in Figure 4.17, while the smaller complexes remain largely unlabelled. As earlier hypothesised, if the fibrils are made up of the smaller homo-oligomeric TatAy^{P2A} complexes, then this would support that hypothesis. It is not possible to conclude from this finding that fibrils are disintegrating but this may provide a feasible justification for the observation of unlabelled smaller round complexes, which were not previously observed in Figure 4.7.

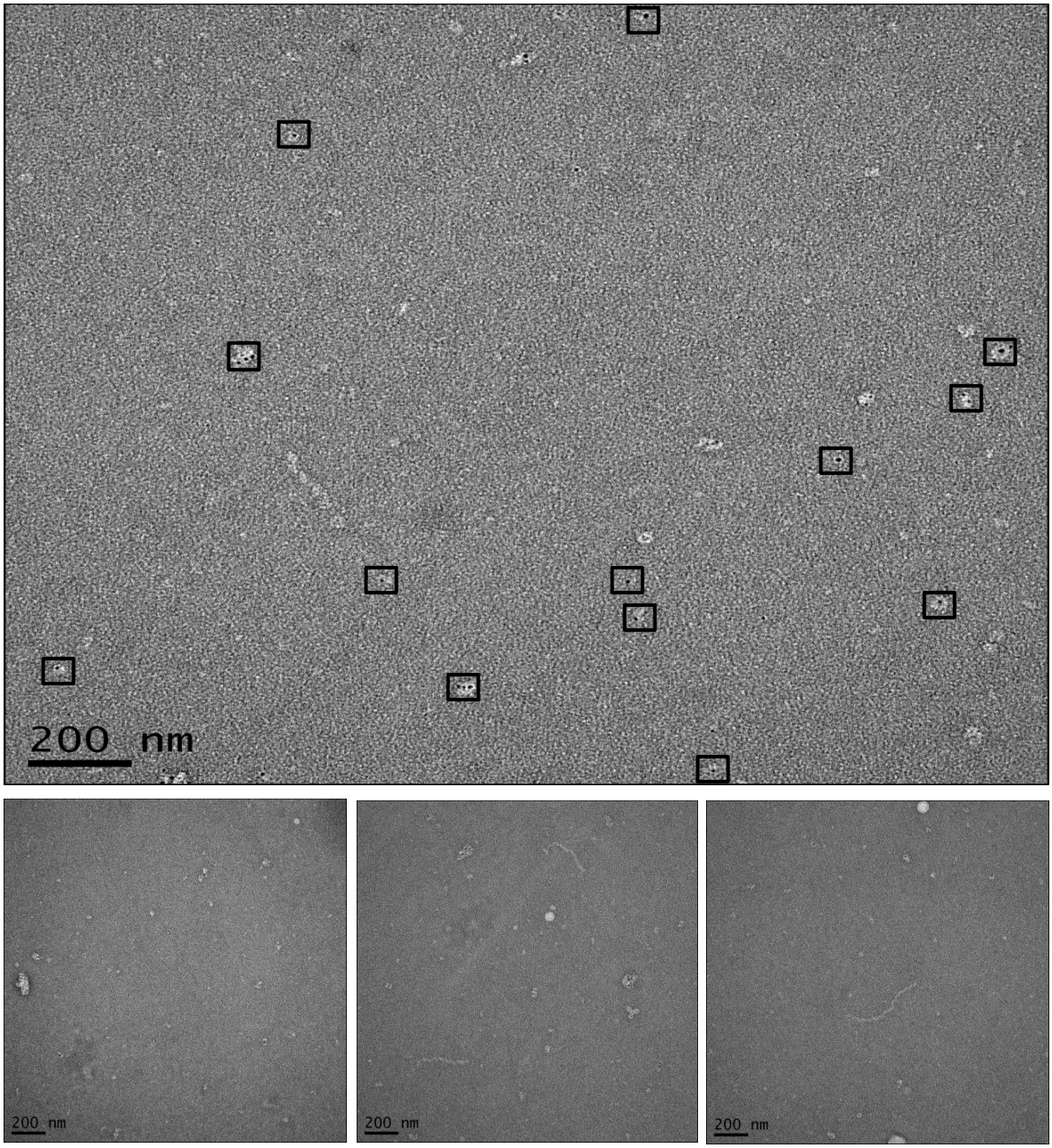


Figure 4.17: anti-His Nano-Gold labelled TatAy^{P2A}Cy image visualised by TEM

TatAy^{P2A}Cy was incubated with H_Nano-Gold overnight and visualised by transmission electron microscopy. The sample was stained with 1% uranyl acetate for better contrast with H_Nano-Gold old. The images were taken at 40000X magnification and a defocus of -1.5 DV was applied. All images show scalebar of 200 nm. Lower percentage of fibrils are seen and of those observed, H_Nano-Gold is not bound by the fibrils. The round TatAy^{P2A}Cy complexes labelled with 40% efficiency.

4.12.4 TatAyCy_S: control H_Nano-Gold labelling

As a control, TatAyCy_S was expressed in *E. coli* and purified using a 5ml Q-Sepharose column (ion-exchange chromatography). Elutions from this purification were then applied to a 5 ml streptactin column (affinity chromatography). The elution from the streptactin column was relatively pure (Figure 4.18), so was directly incubated with H_Nano-Gold at ratio of 10:1 (protein:Nano-Gold). The TatAyCy_S complexes were incubated overnight before purification on a 5ml Superdex200 column. The fractions collected from the gel filtration chromatography were then run on a SDS-PAGE gel and western blotted against anti-strep antibody. The peak elution of TatAyCy_S was then visualised by TEM (Figure 4.19).

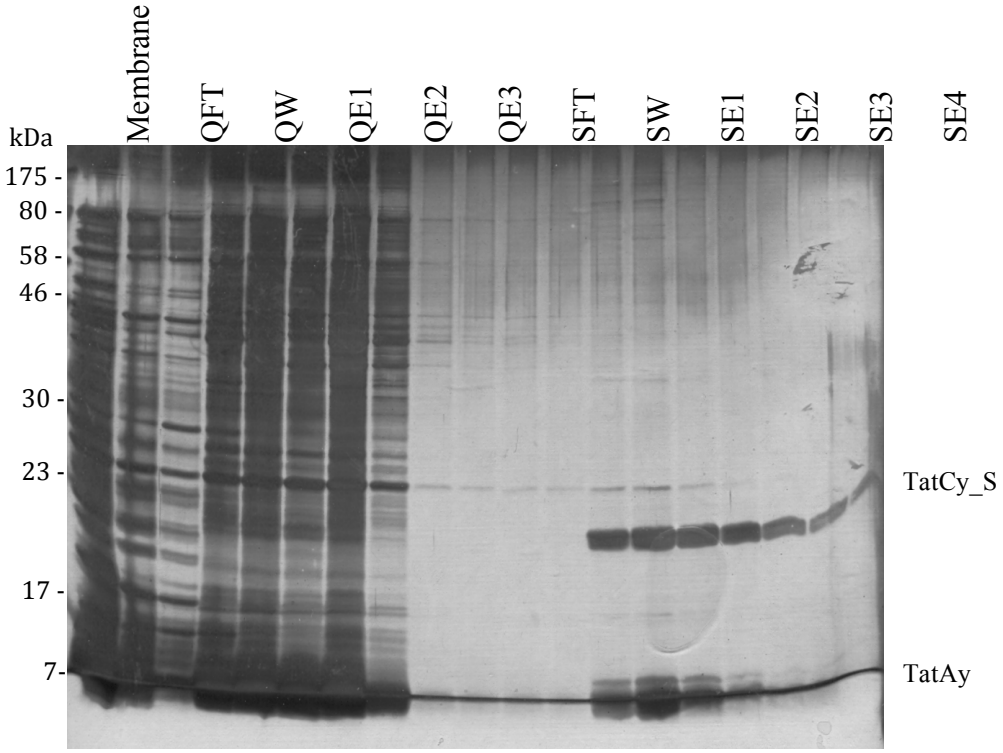


Figure 4.18: SDS-PAGE gel of the TatAyCy_S purification using a StrepTactin column.

The TatAyCy_S complexes expressed in *E. coli* and isolated from the membrane fraction using detergent (2% DDM) solubilisation. The membrane fraction was purified on a 5ml Q-Sepharose and StrepTactin. The flow-through (FT) and wash (W) fractions for each column are labelled as either Q or S. The Q-Sepharose elution (QE) and Streptactin elutions (SE) are shown.

The EM image shows that the TatAyCy_S complexes isolated from *E. coli* are in the range of 8-14nm, similar to the TatAyCy^{WT} complexes isolated from *B. subtilis* that was studied in this chapter. The control experiment shows no non-specific targeting of H_Nano-Gold to the strep-tag of TatAyCy_S.

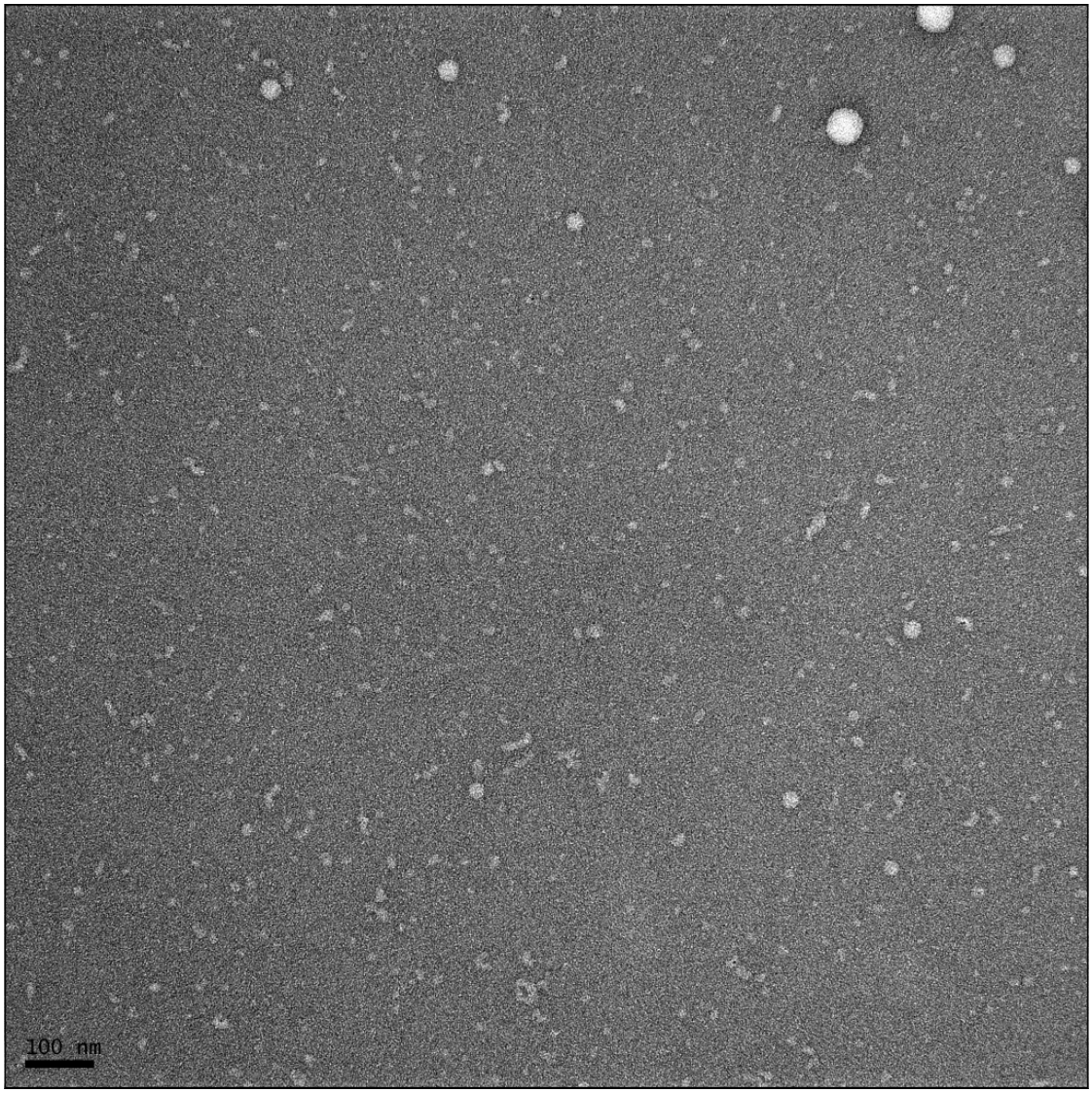


Figure 4.19: TEM images of TatAyCy_S labelled with anti-His Nanogold

The purified TatAyCy_S complex was incubated H_Nano-Gold at a ratio of 1:10 overnight at 4°C. The sample was then separated from free H_Nano-Gold using a 5 ml superdex200 column. Peak elution of TatAy-strep was visualised on a TEM at 40000X magnification and staining with 1% uranyl acetate. The scalebar is shown at 100 nm. The negative control shows no H_Nano-Gold bound to the complexes, however the 1:2 dilution of causes larger aggregating TatAyCy_S complexes to form.

4.13 Discussion

The TatAyCy complex is endogenously expressed in *B. subtilis* with its homolog TatAdCd. The TatAyCy complex differs from the TatABC complex in *E. coli* by multiple factors. The most obvious is the lack of TatB equivalent protein, and so like TatAd, the TatAy is thought to be bifunctional (Barnett *et al.*, 2008). More importantly, although TatAy, TatAd and TatAc are not interchangeable in *B. subtilis*, when expressed in *E. coli,* their specificity is more relaxed and hence forms recombinant complexes (Eijlander *et al.*, 2009; Monteferrante, 2012).

This chapter focused on understanding the structural effect of a single P2A mutation in the TatAy protein in the assembly of the TatAyCy complex. Early purification results showed TatAy co-eluting with the TatCy forming a complex as expected. $TatAy^{P2A}$ bind TatCy_H more abundantly and so the stoichiometry of the TatAyCy is also altered. As observed with TatBC, with 1:1 stoichiometry in *E. coli* (Oates *et al.*, 2005), here it is clear the $TatAy^{P2A}$ is greater than TatCy_H. With $TatAyCy^{WT}$ complexes, the abundance of co-purified TatAy is fairly low as seen in TatAdCd complex in chapter 3. Earlier finding showed that mutation within the first 5 residues of TatA protein influence the binding to TatC (Blaudeck *et al.*, 2005). Similarly, the proline residue in TatAy may also be important in the binding affinity of TatAy to TatCy.

TatAyCy also form two distinct complexes in the resting membrane; TatAy complex and TatAyCy. The size range of these complexes is more homogeneous in size compared to TatA and TatABC, where TatAyCy and TatAy form average complexes of 230 kDa (van der Ploeg *et al.*, 2011). From the gel filtration chromatography the size estimation of the $TatAyCy^{WT}$ complexes approximately 370 kDa and with the $TatAy^{P2A}Cy$ it was >600 kDa (Figure 4.3). The results shown here are not consistent with the earlier publication and are likely to vary with purification methods employed (van der Ploeg *et al.*, 2011). The increasing abundance of the $TatAy^{P2A}$ protein is proportional to the increasing size of $TatAy^{P2A}Cy$ complexes. This shows $TatAy^{P2A}Cy$ complex increase in size as a result of excessive $TatAy^{P2A}$ binding. In addition, the results in this chapter look at the His-tagged $TatAy^{P2A}$ which lead to

unusual cleavage of the His-tag, not observed with Strep-tagged TatAy^{P2A}Cy (van der Ploeg *et al.*, 2011).

The specificity for the TatAyCy complex is stringent where even the PhoD substrate of AdCd cannot be translocated, however when expressed in *E. coli* TatAyCy can freely translocate substrates larger than normal (Barnett *et al.*, 2008). This relaxed specificity of TatAyCy is also seen when replaced by TatAc protein and expressed in *E. coli* (Eijlander *et al.*, 2009).

Structural analysis by EM and AFM of TatAyCy^{WT} produced round complex ranging in size between 8-14 nm in diameter and 4-5nm in height. The Tat complexes were easily differentiated from other aggregation by the distinctive W-shape density profile that is produced. The result of these complexes is consistent with the TatAdCd complexes observed chapter 3. The size estimation of the complexes from the gel filtration shows TatAyCy (340kDa) forms larger complexes that of TatAdCd complexes (270kDa). The TatAy^{P2A} mutation forms long fibril-like structure with a helical organisation. Upon closer inspection, the helical organisation appeared more so as assembly of individual complexes, associated together to form a single fibril.

The density profile of the fibril suggests smaller complex of 4-7 nm, which would be made up of approximately 3-4 TatAy^{P2A} proteins that assemble together. The fibril organisation is largely constituted of the TatAy^{P2A} complexes however; it appears that the TatAy^{P2A}Cy complex would be required initiates the start of this constitutive recruitment event. It is the TatAy^{P2A} protein in the complex that recruits other TatAy^{P2A} proteins continuously. When the sample is concentrated, the frequency of the fibril formation is increased due to increased availability of TatAy^{P2A} (Figure 4.13).

It appears that the helical s results from chirality, which leads to the individual complexes assembling in an end-to-end organisation. Detailed profiling of the fibril shows the complexes forming the fibril have a consistent complex diameter ranging from 9-11 nm. Moreover, density profile of these individual complexes suggests the complexes are assembled in multiple orientations.

There has been previous reporting of the TatA tubule like structure, where the TatA forms tubule within the cytoplasm (Berthelmann *et al.*, 2008a). In this study, the tube formation was not result of a mutation but rather the authors suggest this occurrence may be typical behaviour of the Tat machinery that is disrupted during complex isolation from the membrane. From the size of the round complex and the fibrils observed with $TatAy^{P2A}Cy$, it can be assumed that TatAyCy is required for fibril formation. The requirement for TatC was also disputed by Berthelmann *et al.* (2008a)

Using H_Nano-Gold, specifically targeted to the C-terminus of TatC, which would normally be found on the cytoplasmic side of the membrane, allow complexes that constitute the fibril can be identified. Although, the H_Nano-Gold was not as efficient due to its partial sample coverage, it has provided some insight into the fibril structure. The fibrils are not labelled with the H_Nano-Gold, while the round complexes shown to be TatAyCy were labelled. This finding supports the theory of TatAyCy is the initiator complex that excessively recruits the TatAy proteins.

The expression of mutant TatAd in TatB-lacking *E. coli* cells has shown multiple residues in the periplasmic region are crucial to compensate for the lack of TatB (Barnett *et al.*, 2011). This suggests that the TatB function of bifunctional TatAd may be associated within the N-terminus region. Additionally, the function of TatB is not clearly understood to date. Also this region has been identified to be significant in the TatA protein and its affinity for TatC Blaudeck *et al.* (2005). It has been shown that multiple residues in the periplasmic region of are involved preventing translocation as opposed to a single residue, suggesting periplasmic region is structurally important.

The $TatAy^{P2A}$ mutant changes the structure by reducing the bends in periplasmic region, as result of removing the proline residue and increasing the hydrophobicity by introducing an alanine residue. This results in recruitment of the multiple TatAy complexes as seen in form of fibrils. It can be hypothesised the TatB region of the bifunctional TatAy protein is involved in the regulation of TatAy recruitment during pore-formation. Hence it also questions whether the TatB protein in the TatABC system is responsible for the recruitment of TatA and maintains the interaction

between TatA and TatC. This would agree with the cross-linking studies that show the interaction between TatA and TatB and that with TatB and TatC (Barrett & Robinson, 2005; De Leeuw *et al.*, 2001; Rollauer *et al.*, 2012).

Chapter 5

Interaction of the TatABC complex with the substrate DmsA using SPR

5 Interaction of TatABC complex with substrate DmsA using Surface Plasmon Resonance

5.1 Introduction

The Tat machinery was named for its specificity for the twin arginine residues in the signal peptide motif of the Tat substrates (Berks, 1996; Cristóbal *et al.*, 1999). It has also been shown in previous publications, the mutations within the twin arginine residues; prevents translocation of the substrates by a poorly understood proofreading mechanism (Chaddock *et al.*, 1995; DeLisa *et al.*, 2003). However single residue substitution of the RR motif can be tolerated in *E. coli* (DeLisa *et al.*, 2002; Mendel *et al.*, 2008; Stanley *et al.*, 2000)

DmsA is one of the three subunits of the enzyme, Dimethyl Sulfoxide Reductase, found in *E. coli*. The DmsA and DmsB subunits are both hydrophilic and bound on the cytoplasmic face of the membrane. Additionally, DmsA can transport across the membrane by the TatABC system with its partner, DmsB, by essentially "piggy backing" on to the DmsA protein. DmsA has a 45 residues signal peptide with a SRRGL signal motif (Palmer *et al.*, 2005; Weiner *et al.*, 1998).Within this signal motif the RR residues are located at position R16 and R17 of the signal peptide. The signal peptide has a RR containing N-region, a hydrophobic region in the centre and C-region with the signal peptide cleavage site. It is the RR residues that are thought to be responsible for the targeting and recognition by the Tat system.

In this chapter we aim to understand the substrate-binding event in further detail. The kinetic of the substrate binding has not been previously studied for the Tat substrates and hence the aim of this chapter is to identify how the specificity of the substrate (DmsA) is involved in the binding to the Tat(A)BC (substrate-recognition complex). Surface Plasmon Resonance (SPR) was used as a tool to measure the interactions between the signal peptide (sp) of DmsA and TatABC, in real-time. Using this technique we can understand how the specificity may affect the binding event, in whether binding occurs and if so, the affinity of the interaction. The binding affinity

of the Tat complex *in vitro* is relatively low to that expected *in vivo* due to the presence of detergent during protein purification. The presence of the detergent can drastically influence weak electrostatic interactions that may form during the signal peptide binding. Therefore to reduce the effect of detergent, the purified TatABC complexes were reconstituted into liposomes to mimic the *in vivo* environment.

SPR is a quantitative technique that measures binding kinetics in real time. The principle of SPR is dependent on the mass of sample immobilised to the gold surface of the chip. Subsequent change in the refractive index as a result of bound molecules on the chip surface; a sensogram showing the rate kinetic of the binding is produced.

5.2 Significance of the RR residues in the substrate

The signal peptide of DmsA (spDmsA) was tagged with an YFP protein on the C-terminus. Additionally, the N-terminus of spDmsA was labelled with Strep-tag (S), to give the construct (S_spDmsA^{RR}_YFP). Another construct with the RR residues mutated KK was also used in this study named S_spDmsA^{KK}_YFP.

Previously it has been shown that most Tat substrates undergo cytoplasmic clipping, and, as a result, this leads to the premature clipping of the signal peptide. By introducing a Strep-tag on the N-terminus of the spDmsA, any prematurely clipped signal peptide is not purified with the Strep-tag, when the tag is cleaved off with the signal peptide.

To confirm that the YFP labelled protein is still capable of being transported by the TatABC_S, the constructs were transformed in WT and Δ-Tat *E. coli* cells. The cells were then induced with arabinose for 3hrs, allowing sufficient time for either S_spDmsA^{RR}_YFP or S_spDmsA^{KK}_YFP expression, and transport to the periplasm. To find the localisation of the protein the cells was fractionated into 3 compartments, cytoplasm, membrane and periplasm. The fractionation step was achieved by using lysozyme to disrupt the peptidoglycan layer to release the periplasm component. The periplasm was collected while the spheroplasts were subjected to sonication, to separate the soluble cytoplasm from the membrane by ultracentrifugation. The

membranes were then solubilised in 1% DDM for 1 h to isolate the membrane proteins. The samples were run on a SDS-PAGE gel and stained with silver stain to confirm the efficiency of the fractionation. As shown in the silver stained gel in Figure 5.1 a, the membrane and the cytoplasmic fraction are heavily dense in proteins. The efficiency of fractionation is assessed by the searching for cross contamination of the sample between fractions and periplasm specific bands (20 kDa, 25 kDa and 50 kDa). In some cases, during the lysozyme treatment, the cytoplasm can leak out of the inner membrane, which then can lead to the contamination of periplasmic fraction, showing more protein bands than expected.

A western blot was also performed in complement to determine the localisation of the precursor S_spDmsA_YFP and mature YFP. The localisation was observed by western blots using anti-GFP antibody. Figure 5.1 a shows that in both WT- and Δ-Tat cells, the S_spDmsA^{RR}_YFP protein of 29kDa is present in the cytoplasmic and the membrane fractions. The mature YFP (25kDa) is also present in all fractions when expressed in WT-Tat cells. The presence of the mature YFP in the cytoplasm and the membrane prior transport is a result of premature clipping. However. The mature YFP found in the periplasm of the WT-Tat cells is from cleaving after the translocation of the protein. When S_spDmsA^{RR}_YFP is expressed in Δ-Tat, the premature clipping is seen again at 25kDa, similar to that in WT-Tat (Figure 5.1 b) . However in the absence of the Tat machinery (Δ-Tat), the S_spDmsA^{RR}_YFP protein cannot be translocated, confirming spDmsA is only target to TatABC and not other secretory pathways such as Sec. When the mutated substrate, S_spDmsA^{KK}_YFP was expressed in both the WT and Δ-Tat, in both cases, no YFP was detected in the periplasm (Figure 5.1). This result shows that the lysine substitution in the DmsA signal peptide can block the transport of the YFP protein, with the specificity for transport relying on the RR residues. The S_spDmsA^{KK}_YFP protein is expressed in the cytoplasm and is even found in the membrane fraction with premature clipping forming unfolded YFP as previously seen. Premature clipping in the cytoplasm and at the membrane forms a similar size YFP protein of 25 kDa but the difference between this and that found at the periplasm if the conformation, where in periplasm it is fully folded.

The export assay was able to confirm that in the *E. coli* cells carrying the TatABC machinery, can only export Tat substrates with the twin arginine residues but the binding to the membrane prior export is not well characterised. As shown in the export assay the S_spDmsA^{KK}_YFP precursor is also localised to the membrane showing that the targeting of the signal peptide is not dependent on the SRRXFLK conserved motif itself. The interaction between the TatABC machinery and the substrate therefore needs to be further characterised.

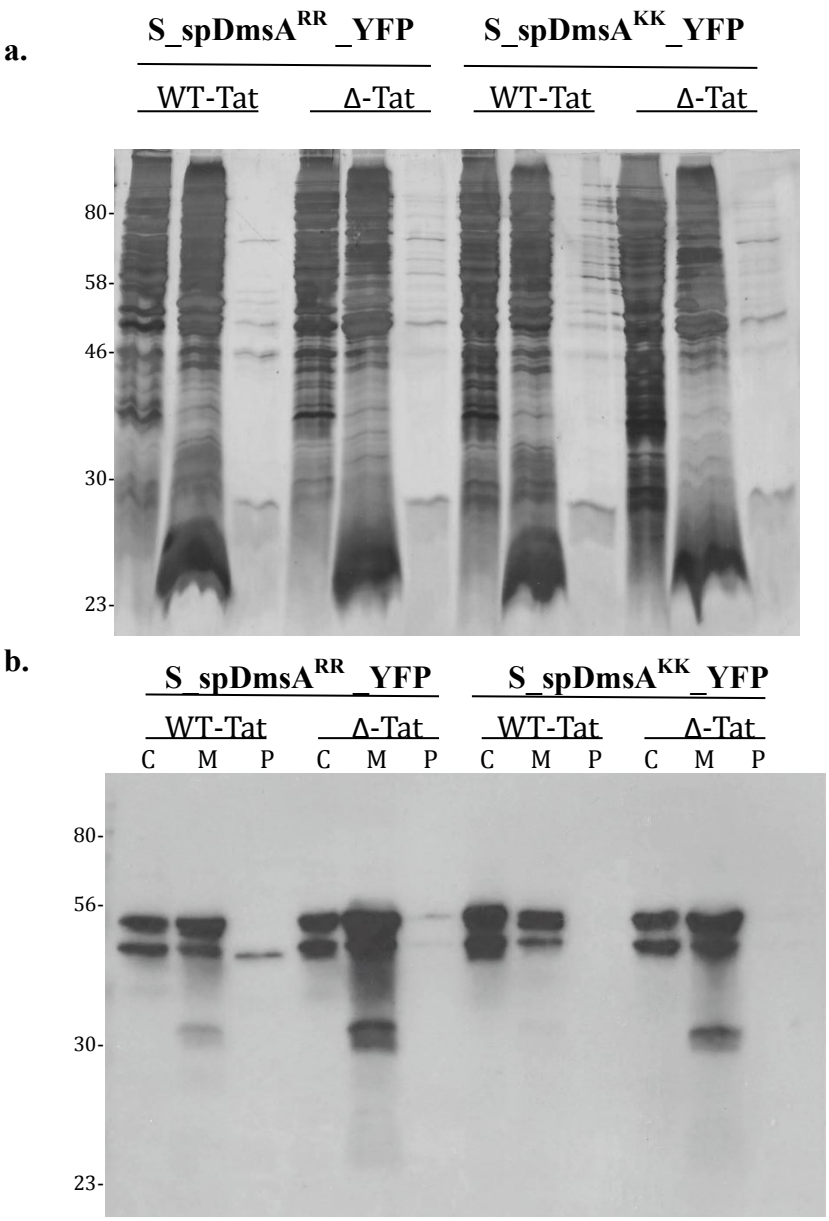


Figure 5.1: Silver stain and Western blot of an export assay for the translocation of S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP in WT- and Δ-Tat *E. coli*

E. coli cells with TatABC (WT-Tat) and without (Δ-Tat) were expressed with either S_spDmsA^{RR}_YFP or S_spDmsA^{KK}_YFP. The cells were fractioned into cytoplasm (C), membrane (M) and periplasm (P) and run on a SDS-PAGE gel. The gel was stained with siver-stain (a) and western blotted against α-GFP (b). The silver stained gel showed the fractions were isolated without cross contamination between fractions. The western blot shows only RR DmsA-YFP can be translocated to the periplasm in WT-Tat cells. The KK DmsA-YFP protein is present in the periplasm. The cytoplasmic and membrane fraction show premature clipping resulting in the following band at 25 kDa and 29 kDa.

5.3 Purification of the substrates

The S_spDmsA^{RR}_YFP or S_spDmsA^{KK}_YFP construct in the pBAD24 plasmid was transformed in Δ-Tat cells to expression in the cytoplasm by restricting the transport of the protein across the membrane to purify protein with the signal peptide. The cell were induced with arabinose for 3 h and then extracted from the cytoplasm by sonication and ultracentrifugation. The cytoplasmic fraction was incubated with avidin to prevent non-specific binding on the column. Then, the fraction was then applied to a 5ml StrepTactin column and washed with buffer to remove unbound proteins before elution. The elution samples were run on a SDS-PAGE gel and stained with Coomassie blue. The protein concentration required to study the binding using SPR must be minimally in the range of 0.5-1mg/ml. Therefore proteins elution from StrepTactin affinity chromatography were not subjected to further purification by gel filtration chromatography, as this would dilute the proteins further. The elution of the S_DmsA_YFP precursor shows only a single band Figure 5.2, as any premature clipping like that observed in Figure 5.1 was purified out as a result of strep-tag cleavage. The western blot using α-strep antibody confirms the single band as the S_spDmsA_YFP.

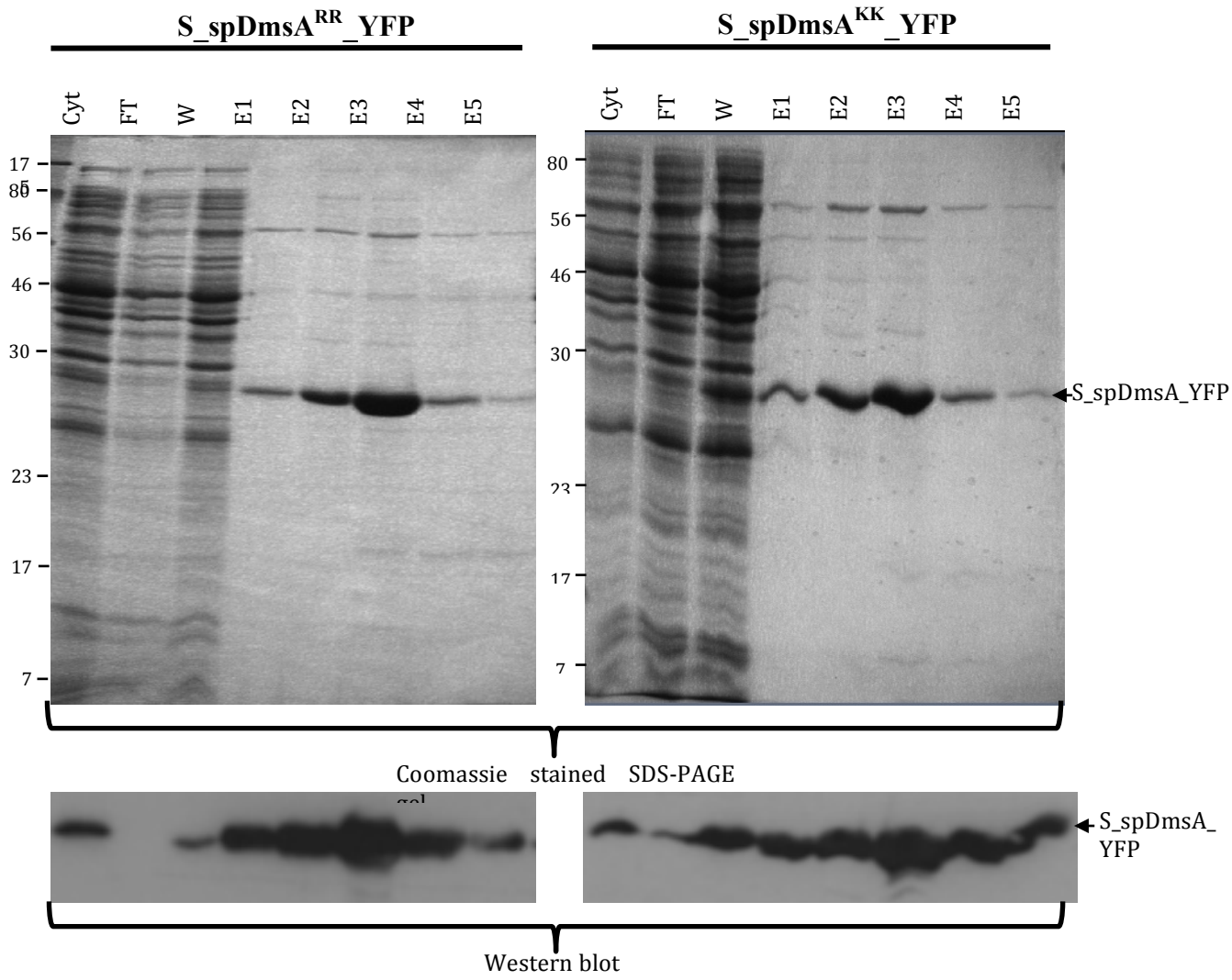


Figure 5.2: Coomassie stained gel of S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP purification on a StrepTactin column with a corresponding western blot against α-strep.

S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP were overexpressed in *E. coli* and isolated from the cytoplasmic fraction. The cytoplasmic fraction was incubated with avidin and then applied to a 5ml StrepTactin column. The flow-through (FT) was collected and the column was washed with 5 column volumes of buffer (W). The protein was eluted in 5 fractions and loaded on a SDS-PAGE gel. The gel was then stained with Coomassie and western blotted against α-strep. S_spDmsA_YFP is eluted at 25kDa.

5.4 Purification of TatABC_S

The TatABC_S complex was expressed from the pBAD24 vector using an arabinose-inducible promoter to allow regulated expression of the membrane proteins in Δ-Tat *E. coli*. As found in chapter 3, inducing the expression of the TatABC proteins during the stationary phase, increases the level of expression. Therefore cells were grown at 37°C until the optical density (OD) reached 0.4 and then induced with 1mM arabinose. The cells were induced for 3 h before isolating the membrane. The membrane fraction was solubilised in 1% DDM before purification on the StrepTactin column. The DDM concentration in the running buffer, for the purification, was reduced to 0.02% (2X CMC). The samples collected from the purification are shown in Figure 5.3. Purification of the TatABC_S complex shows the TatA (17 kDa), TatB (25 kDa) and TatC_S (28 kDa) proteins are co-eluted (Figure 5.3). This purification resulted in higher ratio of associated TatA, which is usually variable during each purification procedure. The *tatA* gene is expressed in a high abundance, but majority of this TatA forms homo-oligomeric complexes and a small proportion of this is associated in the TatABC complex. This is shown by the relative ratio of TatA isolated in the membrane fraction. The dissociation of the TatA protein in the wash fractions also suggests that the interaction between TatA to the TatC_S is fairly weak, which is expected for a complex that change dynamics during protein translocation.

A western blot using antibodies against all three Tat proteins (TatA, TatB and TatC_S) was carried out to show the presence all three proteins in the complex during the purification. The elution pattern seems to follow the same trend for all three proteins, showing that the proteins in the TatABC are co-eluting in a stable complex, where the proteins do not seem to dissociate from the complex. The elution in the Coomassie stained gel shows a fairly pure elution and so the sample was not applied to a gel filtration for further purification, to reduce dilution of the sample. The purification of the TatABC for liposome reconstitution required a high yield of protein, so the peak elution from the StrepTactin column was used directly.

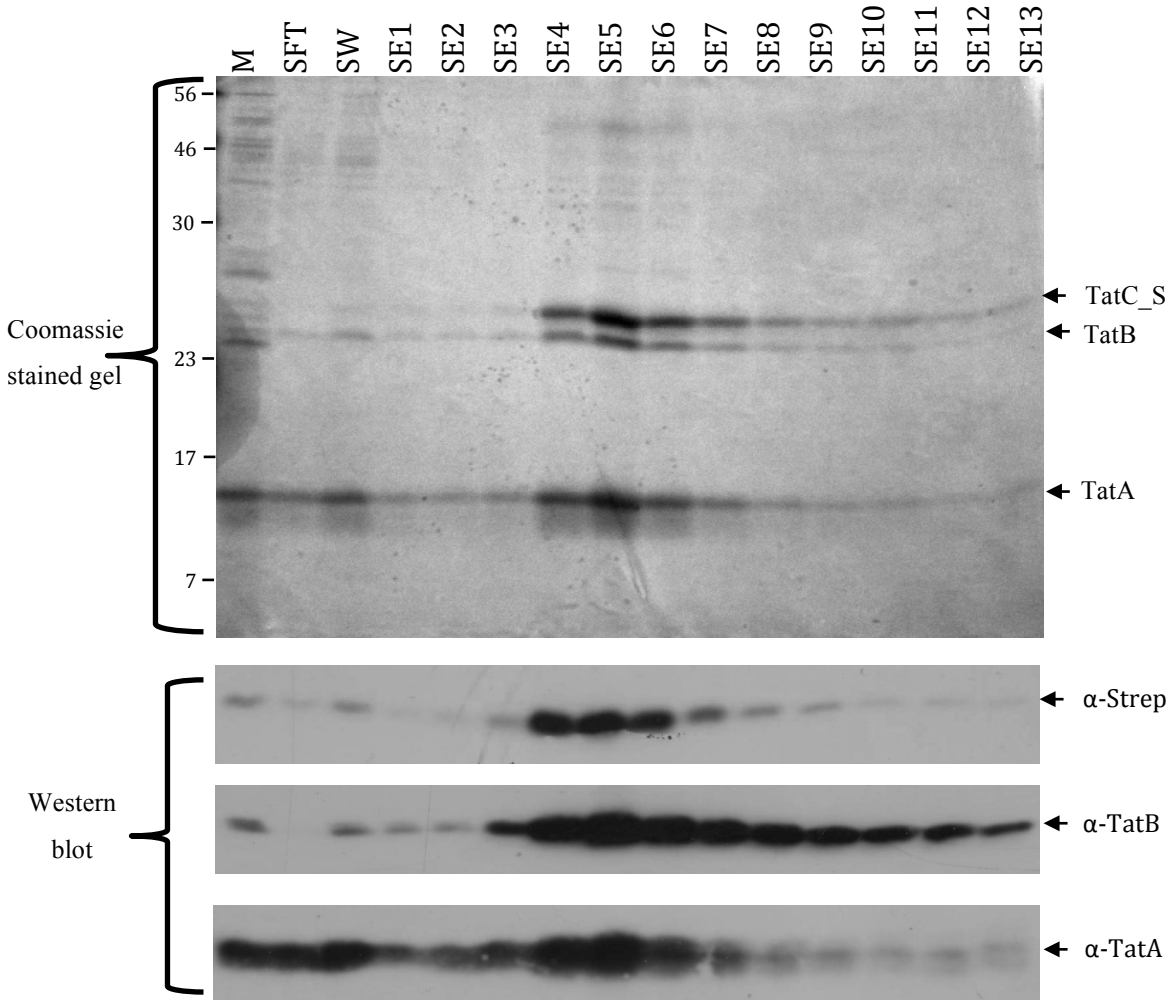


Figure 5.3: Coomassie stained gel and western blot of TatABC_S complexes purification from the *E. coli* membrane fraction

TatABC_S was expressed in Δ-Tat *E. coli* for 3 h and then isolated from the membrane by solubilisation in 1% DDM. The solubilised membrane fraction (M) was applied onto a 5ml StrepTactin column and collected from the flow-through (FT). The column was washed in 5-column volume of buffer (W) before eluting with desthiobiotin containing buffer. The elutions were collected in 1ml aliquots (QE1-13). The fraction were then run on a SDS-PAGE gel and stained with Coomassie. The western blot against α-Strep α-TatA and α-TatB confirm the present of each subunit.

5.5 Liposome reconstitution

Once the TatABC_S complex was purified, it was reconstituted into liposomes made from *E. coli* lipids to keep the complexes similar physiological condition to that *in vivo*. The *E. coli* lipids were rotary evaporated to form a thin layer of lipid film and which was then hydrated in buffer to form multi-lamellar liposomes. The liposomes were then subjected to 5 cycles of freeze-thaw to form larger multi-lamellar liposomes and then extruded through 600 nm and 200 nm membranes in a total of 30 cycles, forming uniform size of liposomes. It is not always possible to get liposomes to complete uniformity, but the purpose of extrusion was to reduce the liposomes size, so that it does not exceed 200 nm. Liposomes larger 200 nm would cause issue in the flow dynamic of the SPR machine. Once the liposomes were formed, Triton-X-100 was added to disrupt the lipid bilayer slowly. Upon gentle agitation, TatABC_S was introduced into the disrupted liposome solution, allowing the protein to be incorporated into the liposome. With protein reconstitution, it is difficult to control the direction of protein insertion, nevertheless by using Triton-X-100, uni-directionality can be promoted during the reconstitution. Excess detergent in the sample was then removed using SM2 Biobeads, which have small pores within the bead, capable of capturing the detergent molecules. The gradual addition of the Biobeads allowed the liposomes to reform, and the proteins to be incorporated without causing aggregation. The proteoliposomes with TatABC_S incorporated (referred to as TatABC-liposomes henceforth) were then isolated from the biobeads and concentrated by ultracentrifugation. The liposomes pellet was then resuspended into a buffer to give the liposomes a final concentration of 2mg/ml.

Liposomes are not very robust and have a short life span and so to assess the integrity of the liposomes after exposure to detergent and the ultracentrifugation step, liposomes were observed by Cryo-TEM. With negative stain microscopy, the heavy metal stains deform the liposomes and leads to flattening of the liposomes; hence Cryo-TEM was used to evaluate the structure of the liposomes. Following protein reconstitution, the liposomes remain stable. The size of the liposomes after extrusion on a 200 nm membrane however resulted in liposomes around 70nm on average (Figure 5.4). This confirmed the liposomes were stable after harsh treatment.

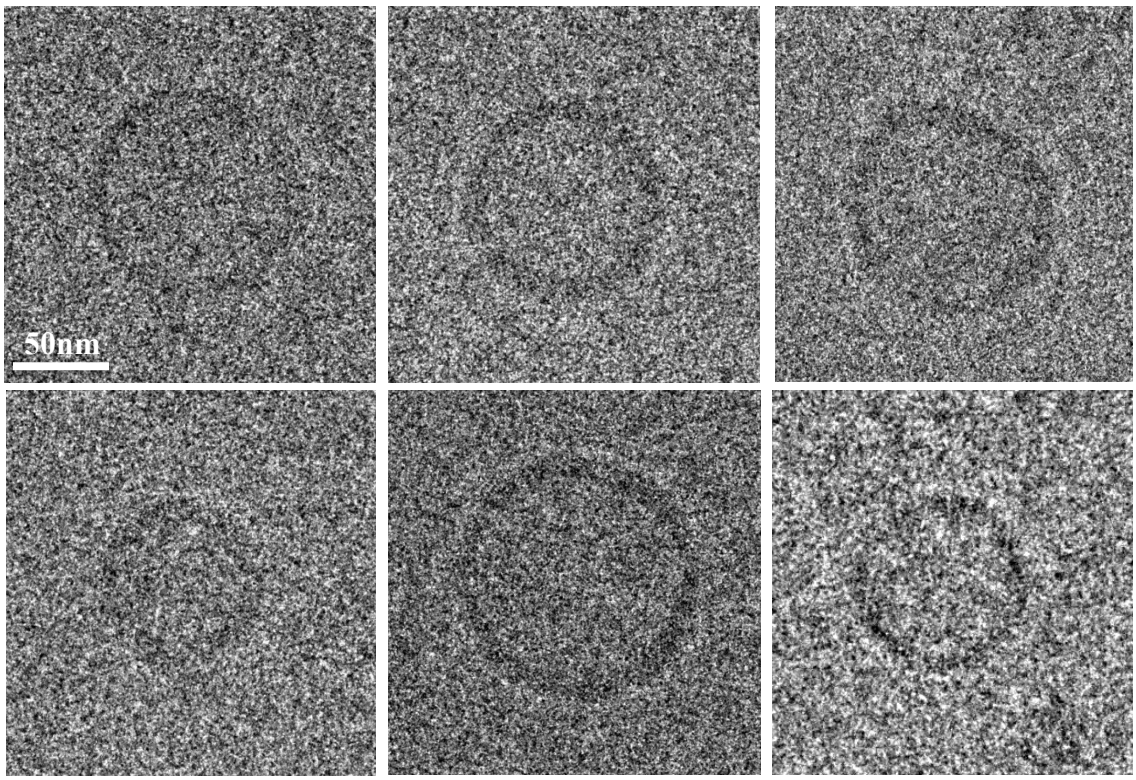


Figure 5.4: Cryo-EM micrograph of liposomes made of *E. coli* lipid after extrusion and centrifugation.

The TatABC-liposomes were extruded trough 200nm with 30 cycles and centrifuged to concentrate the liposomes before imaging the liposomes by Cryo-EM. The liposomes were instantly frozen in ice using liquid ethane and visualised in ice without any stain. The liposomes were images at 40000X magnification.

5.5.1 Sucrose Density Gradient Centrifugation of the TatABC proteoliposomes

To evaluate the efficiency of protein reconstitution the sucrose density gradient centrifugation method was used. The principle relies on the buoyancy of the liposomes, which should separate the free protein from that incorporated into the liposomes. The centrifugation tube was prepared with 2 concentrations of sucrose solution. The first was 40% sucrose with the liposomes suspended, then topped with 20% sucrose followed by liposomes buffer only (0% sucrose). The centrifugation tube had 3 layers; bottom layer containing the TatABC-liposomes or liposomes, 20% sucrose layer and buffer only on the top layer. When centrifuges the liposomes were float to the meniscus of 20% layer during ultracentrifugation while the free proteins remain at the bottom of the tube. This is represented in a schematic of liposome floatation before and after centrifugation (Figure 5.5a).

After centrifugation, a cloudy layer containing the liposomes is formed between the 0-20% sucrose layers (Figure 5.5a.). Aliquots from each layer were run on a SDS-PAGE gel and western blot was used to show large amount of the TatABC_S reconstituted into the liposomes. Some TatABC_S remains free in detergent micelles and is not floated to a lower concentration of sucrose layer because of the high density of the protein. Majority of the TatABC-liposomes are found in the 20% sucrose layer and some found in 0% sucrose on the top. This is likely to be seen with smaller liposomes that have lower density. The control sample with no protein reconstituted has no protein is detected by western blot but the liposomes still floated to the middle 20% layer (Figure 5.5). Although protein reconstitution is successful the concentration of TatABC-liposomes is difficult to determine. This lack of information and variability in liposomes size, limits the kinetic of TatABC interaction to be calculated. Therefore the SPR data presented here does not show quantitative data.

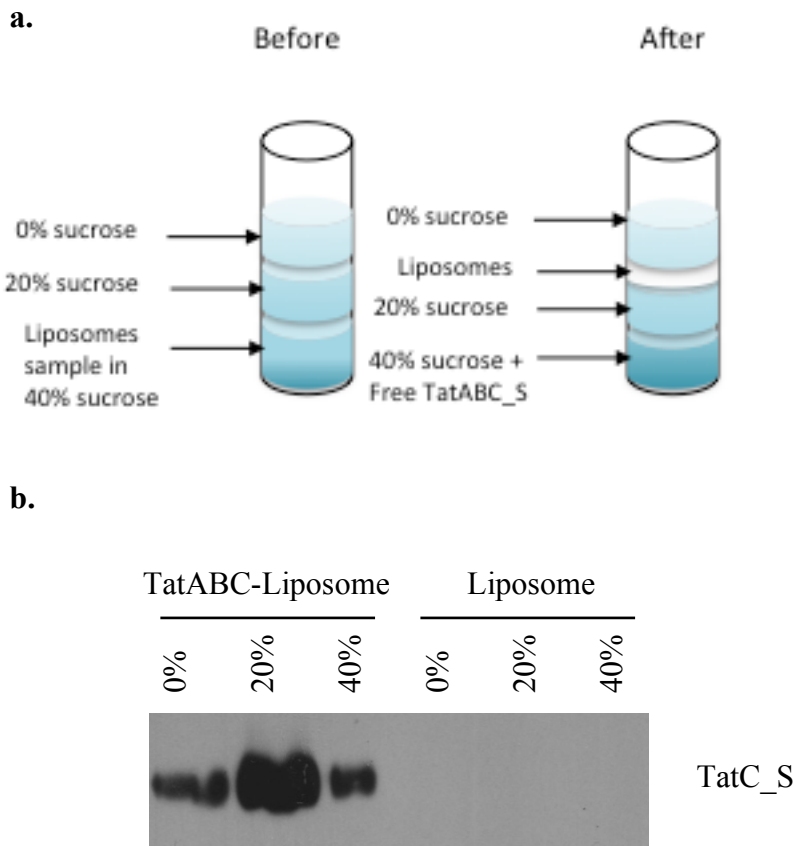


Figure 5.5: Western blot the separation of TatABC-liposomes by sucrose density gradient centrifugation and the schematic diagram of the process.

The liposomes were reconstituted with TatABC_S using lipid disruption method. Upon formation of TatABC-liposomes, the liposomes were separated using different concentration of sucrose of 0%, 20% and 40% as shows as a schematic diagram (a). After centrifugation the aliquots from all layer sucrose was run on a SDS-PAGE and the proteins was identified on a western blot using anti-Strep antibody (b). TatABC_S that was not reconstituted in the liposome is pelleted on the bottom 40% layer, while the liposomes float between the 0-20% layers where majority of the protein is reconstituted. Control liposomes were used in sucrose gradient to confirm the liposomes are made correctly to maintain their buoyancy.

5.6 Substrate binding to using immobilised substrate

5.6.1 Immobilisation of S_DmsA_YFP onto a GLC-chip alliance

Surface Plasmon Resonance (SPR) was used to study the binding of a recombinant Tat-signal peptide to the TatABC_S complex. The construct S_spDmsA^{RR}_YFP was previously constructed to encode a Strep-tag upstream of the DmsA signal peptide, followed by an YFP gene. The signal sequence was mutated to substitute the vital RR residues for KK residues to produce a potentially non-binding construct S_spDmsA^{KK}_YFP. As shown in section 5.2, the substitution of the RR residues to KK blocks the translocation. This section attempts to determine if the mutation results in a change in binding affinity. The TatABC complex is composed of membrane proteins that are isolated using detergent, which makes them difficult use on the SPR system in their isolated form. Initially, the preliminary studies between TatABC_S complexes and the substrate by SPR were performed but due to the high interference from the buffer density, the interaction between Tat and the substrate was masked. Therefore, in an attempt to gain greater difference in density between the ligand and the analyte, liposomes were used overcome the high reading produced by the buffer. For this reason, the Tat complexes were reconstituted into liposomes. Two methods were employed to investigate the binding between the substrate and the Tat complex; firstly immobilisation of the substrate, and secondly, liposome immobilisation method.

The SPR was carried out on the ProteOn XPR36 system, using a GLC chip. This chip is specifically designed for the amine coupling of the proteins, allowing immobilisation on to the chip. The GLC chip has the capacity to bind a single monolayer of proteins. The system is set up to run buffer and samples over the chip. A standard running buffer used in this system is PBS/tween (0.005%), however when liposomes were used, the buffer was exchanged for 20 mM Tris-HCl and 150 mM NaCl, to stabilise the liposomes and the proteins.

First the chip is activated for amine coupling, using the ethyl-dimethylaminopropyl carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) reagent mix, which produces reactive carboxylic acids. The reactive carboxylic acid then forms amine bonds with

any proteins that are passed on to the surface of chip, thus immobilising the proteins. In this case, anti-GFP antibody was the first immobilised on to the chip by amine coupling. The anti-GFP antibody was diluted to 30ug/ml in PBS buffer with the running buffer PBS/tween (0.005%).

The anti-GFP antibody was then used to capture the substrate S_DmsA_YFP. Anti GFP was used as opposed to YFP as GFP and YFP only differ in a single amino acid at position 203, where the threonine (T) is substituted for tyrosine (Y). This substitution has no effect on the structure of YFP, and so the α-GFP antibody can be used to capture the YFP protein. YFP is a mutant a form of GFP, forming a β-barrel structure with the N-terminus and C-terminus on the same face. The anti-GFP antibody is raised against a synthetic peptide corresponding to the amino acids 132-144 of the Green Fluorescent Protein from jellyfish *Aequorea victoria*. The residues that recognise the anti-GFP antibody are on the opposite side to the N-terminus and C-terminus. This allows the substrate to be immobilised and the N-terminus of YFP with DmsA signal peptide is exposed for further binding (Figure 5.6). The immobilised antibodies were then used as bait to capture the S_DmsA_YFP substrate. Once the ligand is bound, the analyte (TatABC-liposomes) is flowed over the surface of the chip. The change in mass on the surface over a period of time produces a sensogram. The process of immobilisation is depicted in Figure 5.6.

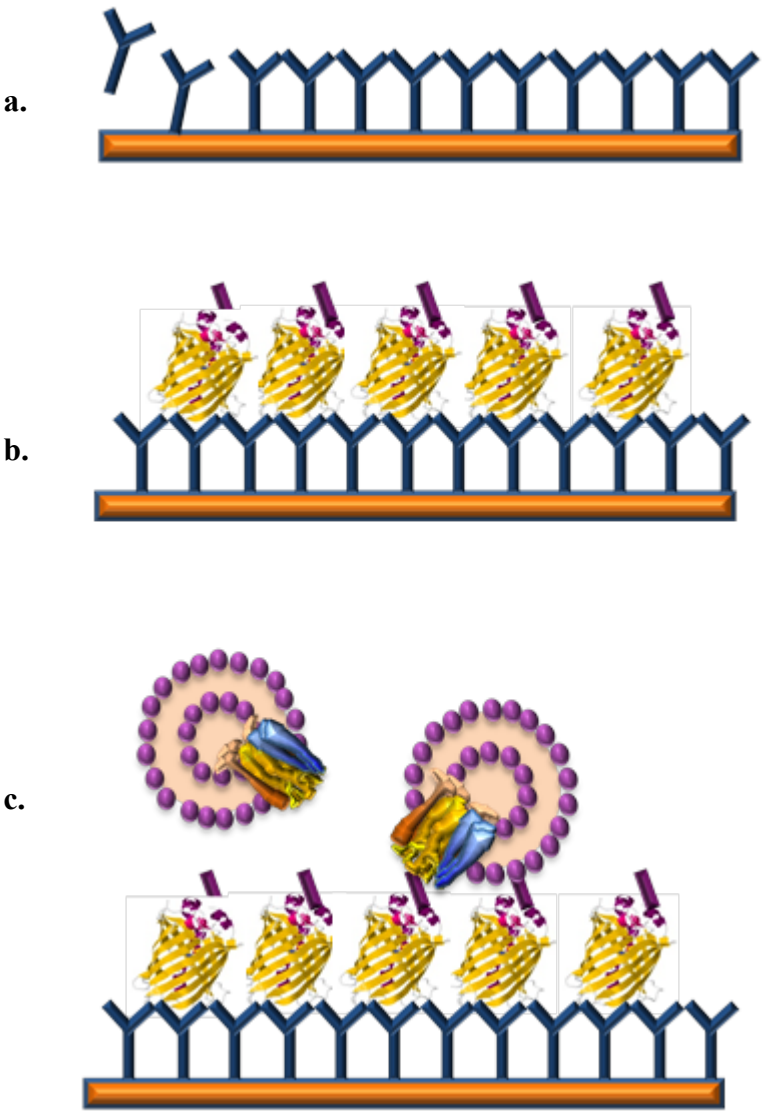


Figure 5.6: Schematic representation of antibody-captured of the S_DmsA_YFP substrate and binding to TatABC-liposomes by SPR

The GLC-chip is activated by a chemical reaction to which the anti-GFP was immobilised through amine coupling (a). The substrate was then immobilised to the same channels of the chip by antibody capture of the anti-GFP to the YFP protein with the DmsA signal peptide (b). Finally the proteoliposomes are then passed transversely across the channel to measure interactions between TatABC and the Strep-DmsA-YFP substrate (c).

5.6.2 Interaction of the TatABC-liposomes to the immobilised substrate

The TatABC-liposomes and empty-liposomes were passed across the chip over 480 seconds at 25 µl/min flow rate. The flow rate of binding was reduced to 25 µl/min when measuring binding to allow the liposomes to form contact with the ligand. The concentration of the TatABC liposomes was varied from 0.2-0.0125 mg/ml with immobilised S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP. Once the data from the sensogram reading was averaged, the buffer was then subtracted from the control channels to produce a binding graph for the interaction (Figure 5.7).

The binding rate for the liposomes produced a linear binding trend in the first 200 seconds and then started to dissociate slowly (Figure 5.7 b). This analyte was passed over the surface for 480 seconds suggesting the initial binding that is observed may result from liposomes binding. The slow dissociation could be a result of loose interactions causing multiple binding of one liposome as it passes across the channel. Overall, the increase from the initial binding of liposomes to subsequent dissociation shows that some liposomes remain bound to the GLC-chip. These are likely to be forming high affinity interactions with the substrate S_spDmsA^{RR}_YFP.

The shallow binding curve for 0.125 mg/ml of liposomes (orange-line) represents lower binding rate with S_spDmsA^{KK}_YFP (Figure 5.7 b). This reduced binding of S_spDmsA^{KK}_YFP to the TatABC-liposomes suggests a small level of difference in binding affinity compared to S_spDmsA^{RR}_YFP. It is difficult to analyse the effect on affinity using this technique, as the kinetics of the binding events cannot be further analysed quantitatively.

In addition, the shallow dissociation curves for S_spDmsA^{RR}_YFP sensorgrams show that where specific interactions do occur, the S_spDmsA^{RR}_YFP dissociates more slowly in comparison to the S_spDmsA^{KK}_YFP (Figure 5.7). This suggests that the interactions between TatABC-liposomes and RR residues occur at a higher affinity than with S_spDmsA^{KK}_YFP.

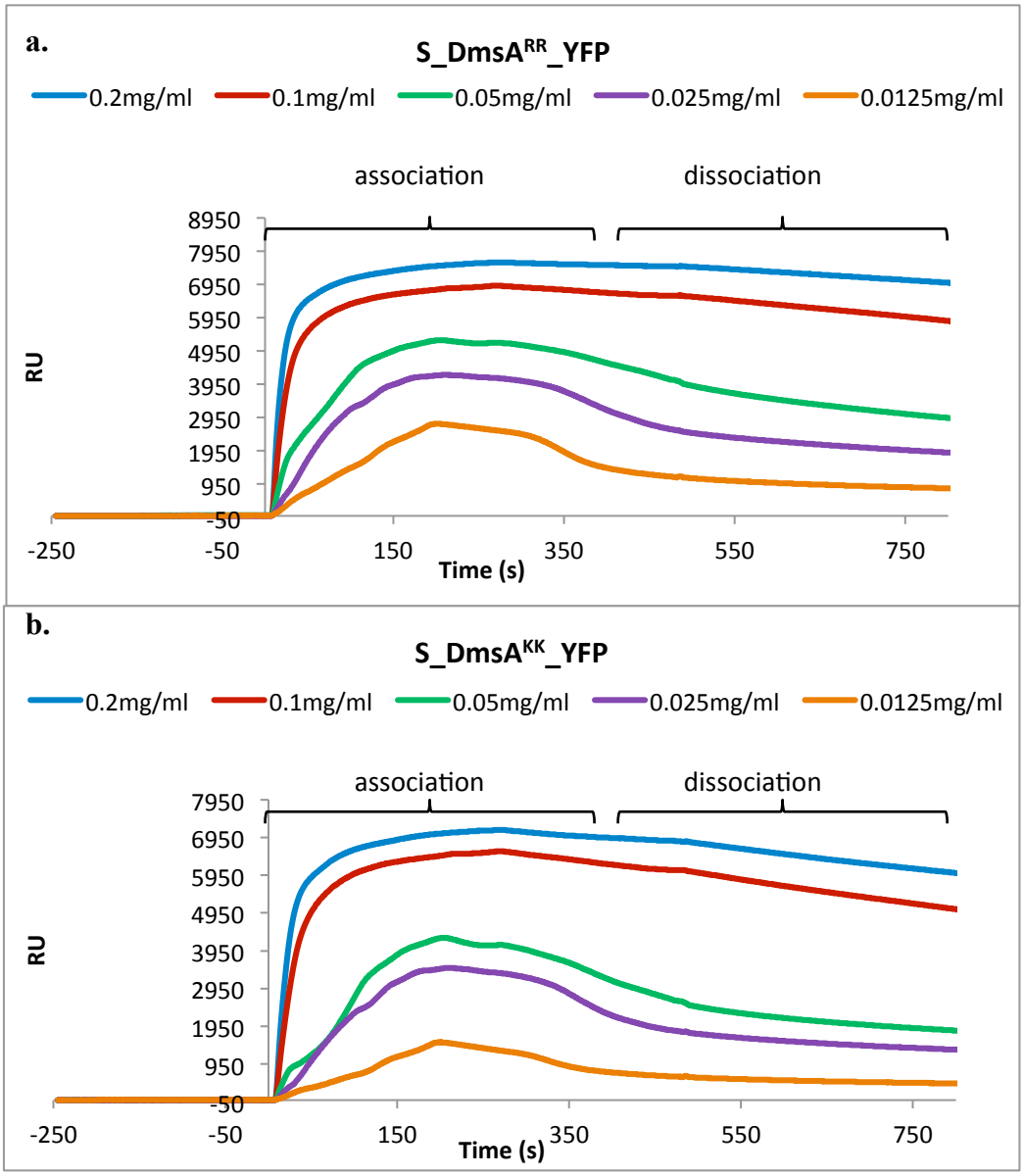


Figure 5.7: The binding of the TatABC-liposomes to immobilised S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP substrate using SPR.

Purified S_spDmsA^{RR}_YFP (a) and S_spDmsA^{KK}_YFP (b) was immobilised to a GLC-chip via anti-GFP antibody capture. A range of TatABC-liposomes (0.2-0.0125 mg/ml) concentrations was then flowed over the immobilised substrate and the binding was measured over 480 seconds at 25 µl/ml flow rate. The data from the sensogram was averaged and the buffer channel was subtracted from the remaining channel to produce normalised binding graph. The GLC chip reached saturation at 0.2 mg/ml. The dissociation of the liposomes was monitored from 460-800 seconds.

The concentration of TatABC-liposome was increased by 2-fold, and as the concentration reaches 0.1 mg/ml, the GLC-chip starts to saturate. When the concentration was further increased to 0.2mg/ml, the increase in binding no longer remains relative to the concentration due to saturation (Figure 5.8). Using lower concentrations in the range of 0.05-0.0125 mg/ml, resulted in a more informative binding data and improved resolution (Figure 5.8). By varying the concentration, it was possible to determine the optimal working concentration for the interaction between ABC-liposomes and S_DmsA_YFP. In addition, it was also clear that TatABC-liposomes binds less readily to S_spDmsA^{KK}_YFP compared to the S_spDmsA^{RR}_YFP. Overall, with increasing concentration of TatABC-liposomes, there is a linear trend that reaches saturation between 0.1-0.2 mg/ml. After 480 seconds, the running buffer was applied across the chip and the dissociation was measured.

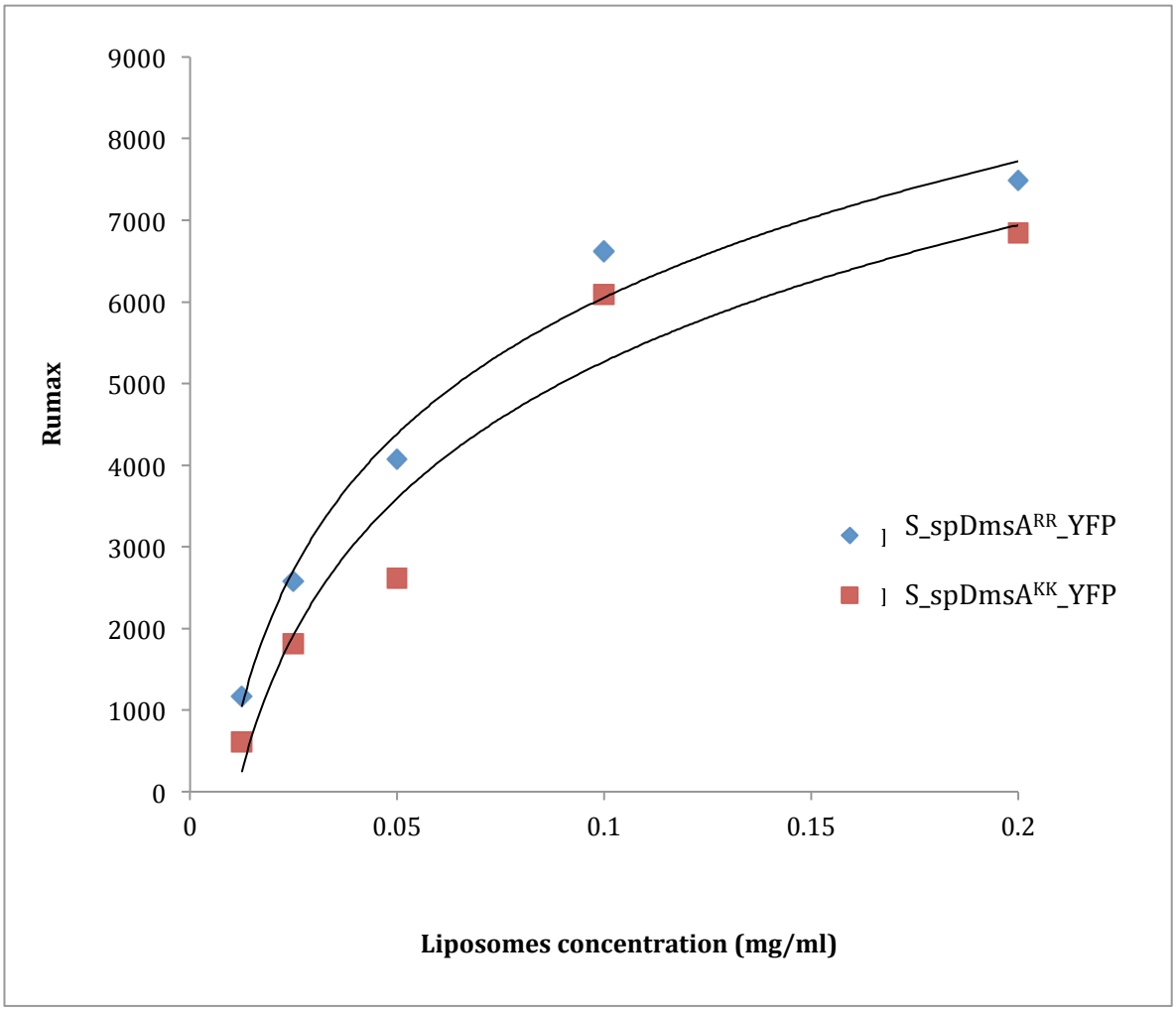


Figure 5.8: The saturation of the GLC-chip with varying concentration of TatABC-liposomes shown by maximum RU at binding equilibrium.

The S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP proteins were immobilised onto the GLC-chip at a concentration of 30 µg/ml. The analyte (TatABC-liposomes) was then passed over the chip at varying concentration, ranging from 0.0125-0.2 mg/ml. The graph for the maximal RU during the association at varied concentrations was produced to show the saturation.

5.6.3 Binding of S_DmsA_YFP substrates to the empty-liposomes

In addition, a control experiment was set up to measure the difference in binding between protein free-liposomes (empty-liposomes) and the TatABC-liposomes. The DmsA-YFP protein were immobilised onto the GLC-chip as described earlier and the empty-liposomes, TatABC-liposomes and buffer was passed over the ligand. The binding to both the S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP was studied by SPR. The earlier findings showed reduced binding of TatABC-liposomes with S_spDmsA^{KK}_YFP (Figure 5.8). Similarly, when the control liposomes (dashed lines) are passed over as analyte, the binding is again reduced to S_spDmsA^{KK}_YFP (Figure 5.9). Therefore this suggests that the reduced binding observed with the KK mutation is more likely due to non-specific interactions rather than specific binding to TatABC_S. The difference in RU between empty-liposomes and TatABC-liposomes in suggests that the concentration estimation for the liposomes is not very accurate as there is less liposome bound to the GLC-chip. This is largely due to the method of liposome preparation during the extrusion process. As the concentration of TatABC-liposome is not comparable to the concentration empty-liposomes, the data are not directly comparable. From this experiment it was evident that binding with S_spDmsA^{KK}_YFP was of lower affinity to the lipid, however this was not due to the presence of TatABC complexes but rather an effect of membrane targeting.

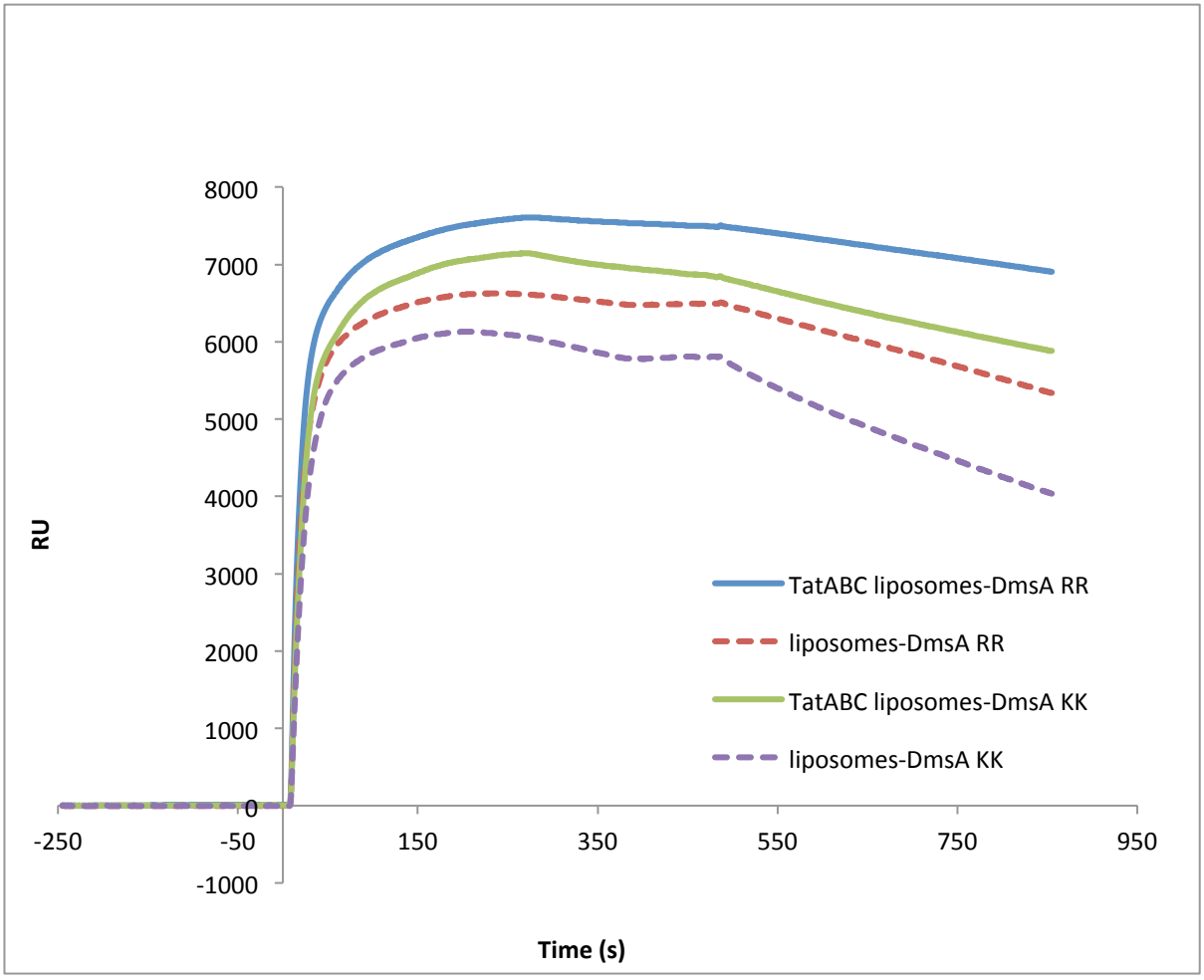


Figure 5.9: SPR sensogram showing the binding of TatABC-liposomes and empty-liposomes to either S_spDmsA^{RR}_YFP or S_spDmsA^{KK}_YFP.

TatABC-liposomes and empty-liposomes were used for the SPR study to measure bind to S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP. The noise from buffer was subtracted from each channel. The binding was measured at 25 µl/ml flow rate over 480 seconds before dissociation. Empty-liposomes have a lower RU compared to the TatABC-liposomes and S_spDmsA^{KK}_YFP binds both liposomes at lower affinity.

5.7 Substrate binding to immobilised TatABC proteo-liposomes

As an alternative method to analyse binding, the liposomes were immobilised on the chip first to confirm similar concentration of liposomes were bound to the chip. This was achieved through the coating of the GLC-chip with a lipophilic C11 reagent (Undecylamine in MES buffer and 10% DMSO) using standard EDC-NHS amine coupling. This relatively new method of immobilising membrane protein is in the final stage of development with the Bio-Rad Research and Development Team (Hemel Hempstead, UK). The liposomes were then bound to the lipophilic GLC-chip to form a monolayer of liposomes. This method of immobilising lipids to the Bio-Rad GLC-chip has not previously carried out with any biological sample in the UK Academic Sector. The TatABC-liposomes and empty-liposomes were immobilised to the chip surface at two concentrations (0.15 mg/mg and 0.3 mg/ml) and the RU of the concentration between both liposomes is comparable (Figure 5.10) showing that similar concentration of liposomes are bound onto the GLC. The RU of 0.3 mg/ml liposomes reached a maximum of 5500 RU and for 0.15 mg/ml a maximum of 5000RU (Figure 5.10). As the concentration of the liposomes was doubled to 0.3 mg/ml, the RU is not doubled suggesting that chip may be fully saturated with a single layer of liposomes. Using this method, the concentration of empty-liposome and TatABC-liposomes are comparable to each other, and if need be, be alterted before measuring the interaction. In addition, by first immobilising the liposome, the variability introduced when using liposomes as an analyte means the different sized liposomes causing variable change in RU. When the S_DmsA_YFP protein is used as the analyte this issue no longer remain as the substrate has a molecular weight of 28 kDa, therefore the change in RU observed to number of substrates bound.

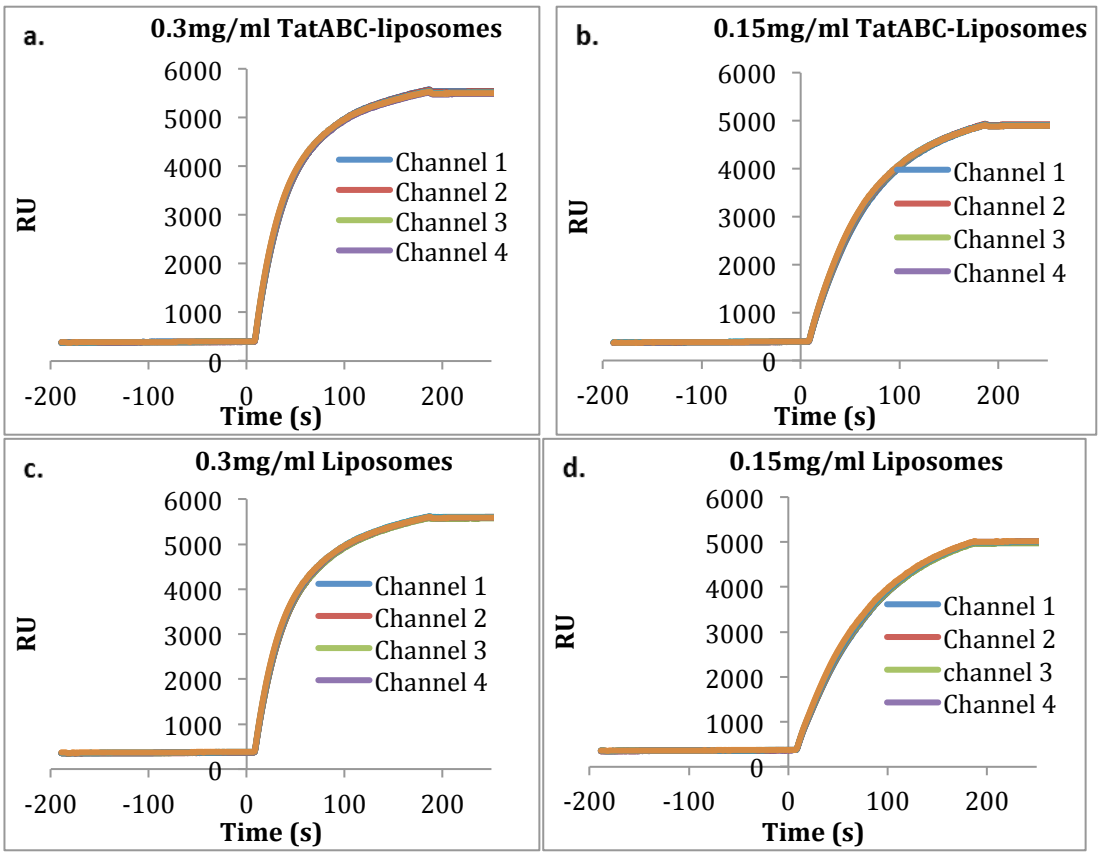


Figure 5.10: Immobilisation of the TatABC-liposome and empty-liposomes onto a GLC-chip

The empty-liposomes (labelled liposomes) and TatABC-liposomes were immobilised onto a GLC-chip via C11 interactions. The immobilisation was achieved at a flowrate of 50 µl/ml over 180 seconds. The liposomes were immobilised on 4 channels, all of which bound the liposomes similarly producing the same trend of association and capture. The RU of both types of liposomes at the concentration of 0.3 mg/ml and 0.15 mg/ml are similar and hence comparable concentration of liposomes have been immobilised to the chip.

The analytes S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP were injected transversely across the liposomes-immobilised channels at a flow rate of 80 µl/min for 300 seconds. The flow-rate was increased compared to that in section 5.6.2, to confirm the binding was not non-specific. With a low flow-rate during association, the chance of interaction also increases, resulting in additional noise from the non-specific binding. Therefore during this step the flow-rate was increased. The sensogram produced from this experiment was interpreted by averaging the graph for repeats of S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP at a fixed concentration of 12.5 mg/ml. The averaged results (Figure 5.11), shows the S_spDmsA^{RR}_YFP has increased interactions in comparison to the S_spDmsA^{KK}_YFP when bound to the TatABC-liposomes. Similary, the binding with with empty-liposomes shows reduced binding with S_spDmsA^{KK}_YFP. Analyses of the dissociation pattern (300-1200 seconds) for the binding of empty-liposomes show significant difference (Figure 5.11). Here, both the S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP substrates dissociate from the liposomes to a similar extent with some substrate that remaining bound to the chip. Whereas, S_spDmsA^{RR}_YFP does not dissociate as efficiently from the TatABC-liposomes. This is shown on Figure 5.11, where dissociation of the substrate reaches on average 400 RU for both S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP, when bound to protein-free liposomes. In comparison the dissociation of S_spDmsA^{RR}_YFP from the TatABC-liposomes remains at 450 RU, while S_spDmsA^{KK}_YFP dissociation reaches the 400 RU as seen with empty liposomes (Figure 5.11). This suggests that although there is no significant difference in the binding of the two substrates, the overall specific interaction that occurs in presence of TatABC_S is of a higher affinity, thus not as easily dissociated.

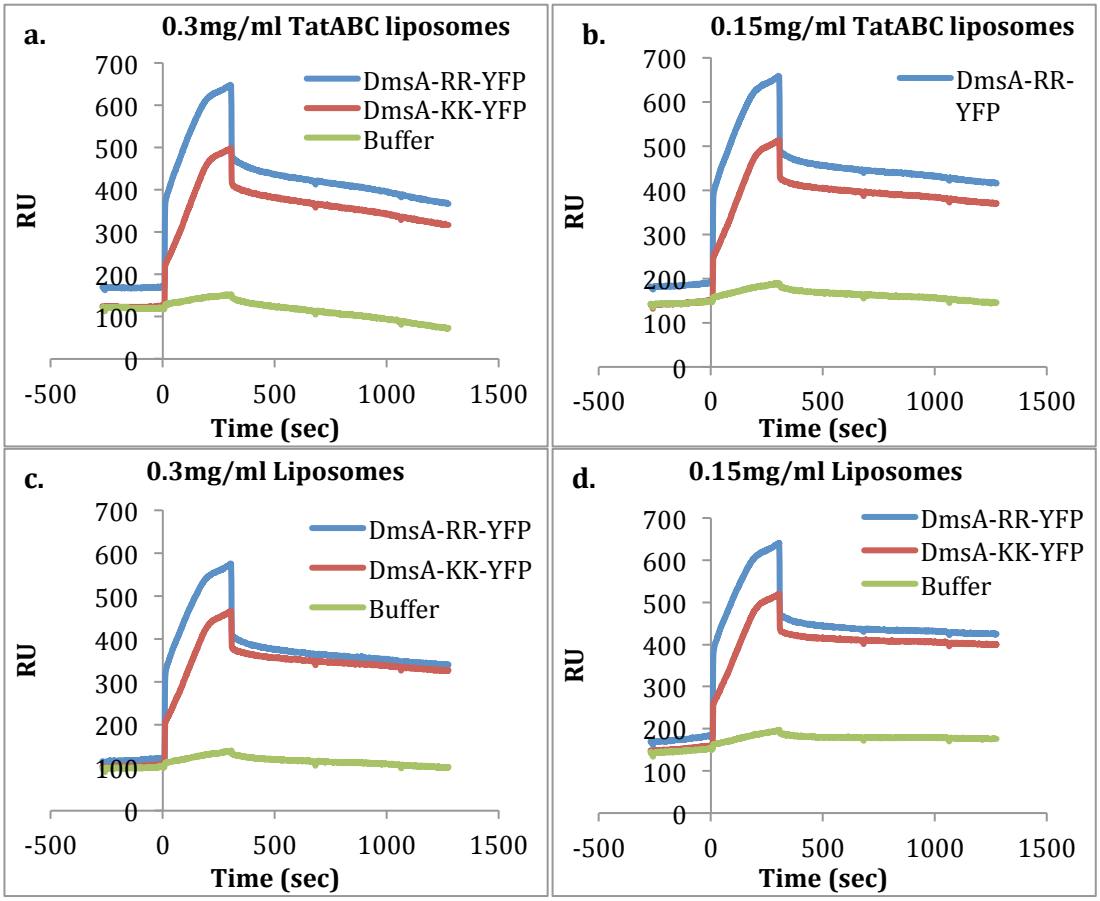


Figure 5.11: Averaged SPR sensogram of the S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP interaction to TatABC-liposomes and empty liposomes.

To study the interaction by SPR, 5uM of the DmsA-YFP substrate added to immobilised liposomes at 0.3 mg/ml and 0.15 mg/ml. The binding of the substrate S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP to immobilised TatABC-liposomes at 0.3 mg/ml (a) and 0.15 mg/ml (b) was measured. The binding to the empty-liposomes (empty liposomes) with 0.3 mg/ml (c) and 0.15 mg/ml (d) is shown below. The analyte was flowed over for 300 second and the dissociation was measured from 300-1250 seconds.

To further analyse the binding data, the non-specific binding produced observed from empty-liposomes was subtracted from the TatABC-liposomes on the averaged data set (Figure 5.12 a and b). The results in Figure 5.12 a and b shows the noise produced whilst buffer was passed over the liposomes. The noise from the buffer masks the interaction so the buffer was then subtracted for the corresponding channel to show only the change in interaction between S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP (Figure 5.12 c and d). Even after the subtraction of the empty-liposomes, the S_spDmsA^{RR}_YFP bound more abundantly to TatABC-liposomes than with S_spDmsA^{KK}_YFP. With 0.15 mg/ml of TatABC-liposomes, S_spDmsA^{RR}_YFP binding is 5RU higher than with S_spDmsA^{KK}_YFP. In addition, the change in concentration of the substrate was proportional to the change in RU, so when concentration is doubled, the RU is also doubled.

The subtraction of the buffer (Figure 5.12 c and d) also shows that there is actually little or no dissociation of the substrate itself and what previously looked like dissociation was an artefact of buffer switching (after the end of the analyte flow). Therefore the apparent change in mass seen at time of dissociation (300 seconds) is a result of change in buffer density from the analyte buffer (100mM Tris-HCl, 150mM NaCl, 1mM EDTA, 2.5mM desthiobiotin) to the running buffer (20mM Tris-HCl, 150mM NaCl).

The results also allowed evaluation of the two methods used in which immobilisation of the liposomes was in this case was shown to be the more efficient approach. Results also show that with this method, higher concentrations of liposomes can be used without leading to saturation of the GLC as previously observed in section 5.6. Also, by using substrates with lower mass, the concentration of substrate can be used with greater control, thereby preventing the true interaction from being masked by the signal noise due from changes in buffer environment.

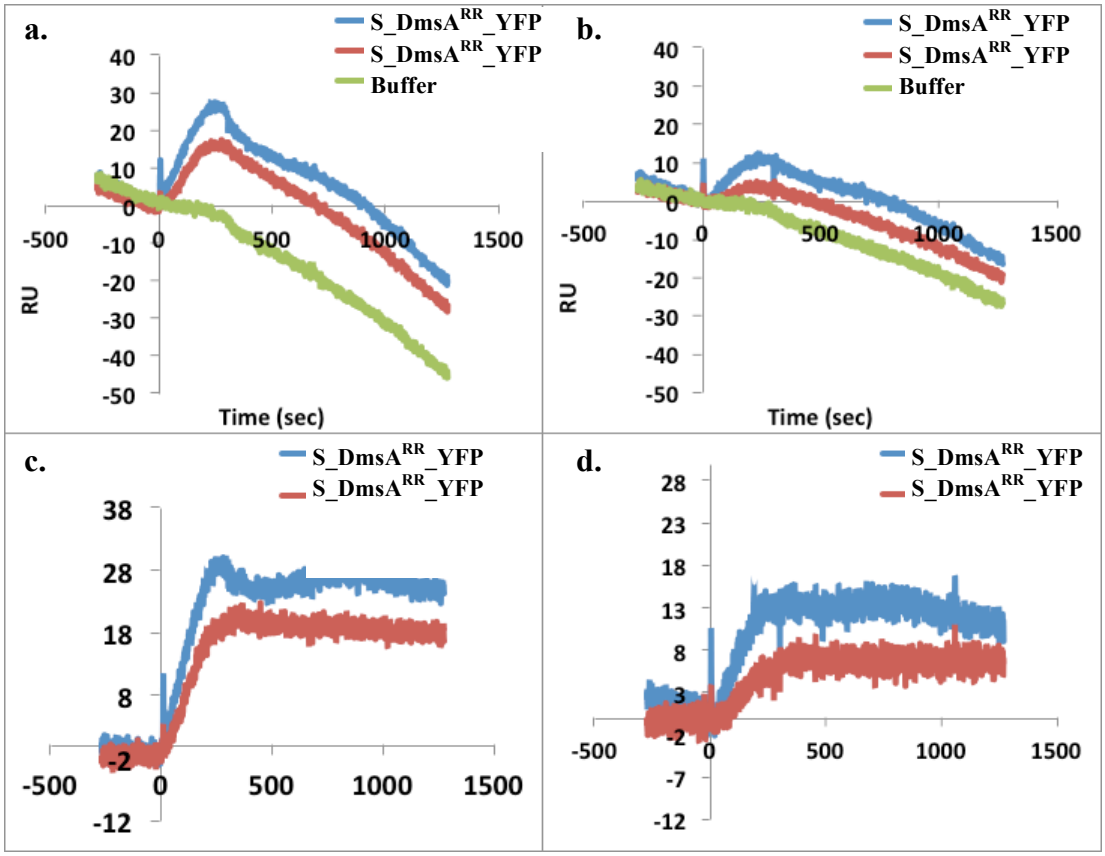


Figure 5.12: The specific binding between S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP with TatABC-liposomes

TatABC-liposomes were immobilised to the GLC-chip at 0.15 mg/ml or 0.3 mg/ml. The analytes S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP was passed across the surface at 5uM. The sensogram, produced in the experiment was further analysed by subtracting the empty-liposomes channel (negative control) to present only specific binding (a and b) with S_spDmsA^{RR}_YFP (blue), S_spDmsA^{KK}_YFP (red) and buffer (green). The background noise from the buffer channel was also subtracted to show difference in binding between S_spDmsA^{RR}_YFP (blue) and S_spDmsA^{KK}_YFP (red) (c and d).

5.8 Discussion

In this chapter, the signal peptide of DmsA was linked to the N-terminus of a YFP reporter protein to be used as Tat-targeted substrate, to investigate the binding specificity. The S_DmsA_YFP constructs was purified from the *E. coli* cytoplasmic fraction, while the TatABC complex was purified from the membrane fraction solubilised in 2% DDM. TatABC_S was then reconstituted into liposomes consisting of *E. coli* total lipids, mimicking the *in vivo* environment. Using inverted membrane vesicles can influence of non-specific binding from other native membrane proteins. Thereby using liposome reconstitution, the specificity for the TatABC_S complex can be studied directly and without influence from other proteins. The interaction between the TatABC_S complex and the S_DmsA_YFP constructs was studied by Surface Plasmon Resonance (SPR) spectroscopy. This interaction between the substrate and the TatABC complex was measured by two methods:

a) Antibody capture of DmsA-YFP
b) Lipophilic capture of the liposomes

DmsA targeting to the TatABC system can be divided into 3 steps; initial chaperone binding by DmsD, membrane targeting and recognition by TatABC system. Early study showed the DmsD protein binding directly to the signal peptide of DmsA (Winstone *et al.*, 2006). In this chapter, we have shown the DmsA can also bind the membrane in absence of the DmsD chaperone, using only purified S_DmsA_YFP. The DmsD chaperone was not included in the study, as it has been shown to reduce the membrane binding affinity of DmsA (Ray *et al.*, 2003). As the *in vitro* experiment did not include other cytoplasmic components of the Tat pathway, the membrane-targeting function must be within the within the signal peptide itself. Although cytoplasmic DmsD binds to the signal peptide of DmsA, it has no role in TatABC targeting (Ray *et al.*, 2003).

More recently, an SPR study has shown DmsD to interact with chimerically swapped regions of the DmsA protein, showing specific binding to the hydrophobic region on the C-terminus of the signal peptide (Fröbel *et al.*, 2012). There are no other cytosolic components involved in DmsA targeting to the Tat complexes, unlike

in the Sec pathway, where multiple chaperone and factors are observed. Therefore the Tat signal peptides must code for membrane targeting within signal peptide itself.

This chapter focuses on understanding the significance of the RR residues in the signal peptide during membrane targeting. There have been previous studies that show that substrates with the RR residues can be targeted to empty lipids. The binding of Tat substrate to a lipid bilayer in the absence of the Tat system was first observed in chloroplast (Hou *et al.*, 2006). Similar binding characteristics were observed in section 5.6.3, where both RR and KK substrates were bound to liposomes with and without the TatABC complex (Figure 5.9). This result supports the finding from a with a study showing TorA-GFP interacting with inverted vesicles from WT and ΔTat *E. coli* membranes (Shanmugham *et al.*, 2006). Although contradicting study has shown the *E. coli* membranes without TatABC, cannot be co-purified with the substrate (Panahandeh *et al.*, 2008). In addition, cross-linking studies suggests lipid binding is not an essential step in membrane targeting (Holzapfel *et al.*, 2007a). Current literature has two opposing views for membrane targeting. First, the membrane binding is an independent of specific targeting to the Tat complex (Hou *et al.*, 2006; Shanmugham *et al.*, 2006; Shanmugham *et al.*, 2012). The second view is that substrate requires the presence of TatABC for membrane binding (Panahandeh *et al.*, 2008; Holzapfel *et al.*, 2007a; Fröbel *et al.*, 2011). With these opposing views of the membrane targeting, the results in this chapter are inclined to the former theory of RR-independent membrane targeting.

Extensive research on the core TatC subunit has identified significant regions within the protein that are associated with a particular function. Previously, the signal peptides have been shown to be incapable cross-linking to TatC, when the twin arginine residues (RR) are mutated to twin lysines (KK), in both *E. coli* and chloroplast (Alami *et al.*, 2003; Gerard & Cline, 2006). This early evidence suggests the specificity is found within the TatC protein. More specifically, the TM1 and the first cytoplasmic loop (Cyt1) of TatC have been associated with substrate binding and recognition process (Holzapfel *et al.*, 2007b; Zoufaly *et al.*, 2012). Although the binding of the signal peptide to the TatC has been characterised, the specificity of the RR residues in membrane targeting, is not understood as clearly. This chapter has shown the specificity of the RR residues is not significant in membrane targeting.

The non-specific binding that was observed with empty-liposomes adds to the previously published data, suggesting membrane binding is the initial step of protein targeting. This membrane-binding step is independent of the presence of the TatABC system. The use of S_spDmsA^{KK}_YFP also shows that the proofreading mechanism is not part of the initial membrane binding stage as it can bind to the liposomes almost as efficiently as S_spDmsA^{RR}_YFP. There is a slight reduction observed in S_spDmsA^{RR}_YFP binding, however this can be attributed to the reversible membrane-binding model (Whitaker *et al.*, 2012). Whitaker *et al.* (2012) used FRET analysis to show the initial binding event is reversible. The initial lipid-binding step is thought to be reversible in chloroplasts although, once bound to the TatABC complex the interaction is irreversible. The S_spDmsA^{RR}_YFP substrate forms both, irreversible interactions with the TatABC-liposomes and reversible interactions with the lipid of the empty-liposomes, resulting in a higher binding RU (Figure 5.12). In contrast, the S_spDmsA^{KK}_YFP can only form reversible interaction with the lipid, so the RU does not reach as high as with S_spDmsA^{RR}_YFP and also dissociates more readily.

The dissociation phase shows irreversible interactions between S_spDmsA^{RR}_YFP and the TatABC-liposome. This was seen in Figure 5.11, where the dissociation phase of S_spDmsA^{RR}_YFP showed no dissociation in TatABC-liposomes. There are some irreversible interactions formed by the S_spDmsA^{KK}_YFP, which has no specificity for TatABC_S. In addition, the S_spDmsA^{KK}_YFP formed contacts with empty-liposomes, further suggesting the membrane-targeting step is an RR-independent event. The S_spDmsA^{RR}_YFP dissociates less frequently with TatABC-liposomes; however, the dissociation rate of S_spDmsA^{RR}_YFP with the empty-liposomes is similar to that observed with S_spDmsA^{KK}_YFP. This finding suggests that once the S_spDmsA^{RR}_YFP rate is membrane-bound. It can interact irreversibly with the TatABC complex, if present.

Membrane binding of S_spDmsA^{RR}_YFP has been shown to bind not only to the membrane, but also integrate into the lipid bilayer. This was shown by partial protection of the signal peptide from proteases degradation (Panahandeh *et al.*, 2008; Schlesier & Klosgen, 2010). This suggests that incorporation of the signal peptide into bilayer occur prior to any specific binding to the TatABC_S complex. In

addition, an early SPR study looked at the targeting of the signal peptide to the lipid membrane, showing membrane targeting was achieved using multiple GFP-labelled signal peptides of TorA, DmsA and OmpA. All substrates were found to bind lipid without in absence of Tat complex. This chapter showed interaction of the signal peptide with empty-liposomes, irrespective of the specificity of S_spDmsA^{RR}_YFP and S_spDmsA^{KK}_YFP, supporting the theory of non-specific membrane binding.

Overall the findings in this chapter and previously published work suggest a region of the signal peptide is responsible for the membrane targeting. Furthermore, it can also suggest the specificity of the RR residues is not involved in the initial membrane binding step. The hydrophobic region of the signal peptide is thought to form a loop within the membrane, which immobilises the peptide to the lipid bilayer. The incorporation into the bilayer gives the peptide protection against the protease. The signal peptide is then thought to associate with the TatC protein in membrane. Here the KK residues substitution shows the substrate is also bound to the empty-liposomes. This finding adds to the evidence that the RR residues are neither significant nor required in the early stage of membrane binding. In addition, it also suggests that substrates may be targeted to membranes and possibly inserted into the membrane without requiring the specificity of the RR residues in signal peptide.

Of the two methods used for SPR, liposome immobilisation using C11 proved to be most efficient. The use of the low mass substrates as the analyte that passes over the GLC-chip proved more beneficial, and prevents saturation of the chip. This was not the case for substrate immobilisation, where the liposomes were used as an analyte, and would easily saturate the chip. The saturation was a result high concentration, and the variability in size of the liposomes, which added ambiguity to the interactions that were made.

Despite this, it is evident from the SPR studies that the DmsA signal peptide is capable of binding the lipid bilayer irrespective of the presence of the TatABC complex. This is a result of the early membrane-targeting step, attributed to the hydrophobic region of the DmsA signal peptide, which helps the signal peptide incorporate into the membrane. There are also implications for the positive charged residues in the N-region of the signal peptide being involved in the membrane-

targeting step, where the translocation by Tat was shown to be dependent on anionic phospholipid (Mikhaleva *et al.*, 1999). Other experiments have also showed that increasing the salinity reduces membrane-binding interaction (Shanmugham *et al.*, 2006). This suggests electrostatic interactions between the anionic phospholipid and the charged residues of the signal peptide, are weakened by increasing salinity.

Frobel *et al.* (2012) recently evaluated the incorporation of the signal peptide into the lipid bilayer. This study suggests signal peptide insertion of the bilayer is a specific step, mediated by the TatC protein and regulated by TatB. However, opposing studies show the membrane proteins are reversibly inserted into the membrane before recognition by TatC (Panahandeh *et al.*, 2008). From the SPR study, it can be concluded that binding to the bilayer occur with out requiring the RR residues, but it is difficult to determine whether the signal peptide is incorporated into the membrane. In a recent study, TatC has been allocated the role of an "insertase" implying the signal peptide insertion is dependent on the interaction with TatC (Froebel *et al.*, 2011). However others argue that the signal peptide is already inserted into the bilayer during the targeting step and then laterally diffuses across the membrane to form interaction with TatC as shown in Figure 5.13 (Whitaker *et al.*, 2012).

a.

b.

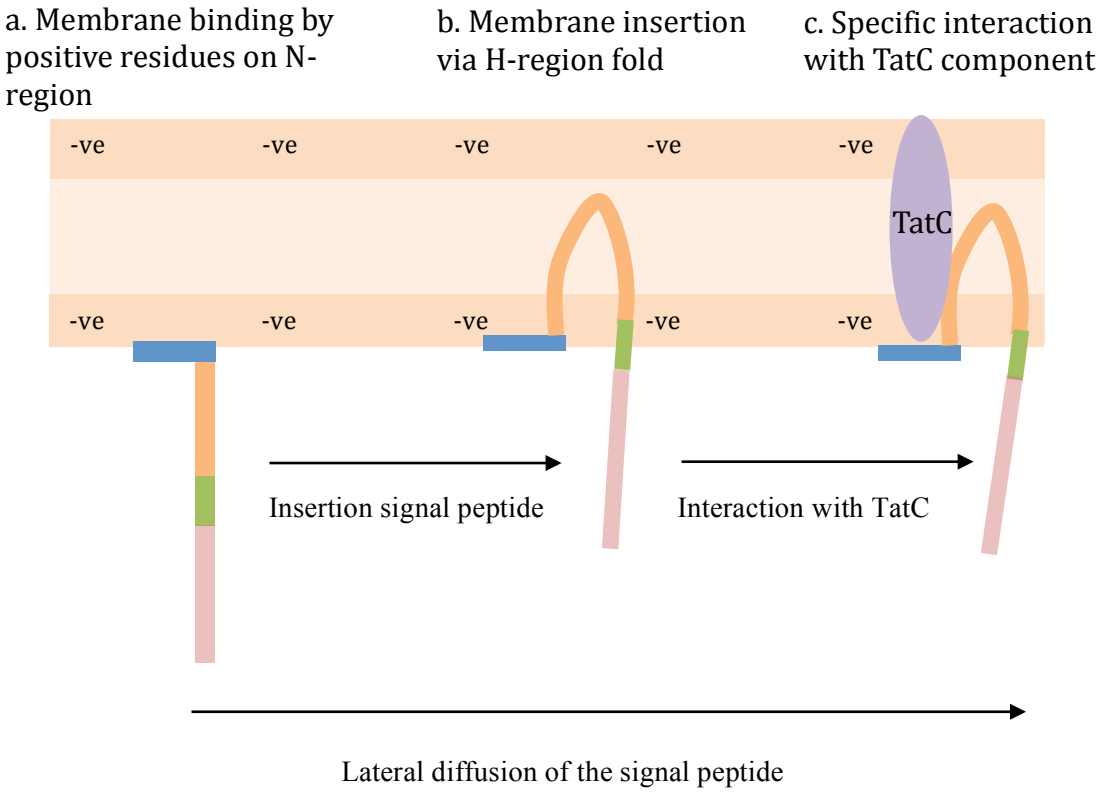


Figure 5.13: Schematic representation of a signal peptide and the step involved membrane targeting of signal peptide to the membrane

The Tat signal peptide has the RR motif in the N-terminus, with a hydrophobic centre and a C-region with the cleavage site (a). The hypothesised membrane targeting and binding step can be divided into 3 sub-steps (b). The initial targeting of the signal peptide to the membrane occurs without any cytosolic components. This is achieved by the electrostatic interaction between the positive residues on the signal peptide and the anionic phospholipids (step 1). The hydrophobic region then forms a loop in the membrane, allowing the signal peptide to be inserted into the membrane (step 2). The final stage is the RR motif specifically interacting with TatC complex.

Chapter 6

Final discussion

6 Final Discussion

The Twin Arginine Translocation (Tat) pathway is found in both chloroplast and bacteria, transporting folded proteins across the membrane. The proteins targeted to the Tat system have a conserved motif containing two highly conserved twin arginine residues. The Tat system is found to vary across Gram-positive and Gram-negative bacteria by the components involved.

The earlier work largely focused on the Tat system in the thylakoid and then progressed onto the TatABC system in the model organism *Escherichia coli* (*E. coli*) (Bogsch *et al.*, 1997; Berks, 1996; Cline, 1992). The lack of the TatB protein in the Gram-positive bacteria has shifted the attention to the model organism *Bacillus subtilis* (*B. subtilis*). In *B. subtilis*, there are two homologous Tat systems present, known as the TatAdCd and TatAyCy systems (Barnett *et al.*, 2008).

The TatAC system of *B. subtilis* varies from the well characterise TatABC system in *E. coli* by multiple factors. The major difference is the lack of variable-sized pore observed with TatA, which is not seen with TatAd, TatE or TatAy complexes (Gohlke *et al.*, 2005; van der Ploeg *et al.*, 2011; Baglieri *et al.*, 2012; Oates *et al.*, 2005). The bifunctional role of the TatAd to compensate for the lack of TatB, suggests a more evolved pathway is present in Gram-positive. Therefore understanding this pathway has now become significant. Especially as the biotechnological application of the Tat system is currently being exploited in *E. coli* and *B. subtilis* (Matos *et al.*, 2012)

The purification of the TatAdCd_S from a pBAD24-based vector yields in low concentration of protein, which is insufficient for further biophysical characterisation. To increase the yield, initial optimisation of the expression investigated by inducing in the stationary phase. The toxicity induced by expression of TatAdCd in *E. coli* is not overcomes by altering the concentration of arabinose (inducer) but can be increased by altering the time of expression as shown in chapter 3.

The use of detergent to isolate the Tat complexes introduces restriction to the type of biophysical studies that are feasible without the interference. To study the structure of these complexes by electron microscopy, the artefact from detergent must be minimised by using minimal detergent concentration. In purification of the TatAdCd _S complex, C12E9 detergent was used. The critical micelle concentration (CMC) of this detergent is very low to allow large-scale purification of TatAdCd complex. Varying the detergent concentration showed aggregation of the TatAdCd complexes if the detergent concentration shifted close to CMC or was 16X above CMC of C12E9 (chapter 3).

The current hypothesised model for the TatABC in *E. coli* shows a 3-step mechanism. Within a resting membrane, the TatABC pathway exists in two separate complexes, the Tat(A)BC substrate-binding complex and the TatA pore-forming complex. The first step is the recognition and binding of the substrate, which occurs via the conserved motif of signal peptide located on the N-terminus of the substrate. This then allows the recruitment of the TatA pore-forming complex in presence of PMF. Finally the protein is translocated through the pore.

This project further looked to structurally characterise the TatAdCd and TatAyCy complexes to understand its similarity to the TatABC system, and, its implication on the mechanism. The TatAdCd _S was expressed in *E. coli* to optimise the purification and allow the complexes to be studied by electron microscopy. There are no current publication showing the structural images of the TatAdCd or the TatAyCy complexes that are shown in chapter 3 and 4. Both the complexes have shown to resemble the TatABC complexes studied under TEM. The TatABC complexes were observed to be round complexes with high heterogeneity in size and the average size of the complexes reaching a maximum of 13nm (Oates *et al.*, 2003). Similarly TatAdCd was observed in the range of 7-13 nm and TatAyCy in the range 8-14 nm with the complexes sizes of 230 kDa and 370 kDa respectively.

The similarity in the size and also in the density profiling of the round complexes formed by TatAdCd_S and TatAyCy_H shows the protein assembling into an ordered round structure with a central cavity, which conserved among the 3 Tat systems. Therefore in a resting membrane, the Tat complex in Gram-positive bacteria

must resemble the TatABC. The lack of TatB does not cause any structural changes to the complex assembly at this resolution however further single particle analysis would give higher resolution information.

It appears that the recognition complexes in *B. subtilis* (TatAdCd and TatAyCy) are similar to that in *E. coli* Tat(A)BC in terms of structure. However, the supposed pore-forming TatAd complex was shown to differ from the TatA complex in multiple ways. First, is the heterogeneity in the size of TatA, which was not observed with either TatAd complexes using EM modelling or with TatAy complexes by BN-PAGE (van der Ploeg *et al.*, 2011; Beck *et al.*, 2013; Gohlke *et al.*, 2005). The variability in the pore-forming complexes TatA and TatAd or TatAy in *B subtilis* may be result of a slightly altered mechanism for substrate transport. This was suggested in finding where TatAd pore was not large enough to accommodate large substrate such as TorA (Beck *et al.*, 2013).

The similarity in the size of the initial TatAdCd and TatAyCy complexes to the TatABC complex suggests a similar mechanism may be used for signal peptide binding and recognition. However the mechanism of transport still remains ambiguous and currently there are two proposed models. These are either the pore-forming model or the recently hypothesised, membrane destabilising model.

In particular, this thesis focuses on understanding the bifunctional nature of TatAy protein. Mutational studies in the TatAd protein showed that mutations within the N-terminus region of the TatAd protein was essential when expressed in *E. coli* cells lacking TatB. The results imply the TatB protein function in TatAd, is located in the N-terminus region of protein (Barnett *et al.*, 2011). The sequence homology between TatAd and TatAy suggests the same is true for TatAy. In support of this, a mutation within this N-terminus of TatAy (P2A) was shown to block translocation activity (van der Ploeg *et al.*, 2011; Blaudeck *et al.*, 2005). Chapter 4 found the mutation within this region lead to accumulation of TatAy protein during the assembly of the TatAyCy complex.

When structurally analysed by EM, the excessive $TatAy^{P2A}$ was shown to form a helical fibril structure in addition to the round $TatAyCy^{WT}$ complexes observed. The

fibrils varied in length from 100-500 nm and the diameter of the fibril was 12nm suggesting it assembled from multiple $TatAy^{P2A}$ complexes. The co-purification of $TatAy^{P2A}$ with TatCy_H suggests some TatCy_H is also present in the fibril structures. Similar observation was made by Berthelmann *et al.* (2008a), showing the requirement for the presence of TatC in tubule formation by TatA. The $TatAy^{P2A}$ Cy complex is hypothesised to seed the formation of the fibril by excessively recruiting $TatAy^{P2A}$ when present in excess.

As the mutation is within the N-terminus region of TatAy, the TatB functional region, it infers TatB may function in the regulation of TatA recruitment once the substrate-TatBC complex is formed. To date, the role of TatB is not clearly understood and the finding from chapter 4 not only shows the region involved in TatA or TatAy/TatAd recruitment but also the function of TatB in *E. coli* likely to be involved in the regulation of the TatA recruitment step.

In a follow up study, the membrane-targeting step of the signal peptide to the TatABC complex studied. The TatAdCd was not the first choice to study the substrate binding interaction using SPR because the DmsA is an the native substrate of *E. coli* (Palmer *et al.*, 2005). TatAdCd can to recognise and transport DmsA-YFP substrate but initial the TatABC system was studied (Mendel *et al.*, 2008). Also, as the TatABC and TatAdCd complexes are similar in size and structure (chapter 3), the preliminary work was carried out on *E. coli* before moving ahead with the TatAdCd complexes in *B. subtilis*.

Membrane binding is one of the first step of signal peptide targeting however here the SPR study (chapter 5) has shown this event of membrane binding, is a non-specific targeting step. This agrees with the earlier binding of signal peptides interacting with lipid in absence of the Tat machinery. In some cases, substrate interaction with the lipid is partially protected from protease, implying the signal peptide might also be incorporating into the lipid independently as well.

The use of DmsA with KK substitution at R16 and R17, showed the targeting with this substrate did not alter the binding to lipid, again emphasising the Tat-independent binding property of the signal peptide. The result here are controversial

to the recent finding of TatC functioning as an "insertase" during membrane targeting, suggesting TatABC is required for this step. There is evidence for both contradictory hypothesises (Shanmugham *et al.*, 2006; Frobel *et al.*, 2012).

The SPR result presented in chapter 5, which shows the interaction between the lipid and the substrate, is Tat-independent. This SPR study does not show whether lipid insertion is actually occurring. Although the direct specificity of TatABC could not be measured using TatABC-liposomes, the specificity was detected during the dissociation phase of the SPR sensogram. The RR containing substrate (S-DmsA^{RR}_YFP) would dissociate less readily from the TatABC-liposomes. Again this would suggest specific binding is the next stage of membrane binding, once the signal peptide laterally moves across the phospholipid. Hence, my interpretation of binding occurrence sides with the transient membrane binding/insertion theory leading to a lateral diffusion of the signal peptide to the TatBC complex where is specifically bound upon recognition (Figure 5.13). It has not been possible to identify the region of the signal peptide responsible for the non-specific targeting to the membrane, however once bound the hydrophobic region in the signal peptide is thought to loop into the membrane (Shanmugham *et al.*, 2006; Maurer *et al.*, 2010; Frobel *et al.*, 2012) .The most likely model would be the positively charged residues on the N-region of the signal peptide are used for the membrane-targeting step

The work presented in this thesis has provided insight into the structure of the TatAdCd and TatAyCy complex. That is supported by the detailed analysis of the bifunctional TatAy mutations to interpret the TatA recruitment role of TatB protein in *E. coli*. Detailed analysis into the structural assembly of multiple mutants in TatAy or TatCy could help further understand the role TatB. Finally, in attempt to investigate the specific targeting, the results presented have shown a non-specific targeting to the membrane is also present in the TatABC pathway as seen with the Sec pathway. In continuation of the SPR study further mutational work need to be carried out to identify the region involved in the initial membrane step.

References

7 References

Akiyama, Y. & Ito, K. (1987) Topology analysis of the SecY protein, an integral membrane protein involved in protein export in Escherichia coli. *The EMBO Journal*, 6 (11): 3465.

Alami, M., Lüke, I., Deitermann, S., Eisner, G., Koch, H.-G., Brunner, J. & Müller, M. (2003) Differential Interactions between a Twin-Arginine Signal Peptide and Its Translocase in Escherichia coli. *Molecular Cell*, 12 (4): 937-946.

Alcock, F., Baker, M. A. B., Greene, N. P., Palmer, T., Wallace, M. I. & Berks, B. C. (2013) Live cell imaging shows reversible assembly of the TatA component of the twin-arginine protein transport system. *Proceedings of the National Academy of Sciences*, 110 (38): E3650-E3659.

Allen, S. C. H., Barrett, C. M. L., Ray, N. & Robinson, C. (2002) Essential cytoplasmic domains in the Escherichia coli TatC protein. *Journal of Biological Chemistry*, 277 (12): 10362.

Bageshwar, U. K. & Musser, S. M. (2007) Two electrical potential–dependent steps are required for transport by the Escherichia coli Tat machinery. *The Journal of cell biology*, 179 (1): 87-99.

Baglieri, J., Beck, D., Vasisht, N., Smith, C. J. & Robinson, C. (2012) Structure of TatA Paralog, TatE, Suggests a Structurally Homogeneous Form of Tat Protein Translocase That Transports Folded Proteins of Differing Diameter. *Journal of Biological Chemistry*, 287 (10): 7335-7344.

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 (5): 2534-2542.

Barnett, J. P., Lawrence, J., Mendel, S. & Robinson, C. (2011) Expression of the bifunctional Bacillus subtilis TatAd protein in Escherichia coli reveals distinct TatA/B-family and TatB-specific domains. ***Arch Microbiol*****, 193 (8): 583-594.**

Barnett, J. P., van der Ploeg, R., Eijlander, R. T., Nenninger, A., Mendel, S., Rozeboom, R., Kuipers, O. P., van Dijl, J. M. & Robinson, C. (2009) The twin-arginine translocation (Tat) systems from Bacillus subtilis display a conserved mode of complex organization and similar substrate recognition requirements. ***FEBS Journal*****, 276 (1): 232-243.**

Barrett, C. M. L. & Robinson, C. (2005) Evidence for interactions between domains of TatA and TatB from mutagenesis of the TatABC subunits of the twin arginine translocase. ***FEBS Journal*****, 272 (9): 2261-2275.**

Beck, D., Vasisht, N., Baglieri, J., Monteferrante, C. G., van Dijl, J. M., Robinson, C. & Smith, C. J. (2013) Ultrastructural characterisation of Bacillus subtilis TatA complexes suggests they are too small to form homooligomeric translocation pores. ***Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*****, 1833 (8): 1811-1819.**

Behrendt, J., Standar, K., Lindenstrauß, U. & Brüser, T. (2004) Topological studies on the twin-arginine translocase component TatC. ***FEMS microbiology letters*****, 234 (2): 303-308.**

Berks, B. (1996) A common export pathway for proteins binding complex redox cofactors? ***Molecular microbiology*****, 22 (3): 393-404.**

Berks, B. C., Sargent, F. & Palmer, T. (2000) The Tat protein export pathway. ***Molecular microbiology*****, 35 (2): 260-274.**

Berthelmann, F., Mehner, D., Richter, S., Lindenstrauss, U., Lunsdorf, H., Hause, G. & Bruser, T. (2008a) Recombinant Expression of tatABC and tatAC Results in the Formation of Interacting Cytoplasmic TatA Tubes in Escherichia coli. ***J. Biol. Chem.*****, 283 (37): 25281-25289.**

Berthelmann, F., Mehner, D., Richter, S., Lindenstrauss, U., Lünsdorf, H., Hause, G. & Brüser, T. (2008b) Recombinant expression of tatABC and tatAC results in the formation of interacting cytoplasmic TatA tubes in Escherichia coli. *Journal of Biological Chemistry*, 283 (37): 25281-25289.

Blaudeck, N., Kreutzenbeck, P., Müller, 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. *Journal of Biological Chemistry*, 280 (5): 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. *The EMBO Journal*, 16 (13): 3851-3859.

Bolhuis, A., Mathers, J. E., Thomas, J. D., Barrett, C. M. L. & Robinson, C. (2001) TatB and TatC form a functional and structural unit of the twin-arginine translocase from Escherichia coli. *Journal of Biological Chemistry*, 276 (23): 20213.

Bowman, L., Palmer, T. & Sargent, F. (2013) A regulatory domain controls the transport activity of a twin-arginine signal peptide. *FEBS Letters*, 587 (20): 3365-3370.

Boy, D. & Koch, H. G. (2009) Visualization of distinct entities of the SecYEG translocon during translocation and integration of bacterial proteins. *Mol Biol Cell*, 20 (6): 1804-1815.

Brüser, T. & Sanders, C. (2003) An alternative model of the twin arginine translocation system. *Microbiological research*, 158 (1): 7-17.

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. *Molecular microbiology*, 43 (6): 1457-1470.

Buchanan, G., Maillard, J., Nabuurs, S. B., Richardson, D. J., Palmer, T. & Sargent, F. (2008) Features of a twin-arginine signal peptide required for recognition by a Tat proofreading chaperone. *FEBS Letters*, 582 (29): 3979-3984.

Casadaban, M. J. & Cohen, S. N. (1980) Analysis of gene control signals by DNA fusion and cloning in< i> Escherichia coli</i>. *Journal of molecular biology*, 138 (2): 179-207.

Chaddock, A. M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R. G., Klösgen, 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. *The EMBO Journal*, 14 (12): 2715.

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 (25): 7396-7404.

Cline, K., Ettinger, W.F., Theg, S.M. (1992) Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *Journal of Biological Chemistry*, 267 (4): 2688-2696.

Cline, K. & Mori, H. (2001) Thylakoid ΔpH-dependent precursor proteins bind to a cpTatC–Hcf106 complex before Tha4-dependent transport. *The Journal of cell biology*, 154 (4): 719-730.

Creighton, A. M., Hulford, A., Mant, A., Robinson, D. & Robinson, C. (1995) A monomeric, tightly folded stromal intermediate on the pH-dependent thylakoidal protein transport pathway. *Journal of Biological Chemistry*, 270 (4): 1663-1669.

Cristóbal, S., de Gier, J. W., Nielsen, H. & von Heijne, G. (1999) Competition between Sec-and TAT-dependent protein translocation in Escherichia coli. *The EMBO Journal*, 18 (11): 2982-2990.

Dabney-Smith, C., Mori, H. & Cline, K. (2006) Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. *Journal of Biological Chemistry*, 281 (9): 5476-5483.

De Keersmaeker, S., Van Mellaert, L., Lammertyn, E., Vrancken, K., Anné, J. & Geukens, N. (2005) Functional analysis of TatA and TatB in Streptomyces lividans. *Biochemical and Biophysical Research Communications*, 335 (3): 973-982.

de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S. B., Müller, M., Sargent, F., Palmer, T. & Berks, B. C. (2002) Oligomeric properties and signal peptide binding by Escherichia coli Tat protein transport complexes. *Journal of molecular biology*, 322 (5): 1135-1146.

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 Letters*, 506 (2): 143-148.

DeLisa, M. P., Lee, P., Palmer, T. & Georgiou, G. (2004) Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway. *Journal of bacteriology*, 186 (2): 366-373.

DeLisa, M. P., Samuelson, P., Palmer, T. & Georgiou, G. (2002) Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *Journal of Biological Chemistry*, 277 (33): 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. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (10): 6115.

Deville, K., Gold, V. A., Robson, A., Whitehouse, S., Sessions, R. B., Baldwin, S. A., Radford, S. E. & Collinson, I. (2011) The oligomeric state and arrangement of the active bacterial translocon. *J Biol Chem*, 286 (6): 4659-4669.

Dilks, K., Giménez, M. I. & Pohlschröder, M. (2005) Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic archaea. *Journal of bacteriology*, 187 (23): 8104-8113.

Dilks, K., Rose, R., Hartmann, E. & Pohlschroder, M. (2003) Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *Journal of bacteriology*, 185 (4): 1478.

Eijlander, R. T., Jongbloed, J. D. H. & Kuipers, O. P. (2009) Relaxed specificity of the Bacillus subtilis TatAdCd translocase in Tat-dependent protein secretion. *Journal of bacteriology*, 191 (1): 196-202.

Eisner, G., Koch, H. G., Beck, K., Brunner, J. & Muller, M. (2003) Ligand crowding at a nascent signal sequence. *J Cell Biol*, 163 (1): 35-44.

Fekkes, P., de Wit, J. G., van der Wolk, J. P., Kimsey, H. H., Kumamoto, C. A. & Driessen, A. J. (1998) Preprotein transfer to the Escherichia coli translocase requires the co-operative binding of SecB and the signal sequence to SecA. *Mol Microbiol*, 29 (5): 1179-1190.

GenericFrobel, J., Rose, P., Lausberg, F., Blummel, A.-S., Freudl, R. & Muller, M. (2012) Transmembrane insertion of twin-arginine signal peptides is driven by TatC and regulated TatB.

Fröbel, J., Rose, P. & Müller, M. (2011) Early Contacts between Substrate Proteins and TatA Translocase Component in Twin-arginine Translocation. *Journal of Biological Chemistry*, 286 (51): 43679-43689.

Fröbel, J., Rose, P. & Müller, M. (2012) Twin-arginine-dependent translocation of folded proteins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367 (1592): 1029-1046.

Froebel, J., Rose, P. & Mueller, M. (2011) Twin-arginine translocation: early contacts between substrate proteins and TatA. *Journal of Biological Chemistry*,

Genest, O., Seduk, F., Ilbert, M., Méjean, V. & Iobbi-Nivol, C. (2006) Signal peptide protection by specific chaperone. *Biochemical and Biophysical Research Communications*, 339 (3): 991-995.

Gerard, F. & Cline, K. (2006) Efficient Twin Arginine Translocation (Tat) Pathway Transport of a Precursor Protein Covalently Anchored to Its Initial cpTatC Binding Site. *Journal of Biological Chemistry*, 281 (10): 6130-6135.

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. *Proceedings of the National Academy of Sciences*, 102 (30): 10482.

Goosens, V. J., Monteferrante, C.G., van Dijl, J.M. (2013) The Tat system of Gram-positive bacteria. *Biochimica et Biophysica Acta (BBA)*,

Gouffi, K., Gérard, F., Santini, C. L. & Wu, L. F. (2004) Dual topology of the Escherichia coli TatA protein. *Journal of Biological Chemistry*, 279 (12): 11608.

Hartl, F.-U., Lecker, S., Schiebel, E., Hendrick, J. P. & Wickner, W. (1990) The binding cascade of SecB to SecA to SecYE mediates preprotein targeting to the E. coli plasma membrane. *Cell*, 63 (2): 269-279.

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 Letters*, 539 (1-3): 61-67.

Higy, M., Gander, S. & Spiess, M. (2005) Probing the environment of signal-anchor sequences during topogenesis in the endoplasmic reticulum. *Biochemistry*, 44 (6): 2039-2047.

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 Letters*****, 497 (1): 45-49.**

Holland, I. B. (2010) The extraordinary diversity of bacterial protein secretion mechanisms. In: eds. ***Protein secretion*****. Springer: 1-20.**

Holtkamp, W., Lee, S., Bornemann, T., Senyushkina, T., Rodnina, M. V. & Wintermeyer, W. (2012) Dynamic switch of the signal recognition particle from scanning to targeting. ***Nat Struct Mol Biol*****, 19 (12): 1332-1337.**

Holzapfel, E., Eisner, G., Alami, M., Barrett, C. M., Buchanan, G., Luke, I., Betton, J. M., Robinson, C., Palmer, T., Moser, M. & Muller, M. (2007a) The entire N-terminal half of TatC is involved in twin-arginine precursor binding. ***Biochemistry*****, 46 (10): 2892-2898.**

Holzapfel, E., Eisner, G., Alami, M., Barrett, C. M. L., Buchanan, G., Luke, I., Betton, J. M., Robinson, C., Palmer, T., Moser, M. & Muller, M. (2007b) The Entire N-Terminal Half of TatC is Involved in Twin-Arginine Precursor Binding. ***Biochemistry*****, 46 (10): 2892-2898.**

Hou, B., Frielingsdorf, S. & Klosgen, R. B. (2006) Unassisted membrane insertion as the initial step in DeltapH/Tat-dependent protein transport. ***J Mol Biol*****, 355 (5): 957-967.**

Huber, D., Rajagopalan, N., Preissler, S., Rocco, M. A., Merz, F., Kramer, G. & Bukau, B. (2011) SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. ***Mol Cell*****, 41 (3): 343-353.**

Ignatova, Z., Hörnle, C., Nurk, A. & Kasche, V. (2002) Unusual Signal Peptide Directs Penicillin Amidase from ***Escherichia coli*** **to the Tat Translocation Machinery.** ***Biochemical and Biophysical Research Communications*****, 291 (1): 146-149.**

Ize, B., Gérard, F. & Wu, L.-F. (2002) In vivo assessment of the Tat signal peptide specificity in Escherichia coli. ***Archives of microbiology*****, 178 (6): 548-553.**

Ize, B., Stanley, N. R., Buchanan, G. & Palmer, T. (2003) Role of the Escherichia coli Tat pathway in outer membrane integrity. ***Molecular microbiology*****, 48 (5): 1183-1193.**

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. ***Journal of bacteriology*****, 183 (5): 1801.**

Jongbloed, J. D. H., Antelmann, H., Hecker, M., Nijland, R., Bron, S., Airaksinen, U., Pries, F., Quax, W. J., van Dijl, J. M. & Braun, P. G. (2002) Selective contribution of the twin-arginine translocation pathway to protein secretion in Bacillus subtilis. ***Journal of Biological Chemistry*****, 277 (46): 44068-44078.**

Jongbloed, J. D. H., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S. & Van Dijl, J. M. (2004) Two minimal Tat translocases in Bacillus. ***Molecular microbiology*****, 54 (5): 1319-1325.**

Jongbloed, J. D. H., Martin, U., Antelmann, H., Hecker, M., Tjalsma, H., Venema, G., Bron, S., van Dijl, J. M. & Muller, J. (2000) TatC Is a Specificity Determinant for Protein Secretion via the Twin-arginine Translocation Pathway. ***Journal of Biological Chemistry*****, 275 (52): 41350-41357.**

Klösgen, 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 Molecular Biology*****, 18 (5): 1031-1034.**

Koch, H. G. & Muller, M. (2000) Dissecting the translocase and integrase functions of the Escherichia coli SecYEG translocon. ***J Cell Biol*****, 150 (3): 689-694.**

Kudva, R., Denks, K., Kuhn, P., Vogt, A., Müller, M. & Koch, H.-G. (2013) Protein translocation across the inner membrane of Gram-negative bacteria: The Sec and Tat dependent protein transport pathways. ***Research in microbiology*****,**

Kuhn, P., Weiche, B., Sturm, L., Sommer, E., Drepper, F., Warscheid, B., Sourjik, V. & Koch, H. G. (2011) The bacterial SRP receptor, SecA and the ribosome use overlapping binding sites on the SecY translocon. ***Traffic*****, 12 (5): 563-578.**

Lakshmipathy, S. K., Tomic, S., Kaiser, C. M., Chang, H. C., Genevaux, P., Georgopoulos, C., Barral, J. M., Johnson, A. E., Hartl, F. U. & Etchells, S. A. (2007) Identification of nascent chain interaction sites on trigger factor. ***J Biol Chem*****, 282 (16): 12186-12193.**

Lange, C., Müller, S. D., Walther, T. H., Bürck, J. & Ulrich, A. S. (2007) Structure analysis of the protein translocating channel TatA in membranes using a multi-construct approach. ***Biochimica et Biophysica Acta (BBA)-Biomembranes*****, 1768 (10): 2627-2634.**

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. ***Proceedings of the National Academy of Sciences*****, 105 (40): 15376.**

Lo, S. M. & Theg, S. M. (2012) Role of Vesicle-Inducing Protein in Plastids 1 in cpTat transport at the thylakoid. ***The Plant Journal*****, 71 (4): 656-668.**

Luirink, J., von Heijne, G., Houben, E. & de Gier, J. W. (2005) Biogenesis of inner membrane proteins in Escherichia coli. ***Annu Rev Microbiol*****, 59 329-355.**

Mangels, D., Mathers, J., Bolhuis, A. & Robinson, C. (2005) The Core TatABC Complex of the Twin-arginine Translocase in ***Escherichia coli*****: TatC Drives Assembly Whereas TatA is Essential for Stability.** ***Journal of molecular biology*****, 345 (2): 415-423.**

Matos, C., 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 reports*, 10 (5): 474-479.

Matos, C. F. R. O., Branston, S. D., Albiniak, A., Dhanoya, A., Freedman, R. B., Keshavarz-Moore, E. & Robinson, C. (2012) High-yield export of a native heterologous protein to the periplasm by the tat translocation pathway in Escherichia coli. *Biotechnology and bioengineering*, 109 (10): 2533-2542.

Maurer, C., Panahandeh, S., Jungkamp, A.-C., Moser, M. & Müller, M. (2010) TatB functions as an oligomeric binding site for folded Tat precursor proteins. *Molecular biology of the cell*, 21 (23): 4151-4161.

McDevitt, C. A., Buchanan, G., Sargent, F., Palmer, T. & Berks, B. C. (2006) Subunit composition and in vivo substrate binding characteristics of Escherichia coli Tat protein complexes expressed at native levels. *FEBS Journal*, 273 (24): 5656-5668.

Mehner, D., Osadnik, H., Lünsdorf, H. & Brüser, T. (2012) The Tat system for membrane translocation of folded proteins recruits the membrane-stabilizing Psp machinery in Escherichia coli. *Journal of Biological Chemistry*, 287 (33): 27834-27842.

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. *Journal of molecular biology*, 375 (3): 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 (3): 331-335.

Mitra, K. & Frank, J. (2006) A model for co-translational translocation: ribosome-regulated nascent polypeptide translocation at the protein-conducting channel. *FEBS Lett*, 580 (14): 3353-3360.

Monteferrante, C. C., et al. (2012) TatAc, the third TatA subunit of Bacillus subtilis, can form active twin-arginine translocases with the TatCd and TatCy subunits. *Appl Environ Microbiol.*, 78 (14): 4999-5001.

Monteferrante, C. G., Baglieri, J., Robinson, C. & van Dijl, J. M. (2012) TatAc, the third TatA subunit of Bacillus subtilis, can form active twin-arginine translocases with the TatCd and TatCy subunits. *Appl Environ Microbiol*, 78 (14): 4999-5001.

Mori, H. & Cline, K. (2002) A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid ΔpH/Tat translocase. *The Journal of cell biology*, 157 (2): 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. *Journal of Biological Chemistry*, 266 (19): 12189.

Oates, J., Barrett, C. M. L., 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. *Journal of molecular biology*, 346 (1): 295-305.

Oates, J., Mathers, J., Mangels, D., Kühlbrandt, W., Robinson, C. & Model, K. (2003) Consensus structural features of purified bacterial TatABC complexes. *Journal of molecular biology*, 330 (2): 277-286.

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 Letters*, 581 (21): 4091-4097.

Pal, D., Fite, K. & Dabney-Smith, C. (2013) Direct interaction between a precursor mature domain and transport component Tha4 during twin arginine transport of chloroplasts. *Plant physiology*, 161 (2): 990-1001.

Palmer, T., Sargent, F. & Berks, B. C. (2005) Export of complex cofactor-containing proteins by the bacterial Tat pathway. *TRENDS in Microbiology*, 13 (4): 175-180.

Panahandeh, S., Maurer, C., Moser, M., DeLisa, M. P. & Muller, M. (2008) Following the path of a twin-arginine precursor along the TatABC translocase of Escherichia coli. *J. Biol. Chem.*, M804225200.

Pop, O., Martin, U., Abel, C. & Muller, J. P. (2002) The Twin-arginine Signal Peptide of PhoD and the TatAd/Cd Proteins of Bacillus subtilis Form an Autonomous Tat Translocation System. *Journal of Biological Chemistry*, 277 (5): 3268-3273.

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

Pugsley, A. P., Francetic, O., Driessen, A. J. & de Lorenzo, V. (2004) Getting out: protein traffic in prokaryotes. *Molecular microbiology*, 52 (1): 3-11.

Ramasamy, S., Abrol, R., Suloway, C. J. M. & Clemons Jr, W. M. (2013) The Glove-like Structure of the Conserved Membrane Protein TatC Provides Insight into Signal Sequence Recognition in Twin-Arginine Translocation. *Structure*,

Ray, N., Oates, J., Turner, R. J. & Robinson, C. (2003) DmsD is required for the biogenesis of DMSO reductase in Escherichia coli but not for the interaction of the DmsA signal peptide with the Tat apparatus. *FEBS Lett*, 534 (1-3): 156-160.

Richter, S. & Brüser, T. (2005) Targeting of unfolded PhoA to the TAT translocon of Escherichia coli. *Journal of Biological Chemistry*, 280 (52): 42723.

Robinson, C., Matos, C. F. R. O., Beck, D., Ren, C., Lawrence, J., Vasisht, N. & Mendel, S. (2011) Transport and proofreading of proteins by the twin-arginine translocation (Tat) system in bacteria. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1808 (3): 876-884.

Rodrigue, A., Chanal, A., Beck, K., Müller, M. & Wu, L. F. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial Tat pathway. *Journal of Biological Chemistry*, 274 (19): 13223.

Rodriguez, F., Rouse, S. L., Tait, C. E., Harmer, J., De Riso, A., Timmel, C. R., Sansom, M. S. P., Berks, B. C. & Schnell, J. R. (2013) Structural model for the protein-translocating element of the twin-arginine transport system. *Proceedings of the National Academy of Sciences*, 110 (12): E1092-E1101.

Rollauer, S. E., Tarry, M. J., Graham, J. E., Jääskeläinen, M., Jäger, F., Johnson, S., Krehenbrink, M., Liu, S.-M., Lukey, M. J. & Marcoux, J. (2012) Structure of the TatC core of the twin-arginine protein transport system. *Nature*, 492 (7428): 210-214.

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. *Molecular microbiology*, 45 (4): 943-950.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular cloning*. Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY.

Sargent, F., Berks, B. C. & Palmer, T. (2002) Assembly of membrane-bound respiratory complexes by the Tat protein-transport system. *Archives of microbiology*, 178 (2): 77-84.

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. *The EMBO Journal*, 17 (13): 3640-3650.

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. *European Journal of Biochemistry*, 268 (12): 3361-3367.

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. *Journal of Biological Chemistry*, 274 (51): 36073-36082.

Satoh, Y., Matsumoto, G., Mori, H. & Ito, K. (2003) Nearest neighbor analysis of the SecYEG complex. 1. Identification of a SecY-SecG interface. *Biochemistry*, 42 (24): 7434-7441.

Schiebel, E., Driessen, A. J. M., Hartl, F.-U. & Wickner, W. (1991) $\Delta\mu_H$+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell*, 64 (5): 927-939.

Schlesier, R. & Klosgen, R. B. (2010) Twin arginine translocation (Tat)-dependent protein transport: the passenger protein participates in the initial membrane binding step. *Biol Chem*, 391 (12): 1411-1417.

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 (5342): 1467.

Shanmugham, A., Bakayan, A., Voller, P., Grosveld, J., Lill, H. & Bollen, Y. J. (2012) The hydrophobic core of twin-arginine signal sequences orchestrates specific binding to Tat-pathway related chaperones. *PLoS One*, 7 (3): e34159.

Shanmugham, A., Wong Fong Sang, H. W., Bollen, Y. J. & Lill, H. (2006) Membrane binding of twin arginine preproteins as an early step in translocation. *Biochemistry*, 45 (7): 2243-2249.

Shimohata, N., Akiyama, Y. & Ito, K. (2005) Peculiar properties of DsbA in its export across the Escherichia coli cytoplasmic membrane. *Journal of bacteriology*, 187 (12): 3997-4004.

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. *Journal of Biological Chemistry*, 275 (16): 11591-11596.

Tani, K., Shiozuka, K., Tokuda, H. & Mizushima, S. (1989) In vitro analysis of the process of translocation of OmpA across the Escherichia coli cytoplasmic membrane. A translocation intermediate accumulates transiently in the absence of the proton motive force. *J Biol Chem*, 264 (31): 18582-18588.

Tarry, M. J., Schäfer, E., Chen, S., Buchanan, G., Greene, N. P., Lea, S. M., Palmer, T., Saibil, H. R. & Berks, B. C. (2009) Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. *Proceedings of the National Academy of Sciences*, 106 (32): 13284.

Tjalsma, H., Bolhuis, A., Jongbloed, J. D. H., Bron, S. & van Dijl, J. M. (2000) Signal peptide-dependent protein transport inBacillus subtilis: a genome-based survey of the secretome. *Microbiology and Molecular Biology Reviews*, 64 (3): 515-547.

Tjalsma, H., Zanen, G., Bron, S. & Van Dijl, J. M. (2002) 1 The eubacterial lipoprotein-specific (type II) signal peptidase. In: Rose, E. D. & David, S. S. eds. *The Enzymes*. Academic Press: 3-26.

Tomkiewicz, D., Nouwen, N., van Leeuwen, R., Tans, S. & Driessen, A. J. M. (2006) SecA supports a constant rate of preprotein translocation. *Journal of Biological Chemistry*, 281 (23): 15709-15713.

Tuteja, R. (2005) Type I signal peptidase: An overview. *Archives of Biochemistry and Biophysics*, 441 (2): 107-111.

van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. & Rapoport, T. A. (2004) X-ray structure of a protein-conducting channel. ***Nature*****, 427 (6969): 36-44.**

van der Ploeg, R., Barnett, J. P., Vasisht, N., Goosens, V. J., Pöther, D. C., Robinson, C. & van Dijl, J. M. (2011) Salt Sensitivity of Minimal Twin Arginine Translocases. ***Journal of Biological Chemistry*****, 286 (51): 43759-43770.**

Veenendaal, A. K., van der Does, C. & Driessen, A. J. (2001) Mapping the sites of interaction between SecY and SecE by cysteine scanning mutagenesis. ***J Biol Chem*****, 276 (35): 32559-32566.**

Walther, T. H., Gottselig, C., Grage, S. L., Wolf, M., Vargiu, A. V., Klein, M. J., Vollmer, S., Prock, S., Hartmann, M. & Afonin, S. (2013) Folding and Self-Assembly of the TatA Translocation Pore Based on a Charge Zipper Mechanism. ***Cell*****, 152 (1): 316-326.**

Walther, T. H., Grage, S. L., Roth, N. & Ulrich, A. S. (2010) Membrane alignment of the pore-forming component TatAd of the twin-arginine translocase from Bacillus subtilis resolved by solid-state NMR spectroscopy. ***Journal of the American Chemical Society*****, 132 (45): 15945-15956.**

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 (1): 93-101.**

Westermann, M., Pop, O. I., Gerlach, R., Appel, T. R., Schlörmann, W., Schreiber, S. & Müller, J. P. (2006) The TatAd component of the Bacillus subtilis twin-arginine protein transport system forms homo-multimeric complexes in its cytosolic and membrane embedded localisation. ***Biochimica et Biophysica Acta (BBA) - Biomembranes*****, 1758 (4): 443-451.**

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. ***Journal of Biological Chemistry*****, 275 (22): 16717.**

Whitaker, N., Bageshwar, U. K. & Musser, S. M. (2012) Kinetics of precursor interactions with the bacterial Tat translocase detected by real-time FRET. *Journal of Biological Chemistry*, 287 (14): 11252-11260.

White, G. F., Schermann, S. M., Bradley, J., Roberts, A., Greene, N. P., Berks, B. C. & Thomson, A. J. (2010) Subunit Organization in the TatA Complex of the Twin Arginine Protein Translocase. *Journal of Biological Chemistry*, 285 (4): 2294.

Whitehouse, S., Gold, V. A., Robson, A., Allen, W. J., Sessions, R. B. & Collinson, I. (2012) Mobility of the SecA 2-helix-finger is not essential for polypeptide translocation via the SecYEG complex. *J Cell Biol*, 199 (6): 919-929.

Winstone, T. L., Workentine, M. L., Sarfo, K. J., Binding, A. J., Haslam, B. D. & Turner, R. J. (2006) Physical nature of signal peptide binding to DmsD. *Arch Biochem Biophys*, 455 (1): 89-97.

Yen, M. R., Tseng, Y. H., Nguyen, E. H., Wu, L. F. & Saier, M. H. (2002) Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Archives of microbiology*, 177 (6): 441-450.

Zimmer, J., Nam, Y. & Rapoport, T. A. (2008) Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature*, 455 (7215): 936-943.

Zoufaly, S., Fröbel, J., Rose, P., Flecken, T., Maurer, C., Moser, M. & Müller, M. (2012) Mapping Precursor-binding Site on TatC Subunit of Twin Arginine-specific Protein Translocase by Site-specific Photo Cross-linking. *Journal of Biological Chemistry*, 287 (16): 13430-13441.